

Diogo Roberto Olsen

Geração de Modelos Digitais Interativos em Alta Definição de Objetos do Mundo Real

Curitiba - PR, Brasil

2020

Diogo Roberto Olsen

Geração de Modelos Digitais Interativos em Alta Definição de Objetos do Mundo Real

Tese apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de doutor em Informática.

Pontifícia Universidade Católica do Paraná - PUCPR

Programa de Pós-Graduação em Informática - PPGIa

Orientador: Edson Emílio Scalabrin

Coorientador: Flávio Bortolozzi

Curitiba - PR, Brasil

2020

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Sônia Maria Magalhães da Silva - CRB-9/1191

O52g
2020
Olsen, Diogo Roberto
Geração de modelos digitais interativos em alta definição de objetos do mundo real / Diogo Roberto Olsen ; orientador, Edson Emílio Scalabrin ; coorientador, Flávio Bortolozzi. – 2020.
[244] f. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Paraná, Curitiba, 2020.
Bibliografia: f. [199-206]

1. Sistemas de varredura. 2. Imagens digitais. 3. Fotografia – Técnica. 4. Informática. I. Scalabrin, Edson Emílio. II. Bortolozzi, Flávio. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título.

CDD. 20. ed. – 004

DECLARAÇÃO

Declaro para os devidos fins que o aluno **DIOGO ROBERTO OLSEN**, defendeu sua tese de doutorado intitulada “**GERAÇÃO DE MODELOS DIGITAIS INTERATIVOS EM ALTA DEFINIÇÃO DE OBJETOS DO MUNDO REAL**”, na área de concentração Ciência da Computação, no dia 29 de junho de 2020, no qual foi aprovado.

Declaro ainda que foram feitas todas as alterações solicitadas pela Banca Examinadora, cumprindo todas as normas de formatação definidas pelo Programa.

Por ser verdade, firmo a presente declaração.

Curitiba, 01 de julho de 2020.



Prof. Dr. Emerson Cabrera Paraiso
Coordenador do Programa de Pós-Graduação em Informática
Pontifícia Universidade Católica do Paraná

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

O Instituto Federal do Paraná viabilizou este trabalho através do afastamento para cursar pós-graduação *stricto sensu*.

Agradeço à minha esposa, minha mãe e meu pai pelo apoio nos momentos de dificuldade. Sem a ajuda de vocês este trabalho não seria concluído.

Aos meus orientadores, coorientadores e professores que me ofereceram um norte durante estes quatro anos: Edson José Rodrigues Justino, Flávio Bortolozzi, Edson Emilio Scalabrin e Manoel de Campos Almeida.

Um agradecimento especial a Flávio de Almeida e Silva, com quem trilhei parte deste caminho.

Aos que contribuíram com este trabalho: Lucas Murbach Pierin, Renato Delbianco, Rafael Cobo.

Aos que contribuíram de forma direta ou indireta, me oferecendo suporte quando necessário.

Resumo

Existem duas formas de digitalização de objetos reais, a fotografia bidimensional, que apresenta alta acuidade visual mas não permite a troca de perspectiva; desenho ou escaneamentos que geram modelos tridimensionais, os quais permitem a troca de perspectivas mas não conseguem uma boa acuidade visual - principalmente para objetos complexos, como os orgânicos. Dentre as técnicas de escaneamento, com ou sem contato, a acuidade visual não é alcançada por problemas como a oclusão. Este trabalho apresenta a proposta de um método de digitalização de objetos reais de forma que seja possível gerar uma visualização interativa — possibilidade de manipular a peça, similar ao processo de manipulação do objeto real, observando-o de diversos ângulos — e com acuidade visual — de forma que os mesmos elementos encontrados na peça real possam ser vistos no modelo adquirido. A proposta baseia-se em adquirir um conjunto de fotografias de todas as perspectivas de um objeto, segundo um intervalo pre-determinado. A acuidade visual é conseguida através do uso adequado de técnicas fotografias, da correção de problemas intrínsecos do processo de fotografia digital e do aprimoramento das imagens adquiridas. O método também prevê a acuidade na aquisição das cores de forma a garantir fidelidade com as cores do objeto real. A interatividade é alcançada pela navegação entre as perspectivas fotografadas, com base na interação do usuário. Adicionalmente a proposta permite aplicar alterações de rotação, translação e escala na visualização. São apresentadas equações que mapeiam a posição da câmera com relação ao objeto durante a etapa de aquisição, considerando as coordenadas internas do scanner. Também é apresentado um conjunto de equações para converter estas coordenadas em outros sistemas de coordenadas, de forma que seja possível identificar qual fotografia melhor representa o ponto de vista de um observador em um sistema de simulação baseado em um ambiente de computação gráfica tridimensional. Assim, foram desenvolvidos dois softwares, um que controla o processo de aquisição fotográfica e outro que processa estas imagens para construir o modelo usado na visualização interativa e com acuidade visual. O modelo gerado apresentou alta acuidade visual, proporcionada por fotografias em 8K, técnicas de fotografia adequadas, processamento de imagens para corrigir erros intrínsecos de fotografia e aprimoramento destas imagens. A interatividade do modelo permite a visualização do objeto virtual com um intervalo angular, na horizontal e na vertical, de 1° entre perspectivas, totalizando até 32.760 fotografias (ou perspectivas) por objeto.

Palavras-chave: Escaneamento, visualização, interatividade, acuidade visual.

Abstract

There are two ways of digitizing real objects, two-dimensional photography, which has high visual accuracy but does not allow the exchange of perspective; drawing or scans that generate three-dimensional models, which allow the exchange of perspectives but do not achieve a good visual accuracy - mainly for complex objects, such as organic ones. Among the scanning techniques, with or without contact, visual accuracy is not achieved by problems such as occlusion. The following paper presents the proposition of a method of real objects digitalization, in a way that it is possible to generate an interactive visualisation — possibility to manipulate the piece, similar to the processes of manipulating the real object, observing it in various angles — and with visual accuracy — in a way that the same elements found in the real piece can be seen in the model acquired. The proposition is based in the acquisition of a set of photos from all perspectives of an object, according to a pre-determined interval. The visual accuracy is obtained by the use of adequate photography techniques; correction of the intrinsic problems of the digital photography; improvement of the images acquired; and the use of metric qualities of those images. The method also foresees the accuracy in the acquisition of the colors in a way that guarantees the precision of the real object's colors. Interactivity is achieved by navigating between the photographed perspectives, based on user interaction. Additionally, the proposition allows to apply movements of rotation, translation and scale in visualization. Equations are presented that map the camera's position in relation to the object during the acquisition step, considering the internal coordinates of the scanner. A set of equations to convert these coordinates into other coordinate systems is also presented, so that it is possible to identify which photograph best represents an observer's point of view in a simulation system based on a three-dimensional computer graphics environment. In this way, two software were developed, one that controls the photographic acquisition process and the other that processes these images to build the model used in interactive visualization. The model generated showed high visual accuracy, provided by 8K photographs, appropriate photography techniques, image processing to correct intrinsic photography errors and enhancement of these images. The interactivity of the model allows the visualization of the virtual object with an angular range, horizontally and vertically, of 1° between perspectives, totaling up to 32,760 photographs (or perspectives) per object.

Keywords: Scanning, visualization, interactivity, visual accuracy.

Lista de ilustrações

1	Interatividade: navegação.	29
2	Propriedades da visualização interativa.	29
3	Processo de digitalização de objetos reais.	36
4	Métodos de aquisição de dados de objetos reais.	37
5	Digitalização de Objetos Reais - Métodos que não geram acuidade visual.	38
6	Digitalização de Objetos Reais - Métodos ópticos.	39
7	Zonas de oclusões nos <i>scanners</i> laser.	40
8	Nitidez - Acutância.	43
9	Nitidez - Resolução.	43
10	Nitidez - Acutância e resolução.	44
11	Correção das cores.	47
12	<i>Color Checker</i>	48
13	Versões 1 e 2 do scanner.	51
14	Versão 3 do scanner.	52
15	Crânios adquiridos na segunda versão do F2S2.	52
16	Scanner Semiesférico de Campo Completo.	55
17	Aquisição do F2S2 com as perspectivas adquiridas e posicionadas na onde foram fotografadas.	56
18	Definição dos eixos do F2S2.	57
19	Sistema de movimentação linear.	58
20	Sistema de movimentação da câmera.	59
21	Sistema de movimentação angular.	59
22	Hardware - Componentes Fotográficos.	63
23	Imagem adquirida com fundo infinito.	63
24	Processo de digitalização de objetos pelo F2S2.	67
25	Semiesfera criada durante o processo de digitalização do objeto.	68
26	Diagrama das etapas de processamento.	69
27	Processo de redução de contexto.	70
28	Processo de segmentação.	71
29	Deslocamento da câmera após instruções GMC.	75
30	Sistema de Coordenadas F2S2.	76
31	Ângulos complementares e colaterais.	77
32	Sistema esférico de coordenadas - com γ definido.	79
33	Conversões entre os sistemas de coordenadas GMC, F2S2 e Esférico.	80

34	Nomes das fotografias adquiridas.	82
35	Criação do Stream2D - Organização das imagens em Streams e Frames.	83
36	Criação do Stream2D - Arquivos organizados e renomeados para visualização.	84
37	Representação gráfica do Stream2D.	86
38	Frame <i>S003F002</i>	87
39	Stream2D com 5 Streams e 20 Frames.	89
40	Sistemas desenvolvidos para visualização do Stream2D.	91
41	Enriquecimento da visualização - Evidenciando partes do objeto escaneado.	92
42	Visualização Interativa - Transformações Geométricas.	93
43	Visualização Interativa - Navegação.	93
44	Alteração entre Frames durante a visualização.	94
45	Visualização estereoscópica - Deslocamento à direita.	96
46	Visualização estereoscópica - Deslocamento à esquerda.	97
47	Visualização estereoscópica - Orientação centralizada.	98
48	Sistema de coordenadas cartesianas.	99
49	Diferenças de posição entre objeto e observador.	101
50	Posição do objeto e observador.	104
51	Conversões entre os sistemas de coordenadas.	104
52	Nível laser instalado para calibração.	106
53	Calibração do eixo X'	107
54	Calibração do eixo Y'	107
55	Distância do eixo X' ao B^a	108
56	Distância mínima do eixo Y' ao fundo infinito.	109
57	Eixo B^a - Deslocamento angular.	110
58	Eixo Z^a - Planejamento.	111
59	Eixo Z^a - Deslocamento angular.	112
60	Aquisição de inseto com raio de 25mm.	114
61	Imprecisão na aquisição de inseto com raio de 25mm.	115
62	Processo de Aquisição.	119
63	Profundidade de Campo.	120
64	Comportamento da profundidade de campo durante a aquisição - Variação γ	121
65	Comportamento da profundidade de campo durante a aquisição - Variação Θ	122
66	Ajuste de foco e profundidade de campo para aquisições do F2S2.	123
67	Foto de inseto com lente 18-55mm.	124
68	Fotos de objetos menores que 1cm.	125
69	Fotos de gafanhoto.	126
70	Efeito da imprecisão do movimento em fotografias.	128
71	Crânio humano.	129
72	Concha.	130

73	Ocorrência de brilho em determinados Frames.	132
74	Reflexo da cor do fundo.	135
75	Objeto causando sombra no fundo infinito.	136
76	Fluxograma do processo de alinhamento.	137
77	Alinhamento com laser de peça anatômica.	138
78	Processo de alinhamento a laser.	139
79	Fita adesiva aparente na fotografia.	140
80	Pata de cavalo.	141
81	Processo de fotometria com luz refletida.	144
82	Processo de fotometria para o F2S2.	145
83	Objetos mistos - luz adequada para partes opacas e reflexivas.	148
84	Problemas no enquadramento.	150
85	Colorchecker.	153
86	F2S2-App.	157
87	Tela de aquisição do F2S2-App.	157
88	Diagrama de atividade do F2S2-App.	158
89	Possíveis colisões.	161
90	F2S2-App - Full Scan.	162
91	F2S2-App - Partial Scan.	162
92	F2S2-App - Single Scan.	163
93	Representação das posições angulares dos Frames para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ; \gamma^0 = 0^\circ; \gamma^p = 90^\circ$	165
94	Representação das posições angulares dos Streams para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ; \gamma^0 = 0^\circ; \gamma^p = 90^\circ$	167
95	Tela inicial do <i>F2S2-BipApp</i>	173
96	Diagrama das etapas de processamento.	175
97	Software ColorChecker Passport.	176
98	Fotografia com correção de cores.	177
99	Distorção causada pelo conjunto óptico.	178
100	Correção da distorção, vinheta e aberração cromática usando a tabela de correções oferecida pela Samsung no software DarkTable.	178
101	Aberração cromática causada pelo conjunto óptico.	179
102	Vinheta causada pelo conjunto óptico.	179
103	<i>F2S2-BipApp</i> - Redução de contexto.	182
104	Processo de segmentação.	182
105	Processo de Segmentação.	183
106	Seleção interativa de sementes.	185
107	Dificuldades na segmentação - transparência e brilho.	186
108	Dificuldades na segmentação - nitidez.	187

109	Suavização de borda.	188
110	Remoção de pixels esparços.	189
111	Remoção de borda.	190
112	Comparação entre modelo gerado pelo F2S2 e reconstrução 3D.	194
113	Comparação entre modelo gerado pelo F2S2 e modelo 3D desenhado.	195

Lista de tabelas

1	Relação de coordenadas de uma aquisição de objeto de 2cm.	115
2	Valores angulares de imprecisão do eixo Z^a	116
3	Valores angulares de correção da imprecisão do eixo Z^a	117
4	Relação de Frames e suas posições para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$	165
5	Coordenadas GMC para uma aquisição de $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$	166
6	Valores arredondados para o eixo Z^a	172

Lista de quadros

1	Quadro comparativo entre sistemas de coordenadas.	105
2	Quadro de valores de entrada, checagens e cálculos realizados para cada modo de aquisição do <i>F2S2-App</i>	163
3	Quadro comparativo entre o Stream2D e modelos 3D baseados em nuvens e malhas de pontos.	196

Lista de códigos

3.1	Exemplo de GMCodes.	61
5.1	Exemplo de configurações salvas em JSON.	164
5.2	Geração das instruções GMC.	167
5.3	Códigos GMC gerados.	170

Lista de abreviaturas e siglas

3D	Tridimensional
4K	3.840×2.160
8K	7.680×4.320
CIE	International Commission on Illumination
CIIM	Centro de Inovação em Imagens Médicas
CMC	Color Measurement Committee
CMOS	Complementary Metal Oxide Semiconductor
CMM	Coordinate Measuring Machine
CNC	Computer Numeric Control
EV	Exposure value
F2S2	Full Frames Semi-spherical Scanner
GB	Gigabyte
HD	Hard Disk
ISO	International Standards Organization
MB	Megabyte
OA	Objetos de Aprendizagem
RGB	Red Green Blue
sRGB	standard RGB
SSD	Solid-State Drive
SVM	Support Vector Machine
UHD	Ultra High Definition

Lista de símbolos

\in	Pertence
\sphericalangle	Ângulo esférico
//	Indica paralelismo
%	Resto da divisão
f	Abertura do diafragma
\forall	Para todo
\mathbb{N}	Conjunto dos números naturais
\mathbb{R}	Conjunto dos números reais
β	Intervalo
Δ	Varição
Δ/Θ	Varição angular entre dois frames para gerar estereoscopia
γ	Ângulo vertical
γ^0	Ângulo vertical inicial
γ^p	Ângulo vertical final
Σ	Somatório
Θ	Ângulo horizontal
Θ^0	Ângulo horizontal inicial
Θ^p	Ângulo horizontal final
Θ^E	Ângulo horizontal referente ao olho esquerdo
Θ^D	Ângulo horizontal referente ao olho direito
φ	Ângulo polar ou zenital
$ArcoCos()$	Função Arco Cosseno
$ArcoSen()$	Função Arco Seno

$ArcoTg()$	Função Arco Tangente
B	Eixo B - Sistema de coordenadas F2S2
B^a	Eixo B^a - Sistema de coordenadas GMC
$Cosseno()$	Valor da função cosseno
$Frame_{id}$	Valor de Identificação do Frame no Stream
H	Eixo H - Sistema de coordenadas F2S2
r	raio de esfera ou círculo
$Seno()$	Valor da função seno
$Stream_{id}$	Valor de identificação do Stream no Stream2D
$Tg()$	Valor da função Tangente
V	Eixo V - Sistema de coordenadas F2S2
X	Eixo X - Sistema de coordenadas cartesianas
X'	Eixo X' - Sistema de coordenadas GMC
Y	Eixo Y - Sistema de coordenadas cartesianas
Y'	Eixo X' - Sistema de coordenadas GMC
Z	Eixo Z - Sistema de coordenadas cartesianas
Z^a	Eixo Z^a - Sistema de coordenadas GMC

Sumário

1	INTRODUÇÃO	27
1.1	Tema	28
1.2	Problema	30
1.3	Hipótese	31
1.4	Objetivos	31
1.4.1	Objetivo Geral	31
1.4.2	Objetivos Específicos	31
1.5	Justificativa	32
1.6	Contribuições e Originalidade	32
1.7	Estrutura da Tese	34
2	CONCEITOS, PREMISSAS E REVISÃO DA LITERATURA	35
2.1	Digitalização de Objetos	35
2.1.1	Visualização interativa a partir de fotografia digital	40
2.1.2	Fotografia com nitidez e acurácia na aquisição de cores	42
2.2	Histórico do projeto	49
3	METODOLOGIA E INFRAESTRUTURA PARA O DESENVOLVIMENTO DA TESE	55
3.1	<i>Full Frames Semi-spherical Scanner</i> - F2S2	55
3.2	Método proposto para o scanner F2S2	66
4	PROPOSTA	73
4.1	Sistemas de Coordenadas, Stream2D e Simulação	73
4.1.1	Sistema de Coordenadas	74
4.1.2	Modelo tridimensional - Stream2D	81
4.1.3	Sistemas de Simulação	89
4.2	Hardware do scanner F2S2	105
4.2.1	Calibragem do scanner F2S2	106
4.2.2	Manutenção, dificuldades e imprecisões enfrentadas	112
4.3	Protocolo de Aquisição	118
4.3.1	Objeto a ser escaneado	119
4.3.2	Configuração do scanner F2S2 e aquisição	133
5	RESULTADOS	155
5.1	Software de Aquisição - F2S2-App	155

5.1.1	Modos de operação	160
5.1.2	Geração das instruções GMC	164
5.2	Software de Processamento - F2S2-BipApp	172
5.2.1	Etapas	175
5.3	Validação	191
5.4	Análise	193
6	CONSIDERAÇÕES FINAIS	197
	REFERÊNCIAS	199
	APÊNDICES	207
	APÊNDICE A – ARTIGO - SEGMENTAÇÃO EM LOTE DE IMAGENS COM FUNDO DE COR ÚNICA	209
	APÊNDICE B – FOTOGRAFIA DIGITAL	223

1 Introdução

Observação de obras de arte, de peças anatômicas, de itens botânicos, além do processo de inspeção de desgaste de componentes mecânicos são apenas alguns exemplos de análises visuais de peças reais que são realizadas cotidianamente em pesquisas científicas, na educação, no sistema produtivo, ou com outros inúmeros fins.

Estes exames visuais são meios para resolução de problemas que vão desde a simples escolha de um consumidor sobre qual produto irá comprar, passando pela educação, diagnóstico de pragas em lavouras, até pesquisas científicas para a cura de doenças humanas, entre outras aplicações.

Porém, nem sempre é possível, ou mesmo interessante, que o observador esteja em contato direto com item a ser estudado. Por exemplo um estudante de nível médio não precisa ser exposto aos riscos de entrar em um laboratório de anatomia humana, onde o estudante estaria em contato com produtos químicos insalubres, para, por exemplo, estudar anatomia do coração humano.

Usando o mesmo exemplo de estudos anatômicos, quando existe a necessidade de que os estudantes entrem em contato com os corpos, como por exemplos nos cursos superiores de medicina, odontologia e enfermagem, estes enfrentam problemas por falta de corpos humanos para estudos de anatomia(LIMA, 2012). A falta de corpos para estudo pode resultar em erros médicos ou impedir pesquisas científicas(BALMANT, 2015).

Pesquisas tem sido feitas para obtenção de modelos visuais anatômicos precisos, como o *Visible Human Project*(ACKERMAN, 1998), no qual dois corpos humanos foram seccionados com intervalos de $1mm$ e $0,33mm$ e as secções foram fotografadas, em ultra alta resolução¹, para gerar bases de imagens de todo o corpo humano. Estas imagens possuem alta acuidade visual, ou seja, permitem a visualização de detalhes com precisão.

Em seu artigo, Ackerman descreve que existem duas formas de representações de imagens: baseadas em fotos e modelos tridimensionais criados por desenhistas. Os modelos tridimensionais sofrem com imprecisões tanto do desenho artístico, quanto do formato de representação da computação gráfica, que se baseia em formas geométricas e fórmulas matemáticas para rerepresentar volumes e estruturas. Porém estas não representam fielmente formas biológicas, que são irregulares por natureza.

Quanto aos conteúdos baseados em fotografias Ackerman afirma que um conjunto de imagens bidimensionais estáticas são limitadas por não oferecer a possibilidade de trocar a perspectiva de visualização da fotografia, como se tem ao manipular manualmente a peça real.

¹ Padrão criado pela NHK <www.nhk.or.jp>.

Ainda, segundo o autor, o entendimento de estruturas biológicas tridimensionais é indispensável para estudos anatômicos, mas seu aprendizado é um desafio. Fotografias são bidimensionais por natureza e forçam o espectador a fazer um exercício mental de construção da visualização tridimensional o que pode acarretar em imprecisões.

A dificuldade de entendimento de estruturas tridimensionais está presente em diversas outras áreas do conhecimento como, por exemplo: artes, na qual a visualização de diferentes perspectivas de uma escultura não é possível através de fotografias estáticas; na biologia em que uma visualização precisa é necessária para identificação de espécies; e engenharias na qual o entendimento de peças e suas relações é fundamental para a mecânica.

1.1 Tema

Este trabalho trata do processo de digitalização de objetos reais, para viabilizar a geração de uma visualização com interatividade - possibilidade de manipular a peça, similar ao processo de manipulação do objeto real, observando-o de diversos ângulos - e com acuidade visual - de forma que os mesmos elementos encontrados na peça real possam ser vistos no modelo adquirido. Além disso, o modelo deve permitir a geração de visualizações estereoscópicas, passando ao espectador a sensação de profundidade.

Interatividade

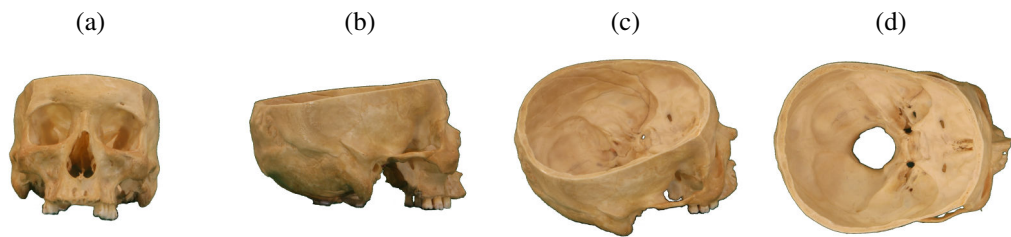
A visualização pretendida deve permitir interatividade com o objeto digitalizado. Neste trabalho a interatividade busca simular a visualização de um objeto real por uma pessoa, de forma que é possível visualizar o objeto por qualquer ângulo (exceto a parte de baixo do objeto), girando o objeto de forma que seja possível enxergar o ângulo desejado. Além disso, é possível aproximar ou afastar o objeto dos olhos para perceber mais ou menos detalhes, também é possível inclinar o objeto para enxergar perspectivas diferentes ou mover o objeto para observá-lo em outras posições.

Desta forma, para prover a interatividade durante a visualização do modelo gerado, esta pode ser dividida em duas formas principais: navegação e transformações geométricas (rotação, translação e escala). Assim é possível manipular a visualização com as características básicas de um modelo tridimensional (GHOSH, 2005).

A navegação permite mudar o ponto de vista para visualizar qualquer ângulo do objeto. Na Figura 1 mostra-se uma sequência com: 1(a) posição frontal; 1(b) navegação para a direita; 1(c) navegação em um ângulo para cima; e, por fim, 1(d) navegação para uma visão superior.

As transformações geométricas não mudam o ponto de vista relativo ao objeto. Estas são aplicadas sobre um determinado ponto de vista. A translação é a capacidade de movimentar a imagem na tela, por exemplo, deslocando o objeto do centro para um dos cantos. A rotação é a

Figura 1 – Interatividade: navegação.

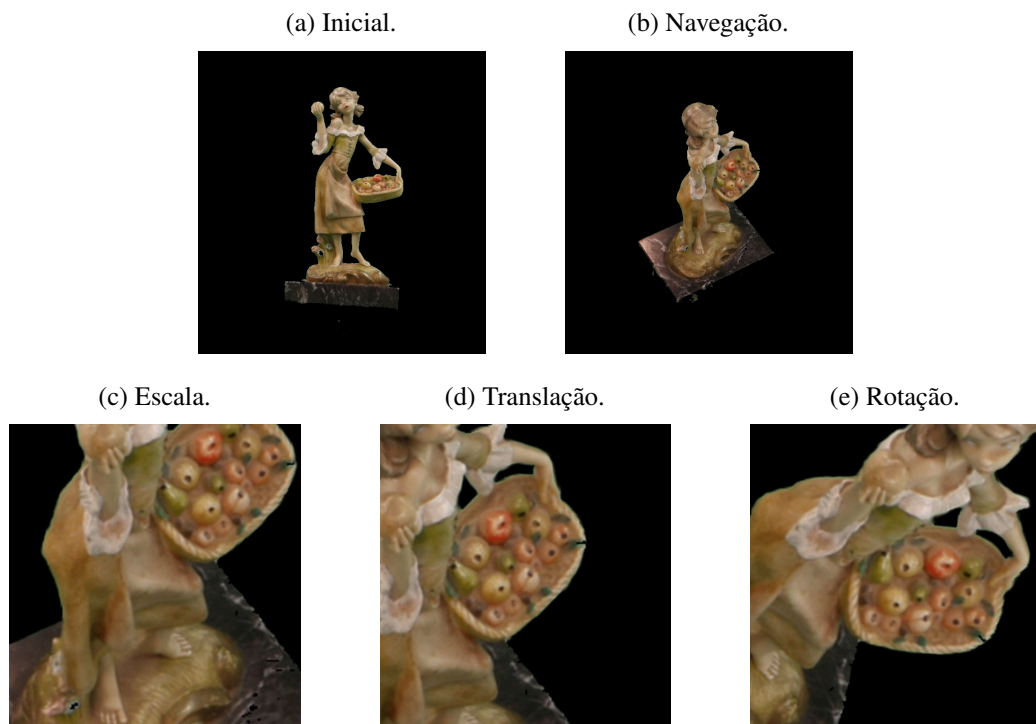


Fonte: Autoria própria.

transformação que muda a orientação de uma visualização, por exemplo, permitindo que um objeto adquirido em pé, possa ser visualizado com qualquer inclinação. A escala é a possibilidade de aproximar (*zoom in*) ou afastar (*zoom out*) a visualização da imagem, deixando-a maior ou menor.

Todas as propriedades da visualização interativa podem ser aplicadas simultaneamente. Por exemplo, considerando a Figura 2, pode-se realizar as seguintes operações: A Figura 2(a) apresenta a um vista frontal da estátua; a Figura 2(b) apresenta uma navegação para cima e para a esquerda; a figura 2(c) transformação geométrica de escala, aproximando a imagem; a Figura 2(d) acrescenta uma translação; e, finalmente, é aplicada uma rotação na Figura 2(e).

Figura 2 – Propriedades da visualização interativa.



Fonte: Autoria própria.

Acuidade Visual

Este trabalho é sobre o processo de digitalização de objetos reais com acuidade visual. Entende-se por acuidade a possibilidade de visualização dos mesmo detalhes vistos na peça real. Assim, cor, forma e texturas devem ser preservadas durante a digitalização das peças.

A acuidade visual deste trabalho pode ser definida por dois componentes principais: nitidez, que trata da clareza de detalhes na imagem; e acurácia na aquisição de cores, que faz com que as cores adquiridas sejam as cores reais do objeto, perante a iluminação durante a aquisição.

A nitidez faz com que formas e texturas sejam preservadas e assim seja possível visualizar os detalhes do objeto digitalizado. Diversas técnicas podem ser aplicadas para aquisição de imagens com nitidez. Uma digitalização com nitidez garante a representação de formas e texturas, o que permitirá a visualização de detalhes da peça real.

A acurácia na aquisição das cores depende da aplicação das técnicas de aquisição de imagens e técnicas computacionais que fazem com que as cores adquiridas representem as cores reais do objeto. Assim é possível gerar uma visualização com acurácia na representação das cores do objeto real.

Aplicando-se em conjunto as técnicas de aquisição de cores com acurácia e digitalização com nitidez, permite que os mesmos detalhes presentes na peça real sejam observados na peça virtual durante a visualização, oferecendo acuidade visual ao modelo.

Cabe ressaltar que quando a visualização ocorre através de telas, é preciso considerar diversos fatores técnicos das telas, como a capacidade de representação das cores, calibração da tela (WEN, 2013), reflexo, iluminação ambiente (CHOI; SUK, 2014), entre outros. Este trabalho trata da digitalização de objetos, sendo que estes fatores não são examinados neste projeto.

1.2 Problema

Como digitalizar objetos reais para gerar modelos interativos e com alta acuidade visual? Modelos tridimensionais que são interativos por natureza, podem ser construídos a partir de escaneamentos por: laser, tomógrafos, ressonância magnética ou sonografias. Além destes, fotografias podem ser usadas para realizar um processo de reconstrução tridimensional - porém nenhum destes apresenta alta acuidade visual.

O modelo também pode ser desenhado por um artista, como é comum nos atlas de anatomia humana (SOBOTTA et al., 2001). Mas neste caso não é possível garantir fidelidade com o objeto desenhado, uma vez que o desenho depende da interpretação e capacidade do artista, além disso, estes exigem um considerável esforço e tempo para realização.

Para tratar o problema da digitalização é necessário estabelecer um processo de digitalização de objetos reais que permita a visualização com interatividade e acuidade visual; um método

de processamento e organização dos dados digitalizados que permita gerar uma visualização com as características desejadas.

1.3 Hipótese

É possível desenvolver um processo de digitalização de objetos reais, baseado em um conjunto de fotografias que representem todos os ângulos de visualização do objeto - considerando um intervalo pré-determinado. Com estas fotografias é possível gerar um modelo que permita uma visualização interativa - ou seja, navegação e transformações geométricas - e possua acuidade visual - nitidez e acurácia na aquisição das cores.

1.4 Objetivos

1.4.1 Objetivo Geral

Desenvolver um processo de digitalização de objetos reais que gere um modelo interativo e com acuidade visual. Por interatividade este trabalho entende a possibilidade de visualizar o objeto de diferentes pontos de vista, como é possível fazer movendo o objeto real em relação ao expectador. Acuidade visual é a capacidade que o modelo possui de representar as mesmas características visuais do objeto real.

1.4.2 Objetivos Específicos

- I Desenvolver um processo de digitalização de objetos reais que permita interatividade, para tanto é necessário que seja fotografado cada ângulo do objeto conforme um intervalo predeterminado;
- II Definir um processo de aquisição fotográfica que permita a acurácia na aquisição das cores, este processo deve ser realizado usando técnicas de calibração da câmera ou das imagens adquiridas;
- III Estruturar um método de aquisição de imagens com garantia de nitidez, para tal devem ser aplicadas técnicas fotográficas adequadas;
- IV Padronizar um procedimento de correção e aprimoramento da acuidade visual das imagens adquiridas;
- V Selecionar um algoritmo de segmentação de imagens para separar o objeto fotografado do fundo da imagem;
- VI Prover um mecanismo de ordenação das imagens adquiridas de forma a viabilizar a posterior geração da visualização do modelo com as características da interatividade e acuidade visual.

Limitações

Este trabalho não trata da geração da visualização. A visualização é trabalho de outro integrante do grupo de pesquisa. Além disto, o processo de escaneamento das peças só adquire a parte superior do objeto, não sendo possível gerar a visualização da parte inferior, a menos que o objeto seja digitalizado novamente com a outra face para cima.

1.5 Justificativa

Atualmente existem diversas tecnologias de exibição de imagens em altas resoluções, com representação de cores próximas às reais, visualizações estereoscópicas, teatros imersivos, pirâmides holográficas, realidade virtual e aumentada. Existe a possibilidade de interação com estes modelos através de sensores como por exemplo acelerômetros e câmeras. Estas tecnologias podem ser encontradas, por exemplo, em cinemas, celulares e videogames.

As tecnologias interativas de exibição de imagens podem auxiliar diversas áreas, incluindo a educação (OLSEN et al., 2018), contornando os problemas citados por Ackerman - de uso de imagens bidimensionais - que prejudicam o entendimento de estruturas tridimensionais. Porém na maioria dos casos são usados modelos tridimensionais, que segundo Ackerman não representam fielmente estruturas complexas como peças biológicas.

Portanto, é necessária uma forma que aquisição de conteúdos que possam ser usados nestas telas, porém, que permitam uma representação adequada de estruturas complexas, mantendo precisão de cor e forma; e possibilitando a visualização de detalhes (como no caso das fotografias). Porém permitindo a interação/manipulação do objeto que está sendo visualizado (como no caso dos modelos tridimensionais).

Visualizações com estas características poderiam ser usadas na educação de diversas áreas, como: medicina, biologia, artes, história, engenharias, design, entre outras. O sistema produtivo poderia aplicar treinamentos mais precisos, como identificação de pragas ou desgaste de peças mecânicas, e a indústria de entretenimento poderia criar aplicativos, jogos ou filmes mais interativos e realísticos.

1.6 Contribuições e Originalidade

Este trabalho não trata da otimização de determinado algoritmo, e sim da definição de um procedimento baseado em um conjunto de técnicas capaz de viabilizar a digitalização de objetos com acuidade visual, provendo um modelo que apresente acuidade visual e interatividade.

As técnicas tradicionais de aquisição/representação de objetos 3D, como desenhos, escaneamentos 3D e reconstrução tridimensional utilizam técnicas baseadas em nuvens ou malhas de pontos e precisam renderizar estas informações para conseguir apresentar o objeto.

Este trabalho adota uma abordagem diferente, baseada em aquisições fotográficas, nas quais, para determinado ponto de vista do objeto, este está integralmente representado, com luzes e sombras. Quando objetos tradicionais são visualizados é necessário processar cada ponto, sua cor e textura; posteriormente são inseridas luzes e sombras no objeto.

No caso deste trabalho é necessário identificar o ângulo desejado pelo usuário e apresentar a foto que melhor representa o ângulo desejado. Trata-se de um processo completamente diferente e mais simples. E apesar de sua simplicidade, o resultado da visualização possui mais acuidade visual, pois os objetos apresentados são digitalizados de forma que todas suas características visíveis a olho nú estejam preservadas no modelo final.

No caso de modelos tridimensionais de objetos reais que foram desenhados por artistas, diversas características originais podem ser perdidas. No caso de escaneamentos tridimensionais tradicionais as limitações do equipamento e do formato de representação destes modelos pode inviabilizar a aquisição com a acuidade visual apresentada neste trabalho.

Ao renderizar objetos tridimensionais tradicionais, aplicando texturas, cores e sombras, por melhor que seja a técnica utilizada, não pode ser comparada com a aquisição deste trabalho, que adquire o objeto já com a iluminação e sombras reais. Como não é necessário renderizar o objeto, as texturas apresentadas neste trabalho não sofrem nenhum tipo de pós-processamento ou deformação, independentemente do ponto de vista visualizado, portando são mais próximas da textura real do objeto.

Escaneamentos a laser possuem dificuldades em representar áreas com elementos difusos, como cabelos e regiões esponjosas, pois estas áreas causam muita oclusão. No caso de fotografias estas áreas são digitalizadas com mais naturalidade. Apesar de nenhuma das técnicas conseguir representar fielmente cada fio de cabelo, no caso da fotografia o volume, luzes e sombras são os reais e não dependem da renderização de uma malha de pontos que pode gerar um volume incorreto para representar o objeto.

Ao digitalizar um objeto com scanners 3D, cada ponto do objeto é adquirido em um determinado instante. Ao digitalizar um objeto com a técnica proposta neste trabalho cada ponto pode ser adquirido várias vezes, em pontos de vista diferentes. Como acontece no mundo real, ao girar um objeto, cada parte deste objeto recebe luz de forma diferente e portanto a aparência deste ponto muda. No procedimento usado neste trabalho estas diferenças são adquiridas também, não sendo necessário processar a imagem para simular estas diferenças. Assim, a interatividade torna-se mais simples e realística.

O processo de digitalização também é mais simples. Não é necessário reconstruir objetos, calcular malhas de pontos ou texturas. Uma vez adquirida determinada perspectiva, esta perspectiva está pronta para ser usada na visualização. Além disto, como várias perspectivas são adquiridas, é possível gerar uma visualização estereoscópica sem a necessidade de grandes cálculos, sendo necessário apenas selecionar dois pontos de vista que viabilizem a estereoscopia.

1.7 Estrutura da Tese

O Capítulo 2 apresenta algumas premissas, conceitos e revisão da literatura, além de um breve relato histórico do projeto. O Capítulo 3 inicia com a descrição do scanner usado neste trabalho e em seguida descreve o método proposto para o desenvolvimento da pesquisa.

No Capítulo 4 está o desenvolvimento deste trabalho. Inicialmente é apresentado um conjunto de equações que permitem a utilização das fotografias em sistemas de simulação baseados em ambientes de computação gráfica tridimensionais. Em seguida o hardware do scanner é debatido. Por último é discutido o processo de digitalização de objetos reais proposto para este trabalho.

No Capítulo 5 são descritos dois softwares, sendo o primeiro responsável pelo controle do scanner durante a digitalização das peças e o segundo o responsável pelo processamento das imagens e construção da estrutura que viabiliza a visualização. Posteriormente são apresentadas a validação e análise do sistema. No Capítulo 6 são apresentadas as considerações finais deste trabalho e são propostos trabalhos futuros.

2 Conceitos, premissas e revisão da literatura

Com o intuito de entender como funciona a digitalização de objetos reais este capítulo apresenta um conjunto de técnicas usadas para este fim. Em seguida apresenta um trabalho que realiza a visualização interativa a partir de fotografias digitais. Então é feita uma revisão de técnicas usadas para adquirir fotografias com nitidez e acurácia na aquisição das cores.

O scanner usado neste trabalho começou a ser desenvolvido antes deste projeto e faz parte de um escopo maior, estando ligado também a outras iniciativas. Assim, este Capítulo também apresenta um histórico do equipamento, usando um texto desenvolvido em conjunto com outros integrantes do grupo de pesquisa para posicionar o desenvolvimento do equipamento e deste trabalho.

2.1 Digitalização de Objetos

Diversas técnicas podem ser usadas para digitalizar objetos reais, desde técnicas automatizadas, até o desenho técnico. Algumas destas técnicas geram modelos interativos, outras geram modelos com acuidade visual. Estas técnicas normalmente geram modelos tridimensionais baseados em nuvem de pontos ou malhas poligonais para representar os objetos digitalizados. A Figura 3 apresenta as etapas essenciais da digitalização de objetos.

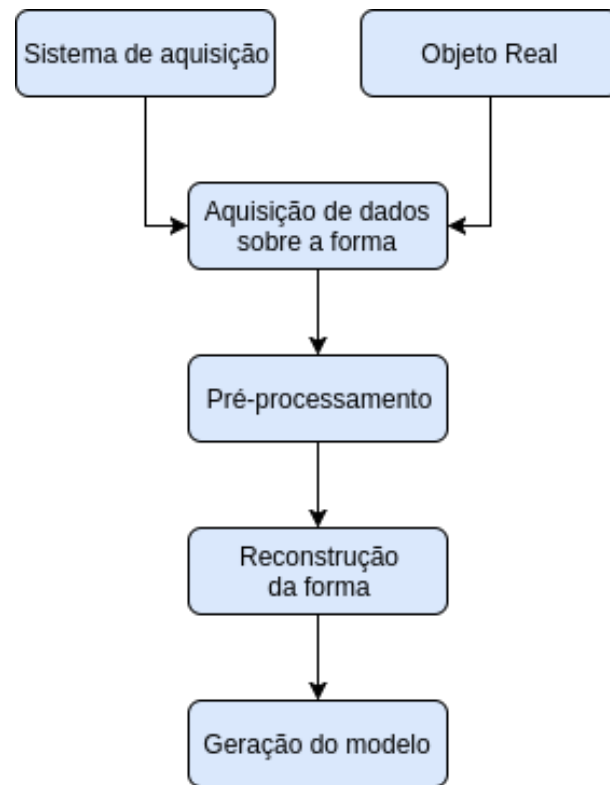
Independente do sistema de aquisição escolhido, o processo de digitalização de objetos reais inicia com a aquisição de dados sobre a forma do objeto. A forma do objeto é medida através de um conjunto de pontos que são armazenados em uma estrutura chamada nuvem de pontos. Cada ponto marca uma posição em um sistema de coordenadas tridimensional. Esta etapa é crucial para a qualidade do modelo resultante.

Após a coleta destes pontos, no pré-processamento, são realizadas etapas de eliminação de ruídos ou pontos associados a erros de medição, dependendo da técnica escolhida. Também é realizado o registro dos pontos, que é a etapa que associa os pontos a seus vizinhos e realiza a suavização das formas (SPIELMAN; TENG, 2004).

O objetivo da terceira etapa é a obtenção de modelos de alto nível de precisão. Para isso são usadas técnicas como triangulação, segmentação e classificação dos pontos para criação de arestas e faces dos modelos. Estas técnicas tentam associar pontos para gerar as superfícies/regiões que compõe o objeto (INTWALA; MAGIKAR, 2016).

A última etapa é a importação dos dados reconstruídos em um software CAD (Computer Aided Design - Desenho assistido por computador) no qual é gerado o modelo tridimensional.

Figura 3 – Processo de digitalização de objetos reais.



Fonte: Adaptado de: (INTWALA; MAGIKAR, 2016).

Após este processo a reconstrução foi concluída e é possível manipular o modelo ou este pode ser exportado para outros formatos e ser usado em: vídeos, jogos, impressões tridimensionais, etc.

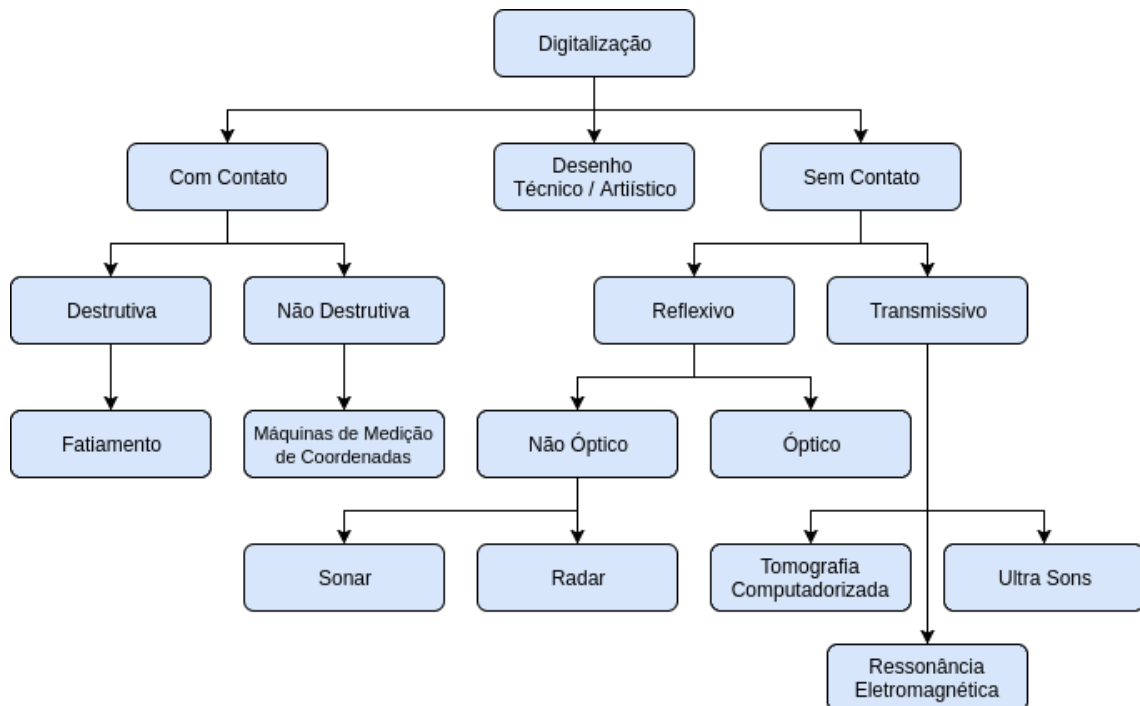
Os sistemas de aquisição de dados sobre as formas dos objetos reais podem ser baseados em escaneamentos por: laser, tomógrafos, ressonância magnética ou sonografias. Além destes, fotografias podem ser usadas para realizar um processo de reconstrução tridimensional, ou o modelo pode ser desenhado por um desenhista técnico ou artista gráfico em um software de modelagem tridimensional. A Figura 4 apresenta os principais métodos.

Como mencionado, uma das formas de digitalização de objetos reais é feito manualmente por desenhistas técnicos. Na indústria, para confecção de componentes mecânicos para reposição de peças desgastadas, um desenhista técnico usa ferramentas como réguas e paquímetros para tomar medidas das peças danificadas e a desenha em softwares CAD, gerando os modelos tridimensionais, que são depois confeccionados em máquinas de torno.

Artistas gráficos também podem desenhar objetos, como nos casos dos atlas de anatomia humana (SOBOTTA et al., 2001). Porém, desenhos dependem muito da capacidade do artista, necessitam de muito tempo e esforço para serem feitos e não existem garantias de que o desenho represente fielmente as cores, formas e texturas do objeto desenhado.

As técnicas baseadas em contato podem ser divididas em duas categorias, as técnicas

Figura 4 – Métodos de aquisição de dados de objetos reais.



Fonte: Adaptado de: (FERREIRA, 2009).

destrutivas baseiam-se em fatiamento dos objetos e aquisição de dados de cada fatia. Um exemplo destas técnicas é (ACKERMAN, 1998). O problema destas técnicas é que o objeto é destruído durante a digitalização. Estas técnicas adquirem informações das camadas internas do objeto, mas não garantem que a visualização externa terá acuidade visual.

As técnicas com contato não destrutivas baseiam-se em máquinas de medição por coordenadas (Coordinate Measuring Machine - CMM). Estas máquinas usam uma sonda de contato que faz a digitalização ponto a ponto das coordenadas da superfície do objeto (HE et al., 2010).

Apesar de estas máquinas possuírem grande precisão na aquisição de informações sobre as coordenadas dos pontos do objeto, por se basearem em contato, não conseguem fazer a aquisição de informações de cor e textura dos objetos. Portanto, não são boas alternativas para geração de visualizações com acuidade visual.

As técnicas de digitalização sem contato utilizam um emissor e um receptor de alguma forma de energia para mapear o mundo real. Considera-se que estas técnicas possuem menos precisão, se comparadas às técnicas com contato (HE et al., 2010; FERREIRA, 2009; INTWALA; MAGIKAR, 2016).

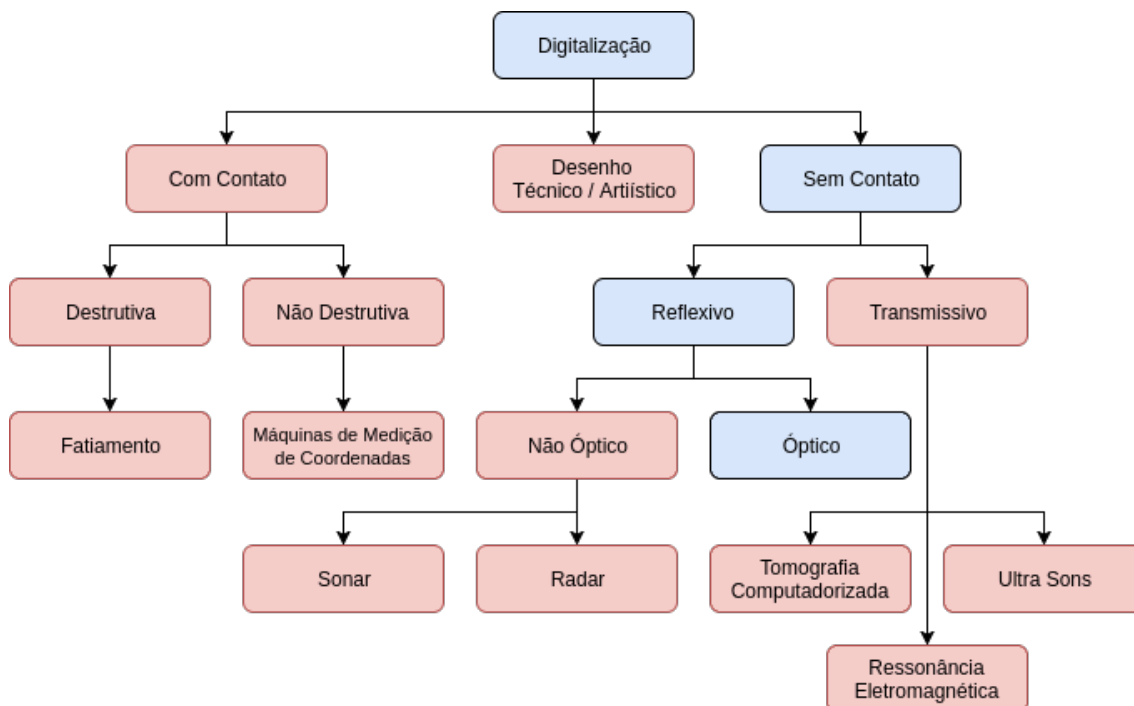
Estas podem ser divididas em métodos transmissivos, quando usam ondas ou campos que atravessam os objetos a serem digitalizados; ou reflexivos, quando usam ondas que refletem no objeto que está sendo digitalizado.

Dentre as técnicas transmissivas estão as: tomografias computadorizadas, que usam ondas eletromagnéticas de raio X; ressonâncias eletromagnéticas, que emitem campos eletromagnéticos e realizam a leitura da oscilação destes campos ao atravessar os objetos; e os ultrassons, que utilizam ondas mecânicas de ultrassom para realizar a digitalização.

Estas técnicas são capazes de digitalizar a estrutura interna dos objetos, porém também não são capazes de adquirir cores e texturas das peças digitalizadas. Novamente esta característica inviabiliza estas técnicas para geração de visualizações com acuidade visual.

Quanto às técnicas sem contato reflexivas, estas podem ser divididas entre ópticas e não ópticas. As últimas são os radares, que usam frequências de rádio, e sonares, que usam ondas mecânicas de som para mapear objetos. Estes equipamentos também não adquirem as cores dos objetos digitalizados, e portanto não geram acuidade visual. Desta forma a Figura 5 apresenta em vermelho as técnicas de digitalização que não podem ser usadas para gerar visualizações com interatividade e acuidade visual.

Figura 5 – Digitalização de Objetos Reais - Métodos que não geram acuidade visual.

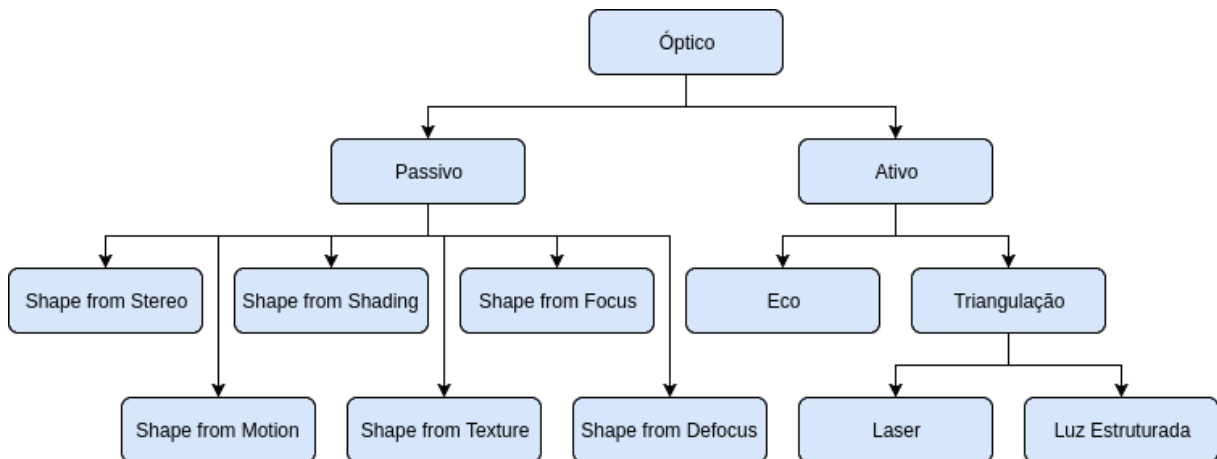


Fonte: Adaptado de: (FERREIRA, 2009).

Quanto aos métodos ópticos, estes podem ser divididos em passivos, quando não necessitam de nenhuma fonte específica de luz para funcionar; e ativos quando projetam uma fonte de luz específica sobre o objeto e realizam a captura desta fonte por sensores. A Figura 6 apresenta os métodos ópticos.

Estes métodos também podem ser divididos em *on line*, quanto geram os modelos tridimensionais durante o processo de aquisição de imagens, ou *off line*, quando podem usar

Figura 6 – Digitalização de Objetos Reais - Métodos ópticos.



Fonte: Adaptado de: (FERREIRA, 2009).

imagens previamente adquiridas para gerar a reconstrução do modelo tridimensional. As técnicas ativas são todas *on line*, enquanto as passivas podem operar das duas formas, dependendo da implementação.

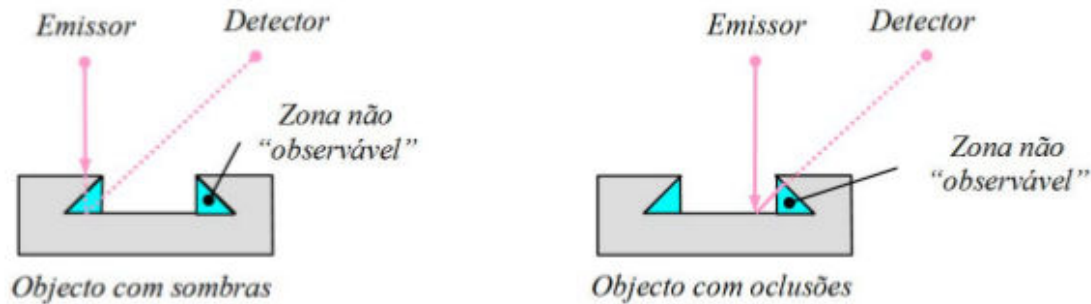
As técnicas ativas usam projeções diretas de algum tipo de luz sobre a peça que está sendo adquirida. Quando o tempo de emissão e recepção do reflexo da luz é usado para realizar os cálculos de distância, estas técnicas são baseadas em eco.

Quando usa-se a posição previamente conhecida do emissor e do receptor para realizar o cálculo da posição de onde a luz emitida refletiu, diz-se que são técnicas baseadas em triangulação. Estas técnicas podem ser baseadas em um determinado padrão de cores, ou em laser. Os *scanners* laser funcionam desta forma.

Estas técnicas possuem problemas para determinar corretamente cantos e escanear peças complexas (INTWALA; MAGIKAR, 2016). Esta característica vem do fato de estas tecnologias serem sensíveis a zonas de oclusão, as quais não podem ser mapeadas pelos sensores dos *scanners* a laser. A Figura 7 apresenta o problema das zonas de oclusão.

Além das zonas de oclusão, os *scanners* laser possuem outras limitações: Não conseguem digitalizar superfícies não reflexivas; não diferenciam os materiais e cores dos objetos; não digitalizam cavidades superiores a determinada profundidade ou inferiores a determinado perímetro (FERREIRA, 2009); apresentam problemas de calibração (LAVRINOV; KHORKIN, 2016); e possuem dificuldade em digitalizar superfícies complexas como superfícies peludas e porosas, comuns em peças orgânicas.

As técnicas ópticas passivas usam fotografias para calcular similaridades e então calcular distâncias entre os pontos de similaridade. Estas técnicas pode usar diversos algoritmos de processamento de imagens para encontrar similaridades entre as fotos, como: alterações de foco

Figura 7 – Zonas de oclusões nos *scanners* laser.

Fonte: (FERREIRA, 2009).

(MALIK; CHOI, 2009; MAHMOOD, 2013), detecção de textura (PAYET; TODOROVIC, 2011), sombras (HUANG; SMITH, 2011; WANG et al., 2008), movimento (CHHATKULI et al., 2016; PARASHAR; PIZARRO; BARTOLI, 2016) e par estéreo de imagens (ROHITH et al., 2013; HOSAM; SUN, 2009).

Porém as técnicas ópticas passivas apresentam problemas para criar modelos tridimensionais, como problemas de escalabilidade, uma vez que podem ser necessárias várias imagens para realização do processo, o que acarreta na necessidade de simplificação de geometrias e inconsistências entre o modelo gerado e o objeto real (INTWALA; MAGIKAR, 2016). Além disso, estes algoritmos são muito sensíveis a erros de aquisição e ruídos presentes nas imagens (TRUCCO; FISHER, 1992).

2.1.1 Visualização interativa a partir de fotografia digital

Além de tecnologias que geram modelos tridimensionais, buscando uma visualização mais realística, (FARIA et al., 2016) desenvolveu um processo de geração de visualizações interativas e estereoscópicas para ensino de neuroanatomia. Esta tecnologia se baseia em um conjunto de fotografias que podem ser visualizadas de forma não linear, passando ao espectador a impressão de estar visualizando interativamente a peça, de forma que é possível visualizar qualquer perspectiva da peça que tenha sido fotografada.

Esta tecnologia se baseia em posicionar a peça a ser adquirida em um suporte e fotografar as diversas perspectivas desejadas, como um intervalo de 5°, 10°, 15° ou 20° entre as perspectivas. Porém, o movimento da peça durante a aquisição é realizado de forma manual, o que pode causar imprecisões no processo, além disso, os intervalos possuem um valor mínimo de 5°, podendo deixar perspectivas interessantes sem serem fotografadas.

O trabalho conclui que o uso da estereoscopia não resultou em vantagem pedagógica. Porém o intervalo mínimo de 5° pode não ser o ideal para criação de estereoscopia. Portanto a estereoscopia gerada pode não apresentar qualidade e este ser o motivo da falta de vantagem pedagógica do uso da estereoscopia. Além disso a estereoscopia usada foi anaglifo, o que

atrapalha o visualizador a perceber cores e formas de modo adequado.

Outro problema da abordagem pode estar no processo fotográfico. (FARIA et al., 2016) explica as técnicas fotográficas usadas, nem se foi utilizado algum recurso de pós-processamento para aperfeiçoar as imagens, corrigir problemas ou mesmo obter métricas sobre a qualidade da imagem adquirida.

Fotografias são muito sensíveis a erros intrínsecos do processo de fotografia e ao uso incorreto de técnicas fotográficas, portanto o processo de aquisição das imagens pode resultar em complicações que implicariam em decréscimo da acuidade visual do modelo final.

Por exemplo, câmeras digitais aplicam filtros que alteram as cores da imagem adquirida, para que esta fique mais atrativa aos consumidores. Este processo porém resulta em cores que não representam fielmente as cores do objeto fotografado. Para o estudo de neuroanatomia, as cores da peça podem ser importantes e (FARIA et al., 2016) não descreve o uso de nenhuma técnica para conseguir acuidade na aquisição das cores, nem nitidez da imagem.

Considerando as similaridades entre o trabalho de Faria e técnica proposta, fez-se um levantamento dos trabalhos que citaram o estudo.

(ZHAO et al., 2020) é um estudo secundário sobre a efetividade da realidade virtual no ensino de anatomia. O artigo afirma que a realidade virtual é uma tecnologia emergente que tem potencial para transformar o ensino na área médica. Se comparados com o ensino tradicional ou digital 2D a realidade virtual melhora a efetividade do ensino de anatomia. Cita especificamente que uma vantagem do trabalho de Faria é o baixo custo para desenvolvimento.

(JACQUESSON et al., 2020) trata de um estudo secundário sobre a efetividade da estereoscopia no ensino de neuroanatomia. A anatomia do crânio e cérebro são particularmente difíceis de entender e ensinar. Segundo o artigo os estudantes estão entusiasmados com as novas tecnologias, porém o impacto nas competências clínicas dos estudantes precisam ser avaliados de forma mais objetiva.

(SHAO et al., 2020) estudo secundário que afirma que a neurocirurgia de tumor na base do crânio é um dos procedimentos cirúrgicos mais desafiadores e delicados. Afirma a realidade virtual é uma das técnicas mais promissoras de planejamento cirúrgico e que deve ser usada no processo de ensino de casos clínicos.

(SATTAR et al., 2019) compara o uso de texto, vídeo e tecnologias imersivas na aprendizagem de escolas públicas e privadas no Paquistão. Segundo o estudo as tecnologias imersivas tiveram os melhores resultados na motivação e competências dos estudantes.

Segundo (FIGUEIREDO; OLIVEIRA, 2019) o treinamento de cirurgias neurológicas por simulação é uma forma de aprendizagem mais rápida e que reduz possíveis erros. O trabalho conclui que simulações com realidade virtual devem ser inseridas no currículo de escolas de residência médica.

(TRIEPELS et al., 2019) avalia se estudo de anatomia em ambientes tridimensionais melhoram o aprendizado de estudantes e conclui que: a maioria dos estudantes prefere estudar em ferramentas tridimensionais; os sistemas baseados em computação tridimensional como realidade virtual e aumentada são mais efetivos para o ensino de anatomia do que métodos tradicionais.

Considerando o uso de fotografias como uma alternativa para geração de visualizações interativas e com acuidade visual, o Apêndice B apresenta um estudo sobre fotografia digital, seus principais problemas, correções, possibilidades de aprimoramento e métricas.

2.1.2 Fotografia com nitidez e acurácia na aquisição de cores

Nesta Subseção são apresentados os conceitos de nitidez e acuidade na aquisição de cores em fotografias. Diversos fatores podem influenciar nestas características e em alguns casos pode ser necessário usar processo após a aquisição fotográfica para conseguir nitidez e acuidade nas cores adquiridas.

Nitidez

A nitidez está relacionada principalmente à acutância das bordas da imagem (ZHANG; ALLEBACH, 2008). Além disso, para que seja possível uma boa definição das bordas, também é necessário uma resolução (quantidade de pixels da imagem) alta para que os pontos possam representar as bordas da imagem (CHOI et al., 2017; ROBINSON, 2016).

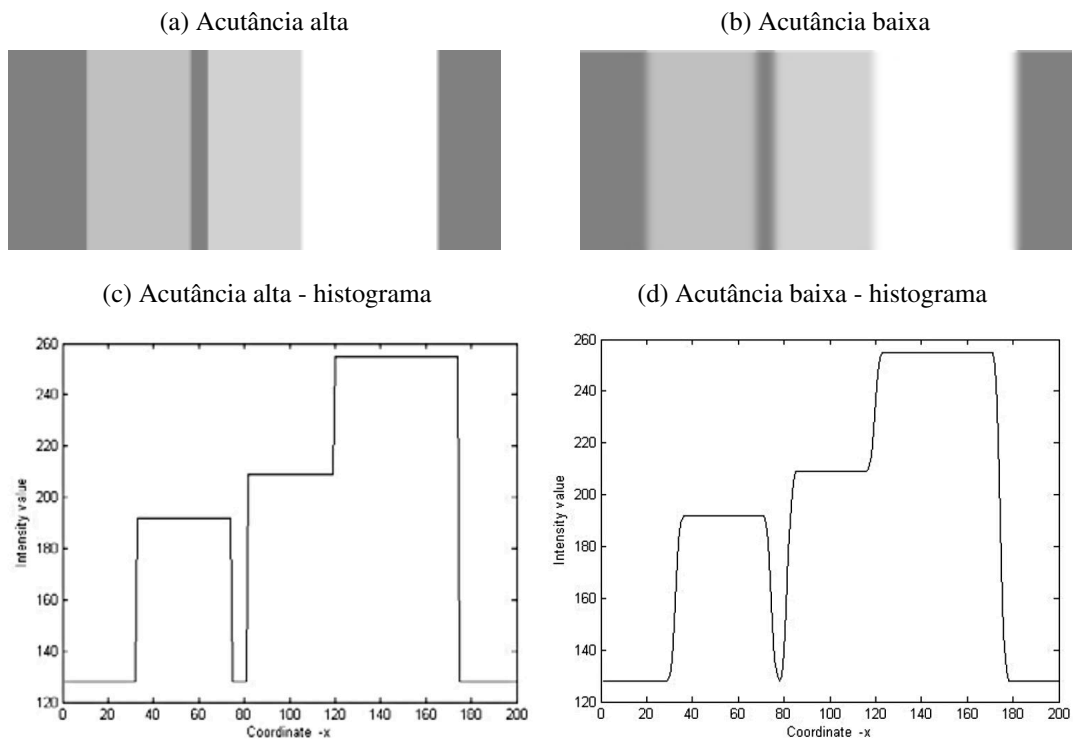
A resolução necessária da imagem - considerando uma foto de boa qualidade - vai depender de vários fatores: distância do objeto até a câmera, nível de detalhe pretendido, iluminação, tamanho do objeto, entre outros. Não existe métrica para quantificar a resolução mínima para se conseguir nitidez na fotografia.

A acutância define a forma de transição entre bordas das imagens (KOIK; IBRAHIM, 2013). Quanto mais breve a transição, maior a acutância e mais nítida será a imagem. Quanto maior a quantidade de pixels usados para delimitar a borda, menor a nitidez e a acutância. A figura 8 apresenta uma imagem com acutância alta 8(a) e outra com acutância baixa 8(b).

Analizando os histogramas é possível perceber que o histograma 8(c) - da Figura 8(a) - possui transições breves, passando diretamente de uma cor para outra, sem cores intermediárias. O gráfico 8(d) - que representa a Figura 8(b) - por sua vez possui uma transição mais lenta, criando cores intermediárias. Percebe-se que na Figura 8(a) é mais fácil delimitar o limite entre as faixas de cor diferente, se comparado com a figura 8(b).

A resolução também interfere na nitidez. Quanto maior a resolução de uma imagem, mais informação pode ser armazenada, permitindo a visualização de mais detalhes do que se comparado com uma imagem de baixa resolução. A resolução mais alta permite a distinção entre detalhes próximos, uma vez que estes não precisam dividir os mesmos pixels. A Figura 9

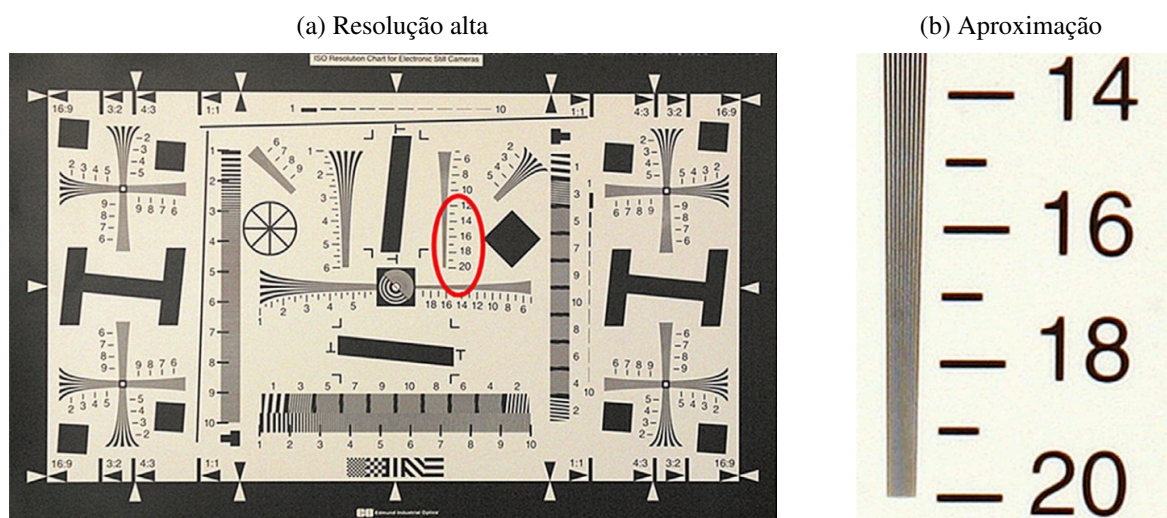
Figura 8 – Nitidez - Acutância.



Fonte: (KOIK; IBRAHIM, 2013).

apresenta duas imagens; a Figura 9(a) foi adquirida em alta resolução e a Figura 9(b) apresenta uma aproximação da área circulada.

Figura 9 – Nitidez - Resolução.

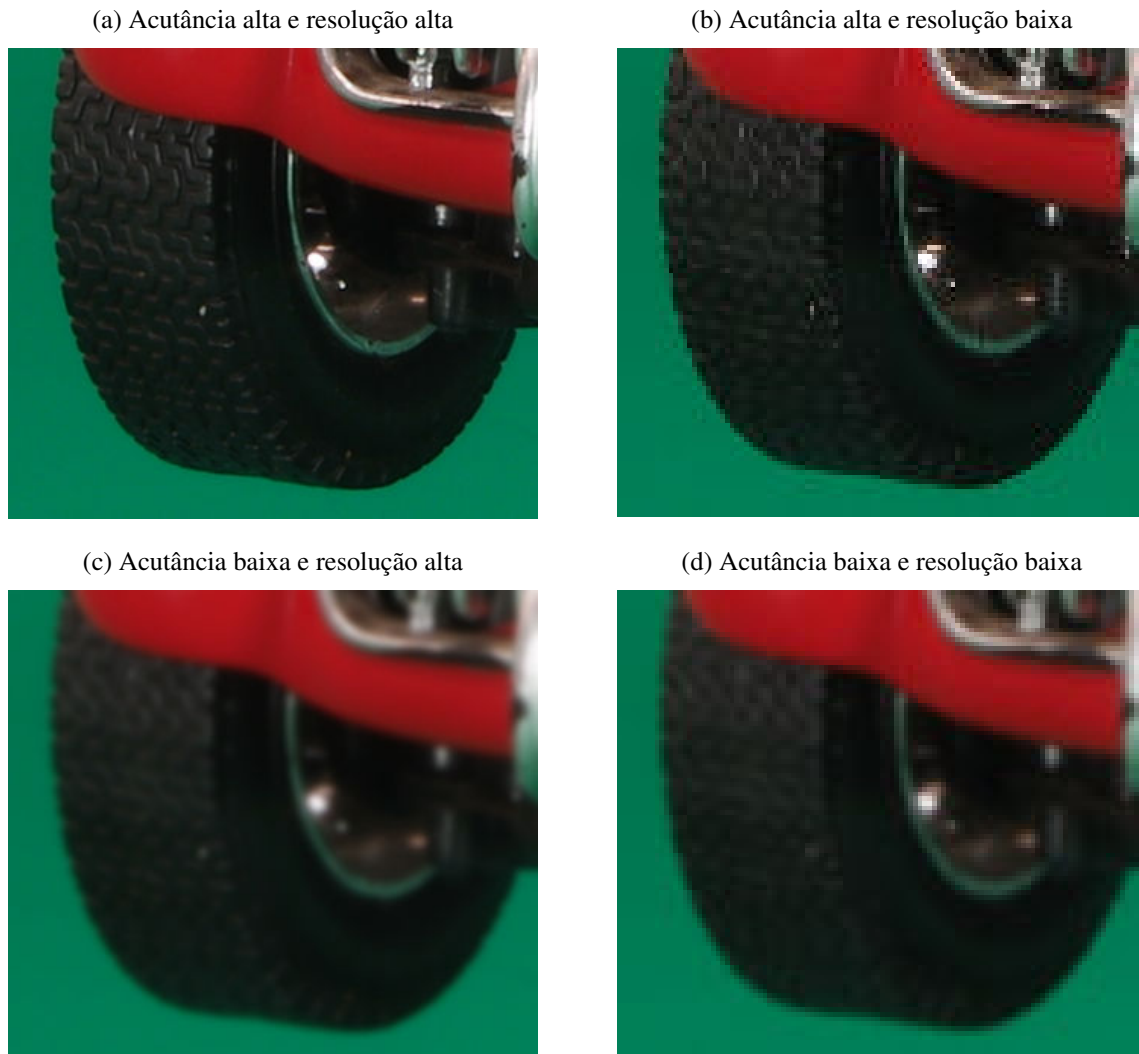


Fonte: (ROBINSON, 2016).

A Figura 10 apresenta quatro imagens: A figura 10(a) possui 350×310 pixels e alta acutância - é a imagem com melhor nitidez; A Figura 10(b) possui 113×100 pixels e alta

acutância; A Figura 10(c) também possui 350×310 pixels, porém baixa acutância; A Figura 10(d) possui 113×100 pixels e baixa acutância - é a pior nitidez entre todas as imagens.

Figura 10 – Nitidez - Acutância e resolução.



Fonte: Autoria própria.

Como mencionado, a nitidez reflete na capacidade de perceber detalhes nas imagens e na definição de bordas (PEDERSEN et al., 2010), portanto é indispensável para o escopo deste projeto. Existem diversos índices na literatura (CHENG; PEDERSEN; CHEN, 2017) para medir quão nítida uma imagem é. Também existe o processo de medição de quão borrada (LI et al., 2016; LIN; KUO, 2011) uma imagem é, o que pode ser entendido como o inverso da nitidez.

Existem métricas de nitidez objetivas, quando um índice de qualidade da imagem é calculado por um método específico, e subjetivas, quando pessoas atribuem escores e criam ranques das imagens mais ou menos nítidas. Entre as técnicas usadas para cálculo de nitidez podem ser usadas técnicas baseadas em (FERZLI; KARAM, 2009):

Cálculo da Variância (ERASMUS; SMITH, 1982): Quando a variância de toda a imagem é

calculada. Quanto maior a nitidez, maior a variância calculada;

Autocorrelação (BATTEN, 2000): É calculada a correlação entre pixels vizinhos. Quanto menor a nitidez de uma imagem, mais pixels são usados para fazer a transição entre as bordas, maior a correlação;

Limiar de Frequência (LAWRENCE et al., 1991): Um intervalo de frequência é escolhido pelo usuário de forma a selecionar as bordas das imagens. Este limite vai sendo ajustado de forma a selecionar corretamente as bordas na imagem. Por precisar de interação humana, este método não é adequado para um grande conjunto de imagens, nem para processos automáticos;

Percepção de Desfoque (MARZILIANO et al., 2004): É aplicado um algoritmo de detecção de bordas na imagem e para cada linha é calculado o comprimento da borda. Posteriormente é calculada a média destes comprimentos. Quanto menor o índice calculado mais nítida é a imagem;

Kurtosis (CAVIEDES; OBERTI, 2004): Estudo aponta alta correlação entre uma análise subjetiva de nitidez e o cálculo do índice de curtose da distribuição do domínio de frequência da imagem (Fourier). Quanto menor o índice calculado, maior a nitidez;

Domínio de Frequência (DE; MASILAMANI, 2013): Utiliza a transformada de Fourier para calcular o domínio de frequência da imagem e através de uma análise da distribuição dos pixels da imagem em relação ao cálculo de domínio da frequência, calcula um índice de qualidade de imagem. Quanto maior o índice, maior a nitidez;

Entropia do Histograma (CHERN; NEOW; ANG, 2001): Considera imagens em nível de cinza. A entropia de um histograma é calculada pela probabilidade de ocorrência de cada nível de cinza de uma imagem. Quanto maior o número de níveis de cinza - e suas probabilidades - maior a entropia. Imagens nítidas possuem mais níveis de cinza - com menor probabilidade cada - e maior entropia;

Derivada (BATTEN, 2000): Usa as derivadas de primeira (Gradiente) e segunda (Laplaciano) ordens na imagem. Quanto maior a métrica calculada maior a nitidez da imagem;

Wavelet (FANG et al., 2014): Usa a transformada Wavelet para calcular quais regiões da imagem estão borradas ou nítidas. Este processo pode ser usado para segmentação de regiões da imagem que não estejam nítidas;

Contorno (NARVEKAR; KARAM, 2011): Calcula o desfoque da imagem pela variação de contraste usando um modelo probabilístico para estimar a probabilidade de detectar desfoque nos contornos da imagem. Após este processo os dados são organizado em um modelo cumulativo de detecção de desfoque; (YUAN et al., 2015) Utiliza uma métrica

que estima a quantidade de contornos, baseada em estatísticas extraídas dos contornos; (SMITH, 2012; CHUNG et al., 2004) Realizam a análise de definição das bordas;

Baixa profundidade de campo (ZHAO et al., 2013): Divide a imagem em várias regiões e utiliza a opacidade de regiões da imagem para determinar se uma imagem está nítida ou não; (RAFIEE; DLAY; WOO, 2012): realiza a detecção de baixa profundidade de campo para realizar a segmentação - usando o algoritmo *graph cut* - das regiões desfocadas.

A nitidez de uma foto resulta de dois momentos: durante a aquisição da imagem (OH et al., 2014; TANG et al., 2016) e, em alguns casos, pode ser realizado um processamento posterior para corrigir ou aprimorar a nitidez (KRASULA et al., 2017; MANTIUK; TOMASZEWSKA; MANTIUK, 2012). Apêndice B apresenta os problemas que podem ocorrer durante o processo de aquisição fotográfica que resultam em perda da nitidez. A seção 5.2 apresenta alternativas para corrigir e/ou aprimorar a nitidez.

Acurácia na aquisição de cores

Para que seja possível gerar visualizações com acurácia na visualização das cores é necessário que as cores fotografadas sejam adquiridas com acurácia. Para que a visualização seja realizada com acurácia também é preciso que a tela/projetor usado esteja exibindo corretamente as cores (JUNG et al., 2016; JUNG et al., 2015; HINCAPIÉ-RAMOS et al., 2015).

Como não faz parte do escopo deste trabalho gerar as visualizações e considerando que as visualizações são dependentes das telas usadas, podem ser usados algoritmos de *color gamut mapping* para que as cores armazenadas sejam representadas com a maior fidelidade possível (LISSNER; URBAN, 2012).

Adquirir digitalmente uma cor consistente com a percepção humana de cores é um desafio, considerando que diversos fatores podem alterar cores de um objeto, como: iluminação, ponto de vista e propriedades do objeto que está sendo fotografado (KOBUS; BRIAN, 2002). Adicionalmente a isso, câmeras digitais possuem problemas intrínsecos ao processo de aquisição fotográfica, como a aberração cromática, discutida no Apêndice B.

Câmeras digitais adquirem imagens e as salvam em diferentes formatos como RAW (NEF - Nikon, CRW - Canon e ARW - Sony), TIF e JPEG. Imagem RAW armazenam informações coletadas pelo sensor da câmera e não a imagem (mapa de pixels) propriamente dita. Estes dados são então convertidos em arquivos de imagem (de diversos formatos) pela própria câmera ou por outros softwares.

A conversão de formatos de imagem ou de espaços de cores acarretam perdas de informações e podem causar alterações nas cores das imagens (LAU; HEIDRICH; MANTIUK, 2011). Por exemplo, uma imagem RAW que use o espaço de cores Adobe RGB¹ conseguirá representar

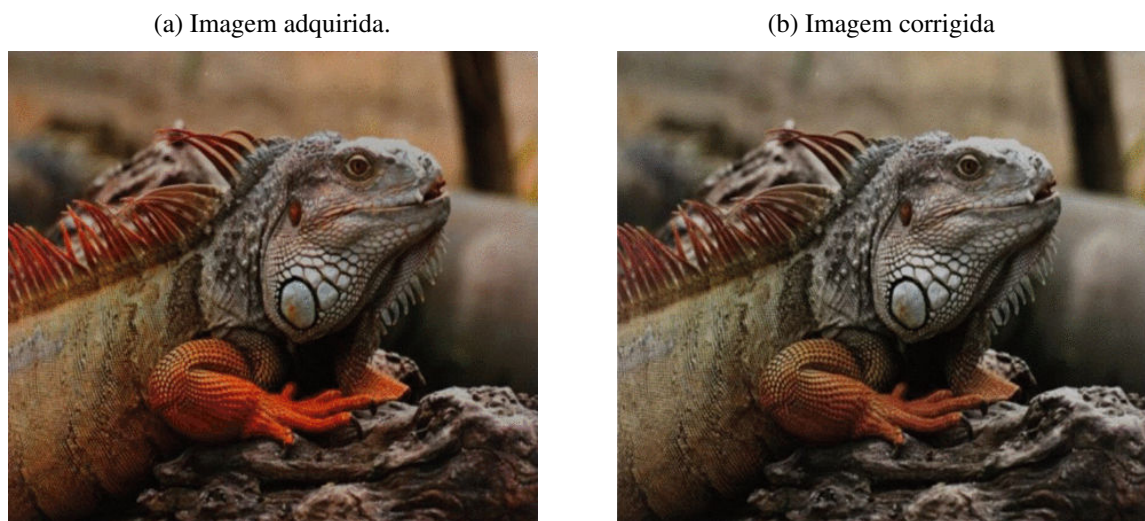
¹ <https://www.adobe.com/digitalimag/adobergb.html>

mais cores do que uma imagem RAW com espaço de cores sRGB².

Além disto, muitas câmeras aplicam automaticamente aprimoramentos nas fotografias para gerar imagens mais agradáveis para os consumidores. Por exemplo, tornando o céu mais azul e o verde de folhas mais vivos. Estes processos não são um problema para fotografias comuns, mas para garantir acurácia na exibição das cores este trabalho precisa eliminar este efeito das imagens.

A Figura 11 apresenta duas imagens. A Figura 11(a) apresenta uma fotografia adquirida por uma câmera. Nota-se que as cores são mais vibrantes e saturadas se comparadas a Figura 11(b), a qual foi corrigida por um algoritmo de calibração de cores (YU et al., 2016).

Figura 11 – Correção das cores.



Fonte: (YU et al., 2016).

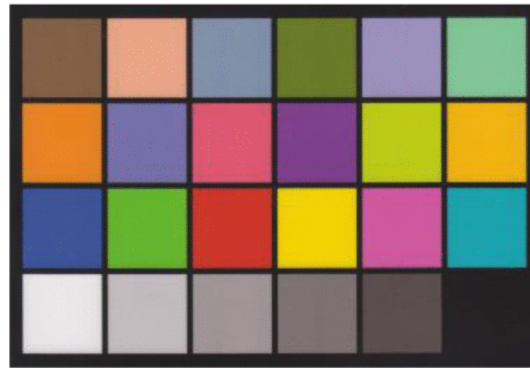
Segundo (YU et al., 2016), para correção de cores, pode ser usada a análise do espectro da luz, usando um *monochromator* ou colorímetros (equipamento que mede o espectro da luz), para então usar funções que corrigem as cores das imagens adquiridas pelas câmeras.

Outra alternativa é usar tabelas de cores conhecidas (*color checkers*, Figura 12) que são fotografadas. Como as cores são conhecidas é possível criar modelos que corrigem as imagens adquiridas pela câmera. Este modelo sendo conhecido, quaisquer fotografias adquiridas pela câmera podem ser corrigidas aplicando o modelo.

A criação dos modelos baseiam-se em amostras de cores RGB que possuem seus valores nos espaços de cores RGB e *XYZ Tristimulus* conhecidos. O espaço de cores tristimulus é baseado nos componentes da visão humana, que possui três receptores (vermelho, verde e azul). As outras cores percebidas pelo olho são misturas destas três cores (STONE, 2005).

² <http://www.color.org/chardata/rgb/srgb.xalter>

Figura 12 – Color Checker.



Fonte: (SULISTYO; WOO; DLAY, 2017).

Para criar os modelos de correção são usadas fórmulas que calculam distâncias entre cores. Com o *color checker* fotografado, é calculada a distância entre a cor adquirida e a cor previamente conhecida da célula do *color checker*, então um modelo de correção pode ser construído. As principais fórmulas de distâncias de cores são (PANT; FARUP, 2011; LIU; XIE; HUANG, 2011):

ΔE CMC : Criado pelo CMC (*Color Measurement Committee*), seu valor é expresso por uma razão com dois parâmetros $l : c$, onde l representa iluminância (*lightness*) e c a cromaticidade (*chroma*). Usualmente se usa uma razão de 2 : 1 como um limiar de diferença aceitável e de 1 : 1 como um limiar de imperceptibilidade;

CIE94 : Criado pelo CIE (*International Commission on Illumination*) em 1994, foi derivado de testes de tolerância de pinturas automotivas. Utiliza o cálculo da distância euclidiana entre duas cores no espaço de cores CIELAB (componentes: L - iluminância; a - cromaticidade verde-vermelho; e b - cromaticidade azul-amarelo). Sua fórmula é representada por: $\Delta E_{ab} = \sqrt{(L_2 - L_1)^2 + (a_2 - a_1)^2 + (b_2 - b_1)^2}$. O limiar de percepção é: $\Delta E_{ab} \approx 2.3$.

CIEDE2000 : Melhoramento do CIE94, onde foram adicionadas correções para uniformizar a tolerância perceptível. Este processo foi necessário uma vez que as pessoas percebem de forma diferente, alterações em cores diferentes. Por exemplo, considerando o CIE94, um $\Delta E_{ab} \approx 2.3$ na cor verde é menos perceptível que o mesmo valor na cor vermelha ou azul.

Além das fórmulas de distâncias entre cores, para construir os modelos de correção de cores baseados em *color checkers* podem ser usados vários métodos, dentre eles:

Look-up tables (KIM et al., 2004): Segundo (YU et al., 2016) atingem melhores resultados que outras técnicas, porém são muito demorados para construir modelos robustos.

Regressões polinomiais (MENDES; CARVALHO, 2005; XIA et al., 1999): Utiliza técnicas de regressão para estabelecer um mapeamento entre as cores adquiridas e conhecidas, calculando assim um fator de correção para as cores;

Redes neurais (ZHAO et al., 2018): Redes neurais podem ser usadas para fazer a correção de cores e balanço de branco em imagens adquiridas por sensores CMOS de alta resolução. Segundo os autores a acurácia é boa, porém consome muitos recursos computacionais.

Existem sistemas de calibração de cores para sistemas multicâmeras. A calibração de cores nestes sistemas é de grande importância, uma vez que se cada câmera adquirir cores diferentes, a visualização destas imagens pode causar fadiga ou desconforto em quem visualiza as imagens (YU; PARK, 2015). Algumas abordagens usadas neste método podem ser baseadas em:

Alinhamento de parâmetros fotométricos (ANIRUDTH et al., 2018): Em sistemas multicâmeras, como celulares com duas câmeras traseiras, são adquiridas duas imagens de uma mesma cena e estas imagens são fundidas para gerar a fotografia final. Para que este processo possa ser feito corretamente é preciso que as imagens sejam equalizadas para não apresentarem distorções entre elas. Neste trabalho o processamento acontece em duas etapas: primeiro é realizado o registro das duas imagens e em seguida é realizado o alinhamento de parâmetros fotométricos (brilho e cor - com equalização do histograma);

Correspondência entre câmeras (GURBUZ; KAWAKITA; ANDO, 2010): Imagens adquiridas de uma mesma cena, a partir de um conjunto de câmeras em posições diferentes podem ser calibradas com o cálculo de correspondências de cores entre as imagens adquiridas. Cabe ressaltar que este trabalho não busca a fidelidade de cores adquiridas com as cores reais do objeto, e sim a equalização de todas as fotos para que estas apresentem as mesmas cores;

Color checker omnidirecional (LI; DAI; XU, 2010): Este trabalho desenvolveu um *color checker* omnidirecional, ou seja, de qualquer ponto em que um sistema multicâmeras adquira imagens, será possível corrigi-las com base no mesmo padrão de cores oferecido pelo color checker;

Histograma (YU; PARK, 2015): São coletadas amostras de pixels das imagens e com estas amostras é realizada uma calibração das cores com base nos histogramas das imagens. Este trabalho trata especificamente de imagens estereoscópicas.

2.2 Histórico do projeto

O grupo de pesquisa denominado Centro de Inovação em Imagens Médicas (CIIM) surgiu em 2018 com o desenvolvimento de dois projetos que correm paralelamente. O primeiro

projeto é o “The Visible Human Table”, onde há a projeção de dois corpos humanos que foram seccionados e fotografados no projeto *The Visible Human Body* (ACKERMAN, 1998).

O segundo foi o desenvolvimento do F2S2, no entanto a primeira versão do scanner é anterior à criação do grupo, este foi um protótipo para verificar a ideia de gerar objetos com visualização 3D e estereoscópica. Os membros desta linha de pesquisa do grupo são: o acadêmico Lucas Murbach Pierin, os pesquisadores Diogo Roberto Olsen (autor deste trabalho) e Flávio de Almeida e Silva.

O F2S2 teve seu início a partir de uma conversa entre o professor Ph.D. Luiz Roberto Aguiar e o professor Dr. Edson José Rodrigues Justino acerca dos problemas encontrados em objetos desenhados, fotografados ou reconstruídos em 3D, os quais acabam ocultando detalhes ou perdendo características como cor e textura.

Assim, foi idealizado o primeiro protótipo do F2S2. Essa primeira versão foi desenvolvida sobre uma estrutura de alumínio com uma base giratória e um plano que a envolvia parcialmente, pois deixava apenas o espaço para a fotografia (Figura 13b). Na segunda versão, foi desenvolvido um plano de fundo na cor verde e foi mantida a base giratória, onde o objeto era posicionado.

Nas duas versões, o deslocamento da câmera era por um trilho curvado. Essa curvatura dava a inclinação da câmera em relação ao objeto. Também foram adicionadas quatro lâmpadas de Led, posicionadas acima e lateralmente ao objeto, além de um flash acoplado à câmera (Figura 13(c)). O sistema anterior utilizava lâmpadas tubulares de Led (Figura 13(a)). O problema das lâmpadas foram a temperatura de cor, isso causava diferenças na aquisição. A Figura 13 apresenta as duas primeiras versões do scanner.

Como a estrutura era aberta, acabava sofrendo a influência externa, principalmente de iluminação, o que prejudicava a aquisição. Assim como a centralização do objeto era realizada colocando-se ele sobre o eixo do motor que fazia o giro, ou seja, estimava-se o centro do objeto e posicionava-o sobre o centro da base, que era o eixo do motor.

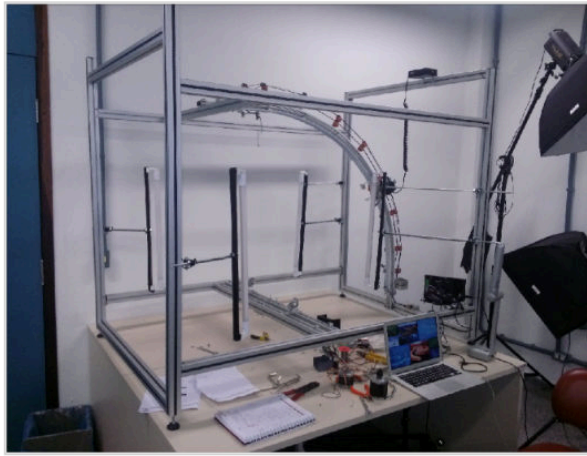
Essa prática não garantia a centralização, por exemplo, um objeto totalmente irregular não se determina visualmente o centro, além de não haver como colocá-lo baseando-se no seu centro geométrico. Isso poderia acarretar um giro elíptico ao se montar a visualização, pois o objeto não giraria sobre seu próprio eixo. Isso comprometeria a formação da visualização 3D, como também não garantiria o intervalo angular necessário para a visualização estereoscópica.

Como será visto na Seção 3.1, na terceira versão do scanner, alguns desses problemas foram contornados, melhorando a aquisição. O scanner foi fechado, deixando de ter a influência da iluminação externa, também foi colocado um conjunto de lâmpadas, com mesma temperatura de cor, acima do plano de fundo e na porta de acesso à câmera. Isso permite que a iluminação seja configurada para diferentes objetos, fundos e câmeras. A Figura 14(a) apresenta o início da construção da terceira versão. Figura 14(b): estrutura da versão 3 do scanner.

Percebeu-se que as aquisições poderiam se tornar Objetos de Aprendizagem (OA) com

Figura 13 – Versões 1 e 2 do scanner.

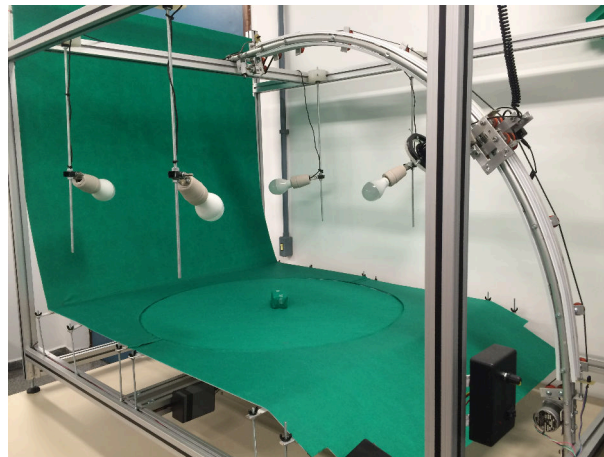
(a) Estrutura da versão 1.



(b) Plano de fundo.



(c) Versão 2 do Scanner.



Fonte: Autoria própria.

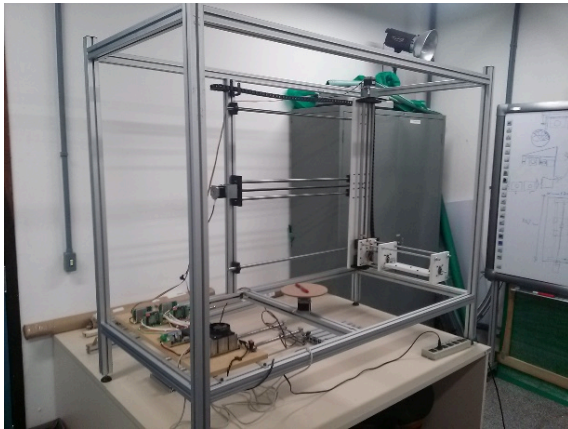
inserções de informações como conceitos, processos, entre outros, em qualquer dos objetos adquiridos. Essa percepção levou ao desenvolvimento de trabalhos voltados a mostrar como o scanner poderia contribuir com a área educacional.

Uma das vantagens das aquisições feitas no scanner é permitir a visualização dos objetos adquiridos em diferentes tipos de dispositivos, como televisões com e sem 3D, óculos de realidade virtual, pirâmides holográficas (técnica conhecida com *Pepper's Ghost*), ambientes imersivos etc. Isso foi apresentado e debatido por (OLSEN et al., 2018).

Como sequência do trabalho supracitado, também houve a apresentação do OA para sua avaliação. Um trabalho que envolveu especificamente os crânios, os quais foram apresentados a professores, profissionais da pedagogia e acadêmicos do Instituto Federal do Paraná – IFPR, no campus de Londrina-PR e também no campus Curitiba-PR. A avaliação se deu por meio de questionários envolvendo questões sobre a visualização estereoscópica e interação com o objeto. A Figura 15 apresenta um dos crânios sobre a base giratória e o outro sobre a mesa.

Figura 14 – Versão 3 do scanner.

(a) Início da construção da versão 3.

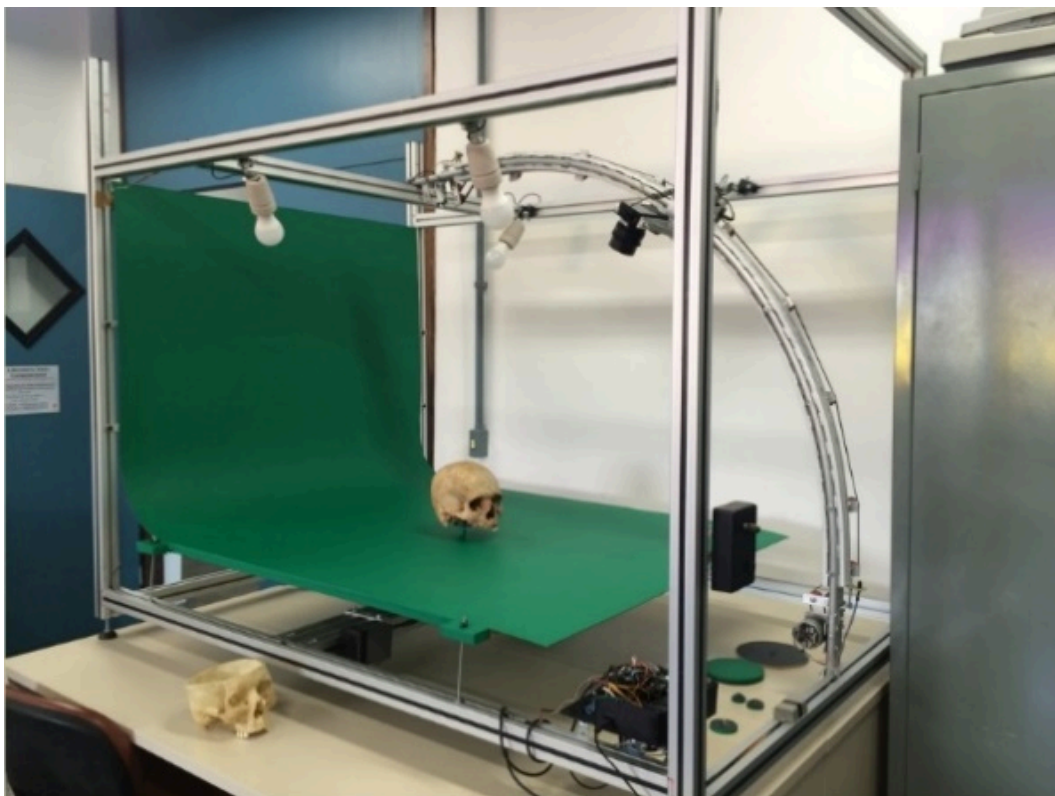


(b) Estrutura do scanner na versão 3.



Fonte: Autoria própria.

Figura 15 – Crânios adquiridos na segunda versão do F2S2.



Fonte: Autoria própria.

Criou-se uma visualização 3D, a qual foi chamada de “visualização interativa”, pois permite visualizar todas as perspectivas de um objeto digitalizado. Além dessa visualização, também foram criadas as noções de proporção e profundidade, ou seja, a estereoscopia. Foram adicionadas informações como nomes de ossos e seu posicionamento no crânio, para atender aos requisitos de ser um OA. O trabalho desenvolvido e o resultado das avaliações são encontrados

em (SILVA. et al., 2019).

Após esses dois trabalhos, focou-se nos ajustes e aquisições do F2S2 terceira versão, pois, como será visto na seção 3.1, houve melhorias no hardware, buscando melhorar a precisão nas aquisições e no posicionamento dos objetos.

3 Metodologia e infraestrutura para o desenvolvimento da tese

Este capítulo é iniciado descrevendo o Scanner usado neste trabalho. Para isso seus componentes são divididos em três grupos: sistema de movimentação; sistema de controle; e componentes fotográficos. Em seguida apresenta-se o método proposto para este trabalho, que envolve três etapas: aquisição, processamento e visualização.

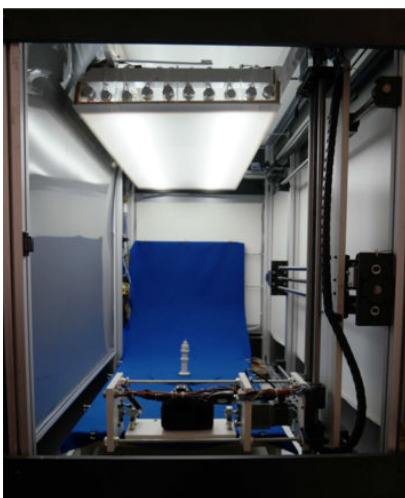
Neste trabalho é realizada uma pesquisa de natureza aplicada, pois visa desenvolver um procedimento baseado em hardware e software capaz de digitalizar objetos reais; com objetivos exploratórios, uma vez que não existem outras técnicas de digitalização de objetos similar à esta; que tem por procedimento pesquisa experimental, visto que analisa diversas variáveis para conseguir determinar a melhor forma de alcançar seu objetivo (GIL, 2008).

3.1 *Full Frames Semi-spherical Scanner - F2S2*

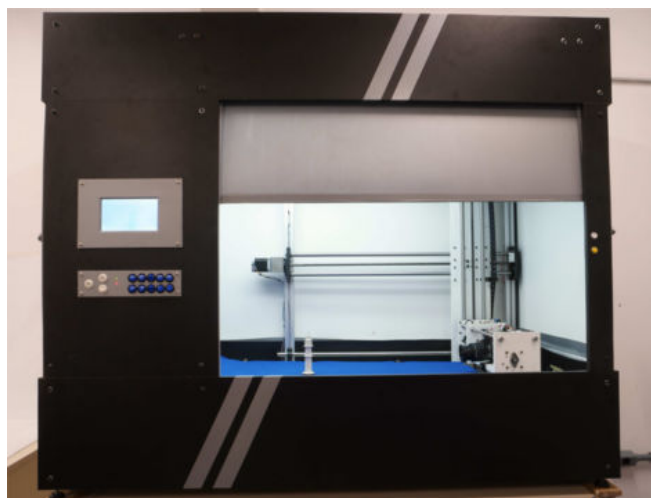
Para que seja possível realizar o processo de digitalização de peças reais de forma automatizada e respeitando a acuidade visual, foi desenvolvido um Scanner Semiesférico de Campo Completo (*Full Frames Semi-spherical Scanner - F2S2*). A Figura 16 apresenta a versão atual do F2S2.

Figura 16 – Scanner Semiesférico de Campo Completo.

(a) Vista interna.



(b) Vista lateral.



Fonte: Autoria própria.

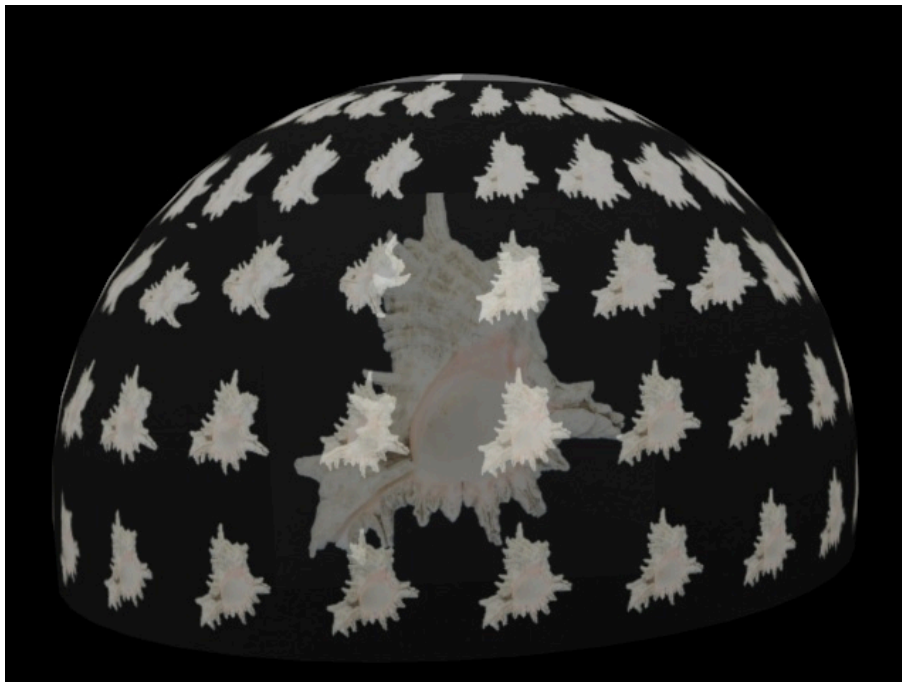
A sigla F2S2 surgiu a partir do nome do Scanner. “*Full Frames*” tem relação com a possibilidade de se adquirir todas as perspectivas de um objeto, aqui também descrito como

todos os ângulos do objeto, respeitando um intervalo angular pré-definido. Além disso, cada perspectiva ou ângulo contém a face completa do objeto, isso se difere de uma aquisição por laser que “fatia” a face do objeto e posteriormente junta essas fatias para reconstruí-la.

O termo “*Semi-spherical*” se refere à possibilidade de iniciar a aquisição em um ângulo, normalmente 0° , posição frontal ao objeto e paralela à base giratória e plano de fundo, terminando em um ângulo de 90° em relação à base giratória. Assim, criou-se o nome a partir dos dois “F” (Full Frames) e os dois “S” (Semi-spherical Scanner), tornando-se “FFSS” e posteriormente “F2S2”.

O F2S2 digitaliza objetos adquirindo um conjunto ordenado de fotografias em UHD. Este conjunto de fotos representa todos os ângulos do objeto digitalizado, conforme um intervalo configurável no *scanner*, com ângulo mínimo de 0.9° entre fotos. Após a aquisição das fotografias, o *framework* do F2S2 realiza o processamento das imagens e a organização destas para que seja possível gerar a visualização interativa. A Figura 17 apresenta um objeto no centro da esfera e as imagens adquiridas na posição onde a câmera estava no momento da fotografia.

Figura 17 – Aquisição do F2S2 com as perspectivas adquiridas e posicionadas na onde foram fotografadas.



Fonte: Autoria de Diogo Olsen e Flávio de Almeida e Silva com design de Felipe Teixeira de Almeida e Silva.

O Hardware do F2S2 pode ser dividido em três grupos de componentes: componentes de movimentação: responsáveis pela movimentação da câmera e do objeto durante o processo de digitalização; componentes de controle: que gerenciam o movimento e os disparos da câmera; componentes fotográficos: que são necessários para viabilizar a aquisição fotográfica com a qualidade almejada. Os três grupos de componentes são apresentado a seguir.

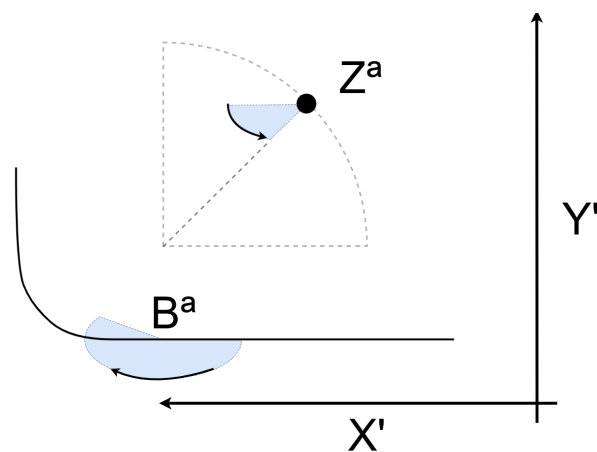
Hardware - Sistema de movimentação dos eixos

O F2S2 possui um sistema de movimentação que faz com que a câmera realize um deslocamento em forma de arco, centralizado no centro geométrico do objeto que está sendo escaneado. Para cada posição da câmera no arco, respeitando um intervalo angular predefinido, o objeto realiza uma rotação completa e a câmera adquire uma fotografia a cada intervalo angular horizontal do movimento do objeto. Desta forma todos os ângulos do objeto, respeitando os devidos intervalos angulares, são fotografados pela câmera.

O sistema de movimentação do F2S2 é composto de 4 eixos, sendo eles: X' : movimentação horizontal da câmera; Y' movimentação vertical da câmera; Z^a inclinação da câmera para que esta centralize o objeto que está sendo fotografado; B^a responsável pela rotação do objeto em torno do próprio eixo.

Optou-se por nomear os eixos com apóstrofo ($'$) e ordinal (a) para evitar a confusão com o sistema de coordenadas cartesianas tridimensionais (com Eixos nomeados por “X”, “Y” e “Z”). Além disso, o apóstrofo representa um movimento linear enquanto o ordinal (a) indica um movimento angular. A Figura 18 apresenta os eixos, seus sentidos de movimentação e um possível o arco formado pelo movimento da câmera. Para controlar o movimento dos eixos foi desenvolvida uma linguagem chamada de GMC.

Figura 18 – Definição dos eixos do F2S2.

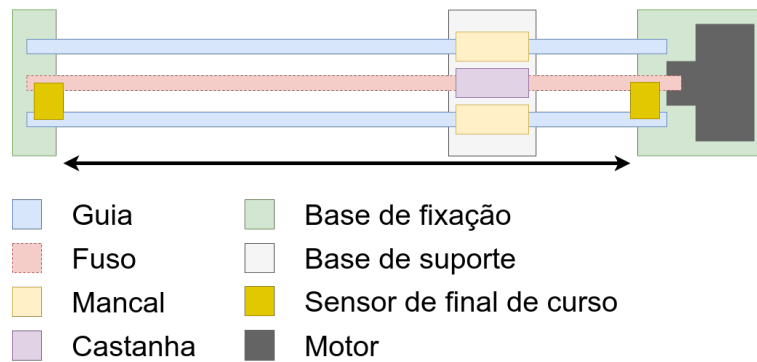


Fonte: Autoria própria.

Os eixos X' e Y' são responsáveis por movimentar linearmente a câmera, respectivamente na horizontal e vertical. Estes eixos foram construídos com base em um sistema de guias e fuso infinito, conforme a Figura 19.

Nestes sistemas existem um fuso infinito ligado ao eixo de um motor de passo. Ao acionar o motor, o fuso gira transferindo energia para uma castanha que está ligada à base de suporte, a qual também está ligada a duas guias paralelas (as duas guias e o fuso são todos paralelos). Desta forma ao transmitir movimento do fuso à castanha/base de suporte, o suporte

Figura 19 – Sistema de movimentação linear.



Fonte: Autoria própria.

só pode se mover acompanhando a direção das guias e fuso, ou seja, o suporte se aproxima ou afasta do motor na direção das guias/fuso. Dependendo do sentido de rotação do motor o suporte se aproxima ou se afasta.

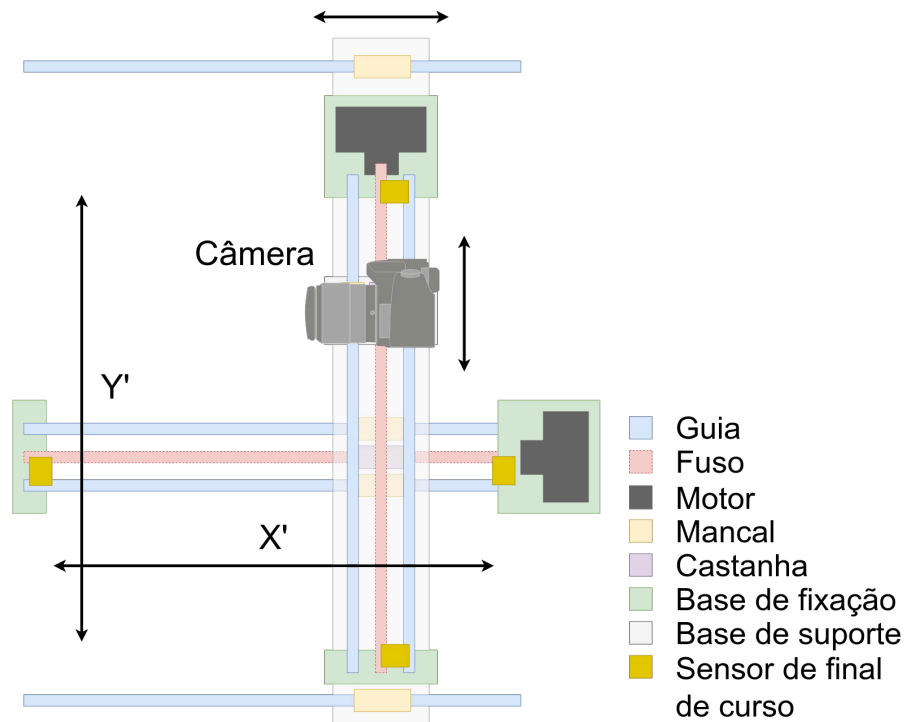
Em sistemas de fuso/castanha existe a conversão de determinada quantidade giros do fuso para uma quantidade de movimento da castanha, por exemplo: a cada volta completa do fuso (360°) a castanha se desloca 2mm. Esta característica da castanha/fuso é chamada de avanço. No final do percurso das guias existe um sensor de final de curso, o qual ao ser pressionado pelo suporte, envia um sinal que informa que o curso da base chegou ao fim. O sinal enviado pelos sensores é interpretado pelos componentes de controle.

Para que a câmera possa se movimentar simultaneamente nos eixos X' e Y' (direção horizontal e vertical), o sistema de movimentação do eixo Y' é montado sobre o sistema do eixo X' , de forma que ao acionar o motor do eixo X' , todo o sistema do eixo Y' é movimentado horizontalmente. A Câmera é acoplada ao sistema de movimentação do eixo Y' de forma que quando o sistema do eixo Y' é acionado, a câmera se movimenta na direção vertical. Também existem duas guias auxiliares ao movimento do eixo Y' para oferecer mais estabilidade ao movimento. Este processo pode ser visto na Figura 20.

Além dos eixos de movimentação linear, existem os eixos de movimentação angular (Z^a e B^a). Estes eixos não alteram a posição da câmera ou objeto quando acionados, e sim realizam a rotação das mesmas. Nestes eixos a base giratória está fixada diretamente ao eixo do motor de passo, de forma que ao acionar o motor, a base gira. Este giro pode ser no sentido horário ou anti-horário. a Figura 21 apresenta os componentes dos eixos de movimentação angular.

Os eixos angulares também contam com sensores para delimitar determinada posição do eixo. Porém no caso dos sistema de movimentação angulares optou-se pelo uso de sensores opticos em vez dos sensores de final de curso mecânicos. Neste caso o sensor detecta quando uma barreira física está a frente do sensor, marcando então a posição. O eixo B^a possui um sensor que informa o ângulo 0° e o eixo Z^a possui dois sensores, sendo que um informa a posição 0° e

Figura 20 – Sistema de movimentação da câmera.



Fonte: Autoria própria.

Figura 21 – Sistema de movimentação angular.



Fonte: Autoria própria.

o outro 90° .

Os eixos X' e Y' conseguem fazer com que a câmera descreva o movimento em forma de arco, ao redor do objeto, porém, ainda é necessário que a câmera aponte para o centro da esfera, ou seja o centro do objeto. Caso contrário, ao mover a câmera, o objeto sairia do campo de visão da mesma. Para realizar o movimento de inclinação da câmera, o eixo Z^a é instalado sobre o eixo Y' .

Assim, o eixo X' move o sistema de movimentação do eixo Y' na horizontal; o eixo Y' move o sistema do eixo Z^a na vertical; o sistema do eixo Z^a inclina a câmera para que esta aponte para o objeto; por fim o eixo B^a rotaciona o objeto para que todos os ângulos do objeto possam ser fotografados.

Cabe ainda ressaltar que os sensores presentes nos sistemas de movimentação lineares e

angulares não mapeiam a posição do sistema em determinado momento, e sim os limites iniciais e finais deste movimento. Para conhecer a posição dos sistemas em determinado momento, calcula-se o avanço dos sistemas e a quantidade de passos realizado pelo motor, assim o estado do sistema é conhecido.

Controle - Linguagem de controle do movimento - GMCodes

O sistema de controle do scanner se baseia em dois sistemas de processamento principais: Um computador que calcula as instruções necessárias para a aquisição e as transfere para o scanner; e um microcontrolador ATmega que está instalado no F2S2, o qual recebe as instruções do computador e as converte em ações que realizam as aquisições. O sistema de controle possui quatro componentes principais:

- **Software F2S2-App:** Instalado no computador, recebe o conjunto de informações do objeto e configuração da aquisição. Então este software calcula o conjunto de instruções na linguagem GMCodes necessárias para realizar a aquisição de forma automática. Após este processo o software se conecta ao scanner usando uma conexão serial por cabo USB ligado ao microcontrolador ATmega e envia as instruções para serem executadas pelo scanner. O F2S2-App será discutido na Seção 5.1.
- **Linguagem GMCodes:** Conjunto de instruções que o scanner pode executar. Podem ser instruções simples, como “mover eixo X para a posição 450” e instruções compostas, como “calibrar o scanner”, que realizam um conjunto de operações simples de forma ordenada. “Calibrar o scanner” significa mover todos os eixos para a posição de origem, ou seja valor zero.
- **Firmware do F2S2:** Software instalado no microcontrolador ATmega que converte as instruções em GMCodes para sinais elétricos que realizam as ações no scanner. O Firmware processa as instruções em GMCodes e se comunica com o F2S2-App enviando mensagens de início e conclusão de cada instrução realizada.
- **Hardware do F2S2:** Além do microcontrolador ATmega e dos sistemas de movimentação, o F2S2 possui outros componentes, como fonte de alimentação, drivers dos motores que respondem ao ATmega e acionam os motores dos sistemas de movimentação.

A parte eletrônica do F2S2 e a linguagem de controle do scanner (GMCodes) foram desenvolvidas pelo prof. Edson Justino em parceria com seus alunos do laboratório. O GMCodes é baseado na linguagem G-Code, a qual é usada em máquinas de Controle Numérico Computadorizado (CNC) e é comum em tornos computadorizados e impressoras 3D. G-Code baseia-se em movimentos geométrico em três dimensões, usando os eixos $G - Code(X, Y, Z)$.

No caso do GMCodes foi necessário acrescentar um eixo à linguagem G-Code, uma vez que o scanner trabalha com 4 graus de liberdade em seus movimentos $GMC(X', Y', Z^a, B^a)$. Além disso, foram necessárias fazer outras duas alterações: a conversão dos movimentos lineares dos eixos do G-Code para os movimentos angulares dos eixos Z^a e B^a do GMCodes; e a inclusão de instruções para a câmera fotográfica (MCodes). Por isso o nome escolhido para a linguagem foi “GMCodes”. O Código 3.1 apresenta um exemplo de instruções GMCodes.

código 3.1 – Exemplo de GMCodes.

```

1  ...
2  G00 B324
3  M989 T3
4  G00 B342
5  M989 T3
6  M999
7  G00 F4000
8  G00 X582 Y65
9  G00 Z30
10 G04 P4
11 G00 F500
12 G00 B0
13 M989 T3
14 ...

```

Fonte: Autoria Própria.

GMCodes se baseiam em um conjunto de instruções executadas na mesma ordem em que são enviadas ao firmware do scanner. Cada instrução é composta de um comando e os parâmetros de configuração, os quais são opcionais. Por exemplo: a linha 02 do Código 3.1 possui o comando de movimentação “G00” e o parâmetro “B234” que faz com que o eixo B^a se movimente até a posição angular de 234° . O Firmware aceita valores inteiros em milímetros para movimentos lineares e em graus para movimentos angulares.

As linhas 2, 4, 8, 9 e 12 apresentam instruções de movimentação dos eixos, sendo que o eixo é informado como parâmetro, seguido pelo valor em que o mesmo deve se posicionar. Este valor pode ser a distância em milímetros da origem do sistema, no caso dos eixos X e Y , ou em ângulos, no caso dos eixos Z e B . A instrução “M989 T3” faz com que o sistema aguarde 3 segundos e então a câmera adquira uma fotografia. Esta pausa de 3 segundos é necessário para estabilizar a peça e câmera que estavam se movendo, para evitar fotos tremidas.

A instrução “M999” realiza a calibração da base (eixo B), fazendo com que esta gire até o sensor optico informar que chegou na posição 0° . As instruções “G00 F” alteram a velocidade dos movimentos do scanner, sendo que quanto maior o valor, mais rápido será o movimento. A instrução “G04 P4” pausa o sistema por 4 segundos para estabilizar a câmera e o objeto.

À todas as instruções do GMCodes foi inserido um mecanismo de retorno que faz com

que o scanner envie um código de confirmação de que a instrução foi iniciada (mensagem: “<START>”) e que foi concluída (mensagem: “<END>”). Assim é possível impedir que as instruções sejam executadas fora de ordem. Considere por exemplo duas instruções: “*G00 X640*” e “*M989 T1*”. Quanto tempo a primeira instrução levaria para ser executada? Depende da posição anterior do eixo X' e da velocidade do movimento (configurada pela instrução *G00F*).

Assim, para garantir que a instrução que fotografa o objeto (“*M989 T1*”) não seja executada antes de que a instrução de movimentação anterior (“*G00 X640*”) seja finalizada, é usado o código de retorno que informa que a primeira instrução concluiu sua execução. Somente após a confirmação de que a instrução de movimentação foi concluída, é enviada e executada a segunda instrução, que tira a fotografia.

Este processo de confirmação de execução das instruções não estava presente nas primeiras versões do GMCodes e isso causou diversos erros, porém após a padronização de que uma instrução só seria enviada após a conclusão da execução da instrução anterior, os problemas foram resolvidos.

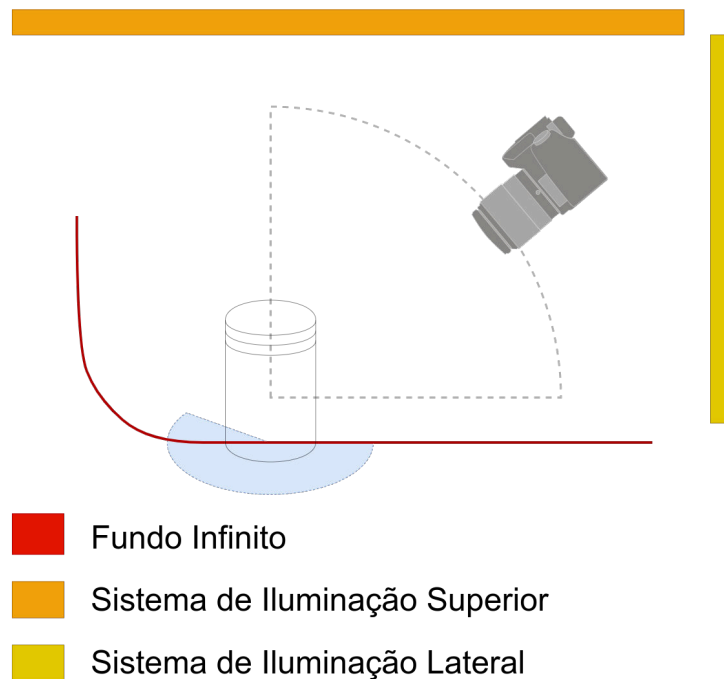
O funcionamento deste processo baseia-se em uma pilha, de forma que cada mensagem “<START>” é adicionada à pilha e cada mensagem “<END>” remove um “<START>” da pilha. Comandos compostos lançam diversas mensagens de “<START>” que são empilhadas e um novo comando só é enviado ao F2S2 após a pilha estar vazia. Comandos simples enviam um “<START>” e o próximo comando só é enviado após a conclusão deste comando com a mensagem “<END>”, independentemente do tempo que a instrução gaste.

Hardware - Componentes Fotográficos

O F2S2 foi idealizado para funcionar como um estúdio fotográfico, de forma que seja possível controlar o processo fotográfico para gerar as fotografias com as características desejadas. Diferentemente do processo de fotografar comum, é necessário considerar que o F2S2 adquire um conjunto de fotografias e estas precisam estar com as mesmas condições de fotometria para que durante a visualização não sejam notadas diferenças nas fotografias, além da movimentação do objeto.

Este processo deve ser feito de modo que sejam garantidas boas condições de iluminação para todas as imagens do conjunto de fotografia. Este procedimento será abordado na Seção 4.3. Além disso, pode ser necessário segmentar o objeto, separando-o do fundo das imagens. Assim, o F2S2 foi construído com os seguintes componentes fotográficos: fundo infinito; sistema de iluminação configurável; câmera. A Figura 22 apresenta um desenho esquemático destes componentes.

Figura 22 – Hardware - Componentes Fotográficos.



Fonte: Autoria própria. Figura da câmera obtida em: Openclipart (2015)

Fundo infinito

Durante a etapa de processamento o objeto que está sendo fotografado é separado do restante da fotografia. Este processo de segmentação é opcional e é discutido na Seção 5.2.1. No que se refere ao hardware do F2S2, este foi equipado com um fundo infinito de cor única. Este fundo trata-se de um plano curvo, recoberto com um material de cor única que ao ser fotografado causa a impressão de que o objeto está posicionado em um plano infinito. A Figura 23 apresenta um objeto fotografado com a técnica do fundo infinito.

Figura 23 – Imagem adquirida com fundo infinito.



Fonte: Autoria própria.

Na Figura 16 o fundo infinito pode ser visto recoberto por um tecido azul. Diversos materiais e cores podem ser usados para recobrir o fundo. A principal consideração a ser feita sobre a cor do fundo, é que esta seja diferente das cores do objeto, uma vez que cores diferentes auxiliam no processo de segmentação. Além disto, usa-se um material de cor única pelo mesmo motivo de facilitar o processo de segmentação.

Sistema de iluminação configurável

O F2S2 possui um sistema de iluminação configurável para proporcionar características de luz adequadas para a fotografia. Neste ponto do trabalho será explicado o hardware que permite esta configuração, na Seção 4.3.2 serão apresentados os procedimentos de medição da luz e configuração do sistema de iluminação.

A primeira consideração sobre a iluminação é que todas as luzes devem possuir a mesma temperatura de cor. Fotos comuns ou artísticas usam diferentes luzes para compor uma cena, contudo, tendo em vista que durante uma aquisição são fotografados diversos ângulos do objeto e para manter o Stream2D com a mesma configuração todas as lâmpadas devem ser da mesma temperatura de cor.

Para garantir o controle total da iluminação durante a aquisição, o F2S2 é fechado, de forma que a iluminação exterior não incida sobre o objeto que está sendo fotografado. Assim apenas as luzes do scanner iluminam o objeto que está sendo escaneado.

Luzes com temperaturas de cor diferentes causavam a impressão de superexposição (excesso de luz, ou seja fotos muito claras) ou subexposição (falta de luz, fotos escuras) durante a visualização. Isso ocorre uma vez que entre fotos diferentes, luzes de temperaturas diferentes incidem no objeto, causando a sensação de mudanças na iluminação. Portanto todo o conjunto de luzes do scanner possui a mesma temperatura de cor.

Outra consideração sobre o sistema de Iluminação é que este deve iluminar com a mesma intensidade todos os ângulos do objeto. Na fotografia convencional, ilumina-se a cena para causar determinado efeito na fotografia. Porém, como no Stream2D estão todos os ângulos de visualização do objeto, é necessário que estes ângulos estejam todos iluminados de forma homogênea. Caso contrário, determinadas fotos apareceriam mais claras ou escuras do que outros.

Para garantir uma iluminação homogênea e com a mesma temperatura de cor foram adquiridas lâmpadas LED no formato tubular, pois estas iluminam de forma mais uniforme se comparadas com lâmpadas bulbo ou *spot*, que por serem pontuais concentram a luz de forma mais centralizada. A temperatura de cor de todo o conjunto de lâmpadas é de 6.000K a 6.500K, que é considerado uma temperatura neutra.

As lâmpadas tubulares estão espalhadas acima (10 Lampadas) e afrente (5 lâmpadas) do objeto, de forma que durante todo o processo de aquisição a câmera passa entre o objeto

e a iluminação. Visando uma melhor difusão da luz incidente da parte superior, foi instalada uma placa acrílica translúcida sobre o sistema de lâmpadas. Cada lâmpada pode ser ligada ou desligada individualmente, permitindo assim a configuração do sistema de iluminação.

Além deste sistema de iluminação principal, existe a possibilidade de ligar dois sistemas de iluminação fixados paralelamente à câmera. Estes sistemas possuem uma luz mais direcional e são usados normalmente em objetos pequenos, uma vez que a câmera acaba sombreando estes objetos por estar muito próxima a eles.

Cabe ressaltar que na fotografia usa-se tradicionalmente flash e não luz contínua. Isso acontece pois a luz contínua acaba perturbando modelos e fotógrafos por causa da sua intensidade. Além disso, os sistemas de iluminação antigos emitiam muito calor, o que inviabilizava a luz contínua para diversos usos. Assim, tradicionalmente, na fotografia optou-se pelo uso de flash.

Como o F2S2 é todo fechado, para impedir que a iluminação externa interfira no processo de aquisição, a iluminação contínua também não ilumina o exterior do scanner, não perturbando quem está operando o scanner.

Porém, no caso do F2S2, a luz de flash possui alguns inconvenientes. O primeiro diz respeito ao tempo de carregamento que alguns equipamentos flash precisam entre cada disparo, o que faria com que o processo de aquisição de até quarenta mil fotos levasse muito tempo. Outra dificuldade é que os equipamentos de flash possuem um determinado número de disparos como sua vida útil. Assim, o número de disparos necessários para o F2S2 (até quarenta mil fotos por aquisição) acarretaria em custos financeiros maiores do que se comparado à iluminação contínua.

Câmera Fotográfica

No que se refere à câmera, esta pode ser substituída por quaisquer modelos, desde que possa ser controlada remotamente. Atualmente, usa-se um cabo USB conectado a câmera, o qual envia um sinal eletrônico à câmera que realiza o disparo. Ao substituir a câmera atual por outro modelo é necessário restabelecer o protocolo de disparo controlado por USB para padrão do hardware da câmera em uso.

Até o momento já foram usados os modelos Samsung NX500 e os modelos fabricados pela Nikon: D7100, D5300 e D5100. Existe uma relação direta entre a qualidade das imagens adquiridas e o modelo da câmera e lentes em uso no scanner. É preciso notar também que conjunto óptico também deve ser adequado à foto que se pretende obter.

Por exemplo: câmeras de cinco megapixels conseguem adquirir imagens de aproximadamente 2560×2048 . Estas imagens não seriam as mais adequadas para uso em um projetor *8K UHD*, uma vez que este consegue projetar imagens com 7680×4320 . Neste caso a imagem apareceria pequena (sem ocupar toda a projeção) ou quadriculada. Neste caso é necessário considerar também a resolução da tela, porém, invariavelmente a tela conseguiria representar muito mais informação do que a contida na fotografia.

O conjunto óptico também precisa ser adequado para fotografar o objeto que está sendo escaneado. Considere usar uma lente de comprimento focal 18mm, que consegue focalizar a uma distância mínima de 25cm. Tentar usar esta lente para adquirir imagens de uma formiga de 0.5cm não geraria bons resultados, uma vez que à 25cm da lente, esta formiga apareceria muito pequena e sem detalhes na fotografia.

As câmeras digitais possuem uma vida útil e esta característica está diretamente relacionada ao obturador da máquina. O obturador pode ser mecânico (mais comum nas câmeras atuais) ou eletrônico. Obturadores mecânicos da Nikon podem durar entre 100.000 a 400.000 fotos(US, 2018). Obturadores eletrônicos por sua vez não apresentam o desgaste mecânico e por isso tendem a durar mais.

Considerando que a Nikon testa seus obturadores para até 400.000 fotos e que uma aquisição pode ter até 40.000, uma câmera Nikon poderia precisar de manutenção no obturador com apenas 10 aquisições. Assim, recomenda-se o uso de máquinas com obturadores eletrônicos. Isso poderia economizar recursos financeiros e faria com que o F2S2 necessitasse de menos manutenção.

Existem no mercado diversos fabricantes e modelos de câmeras, com preços que variam de centenas de Reais a dezenas de milhares. A escolha de um determinado modelo de câmera precisa considerar obrigatoriamente a possibilidade de controle remoto e a quantidade de disparos que o equipamento consegue fazer, visto que existem equipamento testados para apenas 100.000 disparos, o que não representa nem três aquisições completas.

Porém também devem ser considerados o valor do equipamento, possibilidade de troca de lentes e modelos das lentes disponíveis, tamanho das imagens adquiridas, formato das imagens, qualidade das imagens (existem sites especializados que realizam a comparação de modelos de câmeras e lentes seguindo critérios bem definidos para atribuir uma nota à qualidade da imagem gerada pela câmera, como o site <<https://www.dpreview.com/>>), memória de armazenamento, entre outros.

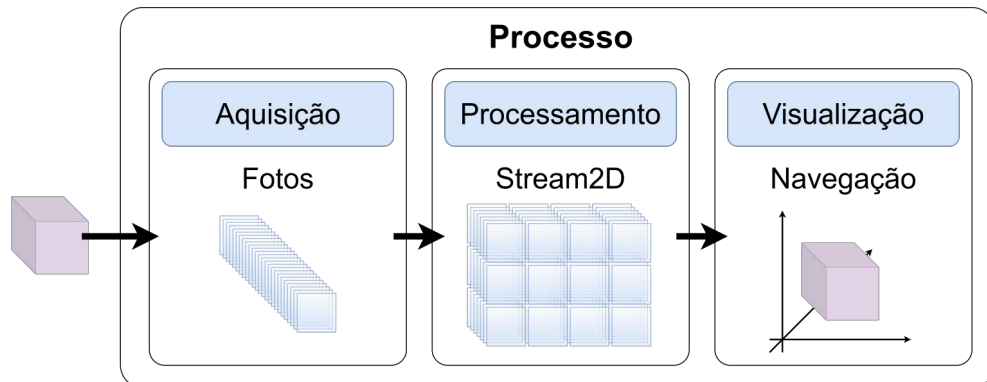
A escolha da câmera e lentes vai depender então do objetivo do uso do F2S2: se para telas com alta resolução ou telas comuns ou de celular; se para aplicações científicas que necessitam de riqueza de detalhes ou para ensino regular, onde a simples visualização do objeto poderá ser suficiente; se para escanear objetos muito pequenos (como insetos), usaria-se lentes macro, ou para objetos maiores usaria-se lentes comuns com distância focal 18mm-55mm.

3.2 Método proposto para o scanner F2S2

Para que seja possível gerar a visualização com as características almejadas, foi desenvolvido um método composto de três etapas: Aquisição, Processamento e Visualização. A Figura 24 apresenta cada etapa, a ordem de execução e o resultado destas etapas. Esta Seção apresenta uma

visão geral deste processo e as relações entre cada uma das etapas: Aquisição, Processamento e Visualização.

Figura 24 – Processo de digitalização de objetos pelo F2S2.



Fonte: Autoria própria.

A etapa de aquisição consiste em fotografar o objeto por todos os ângulos, respeitando um intervalo angular predeterminado entre cada fotografia. Desta forma é criada uma semiesfera em torno do objeto, com todos os pontos de vista fotografados. Assim, como produto do processo de aquisição, é construído um conjunto ordenado de fotografias (chamado de “*pack* de fotografias”) de forma que para cada fotografia adquirida é conhecida a posição relativa da câmera e do objeto que está sendo escaneado. Esta etapa está associada ao Objetivo Específico *I*.

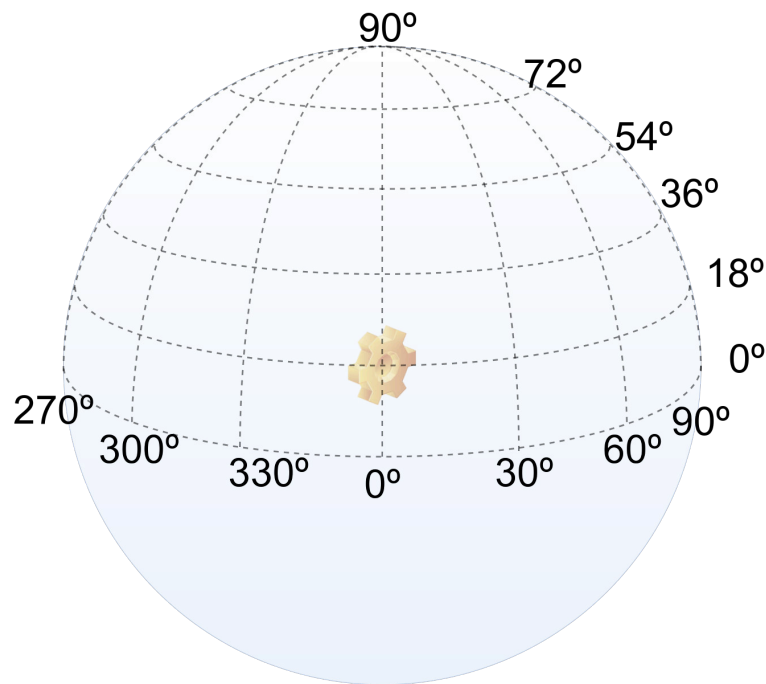
A aquisição é apresentado com profundidade na Seção 4.3 - Protocolo de Aquisição, onde é apresentado um conjunto de técnicas para garantir acurácia na aquisição das cores e nitidez nas imagens, como controle da iluminação, estabilidade do objeto e controle do tempo de disparo da câmera. Esta etapa está associada aos Objetivos Específicos *II* e *III*. O software de controle deste scanner é discutido na Seção 5.1 - Software de Aquisição - F2S2-App.

Para exemplificar este processo, considera-se a Figura 25, na qual a intersecção entre as linhas horizontais e verticais representam o conjunto de fotografias adquiridas durante a digitalização de um objeto. Na figura a engrenagem amarela posicionada no centro da esfera representa o objeto a ser digitalizado. Considerando um intervalo angular de 30° na direção horizontal são mapeadas 12 posições, considerando que a posição do ângulo 0° e 360° são coincidentes.

No exemplo é usado um intervalo angular de 18° na direção vertical, partindo do ângulo 0° até a posição 90° . Desta forma existem 6 posições na direção vertical. Considerando a configuração apresentada, seriam adquiridas 72 fotos do objeto com um intervalo angular na horizontal de 30° e 18° na vertical.

Ao obter o *pack* de fotografias de forma ordenada, primeira foto sendo os ângulos horizontal e vertical iguais a 0° , a segunda com ângulo horizontal sendo 30° e vertical 0° , assim

Figura 25 – Semiesfera criada durante o processo de digitalização do objeto.



Fonte: Autoria própria.

sucessivamente até que se altere o ângulo horizontal em 18° ; então repete-se o processo até concluir as 72 fotografias.

Assim, conhecendo a ordem em que as imagens foram adquiridas e os intervalos angulares pré-definidos, é possível calcular a posição da câmera em relação ao objeto para cada uma das fotografias deste conjunto. O conjunto ordenado de fotografias adquiridas é o resultado final do processo de Aquisição.

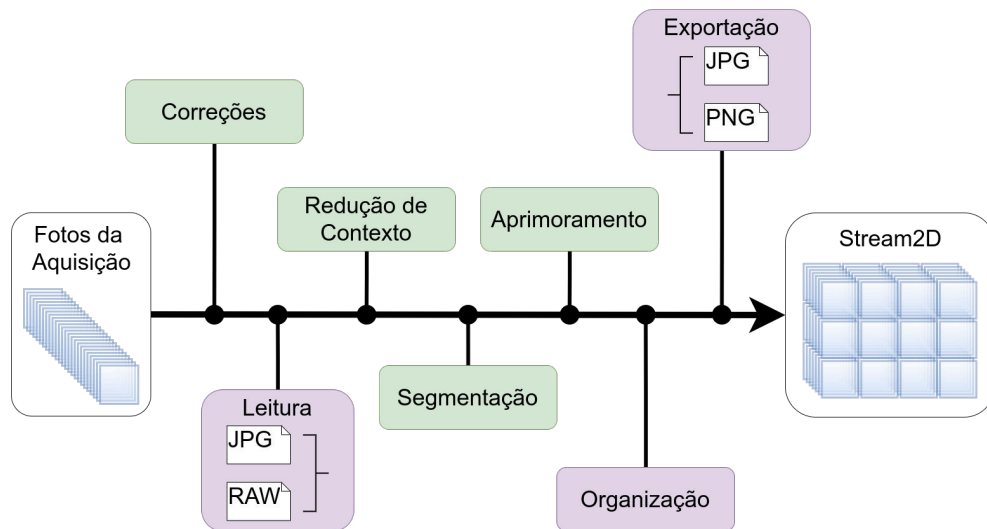
Após a etapa de aquisição, o conjunto de fotografias adquiridas é processada para que possa ser usada na visualização. O processamento destas fotografias é dividido em 7 partes, sendo: Correção, Leitura, Redução de contexto, Segmentação, Aprimoramento, Organização e Exportação. Após o processamento é construído o Stream2D, estrutura que é usada para gerar a visualização.

A Figura 26 apresenta as 7 fases do processamento. As etapas representadas em roxo (leitura, organização e exportação) são obrigatórias e as representadas em verde (correção, redução de contexto, segmentação e aprimoramento) são opcionais e sua execução depende do resultado que se pretende para a visualização.

Correções

O processo de fotografia possui problemas intrínsecos, como a criação de vinheta em fotografias, distorções causadas pela lente e na aquisição de cores que não são fiéis. Dentre estes,

Figura 26 – Diagrama das etapas de processamento.



Fonte: Autoria própria.

a vinheta e as distorções causadas pelas lentes não são dependentes de técnicas ou equipamentos externos, podendo ser corrigidos por software.

Considerando que estas características são originadas pelo hardware das câmeras, alguns fabricantes oferecem softwares capazes de minimizar estes efeitos. Considerando a preexistência destes softwares, optou-se pelo seu uso no processo de correção das imagens, em vez do desenvolvimento de software próprio.

Leitura

Após a correção, o software de processamento chamado de *F2S2-BipApp* recebe um *pack* de fotografias, que podem ser arquivos no formato *.jpg* ou “RAW”(cru em português). “RAW” é o formato de arquivo que salva informações do sensor de câmeras digitais e não uma imagem propriamente dita, sendo que é necessário tratar estes dados para que seja possível gerar a imagem fotografada. Fotografias “RAW” armazenam mais informações do que imagens bitmap e permitem um processamento/edição mais robusto se comparado a imagens *.jpg*.

“RAW” é um formato de arquivos definido pela Adobe Inc.¹, porém não é uma extensão, cada fabricante de câmeras usa o padrão definido pela Adobe, que possui extensão *.dng* (*Digital Negative - Negativo Digital*) e o adéqua para as câmeras que estão sendo produzidas. Assim, cada fabricante possui arquivos “RAW” próprios, como por exemplo: Nikon - *.nef* ou *.nrw*; Canon - *.crw* e *.cr2*; Sony - *.arw*, *.srf* e *.sr2*; Samsung - *.srw*.

Então, esta etapa do processamento consiste em reconhecer a extensão das fotografias e se necessário, no caso de imagens “RAW”, identificar o formato e converter estes dados para formatos de imagens.

¹ <https://helpx.adobe.com/br/photoshop/digital-negative.html>

Redução de contexto

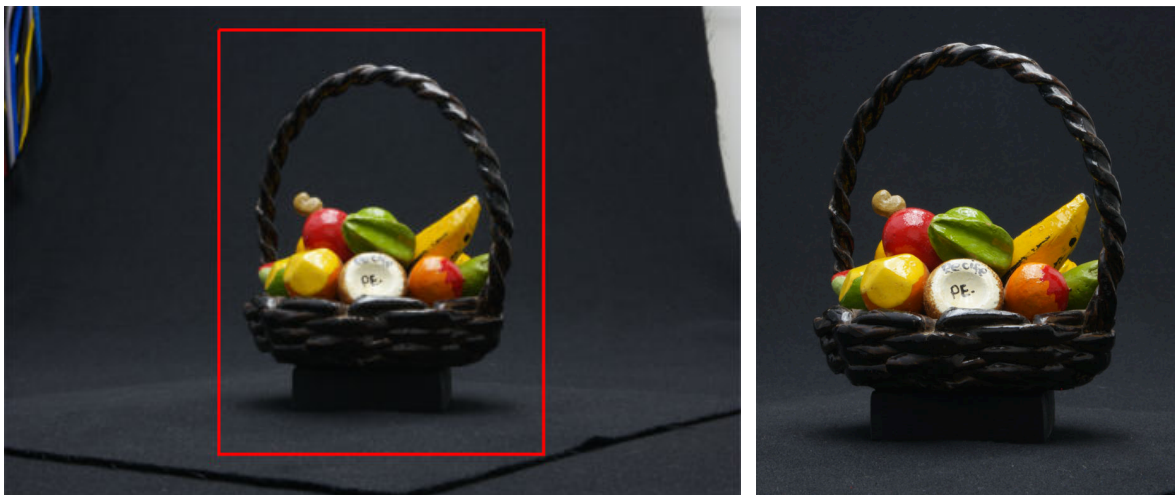
A redução de contexto é responsável por recortar as fotografias de forma que partes das imagens onde não apareçam o objeto de interesse sejam removidos, reduzindo o tamanho dos arquivos e o custo computacional de processamento das imagens, por isso este é o processo executado após o carregamento das imagens. Este recorte da redução de contexto é um retângulo extraído da fotografia original, com coordenadas configuradas pelo usuário do software.

A Figura 27 apresenta o processo de redução de contexto, sendo a Figura 27(a) a figura completa, com uma marcação em vermelho da região onde ocorre o recorte da fotografia; e a Figura 27(b) a fotografia com contexto reduzido.

Figura 27 – Processo de redução de contexto.

(a) Fotografia original com marcação da região a ser cortada.

(b) Fotografia com contexto reduzido.



Fonte: Autoria própria.

Segmentação

A segmentação da fotografia separa o objeto de todos os outros componentes da fotografia. Foi usado um algoritmo classificador SVM semi-supervisionado com duas classes: Fundo e Objeto. O processo de seleção do algoritmo de segmentação, suas configurações e dificuldades são apresentados no Apêndice A - Segmentação em lote de imagens com fundo de cor única. O software de segmentação implementado é apresentado na Seção 5.2 - Software de Processamento - F2S2-BipApp. Esta etapa está associada ao Objetivo Específico V. A Figura 28 apresenta: uma fotografia já com o contexto reduzido 28(a) e a fotografia segmentada 28(b).

Aprimoramento

O aprimoramento é um conjunto de operações que pode ser realizado sobre as imagens para melhorar a visualização ou reduzir problemas da etapa de fotografia ou segmentação. As operações que podem ser realizadas são:

Figura 28 – Processo de segmentação.

(a) Fotografia original com contexto reduzido.



(b) Fotografia segmentada.



Fonte: Autoria própria.

Erosão : Após a segmentação pode acontecer de os pixels que limitam o objeto com o fundo aparecerem no tom do fundo. Isso ocorre pois aqueles pixels representam luz que refletiu ao mesmo tempo do fundo e do objeto. A erosão remove uma camada de pixels mais externos do objeto resolvendo este problema.

Suavização de Borda : Após a segmentação o limite entre o objeto e o fundo pode ficar quadrado. A suavização de borda torna este limite mais suave, melhorando a sensação durante a visualização.

Remoção de pixels esparsos : A segmentação pode manter algumas regiões do fundo na imagem resultante. Para resolver este problema são calculadas as áreas de cada região não conectada após a segmentação. Então é preservada apenas a maior região, eliminando as outras, de forma que apenas o objeto aparece na imagem segmentada.

Os algoritmos de aprimoramento e suas configurações serão apresentados na Subseção 5.2.1. Estes algoritmos só podem ser usados quando for realizada a segmentação das fotografias. Esta etapa e a etapa de correção das imagens estão associadas ao Objetivo específico *IV*

Organização

A etapa de organização é a etapa de construção do Stream2D e consiste em atribuir um identificador único para cada fotografia, de forma que seja possível mapear a posição de qualquer fotografia dentro do *pack*. Esta etapa é obrigatória dentro do processamento das imagens. O processo de construção do Stream2D é demonstrado em detalhes na Subseção 4.1.2 - Modelo tridimensional - Stream2D. Na Seção 4.1 - Sistemas de Coordenadas, Stream2D e Simulação - é apresentado um conjunto de equações que permitem mapear a posição de cada fotografia

desde a etapa de aquisição até a visualização do Stream2D. Esta etapa está associada ao Objetivo Específico VI.

Exportação

A ultima etapa consiste em exportar a fotografia processada e nomeada com o identificador atribuído na etapa de organização para um arquivo que será usado na visualização. As fotografias podem ser exportadas para os formatos “JPG” ou “PNG”, sendo que o formato “JPG” possui o fundo na cor preta e seus arquivos são menores, ocupando menos espaço em disco se comparados ao formato “PNG”. Porém este ultimo permite que o fundo seja transparente, pois além dos canais RGB, possui o canal *alpha*. Imagens com fundo transparente podem ser sobrepostas sobre outros fundos, permitindo mais liberdade na visualização.

4 Proposta

Este Capítulo apresenta como este trabalho foi elaborado, apresentando e discutindo cada etapa do seu desenvolvimento. Inicialmente, na Seção 4.1, são apresentados os sistemas de coordenadas que permitem mapear a posição de cada fotografia adquirida pelo F2S2 em sistemas de coordenadas internas do scanner e no sistema de coordenadas esféricas, além do conjunto de equações que converte coordenadas entre estes sistemas. Estas informações são importantes para permitir que durante o restante deste capítulo, seja possível discutir as etapas deste trabalho.

Os sistemas de coordenadas são essenciais para que seja possível construir o Stream2D, apresentado na Subseção 4.1.2, que é a estrutura que possui todas as informações necessárias para viabilizar a visualização dos objetos em ambientes de simulação. O uso do Stream2D em ambientes de simulação é apresentado na Subseção 4.1.3.

Após apresentar como as imagens adquiridas podem ser usadas na visualização, na Seção 4.2, discute-se como o F2S2 é usado para gerar estas imagens, as principais dificuldades enfrentadas, soluções para os problemas encontrados e como estas dificuldades interferem no resultado final da visualização.

Em seguida, na Seção 4.3, apresenta-se o protocolo desenvolvido para a aquisição fotográfica, discutindo os objetos a serem adquiridos; a implicação das características destes objetos no processo de aquisição e visualização; a configuração do scanner e do software de controle do scanner para uma aquisição e visualização adequadas.

4.1 Sistemas de Coordenadas, Stream2D e Simulação

A visualização dos objetos digitalizados pelo F2S2 baseia-se principalmente em: dadas as posições do observador e do objeto virtual em um ambiente de computação gráfica, identificar qual fotografia adquirida melhor representa o ângulo de visualização do observador para aquele objeto. Por este motivo é fundamental conseguir mapear a posição de cada fotografia adquirida pelo F2S2.

Esta Seção inicia (Subseção 4.1.1) apresentando os sistemas de coordenadas que representam o estado interno do scanner, a posição da câmera em relação ao objeto e em seguida a relação entre objeto, câmera e um espaço tridimensional, além do conjunto de equações que permitem a conversão de um ponto entre estes sistemas de coordenadas.

Usando os sistemas de coordenadas apresentados, na Subseção 4.1.2 é descrito o processo de construção do Stream2D. Na criação do Stream2D cada fotografia recebe um identificador único que permite associar a fotografia a uma posição no espaço tridimensional, além de permitir a conversão de um identificador único de foto para os sistemas de coordenadas apresentados

anteriormente.

Por fim a Subseção 4.1.3 demonstra como o Stream2D pode ser usado em ambientes de simulação baseados em computação gráfica tridimensional. Para isso são definidas: a visualização interativa, visualização interativa estereoscópica, a aplicação do Stream2D em ambientes tridimensionais e por fim é debatido o processo de conversão entre todos os sistemas de coordenadas apresentados.

4.1.1 Sistema de Coordenadas

Para mapear a posição de cada fotografia é necessário entender como as coordenadas GMC, que determinam o estado de cada eixo do scanner em um instante; F2S2, que informam a relação entre a posição da câmera e do objeto; e as coordenadas esféricas, que representam a relação entre o objeto e um ambiente tridimensional; estão associadas e como pontos destes sistemas de coordenadas podem ser convertidos de um espaço para outro.

Assim, esta Subseção apresenta estes sistemas de coordenadas e as equações necessárias para a conversão entre estes sistemas. Esta informação é necessária para que o restante deste capítulo possa referenciar estes sistemas e discutir a implicação de suas características na visualização do objeto virtual.

Sistema de Coordenadas GMC

Como já apresentado na Figura 18, o eixo X' é responsável pelo movimento da câmera na direção horizontal, sendo que no sentido da direita para esquerda e seu valor aumenta. A origem deste eixo é a interseção com o eixo Y' . O eixo Y' é executado um movimento na direção vertical e no sentido de baixo para cima o eixo possui um incremento de valor. Sua origem é no cruzamento com o eixo X' .

Os eixos com movimento linear do sistema de coordenadas (X' e Y') não assumem valores negativos e possuem um limite superior com valor de 650. Infringir estes valores, significaria um colisão com outros componentes do F2S2. Cabe apontar que com $X' = 650$ a câmera está posicionada exatamente sobre o eixo B^a . Além disso, estes valores são referentes ao atual scanner. de forma que na construção de outro equipamento estes valores podem ser alterados. Fisicamente, no F2S2, estes valores representam unidades em milímetros.

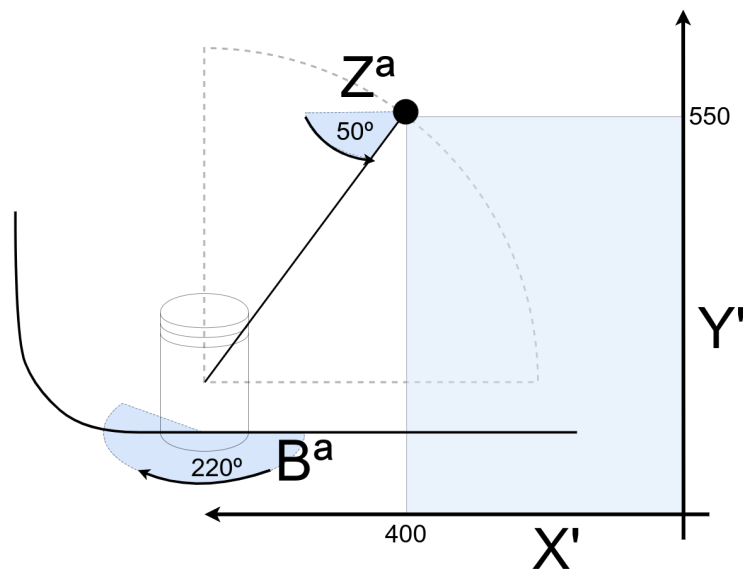
Quanto aos eixos responsáveis pelas movimentações angulares do F2S2, o eixo Z^a é móvel, ou seja, sua posição varia de acordo com os valores dos eixos X' e Y' . Em outras palavras, as coordenadas de X' e Y' representam a posição do eixo Z^a no plano formado por X' e Y' . Com $Z^a = 0^\circ$ sua direção é horizontal e seu sentido é da direita para a esquerda, seu valor máximo é 90° e seu valor aumenta seguindo um movimento no sentido anti-horário, de forma que em seu maior valor sua direção é vertical sentido de cima para baixo. No F2S2 o eixo Z^a é a câmera e sua inclinação faz com que a mesma aponte sempre para o objeto.

Quanto ao eixo B^a , sua posição é fixa na coordenada $X' = 650$ e $Y' = 0$. O sentido do seu movimento é horário. A Equação 4.1 apresenta o conjunto de valores que cada eixo descrito pode assumir neste sistema de coordenadas. Usa-se o conjunto de números naturais (\mathbb{N}) e não reais (\mathbb{R}) pois esta é a precisão conseguida pelo F2S2. Cabe notar que o valor mínimo para o Eixo $Y' = 80$, isto ocorre pois considerando que a coordenada Y' mede o centro da câmera, com valores menores que 80, a câmera se chocaria com o fundo infinito.

$$\begin{aligned}
 X' &= \{X' \in \mathbb{N} : 0 < X' < 650\} \\
 Y' &= \{Y' \in \mathbb{N} : 80 \leq Y' < 650\} \\
 Z^a &= \{Z^a \in \mathbb{N} : 0^\circ \leq Z^a \leq 90^\circ\} \\
 B^a &= \{B^a \in \mathbb{N} : 0^\circ \leq B^a < 360^\circ\}
 \end{aligned}
 \tag{4.1}$$

A Figura 29 exemplifica o deslocamento dos eixos. Neste exemplo considere os seguintes valores: $X' = 400$; $Y' = 550$; $Z^a = 50^\circ$; e $B^a = 220^\circ$. Além disso, também deve ser percebido que o centro do arco formado coincide com o centro geométrico do objeto que está sendo escaneado e não com a origem do eixo B^a . A motivação e o cálculo da posição do centro deste arco é apresentado na Seção 5.1.

Figura 29 – Deslocamento da câmera após instruções GMC.



Fonte: Autoria própria.

Sistema de Coordenadas F2S2

O sistema de coordenadas GMC representa o conjunto de movimentos realizados pelo F2S2 durante uma Aquisição, porém para representar a posição de uma foto com relação a determinado objeto podem ser usados outros sistemas de coordenadas, que representam estas

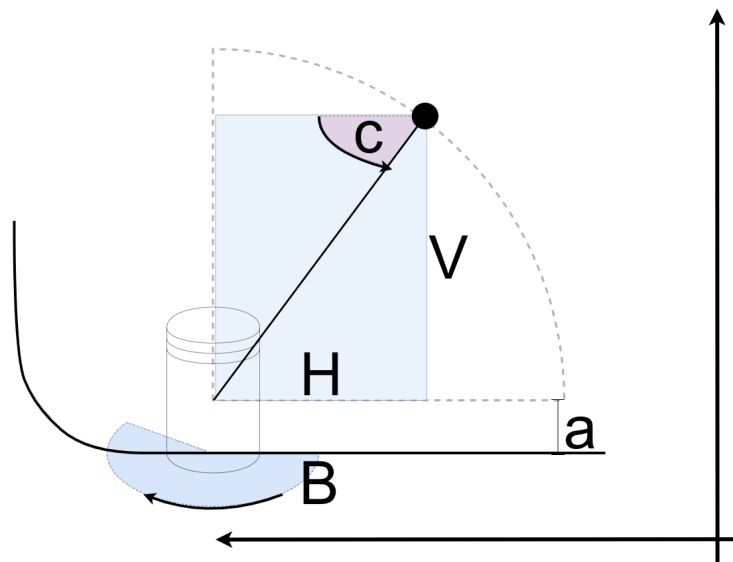
posições de forma mais simples. Isso ocorre pois o sistema GMC usa valores específicos da construção do Scanner. Além disso, a origem do sistema não é no objeto em si, e sim baseada no Scanner.

Para simplificar o processo de mapeamento da posição de cada foto em relação ao centro geométrico do objeto podem ser usados sistemas de coordenadas como o sistema esférico e/ou o sistema de coordenadas F2S2. O sistema de coordenadas F2S2 simplifica o processo de cálculo das coordenadas de cada fotografia de um conjunto gerado pela Aquisição, ao mesmo tempo que permite a compreensão das condições do scanner em determinado instante.

A principal diferença entre estes dois sistemas é que o sistema de coordenadas F2S2 possui como origem o centro geométrico do objeto e não componentes do scanner, de forma que as medidas manipuladas neste sistema são medidas de distância da câmera ao objeto, e não distâncias entre componentes do scanner, como acontece no modelo GMC. Cabe explicar que é possível fazer a conversão de valores do sistema GMC para o F2S2 e vice versa, como será apresentado na seção 4.1.1.

Para representar um ponto no sistema de coordenadas F2S2 são usados 3 componentes: $F2S2 = (H, V, B)$, sendo: H responsável pelo deslocamento na direção horizontal; V representa o movimento vertical; B indica a rotação do objeto sobre o próprio eixo. As três coordenadas usam como origem o centro geométrico do objeto. A Figura 30 apresenta um ponto no sistema.

Figura 30 – Sistema de Coordenadas F2S2.



Fonte: Autoria própria.

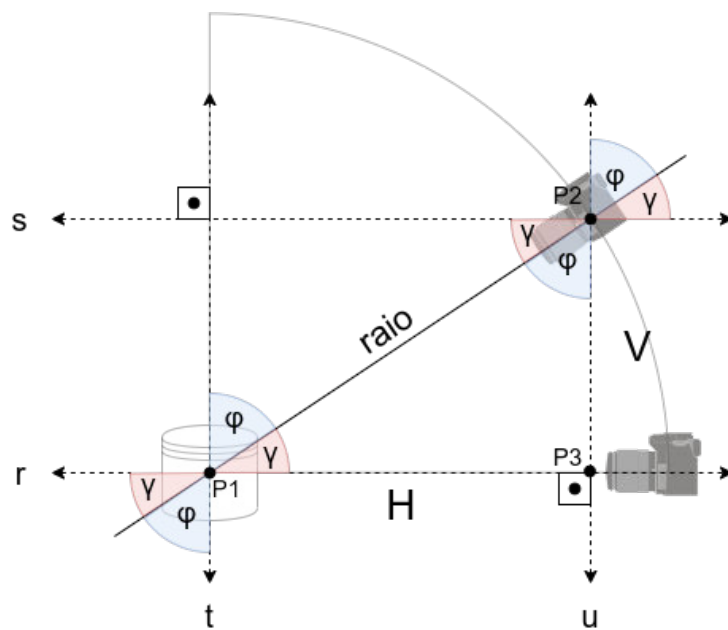
Na figura também aparecem duas outras coordenadas: “ c ” e “ a ”. “ a ” representa a altura da esfera com relação à base do scanner, a qual serve para posicionar o centro da esfera no centro geométrico do objeto. Como já mencionado, a motivação e o processo de cálculo desta medida é apresentado na seção 5.1. A unidade “ a ” não é necessária para compreender o posicionamento

da câmera em relação ao objeto, mas é necessária para realizar a conversão do sistema F2S2 para o GMC.

Outra coordenada presente na figura é “ c ”, a qual representa a inclinação da câmera para que esta aponte para o objeto que está sendo escaneado. Esta coordenada não faz parte das coordenadas do sistema F2S2 pois para determinado par de valores de H e V , só existe um valor possível para “ c ”. Além disto, cabe esclarecer que este valor pode ser calculado por trigonometria.

Considere a Figura 31. Sendo $r//s$ e $t//u$ de forma que a interseção das retas r e t está posicionada no centro geométrico do objeto e a interseção entre s e u está posicionada em determinada posição da câmera, ou seja: u possui um deslocamento horizontal para a direita de distância H em relação à t ; e s possui um deslocamento vertical para cima com deslocamento V em relação a r .

Figura 31 – Ângulos complementares e colaterais.



Fonte: Autoria própria.

Esta condição faz com que os ângulos φ e γ sejam complementares. Além disso, existe a relação de ângulos colaterais na interseção entre as retas r e t e na interseção entre s e u . Assim, ao calcular o valor angular de γ , calcula-se ao mesmo tempo a inclinação da câmera para adquirir a foto do objeto que está sendo escaneado e a variação angular da foto com relação ao ângulo vertical 0° . Para calcular o valor de γ pode-se usar a relação de Tangente, conforme as Equações 4.2 e 4.3:

$$Tg \gamma = \frac{\text{cateto oposto}}{\text{cateto adjacente}} = \frac{V}{H} \quad (4.2)$$

$$\gamma = \text{Arco } Tg (Tg \gamma) \quad (4.3)$$

Sendo:

- $Tg \gamma$: Valor da função Tangente do arco γ .
- *cateto oposto*: Deslocamento V no sentido vertical.
- *cateto adjacente*: Deslocamento H no sentido horizontal.
- $\text{Arco } Tg()$: Função Arco Tangente - Retorna o valor de um ângulo em graus a partir de um determinado valor de tangente.

O sistema de coordenadas F2S2 permite ainda calcular o raio da esfera formada, ou seja a distância entre a câmera e o centro geométrico do objeto. Ainda considerando a Figura 31, esta medida está representada pela hipotenusa do triângulo retângulo formado pelos pontos $(P1, P2, P3)$. Assim, pode-se usar o Teorema de Pitágoras para calcular seu valor, conforme a Equação 4.4:

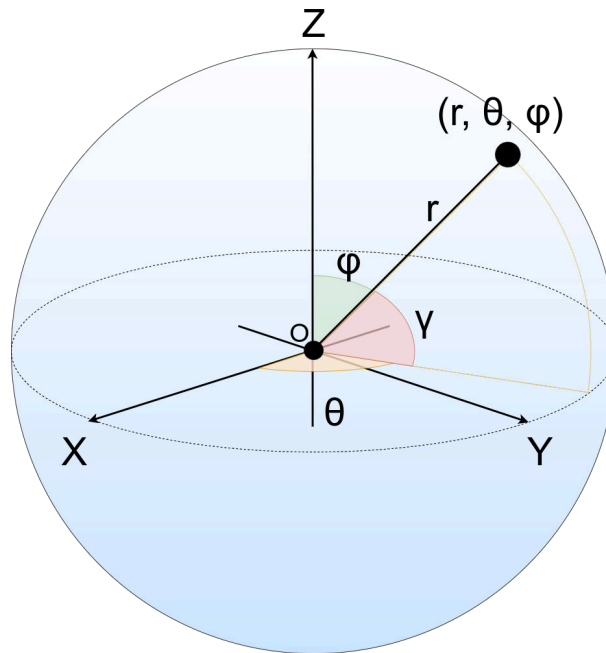
$$\text{raio}^2 = H^2 + V^2 \quad (4.4)$$

Sistemas de Coordenadas Esféricas

O sistema de coordenadas esférico representa um ponto em um espaço esférico usando as coordenadas $P = (r, \Theta, \varphi)$, sendo r o raio da esfera, ou seja a distância entre o ponto P e a origem do sistema; Θ representa a variação angular longitudinal, também conhecido como azimute; φ é o ângulo polar, também conhecido como colatitude ou ângulo zenital.

Como pode ser visto na Figura 32, Θ representa o ângulo formado pelo eixo X e a projeção de P no plano formado pelos eixos (X, Y) . No caso do F2S2, Θ representa a rotação do objeto em seu próprio eixo. φ é o ângulo formado pela linha que une o ponto P à origem e o eixo Z ($\sphericalangle POZ$).

Considerando que a aquisição do scanner inicia com $\varphi = 90^\circ$ e termina em $\varphi = 0^\circ$, para facilitar o processo de cálculo da posição das fotografias em relação ao objeto digitalizado, define-se γ como sendo o ângulo complementar de φ . A Equação 4.5 apresenta o processo de cálculo de γ . Usando γ em vez de φ o valor da variação angular na direção horizontal inicia com $\gamma = 0^\circ$ e seu valor é incrementado ao longo do processo de aquisição até o valor de $\gamma = 90^\circ$.

Figura 32 – Sistema esférico de coordenadas - com γ definido.

Fonte: Autoria própria.

$$\gamma = 90^\circ - \varphi \quad (4.5)$$

Para que o modelos de coordenadas esféricas torne-se mais próximo do modelos da aquisição, é necessário definir os limites para um modelo de semiesfera, em vez de um modelo esférico. Assim, os limites dos valores de r , Θ , φ e γ para a atual versão do scanner são definidos pela Equação 4.6:

$$\begin{aligned} r &= \{r \in \mathbb{N} : 0 < r < 650\} \\ \Theta &= \{\Theta \in \mathbb{N} : 0^\circ \leq \Theta < 360^\circ\} \\ \varphi &= \{\varphi \in \mathbb{N} : 0^\circ \leq \varphi \leq 90^\circ\} \\ \gamma &= \{\gamma \in \mathbb{N} : 0^\circ \leq \gamma \leq 90^\circ\} \end{aligned} \quad (4.6)$$

Conversão entre sistemas de coordenadas

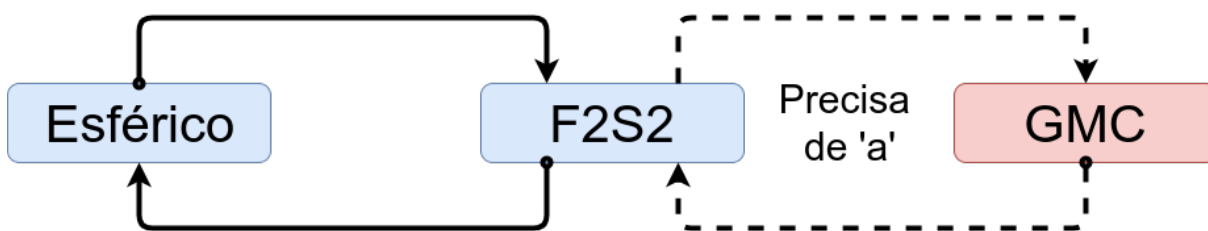
Como mencionado, o processo de aquisição tem como objetivo adquirir um conjunto de fotografias, de forma que para cada fotografia é conhecida a posição da câmera em relação ao objeto. O processo de mapeamento desta posição é realizado através do processo de conversão de sistemas de coordenadas.

Em outras palavras: conhecendo o estado do scanner (sistema de coordenadas GMC) em determinada fotografia, é possível definir a posição da câmera em relação ao objeto (sistemas de

coordenadas esféricas e F2S2) através do processo de conversão de coordenadas de um sistema para o outro.

A Figura 33 apresenta os sistemas de coordenadas já definidos e as conversões mapeadas. Os sistemas de coordenadas em azul apresentam de forma mais direta a posição da câmera em relação ao objeto, se comparado com o sistema GMC (em vermelho). As linhas tracejadas indicam que para realizar a conversão são necessárias informações do processo de aquisição que não estão presentes nos eixos representados pelos sistemas de coordenadas, no caso a variável “ a ” que representa a altura da esfera.

Figura 33 – Conversões entre os sistemas de coordenadas GMC, F2S2 e Esférico.



Fonte: Autoria própria.

Conversão de coordenadas entre os sistemas GMC e F2S2

Considerando os sistemas de coordenadas $GMC = (X', Y', Z^a, B^a)$ e $F2S2 = (H, V, B)$, a Equação 4.7 apresenta a conversão de coordenadas no sistema GMC para o sistema F2S2; a Equação 4.8 apresenta a conversão que retorna as coordenadas do sistema F2S2 para o sistema GMC.

$$\begin{aligned}
 (X', Y', Z^a, B^a) &\Rightarrow (H, V, B) \\
 H &= 650 - X' \\
 V &= Y' - a \\
 B &= B^a
 \end{aligned} \tag{4.7}$$

$$\begin{aligned}
 (H, V, B) &\Rightarrow (X', Y', Z^a, B^a) \\
 X' &= 650 - H \\
 Y' &= V + a \\
 Z^a &= c = \gamma = \text{ArcoTg} \left(\frac{V}{H} \right) \\
 B^a &= B
 \end{aligned} \tag{4.8}$$

Conversão de coordenadas entre os sistemas de coordenadas F2S2 e esférico

Considerando os sistemas de coordenadas $F2S2 = (H, V, B)$ e $Esférico = (r, \Theta, \varphi)$, a Equação 4.9 apresenta a conversão de coordenadas no sistema esférico para o sistema F2S2; a Equação 4.10 apresenta a conversão do sistema F2S2 para o sistema esférico.

$$\begin{aligned}
 (r, \Theta, \varphi) &\Rightarrow (H, V, B) \\
 \gamma &= 90^\circ - \varphi \\
 H &= r \times \text{Sen}(\gamma) \\
 V &= r \times \text{Cos}(\gamma) \\
 B &= \Theta
 \end{aligned} \tag{4.9}$$

$$\begin{aligned}
 (H, V, B) &\Rightarrow (r, \Theta, \varphi) \\
 \gamma &= \text{ArcoTg}\left(\frac{V}{H}\right) \\
 r &= \sqrt{H^2 + V^2} \\
 \Theta &= B \\
 \varphi &= 90^\circ - \gamma
 \end{aligned} \tag{4.10}$$

4.1.2 Modelo tridimensional - Stream2D

A definição de como deveria ser o Stream2D nasceu no laboratório, por um consenso entre os membros do laboratório responsáveis pela aquisição e visualização. Este formato almeja simplificar o processo de visualização. Esta Subseção apresenta os conceitos envolvidos na criação do Stream2D.

Com o objetivo de simplificar a explicação do Stream2D é usado um exemplo com quatro fotografias. Assim o intervalo angular horizontal é de 180° , ou seja, uma foto da frente do objeto e outra das costas. O intervalo vertical é de 90° , gerando uma posição no mesmo nível do objeto (0°) e outra em 90° com o objeto sendo fotografado de cima para baixo.

As quatro fotografias adquiridas seriam nomeadas linearmente pela câmera. Usando o padrão de nomes usado pela câmera do laboratório os nomes das imagens ficariam como representados na Figura 34. Sendo as imagens `__SAM0001.SRW` e `__SAM0002.SRW` fotografias com a câmera na mesma posição horizontal, a primeira capturando a frente do objeto e a segunda as costas. As fotografias `__SAM0003.SRW` e `__SAM0004.SRW` seriam com a câmera em cima do objeto, fotografando de cima para baixo e a diferença entre elas seria um giro de 180° do objeto.

Dada esta configuração, é possível organizar as fotografias de acordo com a posição da câmera em relação ao objeto fotografado. Um grupo de fotografias possui a mesma variação angular na horizontal, sendo o grupo 01 as Figuras `__SAM0001.SRW` e `__SAM0002.SRW` e o

Figura 34 – Nomes das fotografias adquiridas.



Fonte: Autoria própria.

grupo 02 as Figuras `__SAM0003.SRW` e `__SAM0004.SRW`. Quanto à variação angular vertical as imagens do grupo 01 seriam as Figuras `__SAM0001.SRW` e `__SAM0003.SRW` e o grupo 02 as Figuras `__SAM0002.SRW` e `__SAM0004.SRW`.

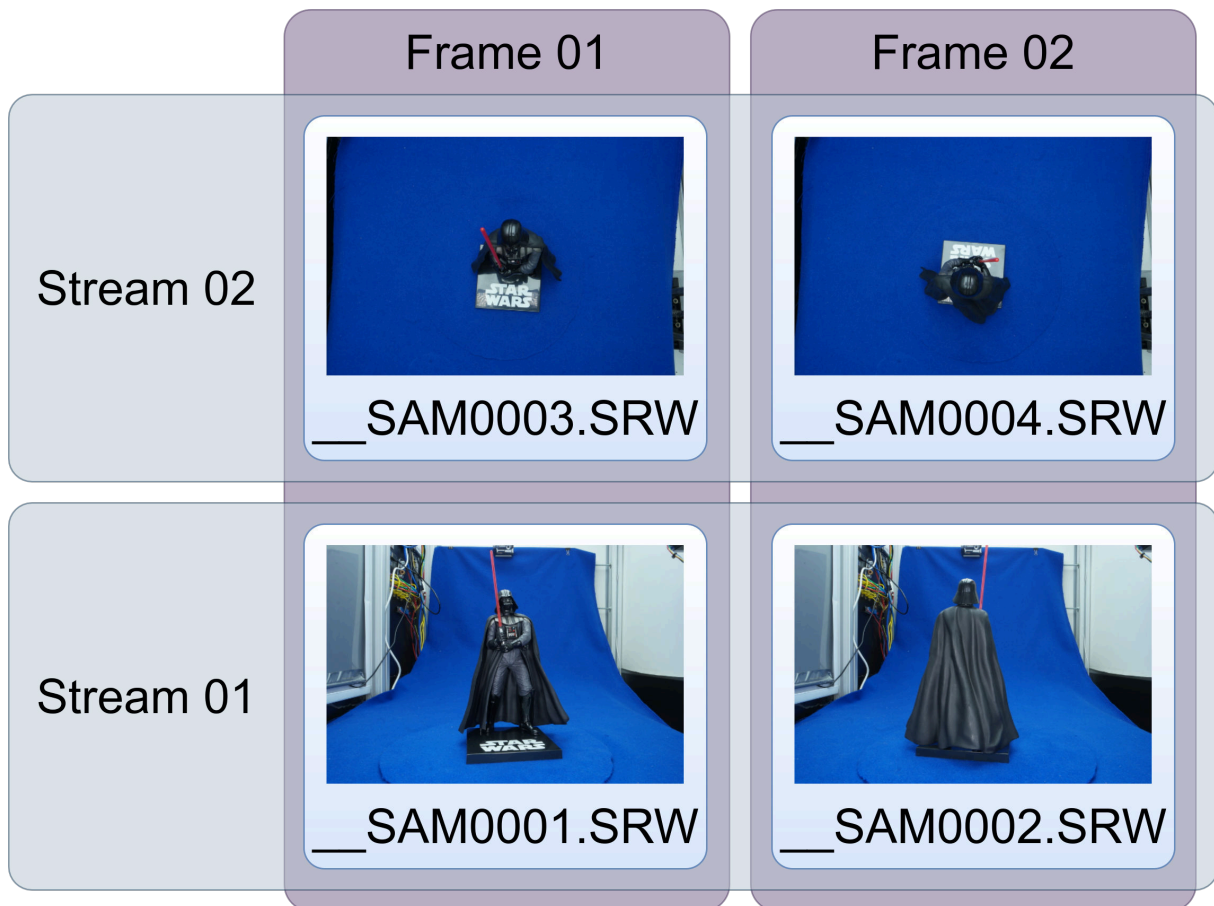
Desta forma, para organizar o conjunto de fotografias resultante da etapa de aquisição no Stream2D foram dados os nomes de Stream e Frames, conforme descrito a seguir. O exemplo de quatro fotografias apresentado seria organizado no Stream2D conforme apresentado na Figura 35.

Stream : Conjunto de fotografias adquiridas com uma mesma variação angular na direção vertical, mas diferentes posições na horizontal.

Frame : Cada fotografia armazenada no Stream2D. Cada Frame está contido dentro de um Stream, o qual identifica a variação angular vertical e o Frame identifica a variação angular na horizontal, dentro do Stream onde o Frame está inserido.

Considerando que o sistema de organização das fotografias é bidimensional, o nome escolhido para a estrutura foi Stream2D. O termo Stream remete a um fluxo contínuo de imagens,

Figura 35 – Criação do Stream2D - Organização das imagens em Streams e Frames.



Fonte: Autoria própria.

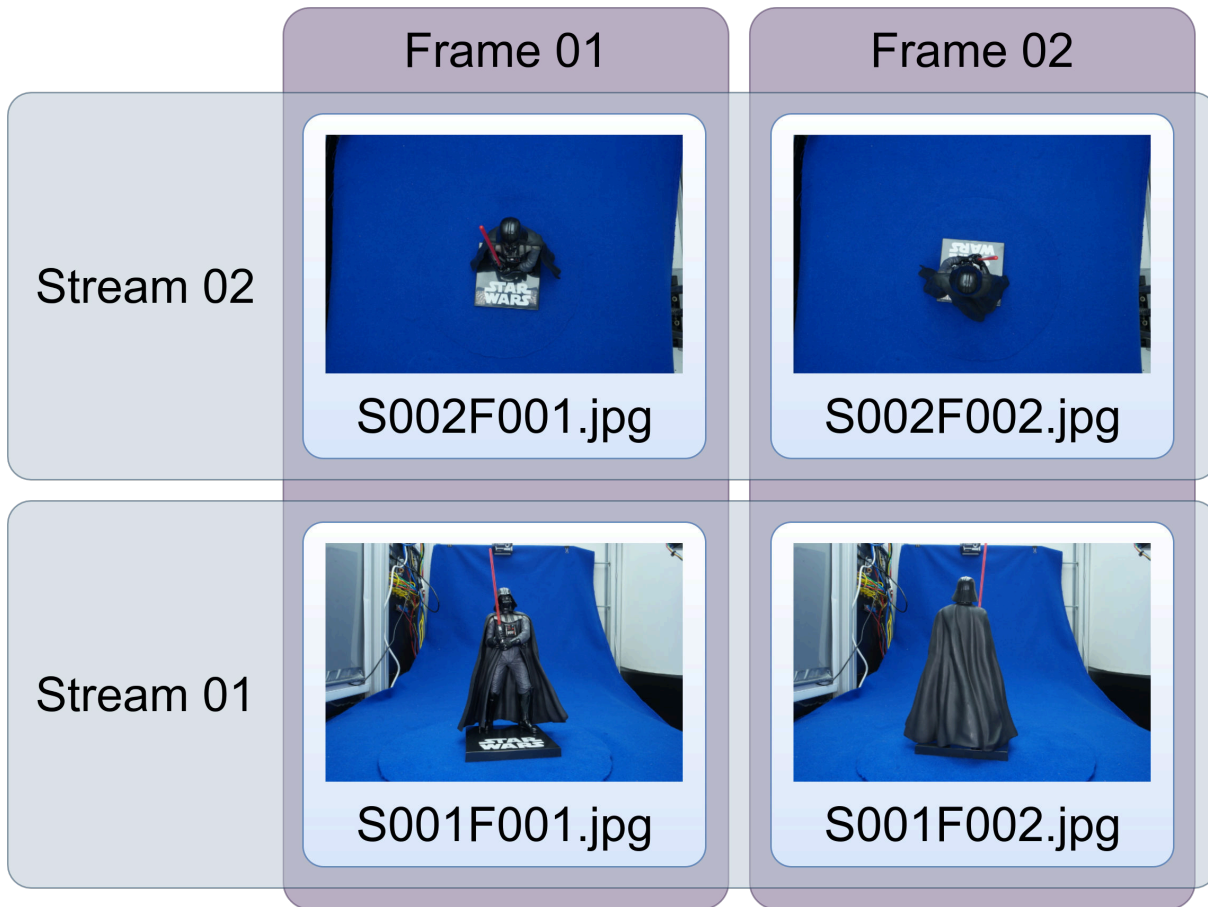
como as fotografias resultantes da aquisição. A Sigla 2D informa que a organização deste stream é bidimensional, ou seja, organizado em Streams (primeira dimensão) e Frames (segunda dimensão).

Isto posto, no Stream2D cada fotografia adquirida recebe um nome que identifica as variações angulares horizontal e vertical em que a fotografia foi adquirida, permitindo assim calcular, a partir do nome atribuído, a posição em que a câmera estava no momento em que a fotografia foi adquirida. A Figura 36 apresenta os novos nomes das quatro fotografias adquiridas, sendo que “S” é a abreviação de Stream e “F” de Frame.

Cabe ressaltar que as extensões dos arquivos estão diferentes nas Figuras 34 e 36 pois o Stream2D é criado no final da etapa de processamento e os arquivos originais são convertidos em outros formatos, como será discutido na seção 5.2.1.

Para calcular a posição angular de um Frame usando o seu nome no Stream2D é necessário conhecer a quantidade de Streams e Frames que compõe o Stream2D, além do intervalo angular vertical da aquisição (Equação (4.12)). Inicialmente calcula-se a variação angular entre Streams e Frames, conforme apresentado nas Equações (4.11) e (4.13).

Figura 36 – Criação do Stream2D - Arquivos organizados e renomeados para visualização.



Fonte: Autoria própria.

$$\Delta\theta = 360 \div \Sigma Frames \quad (4.11)$$

$$\beta\gamma = (\gamma^P - \gamma^0) \quad (4.12)$$

$$\Delta\gamma = \beta\gamma \div (\Sigma Streams - 1) \quad (4.13)$$

Sendo:

- $\Delta\theta$: Variação angular horizontal, entre Frames consecutivos.
- $\Sigma Frames$: Quantidade de Frames por Stream.
- $\beta\gamma$: Intervalo angular vertical da aquisição.
- γ^P : Ângulo vertical final da aquisição.
- γ^0 : Ângulo vertical inicial da aquisição.
- $\Delta\gamma$: Variação angular vertical, entre Streams consecutivos.

- $\Sigma Streams$: Quantidade de Streams no Stream2D.

A aquisição não precisa, necessariamente ser feita de 0° a 90° na direção horizontal, portanto para calcular $\Delta\gamma$ é necessário identificar o intervalo vertical da aquisição ($\beta\gamma$). Além disso é necessário subtrair um (-1) do total de Streams para calcular $\Delta\gamma$, pois o primeiro Stream está no γ^0 e o último no γ^P . Isso resulta em $\Sigma Streams - 1$ intervalos angulares verticais e não em $\Sigma Streams$ intervalos.

Esses procedimentos não precisam ser feitos no cálculo de $\Delta\Theta$ pois a aquisição na direção horizontal é sempre completa (360°) e o primeiro e último Frames estão na mesma posição ($0^\circ = 360^\circ$), de forma que o último Frame do Stream é suprimido do conjunto. Em outras palavras, como o primeiro e último Frames estão na mesma posição, o último Frame adquirido está posicionado no ângulo $360^\circ - \Delta\Theta$. Assim a quantidade de intervalos angulares é $\Sigma Frames$.

Após calcular as variações angulares entre Streams e Frames, com base no nome do Frame no Stream2D é possível calcular a posição angular em que a foto foi tirada com as fórmulas apresentadas nas Equações (4.14) e (4.15).

$$\Theta = (Frame_{id} - 1) \times \Delta\Theta \quad (4.14)$$

$$\gamma = [\gamma^0 + (Stream_{id} - 1)] \times \Delta\gamma \quad (4.15)$$

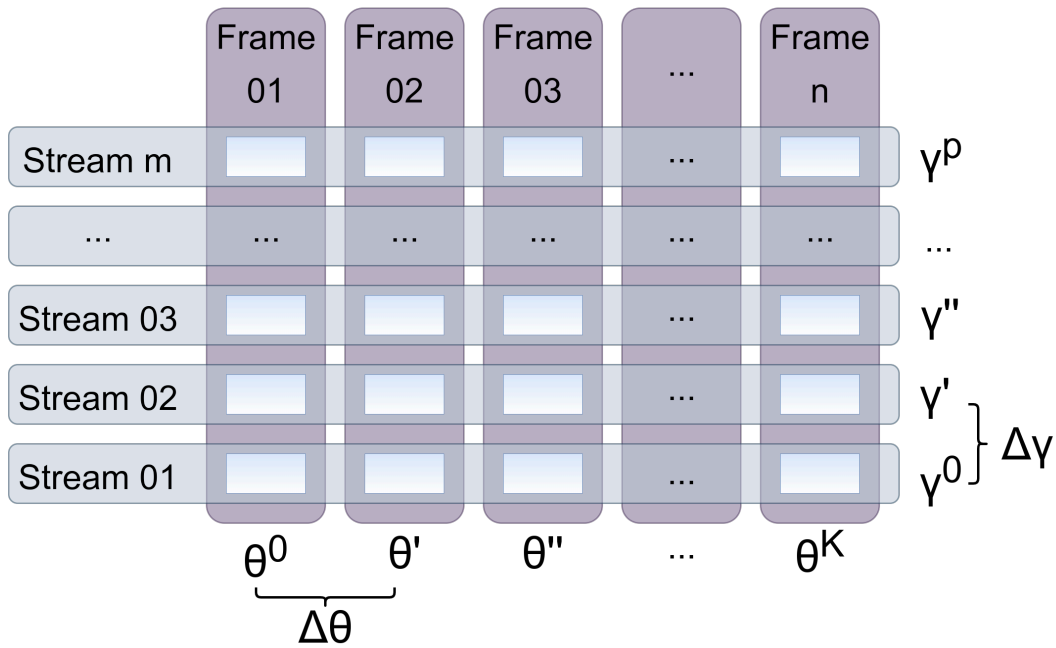
Sendo:

- Θ : Valor angular horizontal do Frame.
- $Frame_{id}$: Valor de Identificação do Frame no Stream.
- γ : Valor angular vertical do Stream.
- $Stream_{id}$: Valor de identificação do Stream no Stream2D

A Figura 37 representa genericamente um Stream2D com m Streams e n Frames. Os valores Θ apresentam o ângulo de determinado Frame dentro do Stream, ou seja a variação angular horizontal. Os valores γ apresentam os ângulos verticais de cada Stream. $\Delta\Theta$ representa a variação angular entre cada Frame e $\Delta\gamma$ a variação angular entre cada Stream.

Para exemplificar o cálculo da posição de um Frame pode ser considerado um Stream2D de 31 Streams por 120 Frames (totalizando 3.720 Frames). Esta aquisição compreende uma variação angular horizontal do 0° ao 90° . Então, para o Frame *S003F002* a posição angular da câmera pode ser calculada com as equações (4.16, 4.17, 4.18, 4.19, 4.20).

Figura 37 – Representação gráfica do Stream2D.



Fonte: Autoria própria.

$$\begin{aligned}
 \Delta\Theta &= 360^\circ \div \Sigma Frames \\
 &= 360^\circ \div 120 \\
 &= 3^\circ
 \end{aligned} \tag{4.16}$$

$$\begin{aligned}
 \Theta &= (Frame_{id} - 1) \times \Delta\Theta \\
 &= (2 - 1) \times 3 \\
 &= 3
 \end{aligned} \tag{4.17}$$

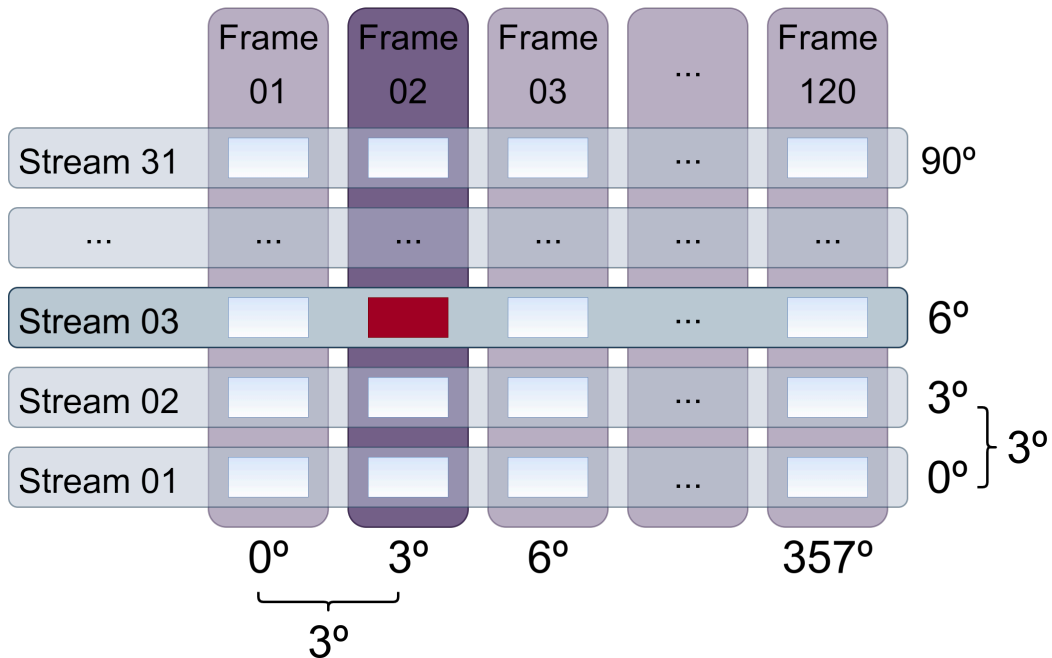
$$\begin{aligned}
 \beta\gamma &= (\gamma^P - \gamma^0) \\
 &= (90^\circ - 0^\circ) \\
 &= 90^\circ
 \end{aligned} \tag{4.18}$$

$$\begin{aligned}
 \Delta\gamma &= \beta\gamma \div (\Sigma Streams - 1) \\
 &= 90 \div (31 - 1) \\
 &= 3
 \end{aligned} \tag{4.19}$$

$$\begin{aligned}
 \gamma &= [\gamma^0 + (Stream_{id} - 1)] \times \Delta\gamma \\
 &= [0^\circ + (3 - 1)] \times 3 \\
 &= 6
 \end{aligned} \tag{4.20}$$

Desta forma a posição angular da fotografia com relação ao objeto do Frame *S003F002* é 6° de inclinação na direção vertical e 3° na horizontal. A Figura 38 representa o Stream2D usado no exemplo e o Frame *S003F002* encontra-se marcado na cor vermelha.

Figura 38 – Frame *S003F002*.



Fonte: Autoria própria.

Como já mencionado, também é possível, a partir de um determinado ângulo de visualização, descobrir qual é o Frame que melhor o representa. Cabe ressaltar que este cálculo pode não ser exato, uma vez que nem todos os ângulos são fotografados, porém é possível descobrir qual o Frame mais próximo, ou seja, que melhor representa a visualização almejada.

Para isso é necessário dividir o ângulo almejado por $\Delta\Theta$, no caso da direção horizontal e por $\Delta\gamma$ no caso da direção vertical. Por último, nos dois casos, soma-se um (+1) para encontrar o número do Frame ou Stream. Isso acontece pois o padrão de nomenclatura adotado para o Stream2D inicia em um e não em zero. As Equações (4.21) e (4.22) apresentam as fórmulas para o cálculo.

$$Frame_{id} = (\Theta \div \Delta\Theta) + 1 \quad (4.21)$$

$$Stream_{id} = \frac{(\gamma - \gamma^0)}{\Delta\gamma} + 1 \quad (4.22)$$

Caso o resultado não seja inteiro, isso significa que não existe um Frame ou Stream na posição exata. Então, caso a parte decimal seja menor que 0.5 usa-se o frame anterior; caso seja maior, usa-se o frame posterior; caso seja igual a 0.5 pode-se usar qualquer um dos frames anterior e posterior.

Para exemplificar o processo de cálculo do Frame usando os valores angulares pode-se usar o mesmo Stream2D com trinta e um Streams e cento e vinte Frames. Considere que o observador está a 7° do objeto na direção horizontal e a 9° na direção vertical. As Equações (4.23) e (4.24) apresentam os cálculos. Os valores angulares de $\Theta = 7^\circ$ e $\gamma = 9^\circ$ com relação ao observador podem ser calculados com trigonometria e o processo de cálculo é apresentado na Subseção 4.1.3.

$$\begin{aligned} Frame_{id} &= (\Theta \div \Delta\Theta) + 1 \\ &= (7 \div 3) + 1 \\ &= 3.333333 \\ &\approx 3 \end{aligned} \tag{4.23}$$

$$\begin{aligned} Stream_{id} &= \frac{(\gamma - \gamma^0)}{\Delta\gamma} + 1 \\ &= \frac{(9 - 0)}{3} + 1 \\ &= 4 \end{aligned} \tag{4.24}$$

Portanto o frame que melhor representa $\Theta = 7^\circ$ e $\gamma = 9^\circ$ neste caso é *S004F003*. Nota-se que na Equação (4.23) o valor de 3.333333 foi arredondado para 3. O Frame *S004F003* foi escolhido em detrimento ao Frame *S004F004*, pois o valor decimal era menor que 0.5, portanto o Frame *S004F003*, que é o anterior, está mais próximo da posição desejada.

Isso pode ser comprovado comparando o valor angular dos Frames. O Frame *S004F003* possui um valor angular de $\Theta = 6^\circ$ e o Frame *S004F004* de $\Theta = 9^\circ$. Percebe-se que $\Theta = 6^\circ$ está mais próximo de $\Theta = 7^\circ$ (diferença de 1°), o qual é o valor desejado, do que $\Theta = 9^\circ$ (diferença de 2°).

Após apresentar as equações que convertem os nomes dos Frames em ângulos e de ângulos para nomes, é possível justificar a opção de salvar os arquivos do Stream2D diretamente no disco em vez de desenvolver um formato próprio de arquivo ou usar outros sistemas, como bancos de dados, para manter estes arquivos.

Persistência do Stream2D

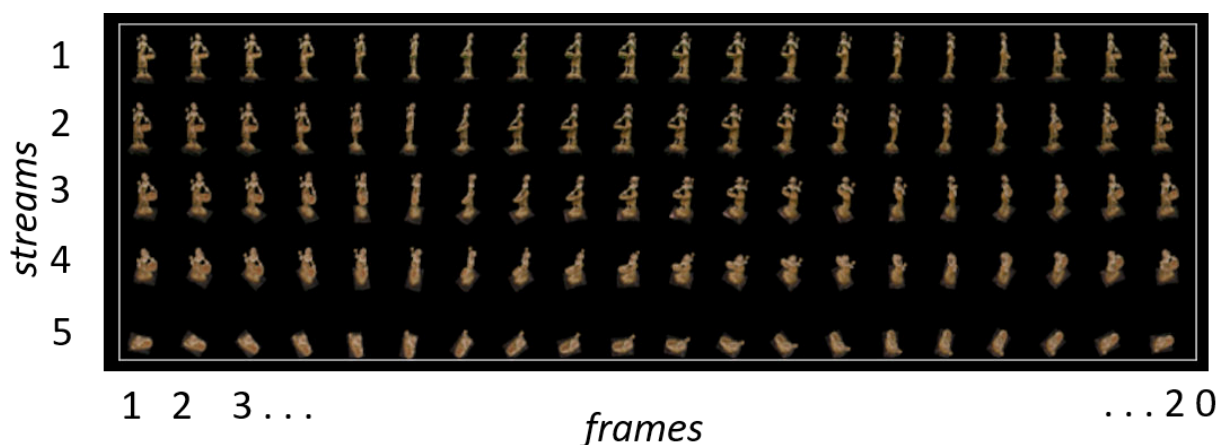
Optou-se pela abordagem de manter as fotografias do Stream2D salvas em disco pois o fato de renomear as imagens com identificadores únicos que mapeiam a posição em que a fotografia foi tirada permite que com cálculos simples um programa descubra qual arquivo deseja usar e acesse diretamente o disco para obter a imagem.

Esse processo reduz o custo computacional do sistema se comparado com acessar outro sistema que seria responsável por entregar a imagem correta, exigindo diversas trocas de contexto

no processador, acessos a memória e até em disco. Considerando que o custo computacional de calcular o nome do arquivo desejado é menos custoso computacionalmente do que atribuir esta tarefa a outros sistemas, essa abordagem torna a visualização mais eficiente.

A Figura 39 apresenta o conjunto de imagens de um Stream2D com 5 Streams e 20 Frames, totalizando cem imagens. A estrutura do Stream2D não impõe limite na quantidade de Frames representados, porém o scanner atual possui uma precisão que permite até cem Streams e quatrocentos Frames, totalizando quarenta mil imagens por aquisição.

Figura 39 – Stream2D com 5 Streams e 20 Frames.



Fonte: Autoria própria.

4.1.3 Sistemas de Simulação

Esta Subseção apresenta como um Stream2D pode ser usado em sistemas de simulação. Um sistema de simulação pode ser qualquer sistema que gere a visualização, desde um jogo ou um software educativo, até um sistema de visualização imersivo. Existem diversos sistemas, compostos por diferentes hardware e software, sendo em em sistemas diferentes podem existir diferenças como inversões de eixos ou de orientação dos eixos. Porém com o formalismo apresentado estas diferenças podem ser facilmente contornadas.

Inicialmente será apresentado como pode ser feita a visualização interativa, selecionando qual Frame deve ser exibido a partir de um determinado ponto de vista. É importante estabelecer que a computação gráfica tradicionalmente usa sistemas de coordenadas cartesianas em seus ambientes. Em seguida é descrito o sistema de coordenadas cartesianas tridimensionais e o processo de conversão entre os sistemas cartesiano, esférico, Stream2D e F2S2.

Com o processo de conversão entre sistemas definido, de forma que é possível mapear um ponto em qualquer um destes sistemas, e com o processo de visualização interativa e estereoscópica também definidos, é possível aplicar um *pack* de fotografias ou um Stream2D para visualizações em ambientes de computação gráfica, gerando a visualização. Por último, são

discutidas algumas considerações sobre os sistemas de coordenadas, suas relações, conversões e limitações. Portanto, a seguir, é definido o processo de visualização usando um Stream2D.

Visualização

O processo desenvolvimento de um software de visualização não faz parte do escopo deste trabalho, sendo responsabilidade de outro membro do grupo de pesquisa. O Autor deste trabalho participou da construção dos visualizadores apresentados nas Figuras 40 e 41, além de desenvolver as definições apresentadas nesta seção sobre como acessar os Frames partindo de coordenadas esféricas, viabilizando a visualização.

Porém, como a visualização não é parte do escopo deste trabalho, não foram desenvolvidas soluções para os desafios como consumo de memória e velocidade de carregamento das imagens. Apesar disso, é importante para este trabalho explicar como a visualização pode ser feita partindo do Stream2D.

Diversos sistemas podem ser desenvolvidos para usar o Stream2D. No Laboratório de pesquisa já foram implementadas abordagens em *Desktop* usando tela ou projetor, sistemas de projeção imersivos *FullDome*, óculos de realidade virtual usando celular e pirâmide holográfica. A Figura 40 apresenta algumas destas abordagens, sendo: 40(a) em óculos de realidade virtual, 40(b) em pirâmide holográfica e 40(c) em tela de TV.

Além disso, com o Stream2D podem ser construídos softwares de visualização que permitam outros tipos de interação, como por exemplo adicionando outros conteúdos hipermedia ou evidenciando partes da imagem. Este processo pode ser visto na Figura 41, na qual existem botões que ao serem clicados evidenciam partes do objeto escaneado.

Visualização Interativa

De forma simplificada o processo de visualização acontece alterando o frame que é exibido. Ao interagir com o Stream2D é possível visualizar quaisquer ângulos do objeto escaneado. Usando alguma entrada de dados, que pode ser teclado, mouse, acelerômetro, entre outros, o sistema converte essa entrada em um movimento no Stream2D. Em outras palavras, ao clicar em um botão o sistema calcula uma nova posição de observação, calcula o Frame que melhor representa o ponto de vista desejado e o exibe na tela. Este processo é repetido até o final da visualização.

Além deste método de visualização, que substitui o Frame que está sendo exibido, podem ser realizadas operações sobre um Frame, como alterações de: escala, posição e inclinação. Assim, as propriedades da visualização interativa foram divididas em dois grupos:

Navegação: Processo de visualização que altera o ponto de vista do observador com relação ao objeto. Neste processo os valores de Θ e/ou γ são alterados e consequentemente o Frame

Figura 40 – Sistemas desenvolvidos para visualização do Stream2D.

(a) Óculos de realidade virtual.



(b) Pirâmide Holográfica.



(c) Televisão.



Fonte: Autoria própria.

que está sendo exibido é substituído. São respeitados os intervalos adquiridos (γ^0 e γ^p) e as variações angulares horizontal e vertical ($\Delta\Theta$ e $\Delta\gamma$)

Transformações geométricas: Operações realizadas em um Frame, que transformam a geometria da imagem que está sendo exibida, porém não alteram os valores de Θ e/ou γ e conseqüentemente não alteram o Frame que está sendo exibido. As transformações geométricas podem ser de:

Escala: Alteram o tamanho da imagem que está sendo exibida, simulando aproximação (*Zoom In*) e afastamento (*Zoom Out*).

Translação: Altera a posição na tela onde a imagem é exibida, causando a sensação de que o objeto moveu-se na tela.

Rotação: Altera a orientação da imagem na tela, inclinando-a no sentido horário ou anti-horário.

Para exemplificar a visualização interativa, a Figura 42 apresenta exemplos de transformações geométricas. Considerando determinado Frame (Figura 42(a)) são aplicadas transformações

Figura 41 – Enriquecimento da visualização - Evidenciando partes do objeto escaneado.



Fonte: Autoria própria.

de: na Figura 42(b) escala, aproximando ou aumentando o objeto; na Figura 42(c) translação, movendo o objeto do centro da tela para o canto inferior esquerdo; e na Figura 42(d) rotação, girando o objeto no sentido anti-horário.

A Figura 43 apresenta três alterações consecutivas de navegação em um Stream2D. A Figura 43(a) apresenta o Stream *S001F001*, com valores angulares de $\Theta = 0^\circ$ e $\gamma = 0^\circ$. Em 43(b) existe uma alteração em $\Theta = 90^\circ$, ou seja, o objeto girou 90° na horizontal, sentido anti-horário. Na Imagem 43(c) ouve uma inclinação na vertical ($\gamma = 45^\circ$) e por fim, em 43(d) visualiza-se o topo do objeto, com $\Theta = 90^\circ$ e $\gamma = 90^\circ$.

A Figura 44 exemplifica o processo de navegação, considerando que o usuário está visualizando o Frame *S001F001*. Usando um teclado o usuário muda para o ponto de vista de 6° na horizontal e na vertical. Usando os cálculos já apresentados identifica-se que o Frame que melhor representa este ponto de vista é o Frame *S003F003*. Neste momento o visualizador altera a imagem para *S003F003*. Ao realizar mudanças consecutivas, tem-se a sensação de que o objeto está se movimentando na tela.

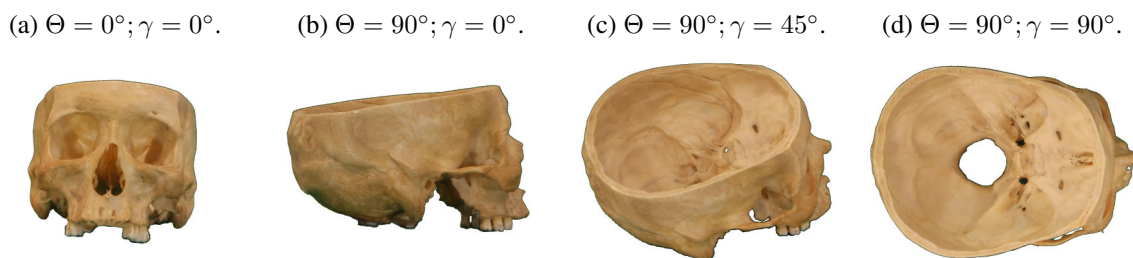
Cabe ressaltar os limites de navegação do Stream2D na direção horizontal e vertical. Na horizontal, ou seja alterando Frames dentro de um mesmo Stream, os Frames funcionam

Figura 42 – Visualização Interativa - Transformações Geométricas.



Fonte: Autoria própria.

Figura 43 – Visualização Interativa - Navegação.

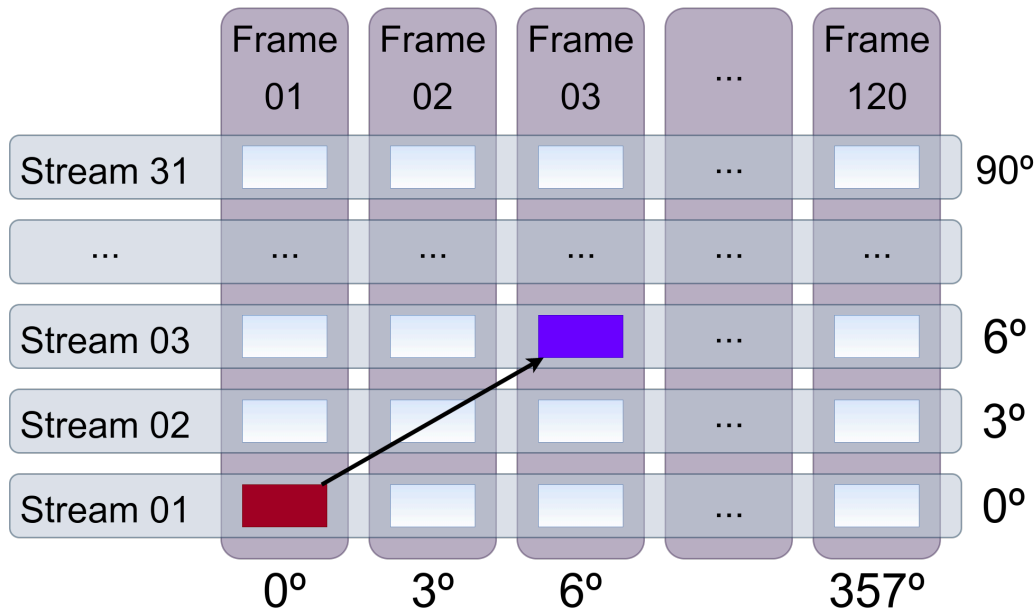


Fonte: Autoria própria.

como uma lista circular. Do último Frame de um Stream, navegando para a direita, passa-se para o primeiro Frame. No caso do exemplo acima, do Frame *S001F120* para o Frame *S001F001*. Do primeiro Frame, navegando para a esquerda, acessa-se o último Frame, ou seja do Frame *S001F001* para o *S001F120*. Este processo permite girar o objeto indefinidamente neste sentido.

Na vertical existem duas situações distintas. Do primeiro Stream (*S001*) para baixo não existem posições digitalizadas. Existe a possibilidade de digitalizar o objeto de ponta-cabeça e então unir as duas aquisições em um único Stream2D. Porém as dificuldade de fixação, oclusão,

Figura 44 – Alteração entre Frames durante a visualização.



Fonte: Autoria própria.

posicionamento e irregularidade de forma, como discutidos na seção 4.3 podem inviabilizar que este processo seja realizado de forma que os Frames estejam corretamente alinhados, principalmente para objetos irregulares.

Por exemplo, ao tentar escanear uma boneca de cabelos compridos e flexíveis, com a boneca “em pé”, seus cabelos cairiam da cabeça para os ombros. Ao tentar inverter o objeto para escanear-la de “ponta-cabeça”, seus cabelos não permaneceriam na mesma posição e estariam orientados dos ombros para a cabeça. Esta mudança seria perceptível nas imagens. Além disso, seria difícil posicionar a boneca sem causar nenhum tipo de inclinação, giro ou irregularidade se comparada à primeira aquisição. Por último, seria necessário usar algum aparato para fixá-la, uma vez que a cabeça é redonda.

Portanto, no caso de um movimento para baixo, partindo do primeiro Stream, existem três possibilidades: ou o movimento é bloqueado; ou é exibida uma imagem informando que aquela posição do objeto não foi adquirida; ou pode-se também realizar um salto de 180° dentro do mesmo Stream e exibir a imagem invertida (de “ponta-cabeça”), como se o objeto tivesse girado até completar o movimento. Este processo seria percebido de forma abrupta e não de forma linear.

Para exemplificar o salto de 180° para baixo partindo do primeiro Stream, considere a aquisição da Figura 44. Considerando o Frame *S001F001* (com valores angulares $\Theta = 0$ e $\gamma = 0$), ao tentar acessar a posição inferior deste Frame, exibe-se o Frame *S001F061* (com valores angulares $\Theta = 180$ e $\gamma = 0$).

Ainda no giro na vertical, partindo do último Stream para cima, o movimento pode ser

bloqueado ou exibe-se a imagem oposta (com uma diferença de 180°) dentro do mesmo Stream. O novo Frame exibido pode ser invertido (de “ponta-cabeça”) para apresentar a sensação de um giro completo ou não. Este processo seria percebido de forma linear. Qual destas abordagens deve ser usada depende da aplicação da visualização e não do modelo do Stream2D.

Para exemplificar o salto de 180° para cima, também considera-se a aquisição da Figura 44. Considerando o Frame *S031F001* (com valores angulares $\Theta = 0$ e $\gamma = 90$), ao tentar acessar a posição superior deste Frame, exibe-se o Frame *S031F061* (com valores angulares $\Theta = 180$ e $\gamma = 90$).

O processo de visualização do Stream2D envolve ainda outro conjunto de desafios. É necessário considerar o volume de dados gerados em uma aquisição. Como podem ser adquiridas até quarenta mil imagens, atualmente com tamanho de até 8K UHD (*Ultra High Definition - Ultra Alta Definição*)¹, cada frame já processado possui, no formato “jpg”, pouco mais de 2Mb, uma aquisição como esta possui mais de 80Gb.

Os computadores regulares atuais não possuem 80Gb de memória disponível para carregar todos os Frames de um Stream2D. Assim é necessário acessar cada imagem, ou determinado conjunto de imagens, do disco. A tecnologia atual de discos (existentes na maioria dos computadores comuns) não permitem um acesso rápido o suficiente para gerar uma visualização sem a sensação de travamento.

É necessário considerar que as tecnologias estão mudando, portanto tanto o tamanho das imagens deve ficar maior, quanto os computadores devem ficar mais rápidos. Assim é necessário considerar a qualidade da visualização pretendida (principalmente: tamanho das imagens e latência na troca entre estas) e a capacidade computacional disponível para realizá-la.

Existem ainda outros desafios, como a possibilidade de transmitir um Stream2D pela internet: que precisaria garantir uma visualização adequada transmitindo os Frames pela internet; o uso em telefones celulares: que possuem menor capacidade computacional; a utilização acadêmica: que precisa garantir acuidade visual e fidelidade com o objeto real em dispositivos diversos e telas não calibradas.

Visualização Estereoscópica

Outra possibilidade de uso do Stream2D é a geração de visualizações interativas e estereoscópicas. A estereoscopia gera a sensação de profundidade e proporção, mesmo em impressos ou telas bidimensionais. Como mencionado em (SILVA, 2020) a estereoscopia humana baseia-se em um par estéreo de imagens. A geração de estereoscopia, os parâmetros e técnicas usados são parte do escopo do trabalho (SILVA, 2020).

O processo de uso do Stream2D para obtenção do par estéreo é similar ao processo de identificação de determinado Frame que representa uma determinada perspectiva durante

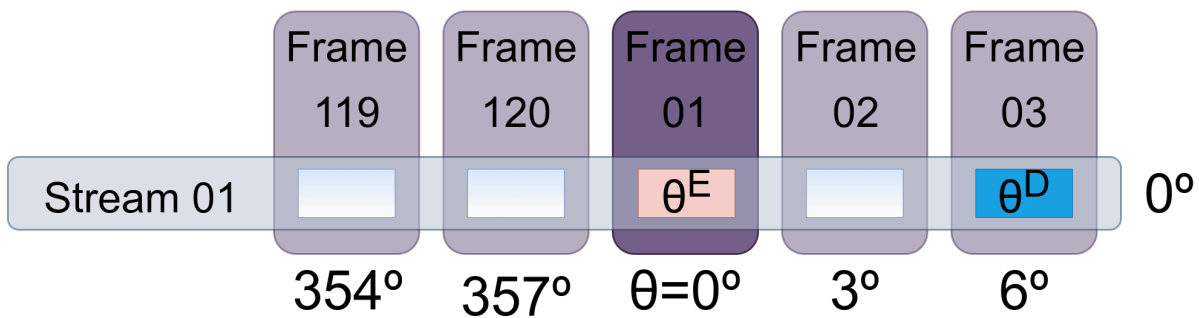
¹ Padronizado pela NHK - www.nhk.or.jp.

a visualização. A diferença no processo consiste em adquirir uma segunda imagem com uma diferença angular de 2° a 8° entre elas, sendo que o valor de 6° é o que gera melhor estereoscopia². Neste trabalho este valor será representado por $\Delta'\Theta$.

Foram determinadas três formas para obtenção do par estéreo: A primeira forma consiste em, considerando determinado ângulo Θ , obter o Frame adjacente à direita de Θ que melhor represente a diferença angular de $\Delta'\Theta$. Θ^E representa a imagem esquerda e Θ^D a imagem direita do par estéreo.

Assim, considerando uma aquisição com $\Delta\Theta = 3^\circ$ e $\Delta'\Theta = 6^\circ$, caso o ângulo de visualização original seja de $\Theta = 0^\circ$ e $\gamma = 0^\circ$, o par estéreo seria formado pelos ângulos $\Theta^E = 0^\circ$ e $\Theta^D = 6^\circ$. Usando as Equações 4.21 e 4.22 é possível determinar que os Frames que melhor representam este par estéreo são os Frames *S001F001* e *S001F003*. A Figura 45 apresenta este caso.

Figura 45 – Visualização estereoscópica - Deslocamento à direita.



Fonte: Autoria própria.

Neste modelo o segundo valor angular (Θ^D), que representa o segundo Frame que compõe o par estéreo, pode ser calculado somando-se $\Delta'\Theta$ a Θ . Porém, existe a possibilidade do valor de Θ^D ser superior a 360° , o que não corresponde ao modelo do Stream2D. Assim, para obter o par estéreo usam-se as Equações 4.25 e 4.26.

$$\Theta^E = \Theta \quad (4.25)$$

$$\Theta^D = (\Theta + \Delta'\Theta) \% 360 \quad (4.26)$$

Sendo:

- Θ^E : Valor angular da imagem esquerda do par estéreo.
- Θ^D : Valor angular da imagem direita do par estéreo.
- %: Módulo ou resto da divisão por 360° .

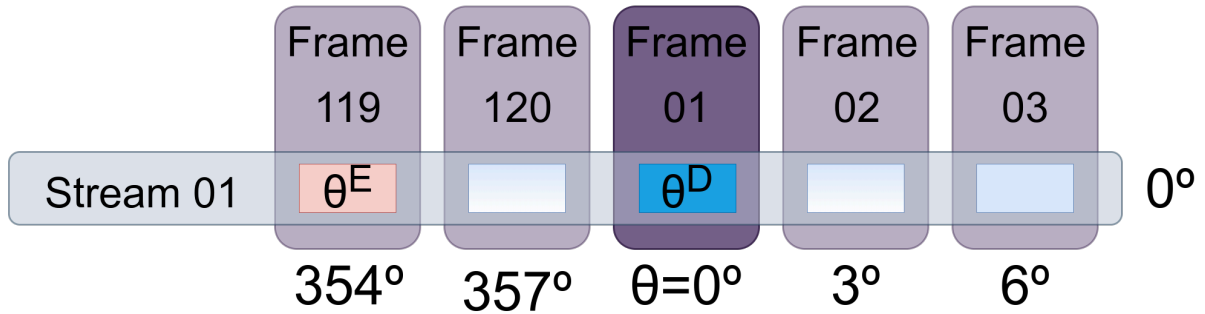
² Estes valores foram determinados no laboratório e são parte do escopo do trabalho de Flávio de Almeida e Silva.

Outra possibilidade de cálculo do par estéreo consiste em realizar o deslocamento à esquerda, de forma que Θ^D represente o ângulo Θ e Θ^E possua um deslocamento angular de $\Delta'\Theta$ para a esquerda, ou seja subtrai-se $\Delta'\Theta$ de Θ . Neste caso o valor angular de Θ^E pode ficar negativo, o que não é representado no Stream2D. Portanto para calcular o par estéreo neste caso usam-se as Equações 4.27 e 4.28. A Figura 46 apresenta graficamente os Frames escolhidos.

$$\Theta^D = \Theta \quad (4.27)$$

$$\Theta^E = \begin{cases} \Theta - \Delta'\Theta & \text{Caso: } \Theta - \Delta'\Theta \geq 0 \\ \Theta - \Delta'\Theta + 360 & \text{Caso: } \Theta - \Delta'\Theta < 0 \end{cases} \quad (4.28)$$

Figura 46 – Visualização estereoscópica - Deslocamento à esquerda.



Fonte: Autoria própria.

A terceira forma consiste em, considerando um determinado ângulo desejado, obter os Frames que melhor representem $\Delta'\Theta$ de diferença angular, sendo um anterior e outro posterior a Θ , de forma que o ponto de vista desejado (Θ) esteja inserido aproximadamente no centro deste intervalo. Existem situações em que, devido ao intervalo de escaneamento $\Delta\Theta$, Θ não esteja posicionado exatamente entre Θ^E e Θ^D .

Para calcular este modelo de obtenção dos Frames compõe o par estéreo, a partir de Θ , Θ^E desloca-se $\Delta'\Theta \div 2$ a esquerda e Θ^D desloca-se $\Delta'\Theta \div 2$ a direita. Considerando as possibilidades destes ângulos assumirem valores negativos e maiores que 360° para, respectivamente, Θ^E e Θ^D ; as Equações 4.29 e 4.30 apresentam como calcular os ângulos destes Frames.

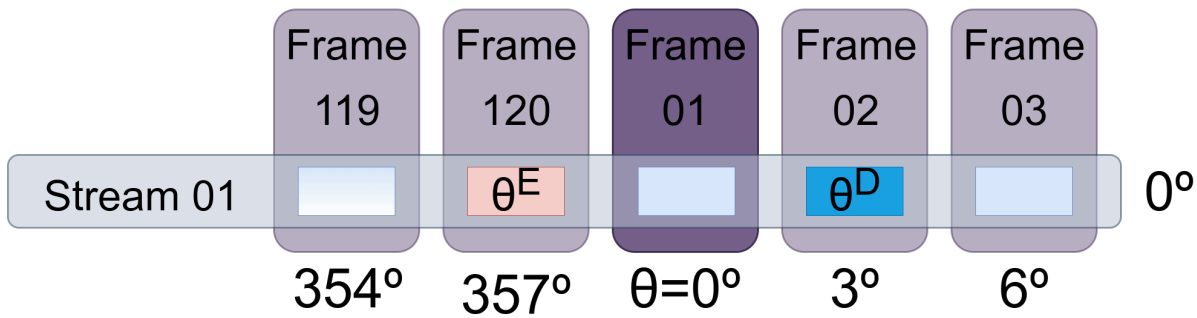
$$\Theta^D = (\Theta + \frac{\Delta'\Theta}{2}) \% 360 \quad (4.29)$$

$$\Theta^E = \begin{cases} \Theta - \frac{\Delta'\Theta}{2} & \text{Caso: } \Theta - \frac{\Delta'\Theta}{2} \geq 0 \\ \Theta - \frac{\Delta'\Theta}{2} + 360 & \text{Caso: } \Theta - \frac{\Delta'\Theta}{2} < 0 \end{cases} \quad (4.30)$$

A Figura 47 apresenta os Frames escolhidos para $\Delta\Theta = 3$, $\Delta'\Theta = 6$ e $\Theta = 0$. Cabe analisar que, considerando a mesma aquisição, mas para $\Theta = 1$ os mesmos Frames seriam

selecionados pelas equações, porém neste caso $\Theta = 1$ não estaria exatamente no centro destes Frames. Mesmo assim, estes seriam os Frames que melhor representariam o par estéreo com orientação centralizada para este caso.

Figura 47 – Visualização estereoscópica - Orientação centralizada.



Fonte: Autoria própria.

Considerando que $\Delta'\Theta = 6^\circ$ é o intervalo angular que melhor gera estereoscopia, os valores para $\Delta\Theta$ que são capazes de gerar intervalos angulares de 6° são: 1° , 2° , 3° e 6° . Valores de $\Delta\Theta = 4^\circ$ ou 5° não geram intervalos angulares de 6° . Portanto estes últimos dois valores de $\Delta\Theta$, ou valores maiores que 6° , devem ser evitados para digitalizações que almejem visualizações estereoscópicas.

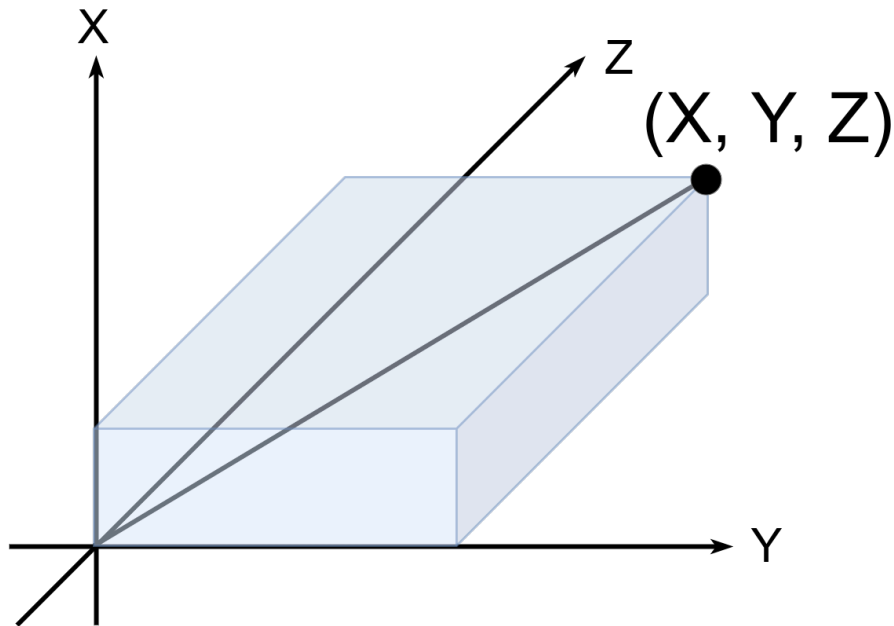
Sistemas de Coordenadas Cartesianas

O sistema de coordenadas cartesianas tridimensional representa um ponto no espaço tridimensional a partir dos seguintes componentes: $P = (X, Y, Z)$, de forma que o valor de cada componente representa o deslocamento linear nos respectivos eixos. A Figura 48 apresenta o sistema de coordenadas cartesianas e um ponto sendo descrito no sistema.

Conversão de coordenadas entre os sistemas Cartesiano e Esférico

O processo de conversão entre os sistemas de coordenadas esféricas e cartesianas pode ser calculado a partir da trigonometria, de forma que a conversão do sistema cartesiano para o esférico é definido pela Equação 4.31 e do esférico para o cartesiano está na Equação 4.33. Vale ressaltar que γ , usado neste trabalho, é o ângulo complementar de φ .

Figura 48 – Sistema de coordenadas cartesianas.



Fonte: Autoria própria.

$$\begin{aligned}
 (X, Y, Z) &\Rightarrow (r, \Theta, \varphi) \\
 r &= \sqrt{X^2 + Y^2 + Z^2} \\
 \Theta &= \text{Arco Tg2}(Y, X) \\
 \varphi &= \text{Arco Cos}\left(\frac{Z}{r}\right) \\
 \gamma &= 90 - \varphi
 \end{aligned}
 \tag{4.31}$$

Sendo:

$$\text{Arco Tg2}(Y, X) = \begin{cases} \text{Arco Tg}\left(\frac{Y}{X}\right) & \text{Caso: } X > 0 \\ \text{Arco Tg}\left(\frac{Y}{X}\right) + 180^\circ & \text{Caso: } X < 0 \text{ e } Y \geq 0 \\ \text{Arco Tg}\left(\frac{Y}{X}\right) - 180^\circ & \text{Caso: } X < 0 \text{ e } Y < 0 \\ 90^\circ & \text{Caso: } X = 0 \text{ e } Y > 0 \\ -90^\circ & \text{Caso: } X = 0 \text{ e } Y < 0 \\ \text{Indefinido} & \text{Caso: } X = 0 \text{ e } Y = 0 \end{cases}
 \tag{4.32}$$

$$\begin{aligned}
(r, \Theta, \varphi) &\Rightarrow (X, Y, Z) \\
\varphi &= 90 - \gamma \\
X &= r \times \text{Cosseno}(\Theta) \times \text{Seno}(\varphi) \\
Y &= r \times \text{Seno}(\Theta) \times \text{Seno}(\varphi) \\
Z &= r \times \text{Cosseno}(\varphi)
\end{aligned} \tag{4.33}$$

Conversão de coordenadas entre os sistemas Cartesiano e F2S2

$$\begin{aligned}
(X, Y, Z) &\Rightarrow (H, V, B) \\
H &= \sqrt{X^2 + Y^2} \\
V &= Z \\
B &= \Theta = \text{Arco Tg2}(Y, X)
\end{aligned} \tag{4.34}$$

$$\begin{aligned}
(H, V, B) &\Rightarrow (X, Y, Z) \\
X &= H \times \text{Seno}(B) \\
Y &= H \times \text{Cosseno}(B) \\
Z &= V
\end{aligned} \tag{4.35}$$

Conversão de coordenadas entre os sistemas Cartesiano e Stream2D

Para realizar a conversão do sistema de coordenadas do Stream2D para o sistema de coordenadas cartesiano ($(Stream_{id}, Frame_{id}) \Rightarrow (X, Y, Z)$) calcula-se a partir do $Frame_{id}$ e $Stream_{id}$ a coordenada esférica com a Equação 4.15 e em seguida usa-se a Equação 4.33 para converter do sistema esférico para o cartesiano. No momento da conversão de Stream2D para esférico é necessário atribuir um valor para o raio da esfera, uma vez que este valor não está presente no Stream2D.

Para realizar a conversão do sistema de coordenadas cartesiano para o sistema do Stream2D ($(X, Y, Z) \Rightarrow (Stream_{id}, Frame_{id})$), calcula-se inicialmente a conversão do sistema cartesiano para sistema esférico usando a Equação 4.31. Em seguida calcula-se o $Frame_{id}$ e $Stream_{id}$ com as equações 4.21 e 4.22.

Este processo permite calcular, qual frame deve ser exibido em qualquer sistema de computação gráfica, baseado em coordenadas cartesianas, onde se conheça a posição do objeto e do observador, como será apresentado na próxima Subseção (Subseção 4.1.3).

Aplicação do Stream2D em ambientes tridimensionais

Para exemplificar como um Stream2D pode ser usado em um ambiente de simulação baseado em coordenadas cartesianas tridimensionais será usado como exemplo um Stream2D composto por 31 Streams e 120 Frames por Stream, considerando que a aquisição foi feita, na vertical, do ângulo 0° ao 90° . Considere que o objeto, ou seja o Stream2D, está posicionado nas coordenadas $Obj = (100, 100, 0)$ e o observador está nas coordenadas $Obs = (60, 60, 33)$.

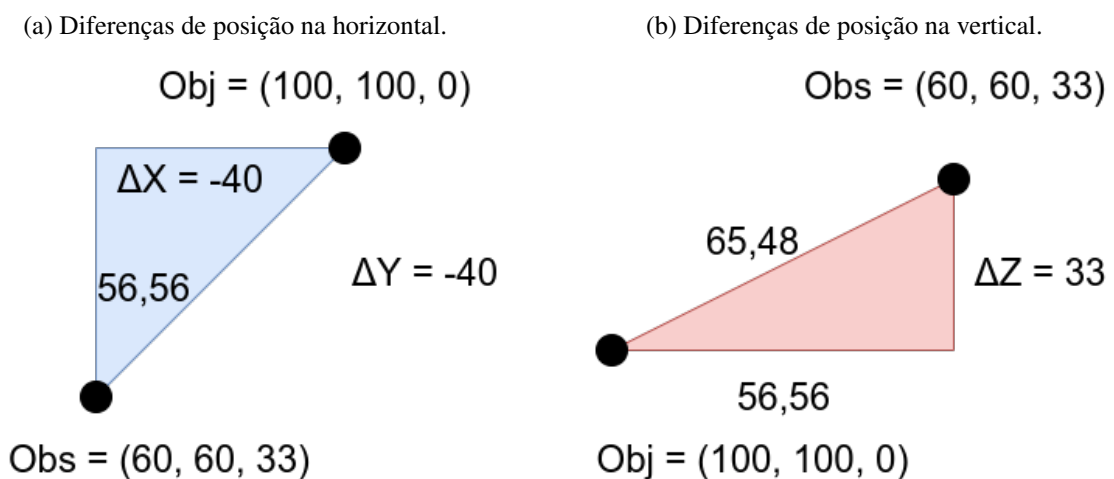
Considera-se também que a orientação dos eixos do objeto e do observador estão alinhados com os eixos do ambiente, ou seja o objeto e o observador não sofreram rotação e translação. Caso isso ocorra, a diferença angular sofrida deve ser adicionada ao modelo para que seja possível calcular o Frame correto.

O primeiro passo para calcular qual Frame deve ser exibido para um observador olhando para o objeto nestas posições calcular a diferença de posição dos objetos no sistema cartesiano, estabelecendo a diferença de posição entre eles. A Equação 4.36 apresenta este cálculo.

$$\begin{aligned}\Delta X &= X_{observador} - X_{objeto} = 60 - 100 = -40 \\ \Delta Y &= Y_{observador} - Y_{objeto} = 60 - 100 = -40 \\ \Delta Z &= Z_{observador} - Z_{objeto} = 33 - 0 = 33\end{aligned}\quad (4.36)$$

A Figura 49 apresenta graficamente as diferenças de posição entre o observador e o objeto. A Figura 49(a) apresenta a diferença na horizontal (Eixos X e Y) e a Figura 49(b) apresenta a diferença na vertical (Eixo Z).

Figura 49 – Diferenças de posição entre objeto e observador.



Fonte: Autoria própria.

Após calcular a diferença de posição no sistema cartesiano entre o objeto e o observador, calcula-se a diferença angular (Θ e γ) no espaço esférico. Isso pode ser feito com trigonometria com a Equação 4.31. O cálculo para estes valores é apresentado na Equação 4.37.

$$\begin{aligned}
(\Delta X, \Delta Y, \Delta Z) &\Rightarrow (r, \Theta, \varphi) \\
r &= \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} \\
r &= \sqrt{40^2 + 40^2 + 33^2} = \sqrt{1600 + 1600 + 1089} \\
r &= \sqrt{4289} = 65,49 \\
\Theta &= \text{Arco Tg2}(\Delta Y, \Delta X) \\
\Theta &= \text{Arco Tg2}(-40, -40) = -135^\circ = 225^\circ \\
\varphi &= \text{Arco Cos}\left(\frac{\Delta Z}{r}\right) \\
\varphi &= \text{Arco Cos}\left(\frac{33}{65,49}\right) = \text{Arco Cos}(0,503) = 59,80 \approx 60^\circ \\
\gamma &= 90^\circ - \varphi = 90^\circ - 60^\circ = 30^\circ
\end{aligned} \tag{4.37}$$

Calculados os valores de $r = 65,49$, $\Theta = 225^\circ$, $\varphi = 60^\circ$ e $\gamma = 30^\circ$, pode-se calcular qual frame deve ser exibido para esta configuração de posição de objeto e observador. Isso pode ser feito com as Equações 4.21 e 4.22. Porém para aplica-las são necessários os valores de $\Delta\Theta$, $\beta\gamma$ e $\Delta\gamma$, que podem ser obtidos pelas Equações 4.11, 4.12 e 4.13. O cálculo do Frame que deve ser exibido nesta configuração é apresentado na Equação 4.38.

$$\begin{aligned}
(r, \Theta, \varphi) &\Rightarrow (\text{Stream}_{id}, \text{Frame}_{id}) \\
\Delta\Theta &= \frac{360^\circ}{\Sigma \text{Frames}} \\
\Delta\Theta &= \frac{360^\circ}{120} = 3^\circ \\
\text{Frame}_{id} &= \frac{\Theta}{\Delta\Theta} + 1 \\
\text{Frame}_{id} &= \frac{225^\circ}{3} + 1 = 75 + 1 = 76 \\
\beta\gamma &= (\gamma^P - \gamma^0) \\
\beta\gamma &= (90^\circ - 0^\circ) = 90^\circ \\
\Delta\gamma &= \frac{\beta\gamma}{\Sigma \text{Streams} - 1} \\
\Delta\gamma &= \frac{90^\circ}{31 - 1} = \frac{90^\circ}{30} = 3^\circ \\
\text{Stream}_{id} &= \frac{(\gamma - \gamma^0)}{\Delta\gamma} + 1 \\
\text{Stream}_{id} &= \frac{(30^\circ - 0^\circ)}{3} + 1 = \frac{30^\circ}{3} + 1 = 10 + 1 = 11
\end{aligned} \tag{4.38}$$

Assim, o Frame que representa a visualização que o observador teria do objeto nesta configuração seria ($Frame_{id} = 76$, $Stream_{id} = 11$). Adicionalmente, caso queira realizar uma observação estereoscópica para a configuração apresentada, pode-se usar um dos modelos de cálculo de par estereoscópico apresentados na Seção 4.1.3.

Para exemplificar este processo será apresentado o cálculo do par estéreo do primeiro modelo, onde o par é composto pelos ângulos Θ^E , já calculado, e a outra imagem (Θ^D) é deslocada 6° deste ponto. Este processo usa as Equações 4.25 e 4.26. O cálculo é apresentado na Equação 4.39. Para este cálculo será considerado $\Delta'\Theta = 6^\circ$.

$$\begin{aligned}\Theta^E &= \Theta = 225^\circ \\ \Theta^D &= (\Theta + \Delta'\Theta)\%360 = (225^\circ + 6^\circ)\%360 = 231^\circ\end{aligned}\quad (4.39)$$

Como o Frame correspondente à $\Theta^E = 225^\circ$ e $\gamma = 30^\circ$ já foi calculado na Equação 4.38 e é ($Frame_{id} = 76$, $Stream_{id} = 11$), para compor o par estéreo basta calcular o Frame que representa Θ^D . Isso pode ser feito como apresentado na Equação 4.38. A Equação 4.40 apresenta o cálculo do Frame correspondente ao olho direito.

$$\begin{aligned}Frame_{id} &= \frac{\Theta}{\Delta\Theta} + 1 \\ Frame_{id} &= \frac{231^\circ}{3} + 1 = 77 + 1 = 78\end{aligned}\quad (4.40)$$

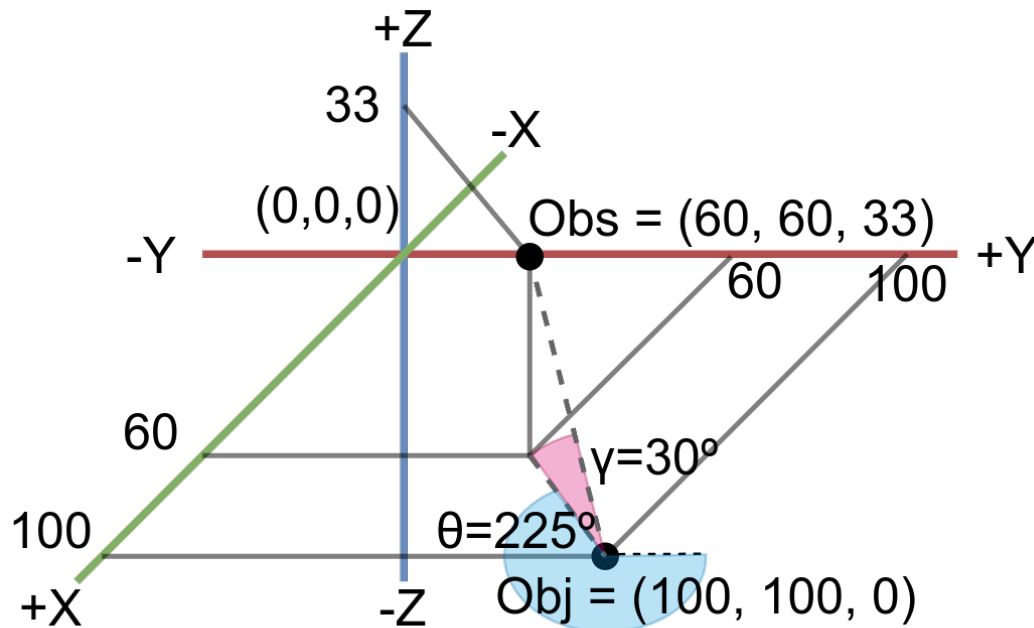
Assim, para uma visualização estereoscópica poderia ser oferecido o par estéreo composto pelos Frames $Frame^E = (Frame_{id} = 76, Stream_{id} = 11)$ e $Frame^D = (Frame_{id} = 78, Stream_{id} = 11)$. A Figura 50 apresenta graficamente a posição do objeto e do observador em um sistema de coordenadas cartesianas.

Considerações sobre sistemas de coordenadas

Apresentado o processo que permite usar as fotografias em ambientes de simulação, as conversões entre os sistemas de coordenadas GMC, F2S2, Stream2D, Esférico e Cartesiano, é possível aplicar as fotografias em ambientes de computação gráfica, bem como a partir de determinado Frame descobrir a posição relativa da câmera ao objeto fotografado. A Figura 51 apresenta os sistemas de coordenadas e as possíveis conversões.

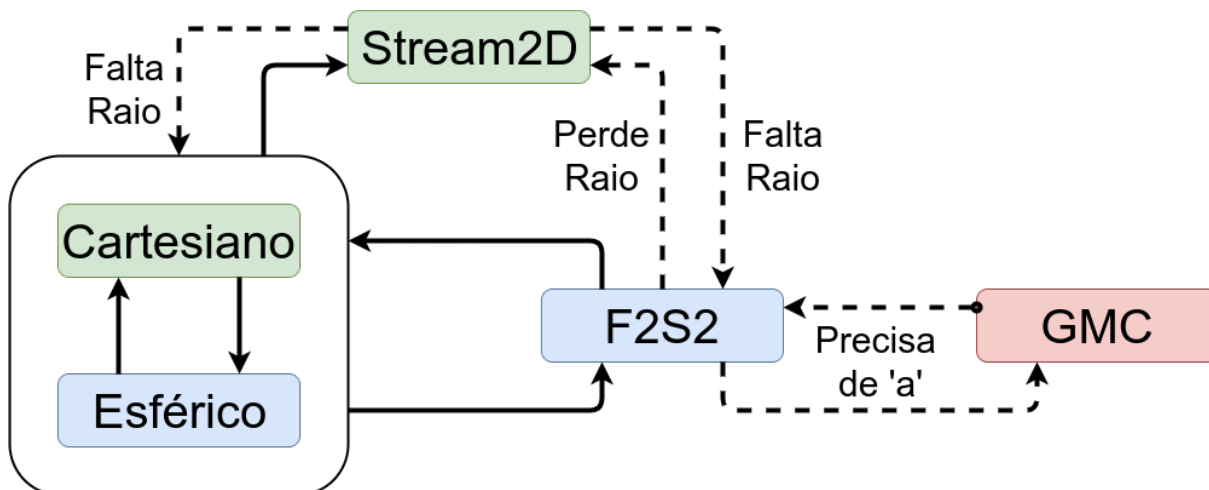
Na Figura 51 o sistema representado em vermelho (GMC) é o que possui as coordenadas referentes ao scanner; os sistemas representados em azul (F2S2 e Esférico) possuem as coordenadas mais adequadas para representar a posição relativa do objeto e da câmera; Os sistemas representados em verde (Cartesiano e Stream2D) são os sistemas voltados para a visualização, ou seja, são usados para representar a posição do objeto e observador no espaço tridimensional e os Frames que devem ser exibidos.

Figura 50 – Posição do objeto e observador.



Fonte: Autoria própria.

Figura 51 – Conversões entre os sistemas de coordenadas.



Fonte: Autoria própria.

As conversões para o sistema de coordenadas do Stream2D perdem as informações referentes ao raio da aquisição, e portanto não seria possível retornar uma coordenada do Stream2D para o sistema F2S2 ou mapear a posição da câmera nos sistemas esférico e Cartesiano. Porém se a conversão foi do sistema F2S2 para Esférico/Cartesiano esta informação poderia ser preservada. A Informação do raio durante a aquisição não é necessária para a visualização, mas pode ser relevante para sistemas de reconstrução 3D ou de cálculos para estereoscopia.

A conversão do sistema F2S2 para as coordenadas GMC dependem da informação

da altura da esfera com relação à base do scanner (variável “a”). Caso esta informação seja armazenada é possível retroceder do Sistema F2S2 para o GMC. Isso pode ser importante para realizar ajustes no scanner, uma vez que a partir de um problema encontrado em uma foto é possível saber a posição da câmera naquele instante.

Fotografias não preservam o tamanho e escala dos objetos. Portanto não é possível medir um objeto fotografado em nenhum dos sistemas de coordenadas. Portanto no momento de gerar uma visualização em um sistema de simulação o tamanho do objeto, ou sua escala, deve ser configurado pelo usuário do sistema. Alguns sistemas de modelagem tridimensional apresentam a mesma característica.

Por fim é apresentado o Quadro 1 comparando os sistemas de coordenadas segundo as características de: quantidade de eixos do sistema; eixos responsáveis pelo movimento horizontal, vertical, rotação do objeto e outros, como o movimento da câmera.

Quadro 1 – Quadro comparativo entre sistemas de coordenadas.

Sistemas de Coordenadas	GMC	F2S2	Stream2D	Esférico	Cartesiano
Número de Eixos	4	3	2	3	3
Horizontal	X'	H	Frame	Θ	X e Y
Vertical	Y'	V	Stream	γ	Z
Rotação	B'	B	Frame	Θ	X e Y
Outros	Z'	a		r	

Fonte: Autoria Própria.

4.2 Hardware do scanner F2S2

Após apresentar o conjunto de operações que permitem mapear a posição de determinada fotografia no espaço e como usar esta informação para gerar a visualização, faz-se necessário estabelecer os limites de precisão do F2S2.

Todo componente eletromecânico possui imprecisões e resolução de movimento e em suas especificações estão presentes estas informações. Porém, quando se desenvolvem sistemas eletromecânicos diversos fatores influenciam no desempenho destes componentes, como: peso, instabilidades na corrente elétrica e mesmo temperatura ambiente, que dilata componentes e endurece graxas lubrificantes.

Assim, optou-se por apresentar o processo de calibragem de cada eixo de movimento do F2S2 e apresentar as dificuldades e imprecisões apresentadas pelo equipamento durante o desenvolvimento deste trabalho. Também são apresentadas alternativas e soluções para problemas encontrados.

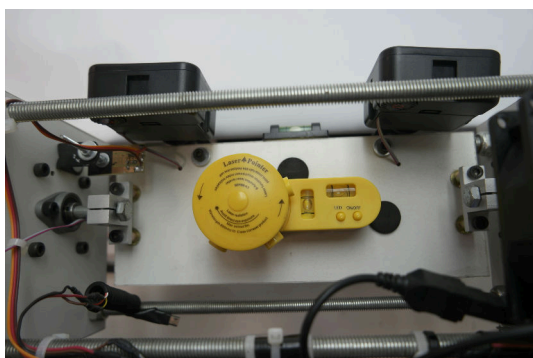
4.2.1 Calibragem do scanner F2S2

O F2S2 passou por um processo de calibragem que visa verificar se os deslocamentos que estão sendo solicitados usando os GMCodes estão sendo executados de forma precisa. Para isso foi usado o *Monitor Serial* do software Arduino IDE³ para enviar instruções em GMCodes de forma manual. Assim, fixava-se uma posição, executava-se um deslocamento e então era feita a medição se o deslocamento foi adequado.

Para facilitar a medição das distâncias e ângulos foi fixado no lugar da câmera um nível a laser (Figura 52). Este equipamento projeta um feixe de laser nos formatos de: ponto central, nível (linha horizontal), prumo (linha vertical) e nível e prumo (formato de cruz). Com os laser marcando determinada posição antes e depois de um deslocamento foi possível medir com uma régua ou transferidor o movimento realizado.

Figura 52 – Nível laser instalado para calibração.

(a) Visão superior.



(b) Visão traseira.



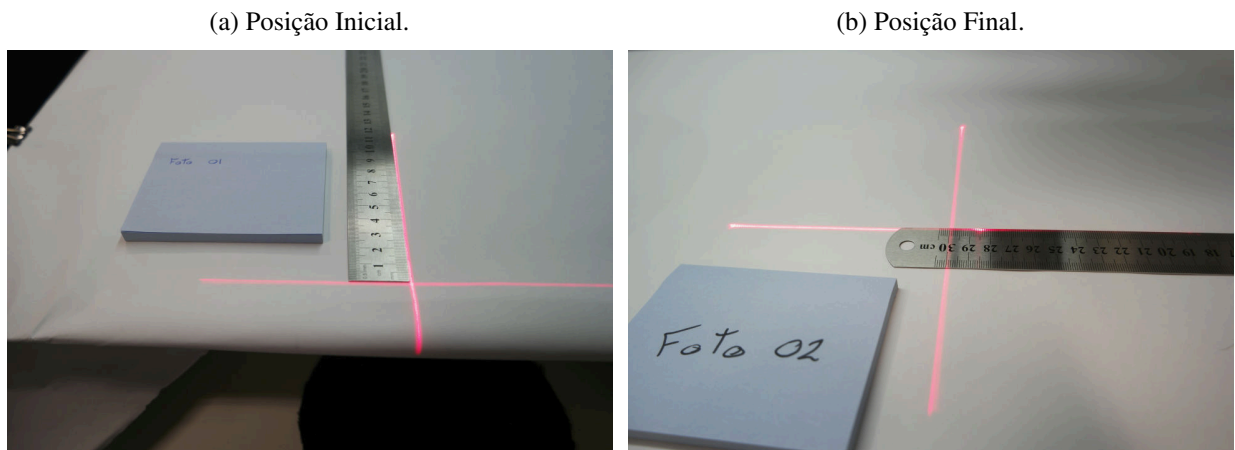
Fonte: Autoria própria.

A correção do scanner consistia principalmente em ajustes finos na posição dos sensores de final de curso dos eixos X' e Y' , e dos sensores ópticos dos eixos Z^a e B^a . Assim foi possível ajustar, por exemplo, se o centro do eixo B^a estava a exatos 650mm da origem do sistema. Caso alguma medida fosse considerada incorreta os sensores eram ajustados para marcarem a posição correta. A calibração ocorreu da seguinte forma:

Eixo X' - deslocamento

Colocou-se o nível para apontar uma determinada posição no fundo infinito; posicionou-se um régua no sentido do eixo X' (Figura 53(a)); moveu-se o eixo X' 30cm (Figura 53(b)). O movimento foi considerado preciso na medição com a régua. O mesmo procedimento foi realizado em vários pontos do fundo infinito e com distâncias diferentes. Em todos os testes o deslocamento foi preciso.

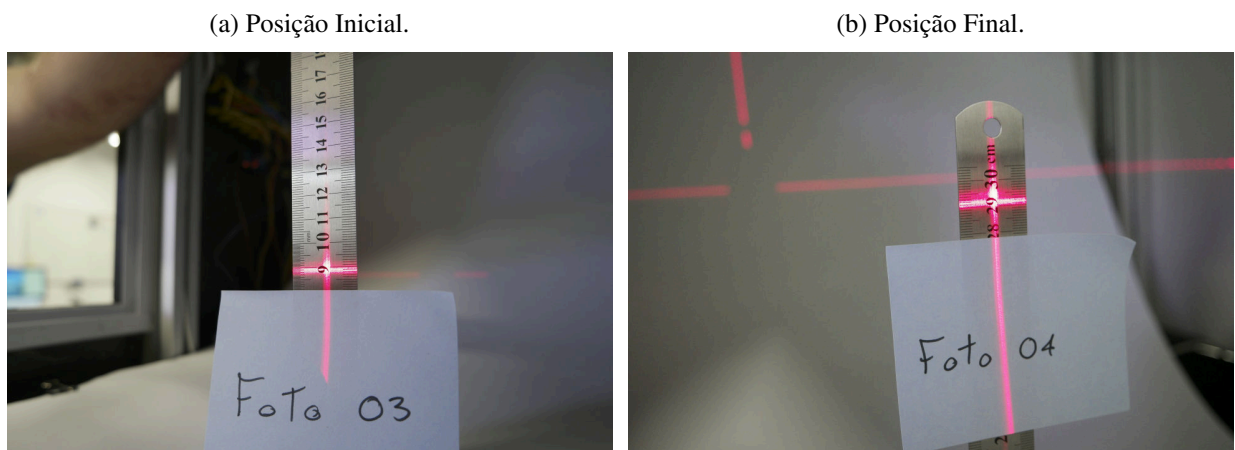
³ <<https://www.arduino.cc/en/Main/Software>> Acesso em: 06 de Março de 2020.

Figura 53 – Calibração do eixo X' .

Fonte: Autoria própria.

Eixo Y' - deslocamento

O mesmo procedimento do teste anterior foi realizado, porém na horizontal, para testar o eixo Y' . Colocou-se o eixo $Y' = 90mm$ (Figura 54(a)); o eixo foi movido para a posição $Y' = 290mm$ (Figura 54(b)). Os testes também consideraram o deslocamento preciso em diversas distâncias e posições.

Figura 54 – Calibração do eixo Y' .

Fonte: Autoria própria.

teste

Distância do eixo X' ao eixo B^a

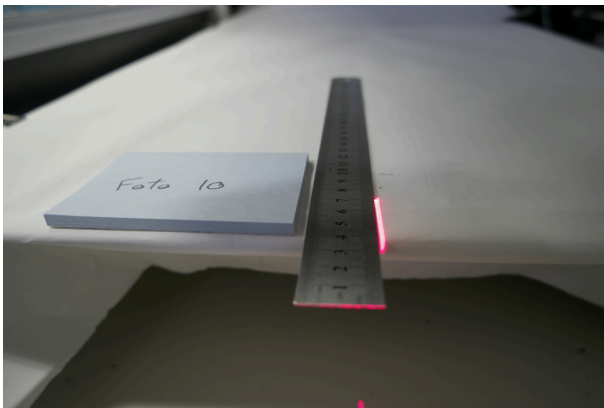
O scanner foi projetado para que o centro do eixo B^a esteja a uma distância de 650mm da posição $X' = 0mm$ na horizontal. Então posicionou-se o eixo $X' = 0mm$, Figura 55(a); foi usado uma régua e uma trena para definir esta posição; instruiu-se o scanner a movimentar o

eixo $X' = 650mm$; esta posição deveria ser o centro do eixo B^a , Figura 55(b). Constatou-se o eixo estava deslocado em 16mm da posição correta.

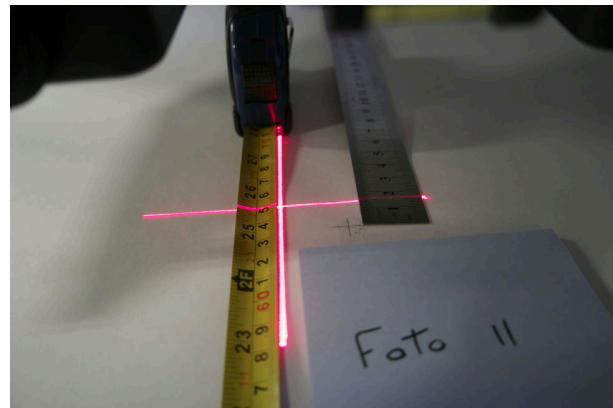
Como foi detectado um problema, o sensor de final de curso do sistema de movimentação do eixo X' foi movido em 16mm. Assim, corrigiu-se o desvio deste movimento. O teste foi realizado novamente e percebeu-se que então o movimento foi preciso. Figura 55(c).

Figura 55 – Distância do eixo X' ao B^a .

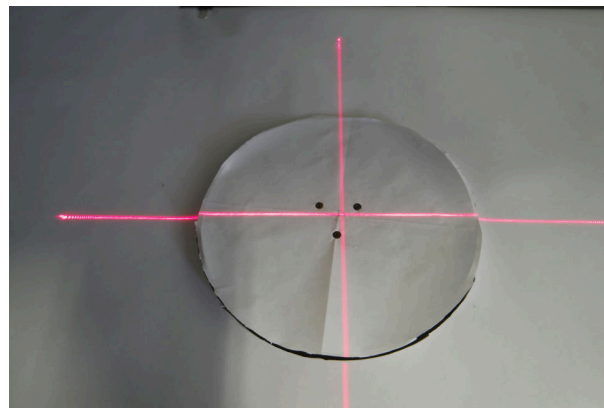
(a) Posição inicial.



(b) Posição final.



(c) Posição corrigida.



Fonte: Autoria própria.

Distância do eixo Y' ao fundo infinito

Para realizar este processo foi alinhada a posição do sensor na câmera (indicado pelo símbolo \emptyset - Figura 56(a)) e o eixo do sistema de movimentação Z^a - Figura 56(b). O objetivo deste teste foi definir a distância mínima da câmera ao fundo infinito para evitar colisões como as apresentadas na Figura - Figura 56(c).

Considerando que o sistema de movimentação do eixo Z^a não pode se chocar com o fundo infinito, usou-se a calibração para determinar a distância mínima que o eixo Z^a pode assumir. O eixo Z^a foi configurado para uma altura de 12cm e então foi sendo movido para

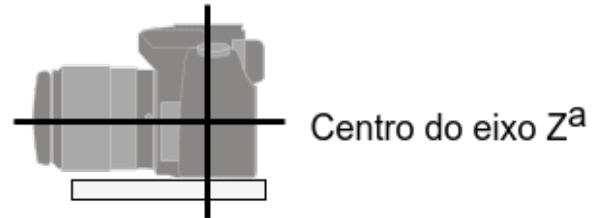
baixo até uma distância segura. A distância considerada segura sem que houvesse uma colisão foi de 85mm (Foto 56(d)).

Figura 56 – Distância mínima do eixo Y' ao fundo infinito.

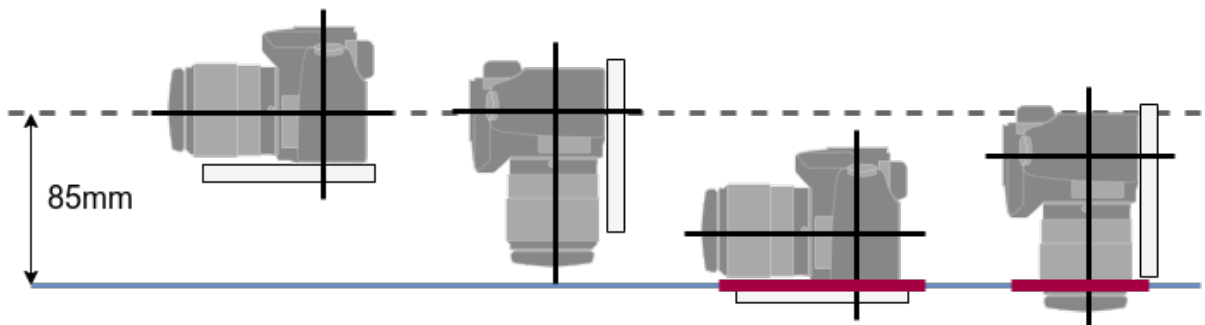
(a) Posição do sensor na câmera.



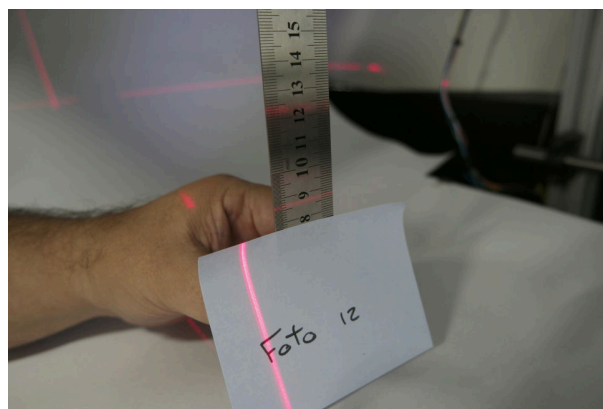
(b) Definição do centro da câmera e eixo.



(c) Colisões.



(d) Realização da calibração.



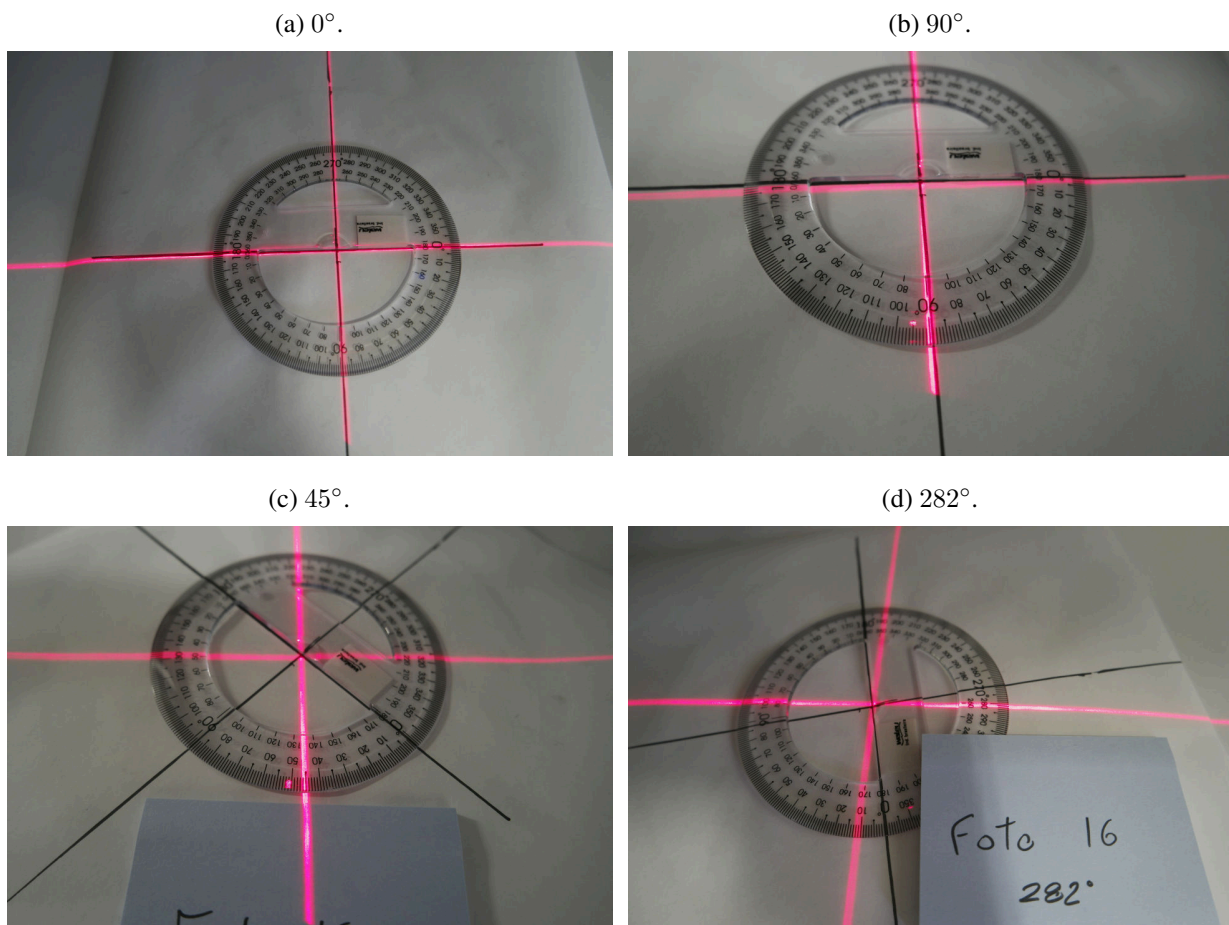
Fonte: Autoria própria.

É necessário estabelecer que diferentes câmeras e lentes apresentam tamanhos diferentes, portanto, ao trocar de câmeras ou lentes pode ser necessário reestabelecer esta distância.

Eixo B^a - Deslocamento angular

Para calibrar o eixo B^a o laser foi posicionado exatamente no centro do eixo B^a , então este eixo foi configurada para a posição de $B^a = 0^\circ$ (Figura 57(a)). Esta posição foi marcada e então foi colado um transferidor centralizado. Posteriormente várias instruções foram passadas ao eixo B^a , como 90° - Figura 57(b); 45° - Figura 57(c); e 282° - Figura 57(d). As movimentações angulares foram conferidas e o sistema apresentou um deslocamento angular preciso. É necessário estabelecer que o sistema foi testado sem peso adicional. Objetos pesados podem alterar a precisão.

Figura 57 – Eixo B^a - Deslocamento angular.



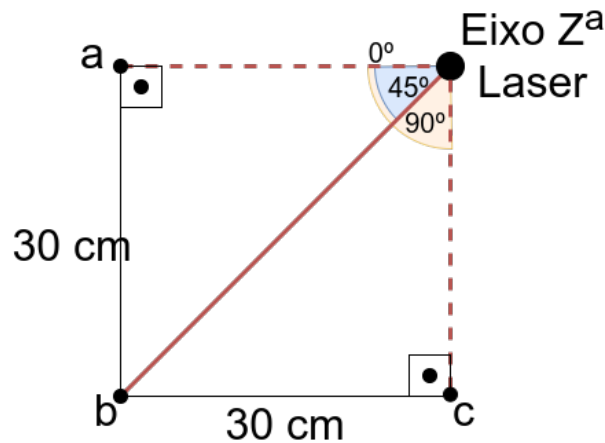
Fonte: Autoria própria.

Eixo Z^a - Deslocamento angular

Para calibrar o eixo Z^a planejou-se o experimento apresentado na Figura 58. O laser foi posicionado a 30cm de altura do plano infinito e uma régua foi posicionada 30cm a frente do laser. Ao configurar o eixo $Z^a = 0^\circ$ o laser deveria iluminar a posição de 30cm da régua (ponto “a”). Com $Z^a = 45^\circ$ o laser deveria iluminar o ponto onde a régua encosta no fundo infinito (Ponto “b”). Na posição $Z^a = 90^\circ$ ele deveria iluminar a posição 30cm antes da régua (esta posição foi

marcada no fundo infinito - ponto “c”). Poderiam ser calculados, usando trigonometria, valores angulares diferentes.

Figura 58 – Eixo Z^a - Planejamento.



Fonte: Autoria própria.

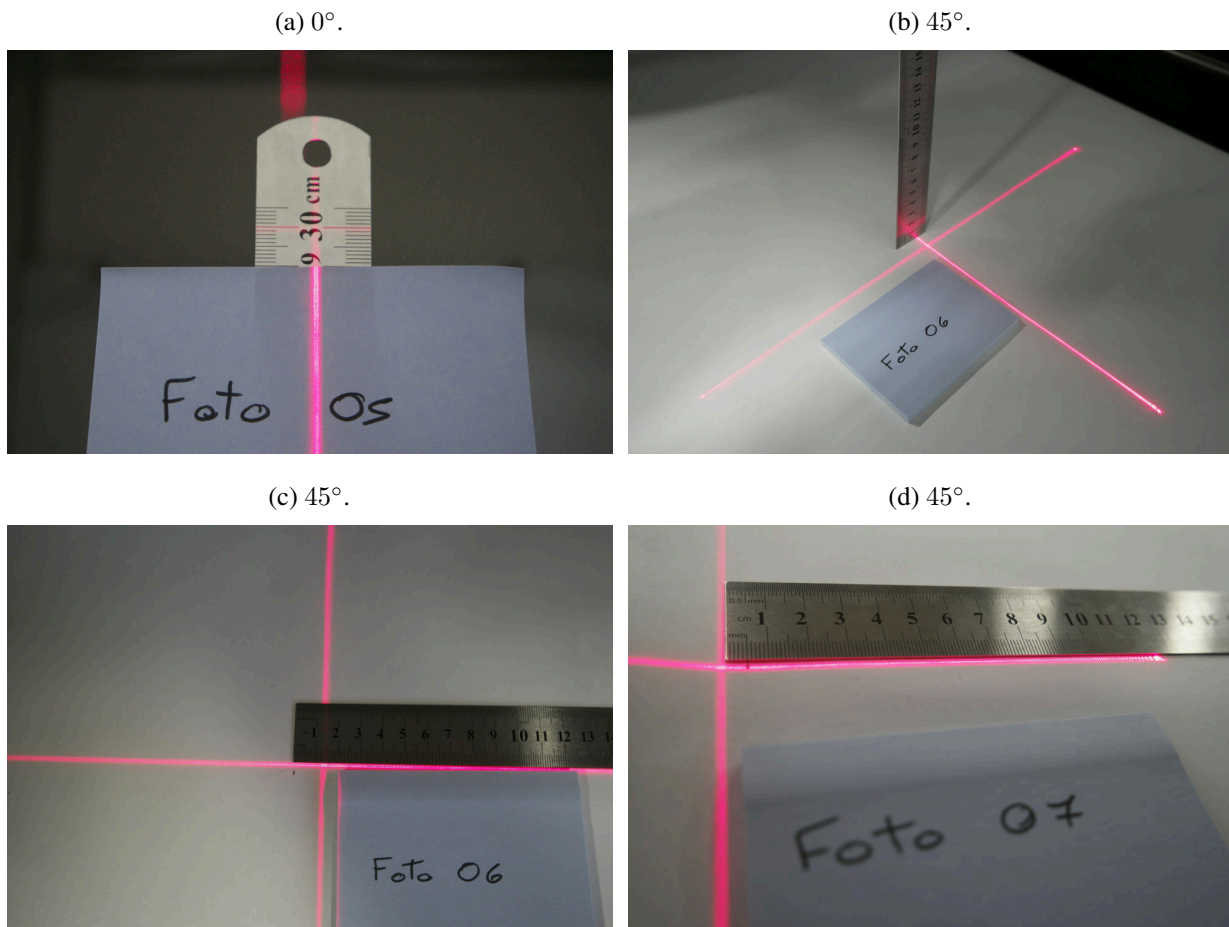
Inicialmente percebeu-se que os dois sensores de final de curso não estavam na posição correta. Usando o laser, um prumo, nível e esquadro os sensores de final de curso foram reajustados para as posições corretas. Então prosseguiu-se para o experimento.

Ao passar uma instrução que usam os sensores de final de curso como referência ($Z^a = 0^\circ$ e 90°), o movimento do eixo é preciso (Figura 59(a)). Porém ao passar instruções que não são limitadas pelos sensores o movimento não é preciso.

Percebeu-se que dependendo do sentido do movimento são percebidas alterações na posição final. Além disso, as posições apontadas não são sempre as mesmas, ou seja, instruindo o sistema à posição $Z^a = 45^\circ$ por três vezes, podem ser apontadas 3 posições levemente diferentes. Este processo pode ser visto nas Figuras 59(b), 59(c) e 59(d), nas quais o valor do eixo era $Z^a = 45^\circ$ e foi constatada uma diferença de mais de 11mm nas posições do laser.

Acredita-se que como em movimentos que decrementam o valor angular, o peso do sistema de eixos é somado à força do motor; enquanto que movimentos que incrementam o valor angular o peso do sistema é subtraído da força do motor. Esta diferença no equilíbrio de forças do sistema de movimento pode ser responsável por, ao instruir o eixo a uma determinada posição, faça com que ele se posicione em posições diferentes. Outros fatores como inércia e atrito também devem estar interferindo no sistema.

Ainda sobre o eixo Z^a , constatou-se que durante o movimento de 0° a 90° existe uma imprecisão de 5° no movimento, distribuído de forma linear. Ou seja o sistema realiza um deslocamento de 16° a cada 17° que deveriam ser realizados. Isso também pode ser atribuído ao peso, inércia e/ou atrito do sistema. Uma solução para o problema de imprecisão do eixo Z^a é apresentado na subseção 4.2.2.

Figura 59 – Eixo Z^a - Deslocamento angular.

Fonte: Autoria própria.

4.2.2 Manutenção, dificuldades e imprecisões enfrentadas

Esta Subseção trata das dificuldades enfrentadas com o F2S2 e foi dividida em três partes: Inicialmente trata-se das manutenções de rotina do F2S2; posteriormente são apresentadas algumas dificuldades que ocorreram com o uso do F2S2; por último são discutidas as imprecisões do sistema e são apontadas soluções para elas.

Manutenção

Todo equipamento fotográfico precisa de limpeza regular, assim, recomenda-se que a lente da câmera esteja sempre limpa e livre de marcas e arranhões. O mesmo vale para o sensor da câmera, que pode sujar quando as lentes são trocadas. Recomenda-se a limpeza sempre que for percebida sujeira na câmera ou nas fotografias, o que pode significar que o sensor está sujo.

O processo de calibração do scanner também deve ser realizado de forma regular. Colisões ou vibrações dos sistemas de movimentação podem afrouçar parafusos e deslocar os sensores da posição correta. Assim, o nível a laser pode ser acoplado no lugar da câmera e alguns testes

podem ser realizados. Caso exista necessidade, ajusta-se os sensores.

Dificuldades

Os sistemas de movimentação lineares do F2S2 são lubrificados com graxa. Em dias muito frios, normalmente após 3 dias sem uso e com temperaturas muito baixas, a graxa solidifica e faz com que os motores não consigam vencer o atrito inicial do movimento, travando todo o sistema.

Nestes casos a solução encontrada foi auxiliar manualmente os motores para que estes consigam iniciar o movimento. Então faz-se alguns movimentos que percorram toda a extensão de todos os eixos para que a graxa se movimente dentro das castanhas e rolamentos, liberando o movimento.

O sistema de movimentação do eixo B^a não é preciso com objetos com peso superior a 3kg ou 4kg, dependendo do formato do objeto. Esta característica limita os objetos que podem ser escaneados. A solução para esta dificuldade seria trocar o motor por um modelo mais forte ou instalar uma caixa de redução no motor. Caixas de redução conseguem diminuir a velocidade do movimento, mas aumentar o torque do motor, o qual conseguiria trabalhar com objetos mais pesados.

O F2S2 está instalado próximo a laboratórios que usam equipamentos com alto consumo de energia. Dependendo do uso destes equipamentos a rede de energia elétrica que alimenta o F2S2 sofre oscilações. Nestes casos as lâmpadas da sala sofrem uma oscilação luminosa e alguns estabilizadores de energia e *nobreaks* chegam a ser acionados.

No F2S2 estas oscilações interferem na iluminação interna, o que pode ser visível nas fotografias adquiridas. Já aconteceram situações em que o Arduino que controla o scanner foi reiniciado e também existem suspeitas de que alguns movimentos do scanner não foram realizados e algumas fotografias deixaram de ser tiradas. Trata-se esta situação como suspeita pois a energia não estava sendo monitorada no momento em que aconteceram.

O problema mais comum é a interferência na iluminação. Uma aquisição do scanner, dependendo da quantidade de fotos, pode levar seis a oito horas. Assim, a instabilidade elétrica pode causar uma quantidade expressiva de retrabalho, para refazer aquisições que foram inviabilizadas por instabilidade no sistema elétrico.

Imprecisões

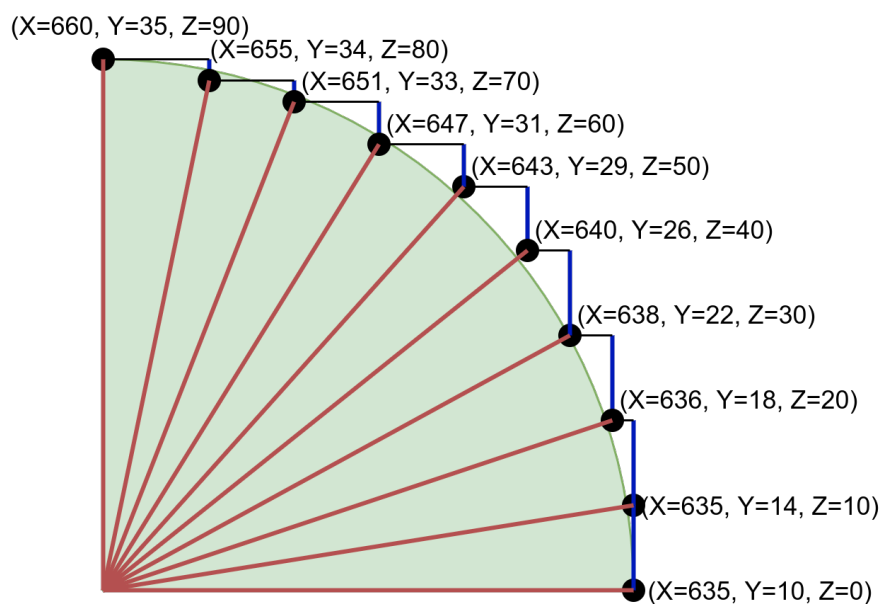
Para discutir o problema da imprecisão do F2S2 será usado como exemplo uma aquisição de um inseto de 2cm de comprimento e de altura. Apenas para fins de comparação, uma mosca possui algo em torno de 0.7cm e uma abelha 1.5cm. Para realizar a aquisição desta peça é necessário usar lentes macro ou lentes com anel adaptador com fator de aumento. No caso do laboratório possuímos anéis adaptadores para lentes com aumento de 2x, 4x, 8x e 10x.

Ao usar estes anéis adaptadores com fator de aumento, a distância entre a lente e o objeto é reduzida. No caso da câmera Samsung nx500 do laboratório, com a lente de 18-55mm e anel com fator de aumento de 10x, a distância do objeto à lente, em que a câmera consegue gerar imagens com foco, é em torno de 2 a 3cm.

Assim, uma possível aquisição para esta configuração teria um raio de 25mm do objeto. Para terminar a configuração da aquisição considere uma aquisição de dez Streams e um Frame por Stream. Considere também que o inseto está posicionado sobre uma base de 80mm de altura. Os parâmetros de configuração de uma aquisição são apresentados na Seção 4.3.

Considerando os parâmetros apresentados, a Figura 60 apresenta um desenho das posições da câmera para esta aquisição. Na Figura em vermelho são apresentados os ângulos que o eixo Z^a assumiria para fotografar o objeto, em azul os deslocamentos no eixo Y' , em preto os deslocamentos do eixo X' e em verde o arco ideal da aquisição. Na figura os pontos onde as retas vermelhas, azuis e pretas se tocam são as posições onde uma fotografia seria tirada. Pode-se notar que como o firmware do F2S2 não trabalha com frações de milímetros, as posições calculadas para as fotografias não são perfeitamente coincidentes com o arco verde, que representa a situação ideal.

Figura 60 – Aquisição de inseto com raio de 25mm.



Fonte: Autoria própria.

A Tabela 1 apresenta as coordenadas calculadas para a aquisição. Na Tabela 1 estão as coordenadas para os eixos X' , Y' e Z^a , assim como os deslocamentos de cada eixo entre cada movimento. O processo de cálculo das coordenadas é apresentado na Seção 5.1.

A Figura 61 apresenta uma estimativa para a imprecisão. O arco em verde representa o formato ideal, o traçado em preto representa as coordenadas calculadas, o traçado em azul

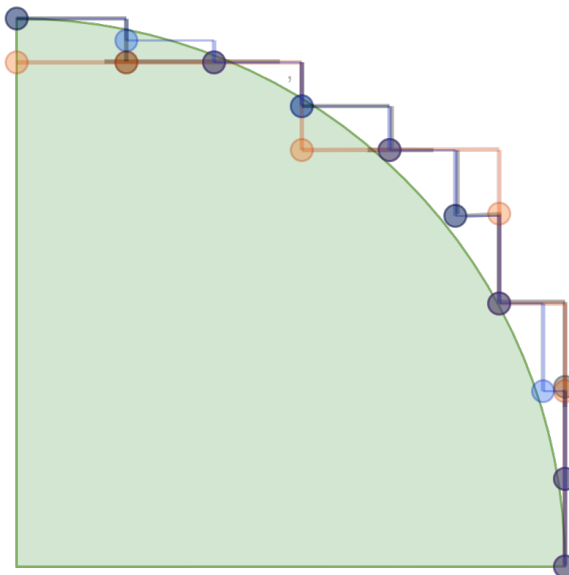
Tabela 1 – Relação de coordenadas de uma aquisição de objeto de 2cm.

# Frame	Coordenadas			Deslocamento		
	X	Y	Z	X	Y	Z
1	635	10	0	–	–	–
2	635	14	10	0	4	10
3	636	18	20	1	4	10
4	638	22	30	2	4	10
5	640	26	40	2	4	10
6	643	29	50	3	5	10
7	647	31	60	4	2	10
8	651	33	70	4	2	10
9	655	34	80	4	1	10
10	660	35	90	5	1	10

Fonte: Autoria própria.

representa o movimento da câmera caso movimentos com deslocamento de até 1mm não sejam realizados até que o somatório do deslocamento seja maior que 1mm, em laranja é apresentado o deslocamento com valor mínimo de deslocamento sendo 2mm. Estes valores são uma estimativa, pois não possuímos instrumentos confiáveis e precisos para realizar esta medição em um caso real. Porém empiricamente a imprecisão do sistema aparenta ser algo em torno de 2mm.

Figura 61 – Imprecisão na aquisição de inseto com raio de 25mm.



Fonte: Autoria própria.

Como pode ser observado na Figura 61, para aquisições com raio pequeno, como no caso do exemplo apresentado, a diferença entre o deslocamento ideal e o realizado pode ser considerável e seria perceptível na visualização do stream.

Este tipo de aquisição não poderia ser executada por câmeras com lentes comuns, como as lentes 18-55mm, uma vez que estas lentes não focam com distâncias menores que 20 a 30cm (forçando um raio de aquisição maior), dependendo do modelo ou fabricante. Este tipo de problema só ocorreria com lentes equipadas com anéis com fator de aumento. Estes anéis diminuem consideravelmente a distância entre a lente e o objeto.

Para resolver este tipo de imprecisão poderiam ser usados fusos mais precisos, com avanço menor e castanhas com sistemas anti-folga. Além disso para diminuir a imprecisão que os motores impõe ao sistema, poderiam ser usadas caixas de redução, que aumentam o torque, mas diminuem a velocidade do movimento. Ajustando os sistemas de movimentação do F2S2 com estes dois equipamentos sua movimentação linear seria muito mais precisa.

No que se refere à imprecisão de movimentos do eixo Z^a , que apresenta uma variação aleatória, além de perder 1° a cada 17° de movimento, somando 5 graus ao final de 90° , aquisições com raio pequeno como os 25mm do exemplo apresentado se tornam ainda mais imprecisas que aquisições com um raio maior. Pela aproximação do objeto à câmera, qualquer imprecisão de movimento do eixo Z^a torna-se perceptível.

A Tabela 2 apresenta variação angular do eixo Z^a acrescida da perda de 1° a cada 17° . Esta variação foi medida diversas vezes usando transferidor, medidor de ângulos e com o laser acoplado no lugar da câmera. Porém nenhum destes métodos apresenta precisão maior que 1° ou confiabilidade na leitura dos dados. Portanto, apesar de empiricamente possuímos uma previsão do comportamento do eixo Z^a , estes valores são aproximados.

Tabela 2 – Valores angulares de imprecisão do eixo Z^a .

Valor da instrução	Valor executado	Erro acumulado
0	0	0
10	10	0
20	19	1
30	29	1
40	38	2
50	47	3
60	57	3
70	66	4
80	76	4
90	85	5

Fonte: Autoria própria.

Considerando o hardware atual do F2S2, foram desenvolvidas duas formas de corrigir esta imprecisão. A primeira forma, ignora-se o erro de até 5° durante toda a aquisição, porém no último Stream, com $Z^a = 90^\circ$, onde o eixo Z^a estaria marcando algo em torno de 85° , corrige-se o valor angular para 90° usando o sensor óptico do eixo.

Esta abordagem faz com que apenas o último Stream sofra um salto perceptível na

visualização. Além disso, ao longo da visualização o objeto sai do centro da tela e se posiciona cada vez mais para baixo na imagem. Dependendo do enquadramento da câmera durante a aquisição pode ocorrer do objeto ser cortado nas fotografias.

A segunda abordagem é acrescentar 1° a cada 17° de movimento angular do eixo Z^a . Porém este processo, para aquisições com vários Streams, torna-se perceptível, pois o objeto vai se deslocando para baixo na visualização e de repente, quando o movimento é corrigido, ele volta a se posicionar no centro da fotografia.

A Tabela 3 apresenta uma aquisição com $\Delta\gamma = 2^\circ$. A primeira coluna contém o valor angular instruído para o eixo Z^a , a segunda coluna uma estimativa do movimento executado e a terceira a variação angular entre cada movimento. A linha em cinza apresenta a correção realizada. Percebe-se que a correção é de aproximadamente 3° sendo que deveria ser 2° . Ou seja, além de corrigir a posição do objeto para o centro da imagem a variação angular é 50% maior que as anteriores. Este processo também é perceptível durante a visualização.

Tabela 3 – Valores angulares de correção da imprecisão do eixo Z^a .

Valor da instrução	Valor executado	$\Delta\gamma$ executado
14	13	≈ 2
16	15	≈ 2
18	18	≈ 3
20	20	≈ 2
22	22	≈ 2

Fonte: Autoria própria.

A solução definitiva para este problema seria a troca do motor do eixo Z^a por um motor mais potente e preciso, e/ou a instalação de uma caixa de redução no motor. Com isso, o atrito e o peso do sistema de movimentação com a câmera teriam menos influência no movimento, promovendo mais precisão.

Existe ainda um problema de origem desconhecida que faz com que algumas fotos sejam adquiridas, sem que seja enviado o comando responsável por isso; e também algumas fotos não são adquiridas mesmo que o comando seja enviado.

No caso de fotografias extras, acredita-se que sejam resultado de ruídos elétricos do sistema de alimentação do F2S2, que acabam por ativar a câmera. Estas fotografias extra são facilmente identificadas com a contagem do número de Frames adquiridos e com uma análise visual do *pack* de fotografias. Fotos extra sempre são tremidas, pois não ocorrem quando o sistema está estabilizado.

Para solucionar este problema de forma definitiva foi instalado um núcleo de ferrite no cabo que se comunica com a câmera. Este filtro estabiliza o sinal recebido pela câmera e impede que a mesma seja acionada acidentalmente por ruídos elétricos ou eletromagnéticos. As fotos

extra eram relativamente raras, acontecendo, aproximadamente, uma foto a cada 4 ou 5 *packs* de 3.720 fotos cada.

No caso de fotos faltantes, a causa é desconhecida. Pode ocorrer por ruídos elétricos, mal contato, instabilidade da rede elétrica, um sinal que seja enviado com menos potência do que deveria por que algum outro componente do scanner estava usando energia, ou mesmo algum procedimento interno da câmera, como a transferência de fotos da memória interna para o cartão SD, que impeça a câmera momentaneamente de adquirir aquela fotografia.

Como as causas (pois podem ser mais de uma) podem ser internas ou externas, não foi possível mapear sua origem. O núcleo de ferrite aparentemente diminuiu a ocorrência de fotos faltantes, mas não resolveu definitivamente o problema, que ocorre aproximadamente uma vez a cada dois ou três *packs* de 3.720 fotos cada (os *packs* mais comuns de serem adquiridos).

Fotos faltantes são um problema mais significativo do que fotos excedentes, pois não existe a possibilidade de manter o *pack* íntegro, como ocorre nas fotos extras, pois não é possível readquirir exatamente a mesma foto, uma vez que a câmera já foi desligada e o objeto retirado do scanner. Até o momento, para contornar o problema, identificamos com uma análise visual qual a foto que está faltando e então duplicamos a foto anterior ou posterior para ocupar o lugar da faltante.

Este processo é necessário para manter a integridade do Stream-2D, que não foi projetado para suportar um número de fotos que seja diferente do estabelecido. Caso a duplicação não fosse realizada, todas as fotos posteriores à faltante estariam na posição anterior, alterando, inclusive, os inícios de Streams, que representam uma alteração em $\Delta\Theta$.

Com este processo a visualização é mantida e a foto faltante possui uma probabilidade baixa de aparecer, e provavelmente nem será percebida por quem estiver visualizando as aquisições. Uma foto em um *pack* de $\Delta\Theta = 3^\circ$ e $\Delta\gamma = 3^\circ$ significa uma probabilidade de $\frac{1}{3720}$. Porém, caso este Frame esteja sendo usado para estereoscopia, ela não ocorrerá (nesta posição) de forma adequada. Neste caso para contornar o problema da estereoscopia, podem ser usadas outras formas de seleção do par estéreo para que a estereoscopia não seja afetada.

4.3 Protocolo de Aquisição

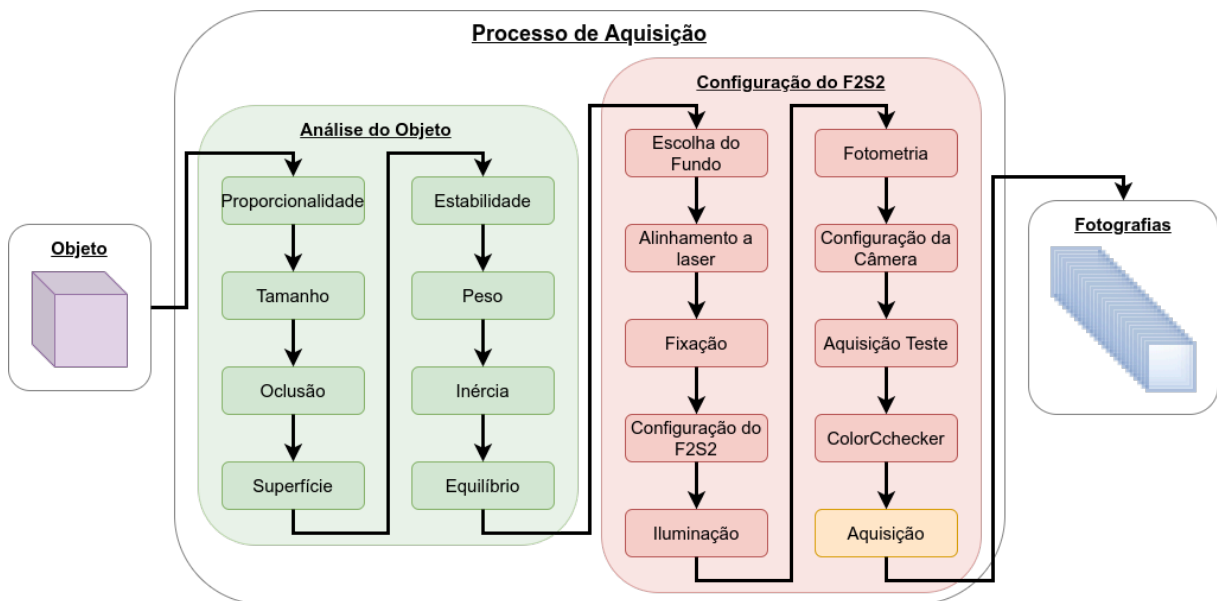
Esta Seção apresenta o protocolo de aquisição. Inicialmente debate-se os objetos a serem escaneados e como as características destes objetos interferem na aquisição e na visualização. Após analisar o objeto, discute-se o material de cobertura do fundo infinito do scanner e como escolher o material e sua cor.

Então debate-se como posicionar o objeto corretamente no scanner. Este posicionamento é essencial para gerar uma boa visualização, pois caso o objeto não esteja corretamente posicionado, ao rotacionar o eixo B^a o objeto realiza, além de um movimento de rotação (sobre o

próprio centro), um movimento de revolução (em torno do eixo B^a). Em seguida apresenta-se forma de fixação do objeto, para mante-lo na posição desejada sem que tombe ou desloque-se. Também discute-se como cada sistema de fixação interfere na peça e na visualização.

Em seguida discute-se os parâmetros da aquisição, iluminação e configuração da câmera. Define-se uma aquisição de teste das configurações para evitar que alguma configuração inviabilize o uso das fotografias adquiridas. Apresenta-se o uso do ColorChecker para aquisições com cores fiéis às cores reais do objeto. Por fim é debatida a aquisição e o resultado do processo. A Figura 62 apresenta o processo de aquisição e suas etapas.

Figura 62 – Processo de Aquisição.



Fonte: Autoria própria.

4.3.1 Objeto a ser escaneado

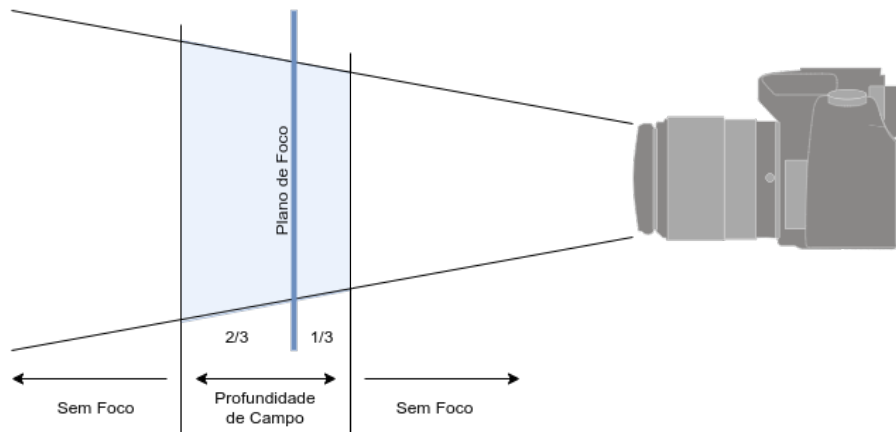
Antes de iniciar uma aquisição é necessário avaliar o objeto que será escaneado. As principais características que devem ser avaliadas são: proporcionalidade, tamanho, forma, textura, superfície, natureza, oclusão, rigidez, estabilidade da forma, peso, inércia e equilíbrio.

Razão entre altura, largura e profundidade

O tamanho das peças influencia diretamente na câmera e na lente que deve ser usada na aquisição. A razão entre a altura, largura e profundidade do objeto influenciam na configuração da câmera, uma vez que altera as necessidades de foco e profundidade de campo.

A profundidade de campo distribui-se a uma proporção aproximada (dependendo da câmera, lentes e configuração) de $\frac{1}{3}$ a frente do plano de foco e $\frac{2}{3}$ atrás do plano de foco. Como pode ser visto na Figura 63. Este é um processo natural da óptica e, portanto, o comportamento intrínseco da fotografia.

Figura 63 – Profundidade de Campo.



Fonte: Autoria própria.

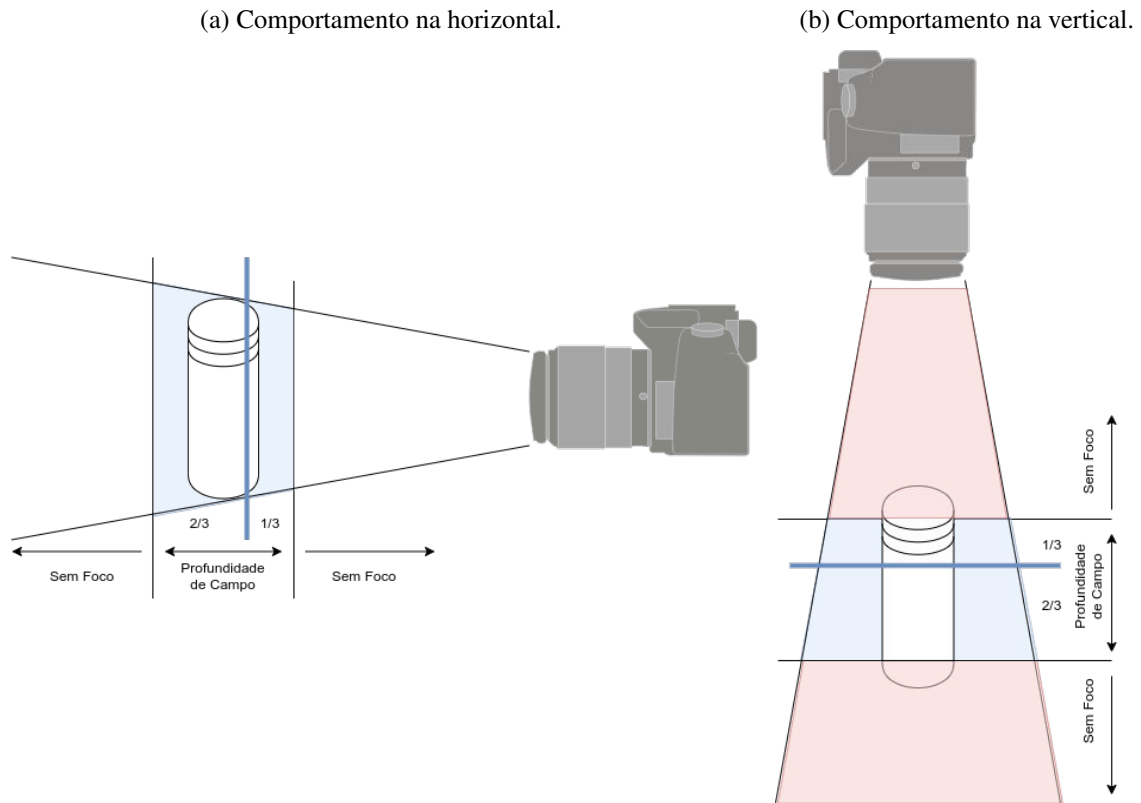
Porém, como o scanner adquire fotografias de diversos ângulos do objeto, e a configuração da câmera precisa ser a mesma para todas as fotografias (conforme explicado na Subseção 4.3.2), o comportamento da profundidade de campo ao longo do *pack* de imagens pode ocasionar dificuldades que não existem na fotografia convencional.

Considere por exemplo a aquisição de uma garrafa térmica com medidas aproximadas de 10cm de largura, por 36cm de altura. Como a razão entre altura e largura é de $\frac{10}{36} = 0,277$, a configuração da câmera deverá contemplar duas condições muito diferentes para a profundidade de campo. Conforme pode ser visto na Figura 64.

Como pode ser visto na Figura 64(a), com a câmera posicionada na horizontal, o objeto está integralmente na região com profundidade de campo, ou seja, aparecerá nítido nas fotografias. Porém ao longo da aquisição, quando a câmera estiver posicionada na vertical (Figura 64(b)), partes do objeto estarão fora da profundidade de campo (área representada em vermelho), fazendo com que estas partes do objeto sejam fotografados sem nitidez.

Caso o objeto do exemplo anterior seja escaneado na horizontal (Figura 65), além de problemas de oclusão do objeto por este estar na horizontal, temos o mesmo problema com a profundidade de campo. Com o objeto posicionado perpendicularmente ao sentido da câmera, ele permanece em foco (Figura 65(a)). Porém, com a variação de Θ (Figura 65(b)), com o objeto posicionado paralelamente ao sentido da câmera, partes do objeto saem da região com profundidade de campo (representada em vermelho) e conseqüentemente são representadas sem foco na fotografia.

No exemplo acima foi considerado uma garrafa com razão de $0,277$ entre altura e largura. Porém existem objetos com razões ainda piores, por exemplo, um cabo de vassoura com espessura de $2,4\text{cm}$ e 120cm de altura possuem uma razão de $\frac{2,4}{120} = 0,02$, ou seja é uma diferença de escala de $1 : 50$. Para minimizar este problema devem ser usadas técnicas fotográficas que maximizem a profundidade de campo, como: valores de f maiores e aumentar a distância entre

Figura 64 – Comportamento da profundidade de campo durante a aquisição - Variação γ .

Fonte: Autoria própria.

objeto e câmera.

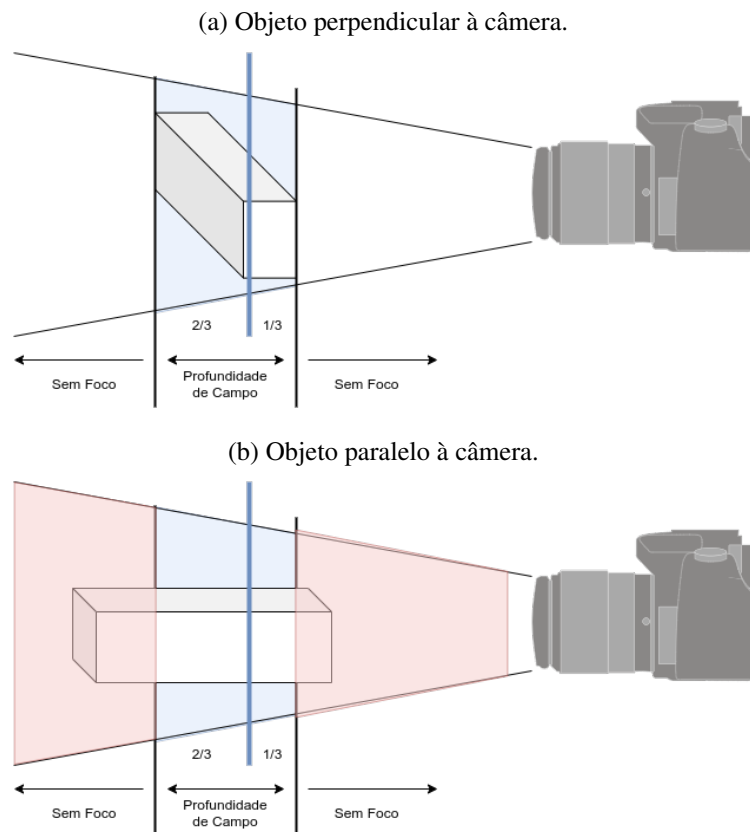
Além disto, ao se configurar o foco para a aquisição deve ser considerada a profundidade de campo e a distribuição desta ao longo do espaço. Em uma fotografia comum (Figura 66(a)), tenta-se colocar o foco alinhado ao início do que se está fotografando, para melhorar o foco deste objeto na fotografia.

Este é o processo usado também no foco automático das máquinas digitais. Ao usar a função de foco automática das câmeras, estas analisam o foco em uma região pré-selecionada pelo usuário e ajustam o foco para que aquela região permaneça no plano de foco.

Porém, como no caso do F2S2 são adquiridas várias fotografias enquanto o objeto é reposicionado, é importante considerar este movimento para configurar o foco. Empiricamente percebeu-se que colocar o foco um pouco atrás do início do objeto traz melhores resultados, pois respeita a distribuição da profundidade de campo. Este processo pode ser visto na Figura 66(b).

Percebeu-se que configurando o foco pensando na profundidade de campo e não apenas no plano de foco, quando objetos irregulares rotacionam, tendem a permanecer dentro da área em foco. Ajustando simplesmente o foco para a primeira face do objeto, quando este rotacionar-se, é mais provável que partes do objeto fiquem sem foco. Porém, por se tratar de uma percepção empírica, é necessário considerar a configuração de foco e profundidade de campo para cada

Figura 65 – Comportamento da profundidade de campo durante a aquisição - Variação Θ .



Fonte: Autoria própria.

objeto que será escaneado.

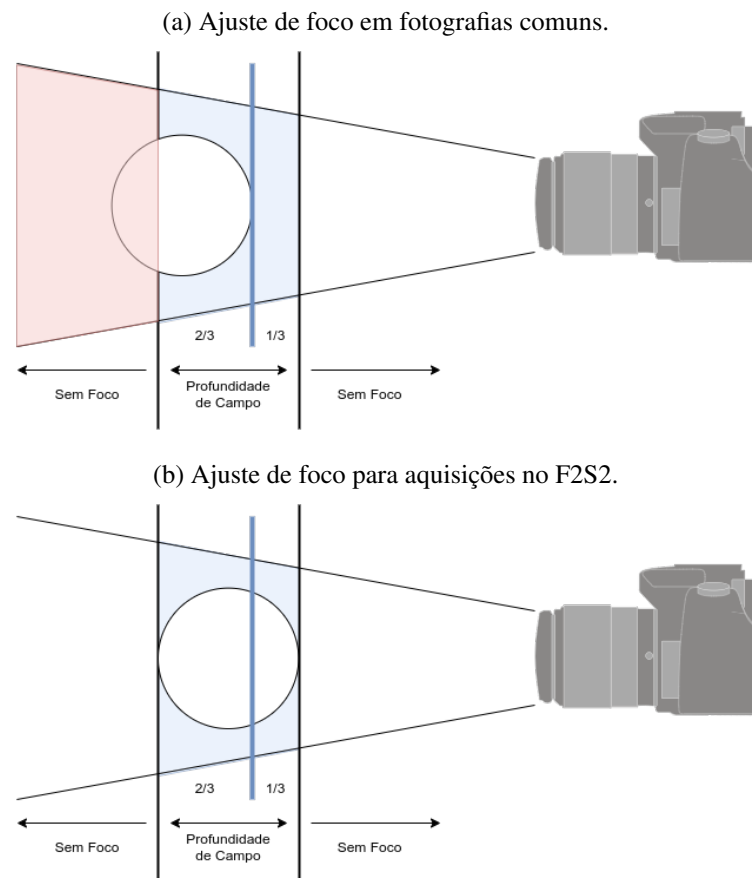
Considerando que uma aquisição pode levar mais de 12 horas e que diversas variáveis podem influenciar na profundidade de campo, como: razão entre as medidas, orientação da peça, tamanho, configuração da câmera, distância da câmera ao objeto, entre outros. A melhor forma encontrada para otimizar a profundidade de campo nas fotografias foi realizar uma aquisição de teste com poucas fotos (algo como cinco Streams e quatro Frames por Stream), analisar o resultado e ajustar o parâmetros necessários, como discutido na Subseção 4.3.2.

Tamanho

O tamanho das peças influencia diretamente na escolha da câmera e objetivas; no raio da aquisição; e na necessidade de base de suporte para a câmera. Para discutir o tamanho das peças escaneadas, estas foram divididas em grupos que levam em conta o tamanho das peças, considerando formas regulares.

É necessário estabelecer que as discussões sobre tamanhos estão relacionadas ao atual scanner, câmera e objetivas. Caso as dificuldades de precisão do scanner, apresentadas na Subseção 4.2.2, sejam contornadas e seja acoplada uma nova câmera e objetivas, os limites inferiores

Figura 66 – Ajuste de foco e profundidade de campo para aquisições do F2S2.



Fonte: Autoria própria.

apresentados não seriam mais válidos. Além disto, caso um scanner maior seja construído, os limites superiores de tamanho também não seriam determinantes. Ou seja, o modelo proposto pode ser extrapolado para objetos com tamanhos diferentes dos aqui apresentados. Além disto, as medidas são referenciais e não valores exatos.

objeto > 35cm : Considerando que o F2S2 consegue realizar aquisições com o raio de até 65cm e que a câmera precisa de algo em torno de 20 a 30cm de distância do objeto para conseguir enquadrá-lo e focá-lo, objetos com aproximadamente 35cm estabelecem o limite máximo para o tamanho dos objetos em uma aquisição completa. Objetos maiores podem ser adquiridos, porém, dependendo das dimensões, a câmera não conseguirá enquadrá-lo completamente, nem se configurar para que o objeto esteja em foco.

16cm < objeto < 35cm : Com estes objetos o scanner e a câmera conseguem realizar a aquisição completa sem dificuldades no que se refere às distâncias do objeto, foco, enquadramento, precisão do movimento. Também não são necessárias bases de suporte para elevar o objeto. Constituem o tamanho ideal de objetos para serem adquiridos pelo F2S2.

4cm < objeto < 16cm : Itens com estas dimensões podem ser adquiridos sem dificuldades,

porém precisam de uma base de suporte para elevar o objeto. A esfera formada durante a aquisição deve estar alinhada como o centro geométrico do objeto. Como a câmera não pode ficar a uma altura inferior de 8.5cm do fundo infinito (ver Figura 56c), o objeto deve ser suspenso com uma base de suporte para que seu centro respeite este limite. O processo de uso das bases de suporte é explicado na Subseção 4.3.2.

1cm < objeto < 4cm : Além da necessidade de uso de base de suporte, para objetos destas dimensões, são percebidas dificuldades no processo de alinhamento do objeto sobre o eixo B^a (discutido na Subseção 4.3.2) e de fixação do objeto para que este não se desloque com o movimento do eixo B^a (Subseção 4.3.2). A Figura 67 apresenta um inseto de 4cm fotografado sobre uma base de suporte.

Figura 67 – Foto de inseto com lente 18-55mm.



Fonte: Autoria própria.

Objeto < 1cm : Objetos menores que um centímetro constituem o maior desafio para o atual scanner, câmera e objetivas. Os principais desafios de aquisição deste tipo de objetos são:

Objetivas : Lentes comuns, 18-55mm, possuem uma distância mínima do objeto, para que consiga fotografar com foco, de aproximadamente 20cm . Neste caso um objeto menor do que um centímetro ocuparia uma região muito pequena da fotografia, inviabilizando a visualização com acuidade visual e aplicação de zoom. A Figura 68(a) apresenta este problema.

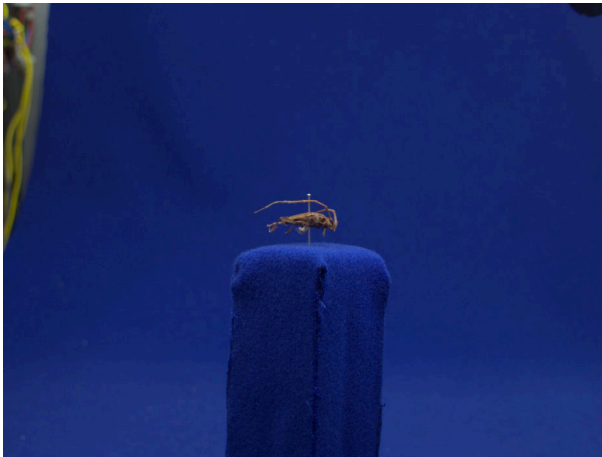
Para objetos deste tamanho o ideal é o uso de lentes macro⁴ com taxa de ampliação 1:1 . Estas lentes conseguem fotografar com distâncias pequenas entre a lente e o objeto e maximizar a profundidade de campo. Além disto, a taxa de ampliação 1:1

⁴ <<https://www.sony.com.br/electronics/lentes-fotografia-macro>> Acesso em: 26 Mar. 2020

faz com que, caso o objeto esteja próximo o suficiente da lente, este seja capturado com uma proporção de 1:1 entre o tamanho real e o representado no sensor da câmera. Assim mais detalhes seriam capturados. A Figura 68(b) apresenta um inseto de 6mm fotografado com lente macro 1:1.

Figura 68 – Fotos de objetos menores que 1cm.

(a) Inseto com lente 18-55mm



(b) Inseto com lente Macro 1:1



Fonte: Autoria própria.

Profundidade de campo : A profundidade de campo é um desafio na fotografia de objetos pequenos. No caso do F2S2, como o objeto/câmera se movimentam para adquirir fotografias de outros ângulos, a profundidade de campo se torna crítica.

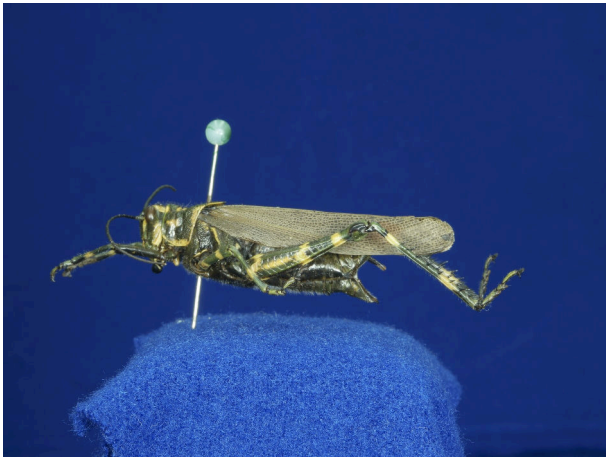
A Figura 69 apresenta fotografias de um gafanhoto adquiridas no scanner. A Figura 69(a) apresenta o gafanhoto perpendicular à câmera, de forma que todo inseto está mais próximo do plano de foco e por isso possui maior probabilidade de estar dentro da profundidade de campo (ver Figura 65(a)). Mesmo neste caso, ao ampliar a fotografia (Figura 69(b)) percebe-se que a pata do animal que está em primeiro plano está nítida, enquanto que a outra pata está sem foco. Mesmo a pequena distância entre duas patas de um gafanhoto foi o suficiente para retirar a segunda da região nítida.

Na Figura 69(c) pode-se perceber que a cabeça do animal está nítida, porém outras partes do corpo do inseto mais afastadas do plano de foco, como as patas traseiras, tornam-se desfocadas (ver Figura 65(b)).

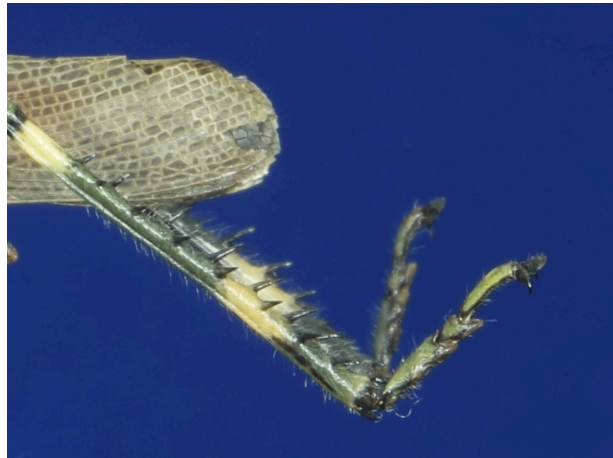
Imprecisão do movimento : Para objetos muito pequenos a imprecisão de movimentos angulares ou lineares (conforme discutido na Subseção 4.2.2) se tornam mais visíveis do que para objetos maiores. Com objetos de menos de um centímetro a uma distância de 2cm da câmera, um erro no deslocamento de 1mm será perceptível na visualização. Considere que uma lente macro 1:1 representa um objeto de 10mm de comprimento por 10mm de altura com as mesmas dimensões no sensor da câmera. Um erro de

Figura 69 – Fotos de gafanhoto.

(a) Fotografia completa.



(b) Ampliação das patas traseiras.



(c) Gafanhoto com o plano de foco alinhado à cabeça do animal.



Fonte: Autoria própria.

1mm no deslocamento neste caso representa 10% do tamanho do objeto e de sua representação no sensor. Isto projetado em uma tela também representará 10% do tamanho do objeto.

A Figura 70 apresenta a diferença que a imprecisão de movimento do F2S2 pode

causar nas fotografias e a percepção de uma imprecisão em fotos de objetos pequenos muito próximos à lente. A Figura 70(a) apresenta um objeto pequeno próximo à lente macro 1:1. As figuras em azul e vermelho representam um pequeno deslocamento, como 1mm em um objeto de 1cm. A Figura 70(b) apresenta a projeção destes objetos no sensor da câmera.

A Figura 70(c) representa um objeto pequeno (como 1cm) afastado da lente (como no caso de uma lente comum 18-55mm) e a Figura 70(e) representa um objeto maior (como 10cm) afastado de uma lente comum. As Figuras 70(d) e 70(f) representam suas respectivas projeções no sensor. É possível perceber pelas simulações das projeções que a diferença mais perceptível entre as projeções azul e vermelha ocorre na projeção 70(b).

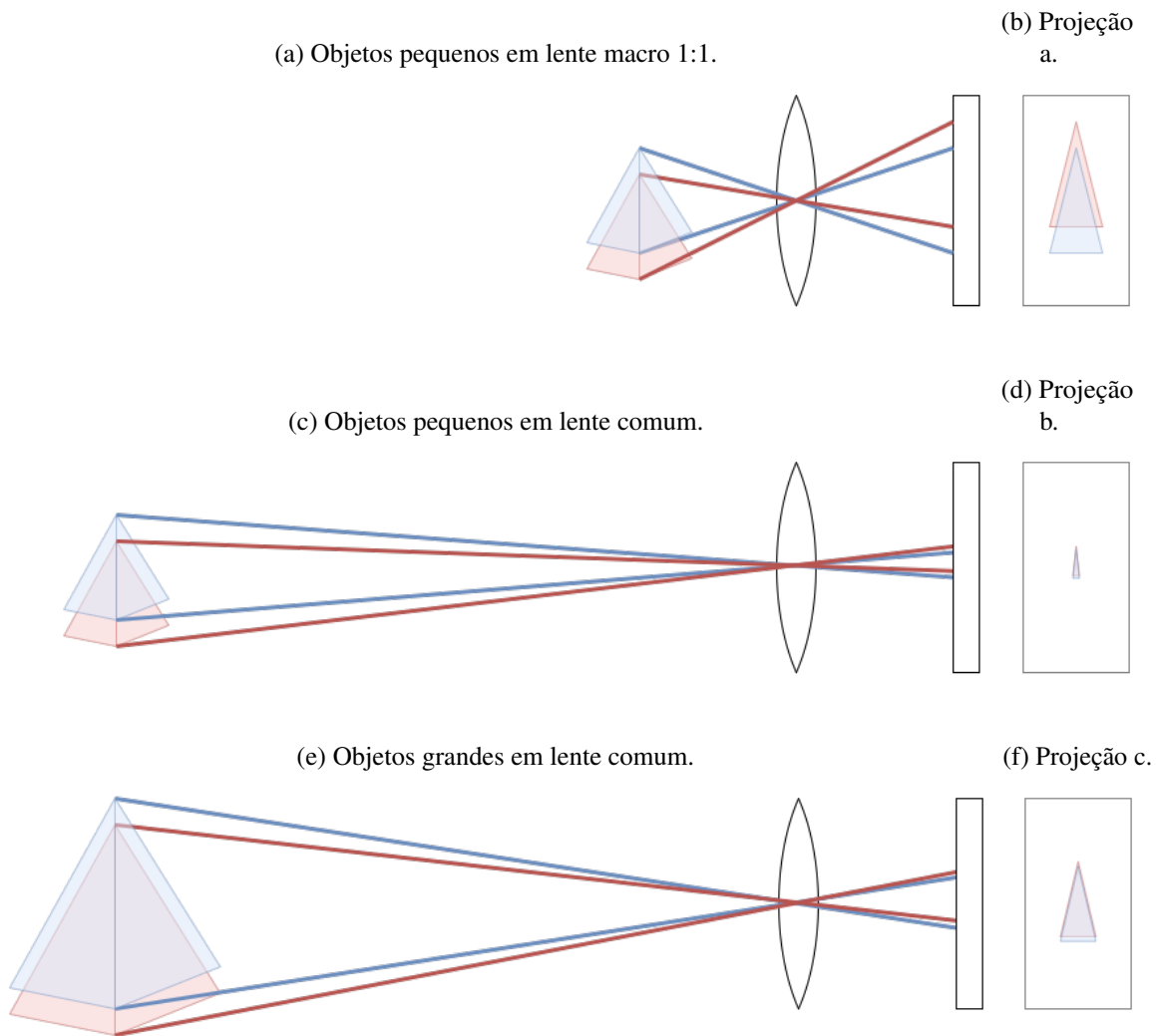
Isto ocorre pois nos outros casos a diferença de movimento é menos representativa na fotografia. No caso da Figura 70(d) a diferença de movimento não é tão representativa, pois o objeto é representado por menos pixels em comparação à Figura 70(b); como a representação do objeto é menor, a representação do deslocamento também é menor. No caso da Figura 70(f), como o objeto é maior, o deslocamento é proporcionalmente menor e conseqüentemente a percepção na projeção também é menor.

Base de suporte : A Base de suporte eleva o objeto que será escaneado para que a câmera não colida com o fundo infinito. Para objetos maiores esta base pode ser, por exemplo, um cubo 8,5cm de lado. Porém para objetos menores de 1cm este cubo de 8,5cm não resolveria o problema, pois como a câmera deve estar aproximadamente a 2 ou 3cm do objeto, ela colidiria com a base. Uma base cilíndrica de 8,5cm de altura e 2cm de raio poderia resolver o problema, mas é necessário garantir a estabilidade desta base para que a mesma não vibre e/ou tombe durante a aquisição.

Fixação : Todos os objetos devem ser posicionados e fixados para que não se movam durante a aquisição, como será discutido na Subseção 4.3.2. Porém para objetos muito pequenos, qualquer mecanismo de fixação pode ficar aparente na fotografia. Além disto estes objetos podem ser muito frágeis, como asas de uma borboleta, o que impede a fixação do objeto com cola, fita adesiva ou grampos.

Os insetos que foram escaneados já estavam alfinetados, como pode ser visto nas Figuras 68 e 69. O alfinete fica aparente na aquisição e normalmente não é possível segmentá-lo nas fotografias, pois ele causa oclusão de partes do inseto. Assim, ao segmentar o alfinete, o inseto aparentaria estar seccionado. Porém, se estes insetos não tivessem alfinetados e não fosse permitido alfinetá-los (para preservar o item), este processo poderia ser um desafio.

Figura 70 – Efeito da imprecisão do movimento em fotografias.



Fonte: Autoria própria.

Forma e Oclusão

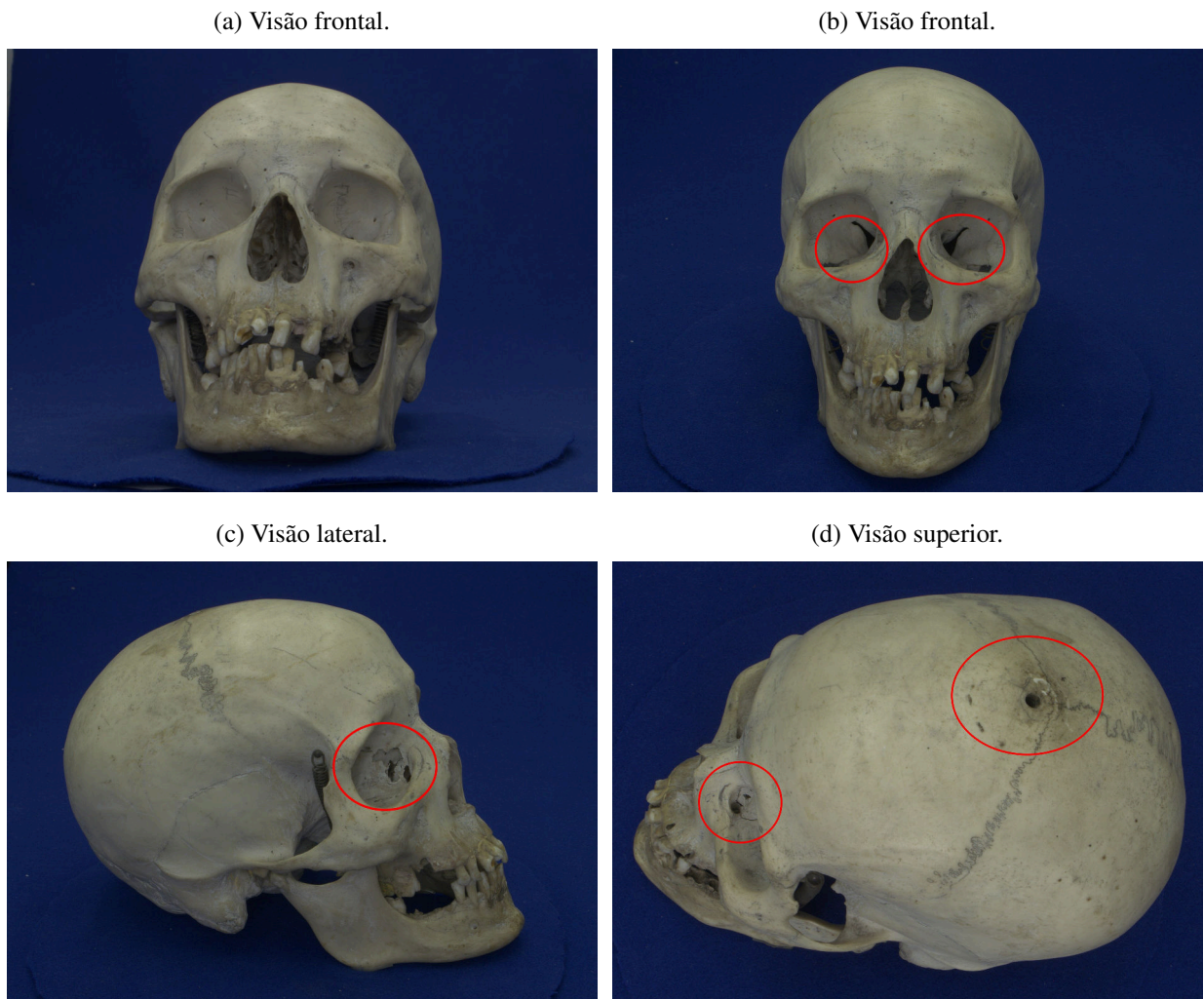
A forma dos objetos escaneados também influencia no processo de aquisição. Para debater isso os objetos foram divididos com base nas seguintes características: concavidade, convexidade e proporcionalidade das razões entre altura, largura e comprimento. Estas características influenciam na oclusão.

Objetos regulares, neste contexto significam: sem concavidade, convexidade e com proporcionalidade; necessitam de menos fotografias para representar toda a sua superfície. Considere por exemplo uma esfera, com apenas duas fotos, posicionadas em uma reta, mas em sentidos opostos e com a esfera entre as fotos, é possível adquirir toda a superfície da esfera.

Por outro lado, objetos irregulares necessitam de mais fotos para representar toda a sua superfície. Considere o crânio da Figura 71. Na visão frontal da Figura 71(a) não é possível identificar as fissuras ósseas da parte interna do globo ocular, visíveis na Figura 71(b); fissuras

internas do globo ocular 71(c); e o orifício no topo do crânio 71(d).

Figura 71 – Crânio humano.



Fonte: Autoria própria.

A Concha da Figura 72 apresenta concavidades, mas também apresenta convexidades. As protuberâncias da concha causam oclusão de superfícies, portanto também são necessárias mais fotos para representar toda sua superfície.

Cabe ressaltar que para objetos irregulares existem partes da superfície de não podem ser fotografadas pelo scanner, independente da quantidade de fotos. Por exemplo a superfície de partes internas da concha só pode ser fotografada se a câmera for inserida dentro da concha. Nenhum ângulo consegue visualizar estas estruturas do plano da esfera formada durante a aquisição.

A desproporcionalidade entre as razões da altura, largura e profundidade de um objeto podem interferir na profundidade de campo durante a aquisição. Assim, podem exigir mais fotografias para representar integralmente o objeto que está sendo digitalizado.

Figura 72 – Concha.



Fonte: Autorial própria.

A quantidade de fotos necessárias para um determinado objeto depende então da complexidade de sua forma. Quanto mais concavidades, convexidades e desproporcionalidades um objeto apresentar, mais fotografias serão necessárias. Além disto, depende também do objetivo da visualização, alguns usos aceitam a oclusão de partes do objeto.

Foram escaneados objetos com concavidades, convexidades e desproporcionalidade, como plantas e a concha da Figura 72, com uma variação angular $\Delta\Theta = 3^\circ$ e $\Delta\gamma = 3^\circ$, pois estes intervalos são adequados para a visualização estereoscópica. Nestas aquisições não foram percebidos pontos do objeto que ficassem sem representação nas imagens e fosse possível vê-los a olho nú com o objeto em mãos.

Isto considerando apenas pontos de vista que fazem parte da semiesfera formada pela aquisição, a semiesfera inferior necessita de outra aquisição. Assim, empiricamente, estes intervalos angulares de 3° aparentam ser suficiente para representar todo o objeto. Porém, ressalto que podem existir objetos e aplicações que necessitem de uma aquisição mais detalhada.

Superfície

Os objetos escaneados podem possuir diversos tipos de superfícies com características diversas que podem interferir no processo de aquisição, processamento e visualização. Inicialmente as superfícies dos objetos foram analisadas quanto a suas características de textura, por exemplo: se eram lisas, polidas, felpudas, rugosas e porosas.

Porém, percebeu-se que estas características poderiam ser resumidas segundo o compor-

tamento da superfície perante a luz. Então definiu-se quatro grupos para os objetos:

Opaco : objetos que ao refletirem a luz não apresentam brilho nem reflexo. São objetos que ao serem iluminados não permitem identificar a posição da fonte luminosa e tentem a suavizar sombras. Estes objetos tendem a ser ásperos, rugosos e porosos. Exemplos destes materiais são: papelão e cortiça.

Reflexivo : ao refletirem a luz é possível identificar brilho, formas e a posição da fonte de luz pode ser identificada. Sombras e diferenças luminosas tendem a ser mais acentuadas do que em objetos opacos. Estes objetos tendem a ser lisos e polidos. Como exemplos podem ser citados aço escovado e plástico polido.

Espelhado : a luz refletida tente a criar imagens e formas nítidas. Além de ser possível identificar sombras, também é possível visualizar o entorno do objeto em seu reflexo. Tendem a ser superfícies polidas. Exemplo: espelho, aço polido e cromados.

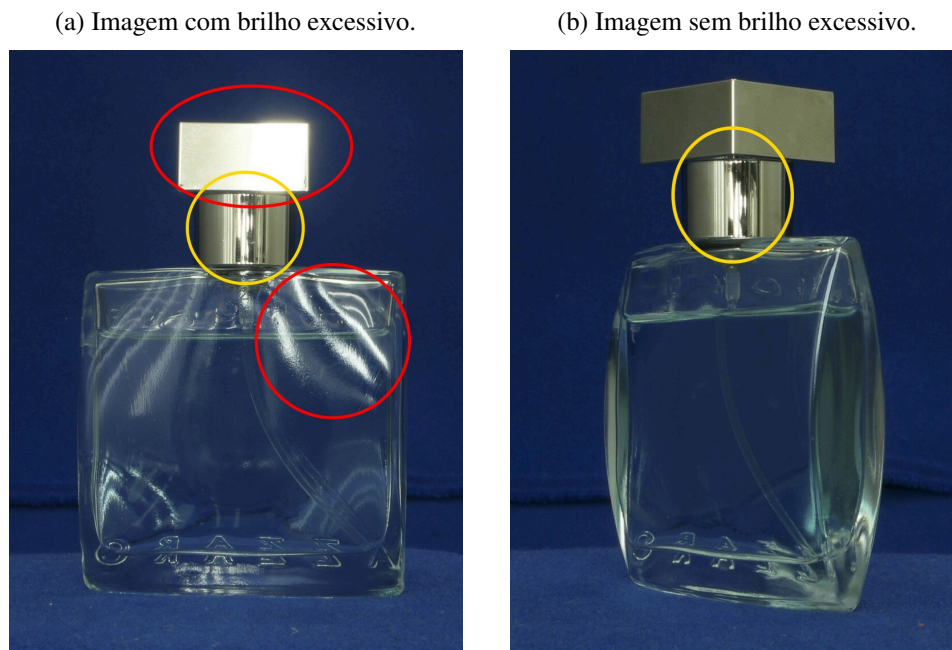
Translúcido : objetos que permitem a passagem de luz, permitindo ver através do objeto. Estes objetos podem ter cor, como vidros coloridos e tecidos transparentes, ou não, como vidros e acrílicos incolores. Podem apresentar brilho, como no caso do vidro, ou não.

O maior desafio é referente à objetos, ou partes do objeto, que sejam reflexivos, espelhados ou translúcidos. Na fotografia de produtos o brilho, reflexos e transparências são manipulados usando a iluminação e a posição da câmera. Porém no processo de aquisição do F2S2 isto não é possível. Como a iluminação deve ser a mesma para todas as fotografias, ao corrigir o brilho, reflexo ou transparência para determinada posição, ao mover o objeto ou a câmera este ajuste é perdido.

A Figura 73 apresenta, marcado em vermelho, a ocorrência de brilho excessivo em uma determinada posição do objeto (Figura 73(a)) e a ausência de brilho (Figura 73(b)) em outro Frame do mesmo Stream. Em amarelo é apresentado reflexo de partes do scanner. Considerando a quantidade de fotografias que o F2S2 obtém em uma aquisição, não foi possível prever todos os casos de brilho e reflexo que poderiam acontecer.

Em outras palavras, dependendo do objeto, não será possível eliminar brilhos, reflexos ou segmentar corretamente transparências para todas as fotografias. Considere por exemplo uma aquisição de esfera de vidro transparente ou uma esfera de aço cromada. O vidro invariavelmente vai permitir que sejam percebidos nas fotografias reflexos de partes do scanner que estão atrás do objeto. O cromado vai refletir a câmera e partes do scanner que estão na frente do objeto. O mesmo vai acontecer com a cor escolhida para o fundo, de forma que em ambos os casos as peças tendem a assumir a cor do fundo do scanner, pois esta será a cor da luz mais presente dentro do scanner.

Figura 73 – Ocorrência de brilho em determinados Frames.



Fonte: Autoria própria.

Na Subseção 4.3.2 serão discutidas as configurações ideais de iluminação para estes grupos de objetos. Além destas características, objetos felpudos podem ser um desafio na etapa de processamento e segmentação. Isto será discutido na Subseção 5.2.1.

Rigidez e estabilidade da forma

Os objetos que podem ser escaneados precisam ser rígidos o suficiente para que mantenham sua forma para a aquisição. Alguns problemas referentes a isso foram identificados principalmente com peças anatômicas. Órgãos como intestino precisam passar por tratamentos capazes de atribuir ao tecido a rigidez necessária para o órgão possa ser fotografado. A técnica mais usada neste caso é a glicerinação, porém esta técnica altera as cores reais da peça. Caso a peça seja escaneada sem este tratamento, simplesmente colocada sobre uma superfície, a aquisição não agregará muitas vantagens se comparada a uma fotografia comum, pois não será possível visualizar a peça sobre diversos ângulos.

Outros objetos podem mudar de forma durante a aquisição, por exemplo folhas ou flores retiradas de uma planta podem não suportar uma aquisição de 12 horas, fazendo com que ao visualizar os últimos Streams, possa ser percebida a deterioração da peça. O mesmo problema acontece com objetos congelados, que descongelariam durante o processo de aquisição.

Alguns objetos possuem partes móveis e pendulares, como uma vara de pesca que possui uma chumbada e anzol móveis. Estes objetos necessitam de mais tempo entre cada fotografia, para que após o objeto girar sobre o eixo B^a , suas partes móveis tenham tempo de pararem

qualquer movimento que tenham realizado por influência do movimento do eixo. Além disso, cabelos de uma boneca podem ser desarrumados durante o giro da peça e isto será percebido na visualização.

Também não é possível escanear objetos que soltem líquidos, como órgãos recém retirados de organismos vivos, pois os fluídos poderiam estragar o scanner e a deterioração do material seria perceptível. Peças imersas em meio líquido, como peças anatômica em formol, também não são compatíveis com o scanner, pois ao girar o recipiente, a peça dentro do líquido não sofreria o mesmo movimento.

Peso

O F2S2 possui estabilidade no movimento com peças de até 3Kg, somando-se a base (se necessário) e a peça. Itens mais pesados forçam o motor e fazem com que o movimento, ou não seja executado, ou ocorre de forma abrupta.

Inércia/Vibração

Objetos com pouca rigidez e/ou uma distribuição de peso muito irregular sofrem com a inércia para pararem seu movimento de rotação. Algumas peças após sofrerem o movimento do eixo B^a , permanecem vibrando por causa da inércia. Para resolver este problema é necessário realizar um movimento mais lento e/ou aguardar a vibração cessar antes de adquirir cada fotografia.

Equilíbrio estático

Outro ponto importante é o equilíbrio estático da peça. Neste caso é importante que o objeto consiga realizar a rotação forçada pelo eixo B^a , sem tombar ou se mover. Para algumas peças isso pode ser conseguido com fita adesiva, espetos ou grampos. Novamente peças muito delicadas podem exigir mais atenção, pois podem ser danificadas por estes sistemas de fixação.

4.3.2 Configuração do scanner F2S2 e aquisição

Após analisar o objeto que será escaneado, esta subseção apresenta o preparo e configuração do F2S2 para a aquisição. Primeiramente escolhe-se a cobertura do fundo infinito, então é realizado o posicionamento e fixação do objeto. Em seguida são definidas as configurações da aquisição, fotometria, iluminação e configuração da câmera.

Após estas etapas, definiu-se realização de uma aquisição de teste, para verificar se os parâmetros definidos estão corretos. Apresenta-se então o uso do ColorChecker, para realizar uma digitalização com acurácia na representação das cores e por fim, apresenta-se a aquisição.

Escolha do material e da cor do fundo infinito

O fundo infinito pode ser recoberto com vários materiais e cores. No laboratório já foram testados: tinta, película reflexiva ⁵, tecido, plástico, papel e chapa de madeira com fórmica. As cores já testadas foram: branco, preto, cinza, verde, azul e vermelho. Neste trabalho as imagens onde o fundo aparece na cor branca, foram fotografadas com fundo de papel; nas imagens com fundo verde o material usado foi a tinta; as imagens com fundo azul, cinza e preto eram com fundo em tecido.

A Película reflexiva é composta por esferas de vidro oco e refletem a luz incidente com grande eficiência. Inicialmente pensou-se que este material ofereceria uma iluminação do sentido do fundo para o objeto, permitindo que mais detalhes fossem vistos e obtendo uma iluminação mais homogênea. Porém, o material acaba funcionando como um prisma, decompondo a luz em várias cores, dificultando a segmentação; causa um brilho intenso que pode ofuscar detalhes dos objetos; a homogeneidade da iluminação impede a formação de sombras, dificultando a sensação volumétrica do objeto. Assim, o material não foi mais utilizado.

O uso de coberturas plásticas não foi adequada, pois o plástico causa muito brilho nas fotografias. As coberturas com tinta foram adequadas, porém não foram usadas pela dificuldade de trocas de cores. O mesmo problema foi enfrentado com as lâminas de madeira com fórmica. Considerando as dificuldades apresentadas por outros materiais, o tecido e o papel acabaram sendo os materiais mais usados. O papel apresenta problemas de amassar e sujar mais facilmente e de forma mais aparente do que o tecido, no entanto isso não inviabiliza seu uso.

No que se refere à escolha da cor, deve-se pensar inicialmente no processo de segmentação. Para imagens que não serão segmentadas para extrair o fundo, sugere-se cores neutras como branco, cinza ou preto, pois estas cores diminuem a percepção de reflexos nas imagens. Para objetos que serão segmentados a escolha do fundo deve se basear na diferença de cor do objeto, de forma que o objeto não possua a mesma cor do fundo. Por exemplo: caso o objeto seja amarelo, pode-se usar o azul ou verde, pois estas cores não estão presentes no objeto.

Também deve ser considerada a reflexividade e transparência dos objetos. A Figura 74 apresentam reflexos da cor do fundo no objeto. Alguns materiais muito reflexivos, espelhados ou translúcidos tendem a assumir a cor do fundo. A Figura 74(a) apresenta uma peça com base de madeira envernizada. Ao ampliar a fotografia 74(b) o reflexo azul do fundo fica evidente. Assim, para estes objetos cores neutras, como branco, cinza e preto, são mais adequadas.

Além disto, quanto mais reflexivo for o objeto, mais interferência da cor do fundo o objeto sofrerá. Considere uma esfera cromada ou um espelho, qualquer que seja a cor do fundo, esta cor será refletida pelos objetos. Assim, nem sempre é possível eliminar este efeito. Nestes casos devem ser escolhidas cores neutras, que minimizem a alteração da cor do objeto, como

⁵ <https://www.3m.com.br/3M/pt_BR/3m-do-brasil/todos-os-produtos-3m-do-brasil/?N=5002385+3290422329&preselect=8710677+8710896+3293786499&rt=rud> Acessado em 30 de Março de 2020.

Figura 74 – Reflexo da cor do fundo.

(a) Imagem completa.



(b) Imagem ampliada - Reflexo do fundo no verniz.



Fonte: Autoria própria.

branco, cinza ou preto. Porém nem sempre estas cores podem ser usadas, considere um objeto que possui as cores branco, cinza e preto na sua composição. Nestes casos, para viabilizar a segmentação, pode ser necessário usar uma cor contrastante apesar do reflexo.

Pode-se concluir então que a cor ideal para o fundo infinito é o branco, uma vez que esta não absorve muitos comprimentos de luz e portanto não alteraria a cor do objeto. Porém o fundo infinito de cor branca é o mais sensível à sombra causada pelo objeto no fundo. Este efeito pode ser percebido na Figura 75. Esta sombra pode atrapalhar significativamente o processo de segmentação. Considere que a sombra projetada no fundo possui cores similares à do objeto que está sendo fotografado.

Portanto, a escolha do material e cor do fundo infinito depende de cada objeto e suas características. Objetos menos reflexivos e menos transparentes permitem mais liberdade da escolha do fundo. Objetos que não causam muita sombra podem usar fundos mais claros. O sistema de Iluminação também pode ser configurado de forma a minimizar estes efeitos.

Posicionamento do Objeto

Inicialmente demoramos para perceber que o posicionamento do objeto alinhado ao eixo B^a poderia causar diversos problemas. Dificuldades fotográficas como falta de foco (causada pela diferença de distância do objeto à câmera) e fotografias tremidas (causadas pela maior quantidade de movimento, uma vez que o objeto sofria revolução em torno do eixo e não rotação sobre seu eixo) eram atribuídas à configuração da câmera, porém eram causadas pela descentralização do objeto.

Além disto, a falta de alinhamento atrapalha no processo de visualização, pois ao navegar pelas imagens o objeto se move na tela. Este processo é ainda mais crítico para o processo de

Figura 75 – Objeto causando sombra no fundo infinito.



Fonte: Autorial própria.

criação de estereoscopia, pois faz com que o intervalo angular entre cada foto do par estéreo não seja respeitado.

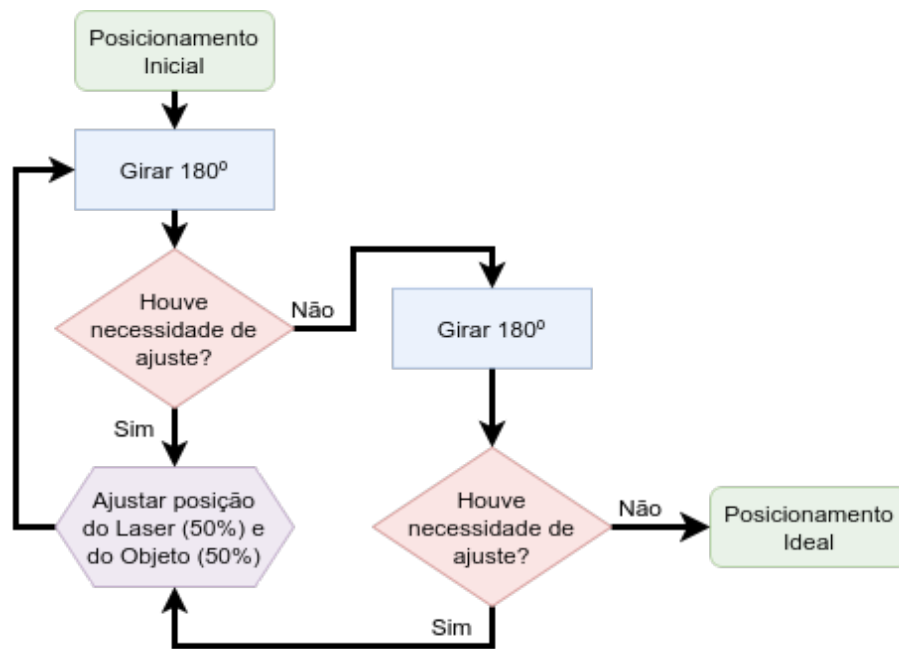
Para resolver o problema pensou-se em marcar com um laser o centro do eixo B^a e construir suportes de madeira para alinhar o objeto. Tentou-se prestar mais atenção ao processo de alinhamento de forma manual e também observar pela câmera se o objeto estava centralizado. Porém nenhuma destas técnicas oferecia um resultado satisfatório. Então foram adquiridos três níveis a laser e com estes níveis foi desenvolvida uma técnica capaz de alinhar o centro geométrico do objeto ao centro do eixo B^a .

Alinhamento a laser

Para alinhar o objeto ao eixo B^a é necessário considerar que o objeto deve ser alinhado pelo seu centro geométrico e não centro de massa ou gravitacional. Para objetos em que o centro de gravidade está distante do centro geométrico, a probabilidade de tombamento durante a aquisição aumenta. A identificação do centro geométrico pode ser um processo simples em objetos regulares, como esferas ou cubos, porém pode ser complicado em objetos irregulares, como a concha da Figura 72, ou muito pequenos, como a abelha da Figura 68b. Cabe notar que para objetos que necessitem de base de suporte, a base não precisa ser alinhada, apenas o objeto.

O processo de alinhamento a laser pode ser definido pelo fluxograma da Figura 76. Inicialmente posiciona-se o objeto sobre o prato giratório do eixo B^a , preferencialmente o mais centralizado possível. Posiciona-se dois lasers paralelos ao sentido da câmera, de forma que ambos toquem a parte mais externa do objeto, delimitando seu comprimento. Um terceiro laser é posicionado perpendicularmente aos dois primeiros, de forma que demarque o ponto mais afastado da câmera. Este processo pode ser visto na Figura 77, na qual as marcações em vermelho apontam a região em que o laser atinge o objeto.

Figura 76 – Fluxograma do processo de alinhamento.



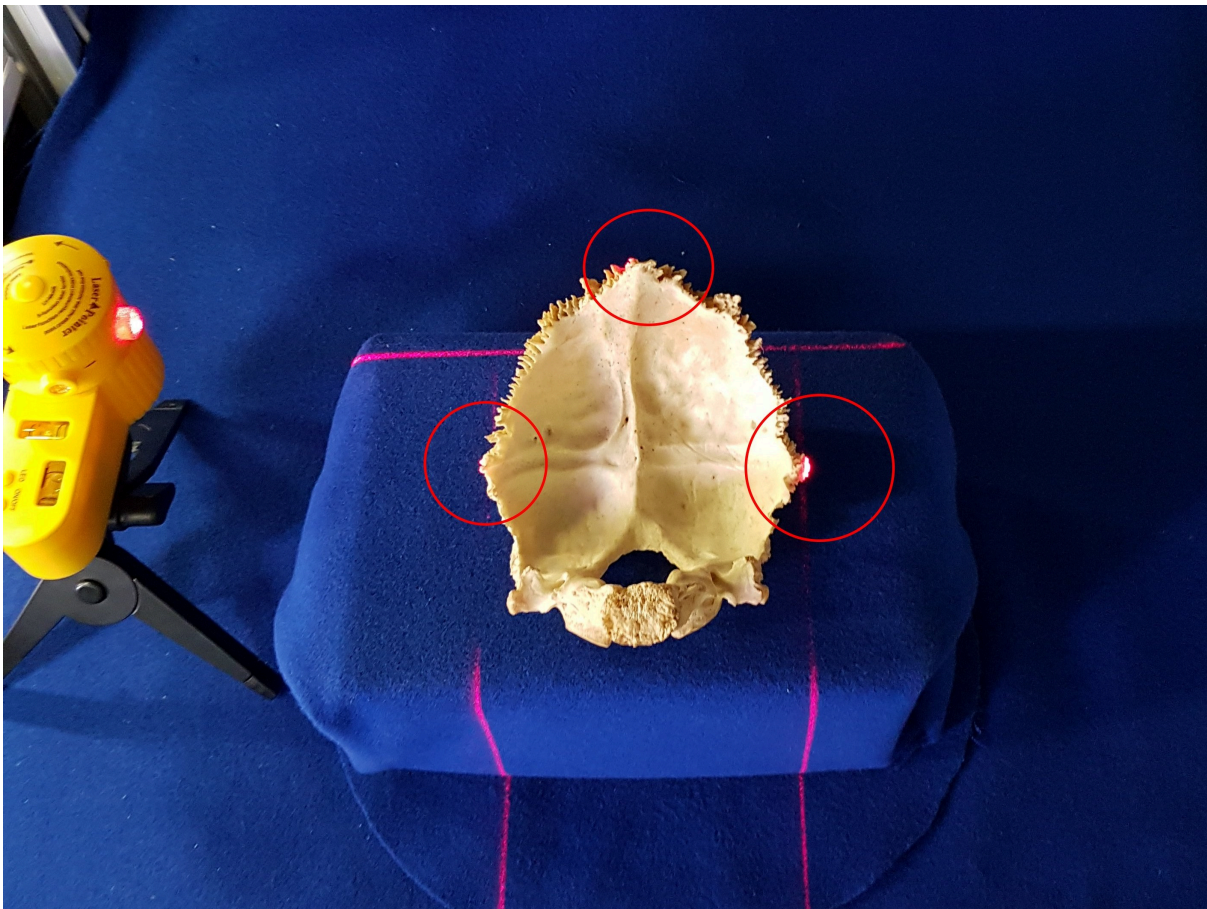
Fonte: Autoria própria.

Após o posicionamento inicial (Figura 78(a)), o eixo B^a gira 180° (Figura 78(b)). Verifica-se então se o objeto ultrapassou os limites definidos pela laser. Caso isto tenha ocorrido, ajusta-se a posição dos lasers e do objeto de forma que metade do erro de posicionamento seja corrigido alterando a posição do objeto e metade a posição do laser.

Por exemplo: considere que após o giro de 90° o objeto tenha ultrapassado em 4cm para a direita a posição marcada pelo laser. Deve-se então reposicionar o objeto 2cm para a esquerda e o laser 2cm para a direita, de forma que o laser redefina a posição do objeto. Este processo pode ser visto nas Figuras 78(c) e 78(f). O procedimento é realizado para os três lasers a cada movimento.

A técnica corrige parte do erro a cada novo reposicionamento do objeto e do laser. Não adianta mover apenas o objeto ou apenas o laser, pois assim um dos dois continuará com a posição incorreta. Ao reposicionar ambos simultaneamente, a cada novo ajuste a quantidade de erro tende a diminuir, de forma que o objeto é reposicionado cada vez mais alinhado ao centro do eixo B^a .

Figura 77 – Alinhamento com laser de peça anatômica.



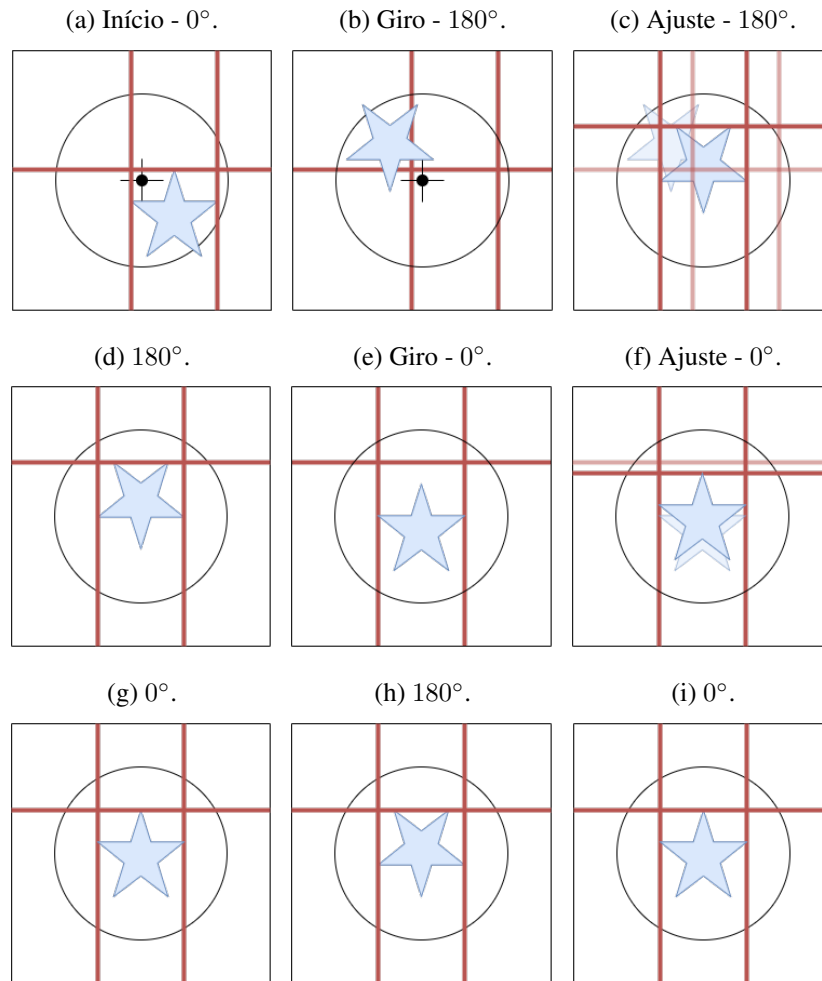
Fonte: Autoria própria.

Esta sequência de passos é repetida até que ao rotacionar o objeto, este não ultrapasse nenhum dos três níveis laser posicionados; e que ao rotacionar o objeto este sempre seja atingido pelo laser em seu limite mais externo (Figuras 78(g) 78(h) 78(i)). Quando estas condições forem verdadeiras para uma rotação de 0° e 180° , tanto para o laser paralelo à câmera, quanto para o perpendicular, significa que o objeto está com seu centro geométrico alinhado ao centro do eixo B^a .

A técnica pode ser aplicada usando dois, três ou quatro lasers. No caso de dois níveis, um deve estar paralelo à câmera e o outro perpendicular. No caso de três unidades recomenda-se usar dois paralelos à câmera e um perpendicular. No caso de 4 unidades, um par deve estar paralelo e outro perpendicular.

Cabe notar que este processo pode ser realizado sem a fixação do objeto, pois este será movido diversas vezes durante o processo. Após os lasers demarcarem corretamente a posição em que o objeto está centralizado, o objeto é retirado para que possa ser definitivamente fixado na posição demarcada pelos lasers.

Figura 78 – Processo de alinhamento a laser.



Fonte: Autoria própria.

Fixação do objeto

Assim como o processo de alinhamento do objeto, o processo de fixação deste também aparenta ser simples. Porém pode ter grandes implicações na aquisição e visualização. Uma fixação inadequada pode inclusive danificar a peça que está sendo escaneada, seja por tombamento, que pode quebrar a peça, ou pelo uso inadequado de ferramentas de fixação.

Nas primeiras aquisições realizadas, quando o processo de alinhamento a laser também não era realizado, percebeu-se que ao longo da aquisição algumas peças se movimentavam sobre o disco do eixo B^a . Isto acontecia pois ao girar o eixo B^a , a força centrífuga poderia vencer o atrito da peça com o prato ou base de suporte e mover a peça para longe do centro do eixo B^a . Esta força pode até derrubar o objeto.

Ao perceber este efeito, o objeto passou a ser colado com fita adesiva dupla face ou cola bastão, dificultando a movimentação do objeto. Porém esta técnica não resolve todos os problemas. Assim foram definidas algumas ferramentas que podem ser usadas e apresentadas as

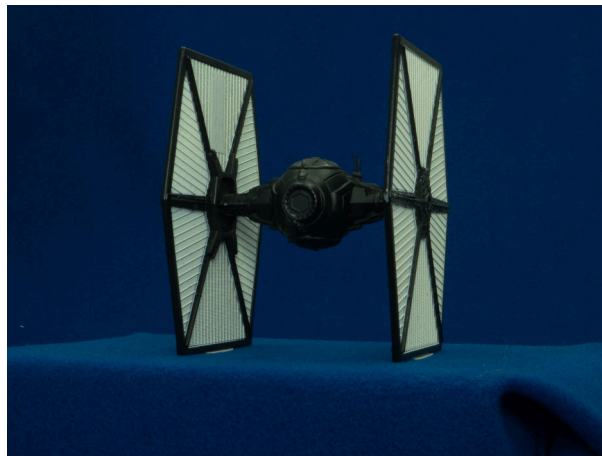
dificuldades de uso de cada uma:

Fixação do objeto com fita adesiva dupla face

A maioria dos objetos pode ser fixada com fita adesiva dupla face, porém esta técnica pode não ser adequada em alguns casos. Objetos frágeis podem ser danificados pela fita. Para objetos com altura muito maior que o comprimento e profundidade, a fita adesiva pode não ser suficientemente forte para estabilizar o objeto. Por exemplo: um lápis não ficará em pé usando apenas a fita adesiva. Para objetos com uma área muito pequena de contato com o prato giratório, a fita pode ficar aparente, como apresentado na Figura 79.

Figura 79 – Fita adesiva aparente na fotografia.

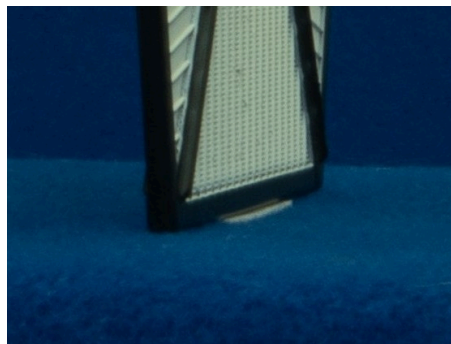
(a) Imagem completa.



(b) Imagem ampliada - esquerda.



(c) Imagem ampliada - direita.



Fonte: Autoria própria.

Fixação do objeto com cola bastão

Para aquisições com o fundo infinito recoberto com papel, pode ser usada cola bastão para auxiliar na fixação do objeto. Porém é necessário expor que a cola pode deixar resíduos no objeto. A cola também não fixa fortemente o objeto, apenas aumenta um pouco o coeficiente de atrito entre a peça e o papel.

Fixação do objeto com agulha (insetos) ou espetos

Para algumas peças a melhor forma de fixação são agulhas ou espetos. Estas ferramentas danificam a peça, porém, dependendo da peça isso pode ser aceitável. O maior problema desta técnica é que os alfinetes podem ficar aparentes e dependendo de sua posição pode ser impossível segmentá-los posteriormente de forma que estes ficaram aparentes na visualização.

Os insetos que foram digitalizados eram armazenados fixados por alfinetes, então foram usadas as mesmas alfinetes na fixação destes no scanner. Pode-se notar que neste caso os alfinetes ficam aparentes, causam oclusão nas peças e por este motivo não é possível segmentá-los. Porém como a instrução dos entomologistas era de não remover os alfinetes os insetos foram digitalizados assim.

A Figura 80 apresenta uma pata de cavalo. Neste caso, devido ao peso e altura da peça, para manter o item em pé, foi necessário parafusá-lo em uma base recoberta com o mesmo tecido usado no fundo infinito do scanner. Desta forma nem os parafusos, nem a base ficam aparentes, porém a peça foi danificada.

Figura 80 – Pata de cavalo.



Fonte: Autoria própria.

Fixação do objeto por garras ou grampos

Por último pode ser necessário usar garras ou grampos para fixar o objeto. O grande problema desta abordagem é a oclusão que a garra faz no objeto, de forma que o grampo ficaria aparente na visualização.

Base de Suporte

Como já mencionado, para objetos menores do que $16cm$ é necessário o uso de uma base de suporte. Neste caso a base precisa ser fixada sobre o prato giratório e a peça sobre o prato. A base deve ser recoberta pelo mesmo material usado no fundo infinito do scanner, permitindo assim a segmentação da base. O tamanho mínimo da base deve ser de: $\frac{\text{tamanho do objeto}}{2} \geq 85mm$ para que o centro do objeto esteja a no mínimo 85mm de altura.

Seleção dos parâmetros de configuração da aquisição

Após fixar o objeto é necessário estabelecer os parâmetros de configuração do scanner. Estes dados são as configurações para o software que controla o F2S2, o processamento destas informações será apresentado na Seção 5.1. Neste momento o usuário do scanner precisa apenas configurar o software com informações relativas à aquisição.

Número e distribuição das fotos

A quantidade e a distribuição espacial das fotos pode ser configurada de acordo com o objetivo da aquisição. Podem ser configurados: ângulo vertical inicial (γ^0); ângulo vertical final (γ^P); variação angular vertical ($\Delta\gamma$); variação angular horizontal ($\Delta\Theta$). Com estas variáveis é possível desde adquirir apenas uma foto em uma determinada posição, até 101 Streams com 400 fotos cada, oferecendo uma variação angular de 0.9° , totalizando 40.400 fotos do objeto.

Assim, a quantidade e distribuição das fotos depende do objetivo da aquisição. Para uma aquisição de um produto para publicidade na internet, 40.400 fotos pode ser um excesso, principalmente por que o usuário precisaria aguardar o carregamento destas fotos. Neste caso, 4 Streams com 10 Frames cada ($\Delta\Theta = 30^\circ$ e $\Delta\gamma = 36^\circ$) poderia ser o suficiente para apresentar o produto. Quarenta imagens *jpg* de 1000x1000 pixels ocupam algo em torno de $5Mb$, viabilizando esta quantidade de imagens para uso em publicidade na internet.

Para atividades de ensino poderiam ser usadas mais imagens e em um tamanho maior. Estudantes do ensino médio poderiam estudar arte ou anatomia com algo entre 300 fotos, sendo 10 Streams com 30 Frames cada ($\Delta\Theta = 12^\circ$ e $\Delta\gamma = 10^\circ$). Considerando imagens *jpg* $2K$ que possuem algo em torno de $2Mb$ cada, esta aquisição ocuparia algo em torno de $6000Mb$.

Outras atividades, como pesquisa científica podem precisar de mais imagens e de tamanhos maiores. Uma aquisição de 900 imagens ($\Delta\Theta = 6^\circ$ e $\Delta\gamma = 6^\circ$) em $6K$ ocupa aproximadamente $8Gb$ de espaço em disco. Este é o valor mínimo para se obter uma visualização

estereoscópica, uma vez que são necessário 6° entre as fotos. Porém no laboratório, para visualização estereoscópica, costuma-se usar $\Delta\Theta = 3^\circ$ e $\Delta\gamma = 3^\circ$, totalizando 31 Streams e 120 Frames (3720 Fotos) que no formato *RAW* ocupam aproximadamente 100Gb.

Raio da esfera

O Raio da esfera formada depende do tamanho do objeto e do conjunto óptico da câmera. É necessário levar em consideração a profundidade de campo, que diminui ao aproximar a câmera do objeto. Então, deve-se usar sempre o maior raio que permita um bom enquadramento do objeto nas fotografias. Para objetos pequenos é necessário aproximar a câmera para que seja possível enquadrar adequadamente o objeto.

No caso do scanner, um bom enquadramento significa que ao longo do processo fotográfico nenhuma parte do objeto seja cortada e que nas fotografias sejam usadas a maior quantidade de pixels para representar o objeto e não o fundo infinito. Para isso recomenda-se fazer uma aquisição de teste com poucos Frames, viabilizando uma análise do resultado do enquadramento.

Altura da esfera

A altura da esfera com relação ao fundo infinito possui por objetivo centralizar a esfera formada com o centro geométrico do objeto. Este processo possui duas possibilidades: a primeira é para objetos maiores de 16cm , pois não é necessário base de suporte para elevar o objeto. Neste caso a altura da esfera será a metade da altura do objeto. No caso de objetos menores, deve ser usada a base de suporte para elevar o centro do objeto para uma altura mínima de 8cm .

Tempo entre fotos

Como já apresentado, após mover o objeto e/ou a câmera é necessário aguardar um tempo para que estes parem de se mover e vibrar. Este valor pode ser configurado para a aquisição e depende principalmente da estabilidade do objeto. Os valores mais usados no laboratório são 3 segundos entre Frames, ou seja após mover o objeto; e 4 segundos entre streams, ou seja após mover a câmera.

Fotometria

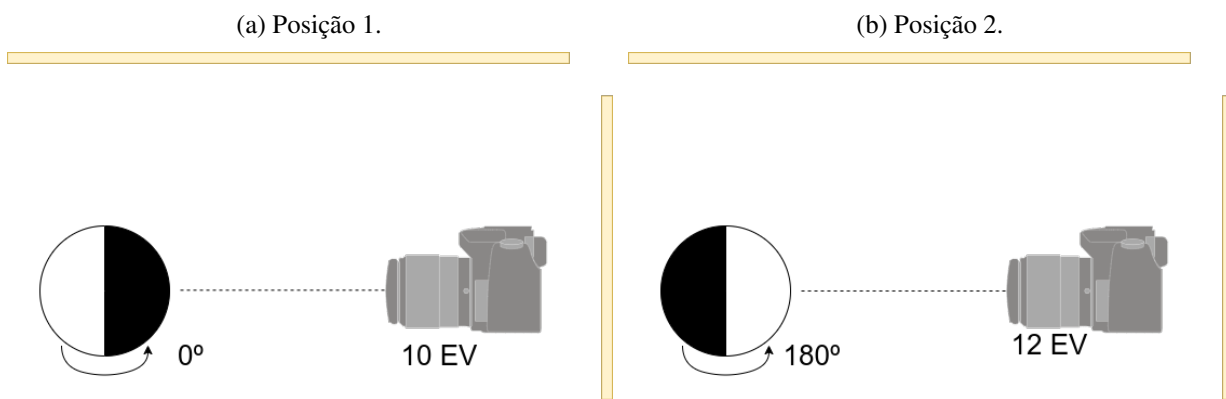
O processo de controle da iluminação inicia-se com o uso do fotômetro de luz incidente. Como o scanner adquire um conjunto de fotografias de diversas posições, usa-se o fotômetro para estabelecer a quantidade de luz necessária para cada objeto e principalmente a relação entre a luz em diversas posições durante o processo de aquisição.

Em fotografias cotidianas o fotômetro interno da câmera é usado para estabelecer os parâmetros de exposição para determinada foto (a abertura do diafragma, a sensibilidade ISO e a velocidade do obturador). Este processo pode fazer uma medição pontual ou de regiões de uma

cena. Porém este processo não é o suficiente para o scanner, uma vez que outras fotografias de outros ângulos serão tiradas.

Usar apenas o fotômetro interno, que funciona com luz refletida, não contempla todo o conjunto de fotografias de uma aquisição. Fotômetros de luz refletida são muito sensíveis a mudanças na cena que está sendo fotografada. Considere por exemplo um objeto com um lado preto e opaco e outro lado branco e reflexivo. Ao fazer a medição desta cena em um determinado ângulo obtém-se um determinado valor da quantidade de luz refletida, ao mover a câmera ou o objeto, obtém-se outro valor. Este processo pode ser visto na Figura 81.

Figura 81 – Processo de fotometria com luz refletida.



Fonte: Autoria própria.

Assim, como a fotometria por luz refletida é sensível ao objeto que está sendo fotografado e pode ser muito sensível a variações do índice de reflexão da luz de determinadas partes do objeto, além de não contemplar outros ângulos que o scanner usaria para obter imagens, foi necessário desenvolver uma técnica de medição da luz que contemplasse de forma adequada o processo de aquisição fotográfico do F2S2.

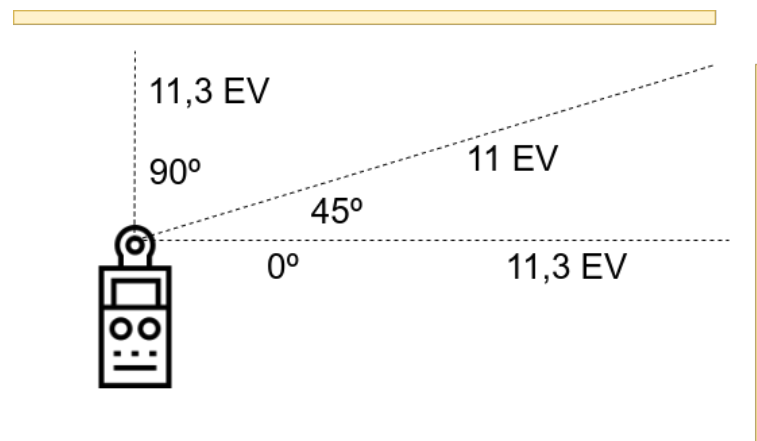
A técnica de fotometria desenvolvida para o F2S2 considera que o scanner adquire fotografias de vários ângulos do objeto. Desta forma o fotômetro deve ser usada para garantir que as condições luminosas sejam satisfatórias e equânimes para todos estes ângulos. Para contemplar esta característica do processo de aquisição foi usado um fotômetro de luz incidente, porém apenas esta alteração de equipamento não resolve todo o problema.

Fotômetros de luz incidente medem a intensidade luminosa que alcança determinado ponto no espaço, oriundos de uma determinada direção. Este tipo de medição não é sensível ao objeto que será fotografado, uma vez que não mede a luz refletida e sim a incidente. Trata-se de uma medição mais efetiva do que a de luz refletida, uma vez que a medição de luz refletida tende a normalizar todos os objetos para uma intensidade luminosa referente ao cinza médio (cinza 18%), enquanto que a fotometria por luz incidente tende a manter o preto e o branco em mais conformidade com realidade.

Assim, para que ao visualizar o Stream2D adquirido não sejam percebidas variações na iluminação, o processo de mensuração da luz para o F2S2 consiste em verificar se a intensidade luminosa que alcança a região em que ficará o objeto é constante, de forma que ao longo de todo o processo de aquisição a iluminação esteja dentro de uma variação considerada aceitável. No contexto do Scanner, considerando a possibilidade de ajuste da iluminação, estabeleceu-se que $\frac{1}{3}EV$ é o limite aceitável de variação de iluminação.

Então, a técnica desenvolvida para ajuste da iluminação baseia-se em medir a luz incidente com o fotômetro posicionado no lugar onde seria posicionado o objeto e apontado para o percurso que a câmera faria. Este processo pode ser visto na Figura 82. Inicia-se medindo a luz na horizontal e gira-se o fotômetro até a vertical, fazendo medições intermediárias. Caso em algum ponto deste percurso a iluminação possua uma variação maior que $\frac{1}{3}EV$, a iluminação do F2S2 deve ser reconfigurada para equalizar este valor.

Figura 82 – Processo de fotometria para o F2S2.



Fonte: Autoria própria.

Cabe ressaltar que os equipamentos fotográficos normalmente medem a intensidade com precisão de $\frac{1}{3}EV$ ou $\frac{1}{4}EV$. Portanto a variação máxima de $\frac{1}{3}EV$ é a menor variação medida pelos equipamentos. Empiricamente, considerando que esta variação será distribuída entre várias fotos, durante a visualização esta diferença luminosa é imperceptível.

$$\forall \Theta / Lux(\Theta) = Lux(\Theta') \pm \frac{1}{3}EV$$

Além disto, fotômetros servem como um referencial para ajustar a câmera e não como um valor exato que deve ser seguido obrigatoriamente. Inclusive é recomendado ajustar os parâmetros de exposição da câmera para absorver mais ou menos luz dependendo do objeto que será fotografado.

Iluminação

Diferentemente da fotografia comum, que prepara a iluminação para uma única posição da câmera em relação ao objeto, no caso do F2S2 a iluminação deve ser configurada para todas as

fotografias da aquisição. Além disso, na fotografia, a luz é responsável por criar sombras, as quais oferecem parte da sensação de profundidade. Considerando isso, não é possível simplesmente adicionar a maior quantidade de luz para todos os casos.

A computação gráfica usa parte da capacidade computacional para criar sombras realistas, oferecendo naturalidade a cenas tridimensionais. No caso do scanner este efeito é garantido pela iluminação e após a aquisição não consome recursos computacionais para conseguir exibi-lo, uma vez que a foto já foi adquirida com este efeito. Além disso, o efeito da luz e sombras são reais, ou seja, não estão suscetíveis a falhas ou imprecisões.

Trata-se de uma grande diferencial com relação à modelos tridimensionais que possuem custo computacional para renderizar o objeto e posteriormente utilizam mais recursos computacionais para criar sombras para os objetos desenhados. Nos objetos adquiridos pelo F2S2 as sombras reais já estão presentes no modelo sem custo computacional extra.

Por outro lado, o processo de fotografia apresenta dificuldades intrínsecas relativas à brilho e reflexo. Estas características podem ser minimizadas com um iluminação homogênea e difusa, ou seja, que não é originária de um ou poucos pontos de muita intensidade e sim de grandes áreas bem distribuídas espacialmente e de baixa intensidade. O F2S2 possui um sistema de iluminação com 14 lâmpadas LED tubulares, destas 10 unidades estão protegidas por uma placa acrílica translúcida de forma que a luz é emitida de forma difusa.

O processo de controle de iluminação do scanner permite o acendimento individual das lâmpadas, permitindo, por exemplo, intercalar lâmpadas acesas e apagadas, homogeneizando ainda mais a iluminação e diminuindo possíveis brilhos. O ajuste da intensidade luminosa é avaliado com um fotômetro e explicado na Subseção 4.3.2.

Translúcidos

Ao fotografar objetos translúcidos é comum a acentuação de brilho, que pode comprometer partes da fotografia, omitindo detalhes do objeto. Este processo pode ser suavizado diminuindo a quantidade de luz e maximizando a homogeneidade - diminuindo para zero EV a diferença de intensidade luminosa durante a aquisição. Porém, é preciso levar em consideração o objeto como um todo e o conjunto de fotografias que serão adquiridas.

Reflexivo e espelhado

Objetos reflexivos e espelhados apresentam brilho muito acentuado e tendem a assumir a cor do fundo infinito. Para objetos espelhados não é possível contornar estas características controlando a iluminação, porém é possível minimizar o brilho oferecendo uma iluminação homogênea (Variação da intensidade luminosa menor que $\frac{1}{3}EV$) e com intensidade adequada. Por intensidade adequada entende-se uma intensidade que permita uma configuração de exposição compatível com a câmera (sensibilidade ISO e velocidade do obturador), componentes ópticos

(abertura do diafragma) e com a aquisição, no que se refere ao tempo de obtenção da fotografia (velocidade do obturador) e o tempo entre fotos 4.3.2.

Sobreposição e sombra

Para objetos com muita sobreposição e sombras, como o boneco da Figura 34, é necessário ajustar a iluminação com mais intensidade. Neste caso específico, como a capa do boneco tende a causar sombras no seu corpo, a luz deve ser direcionada mais fortemente de maneira frontal, usando $\frac{1}{3}EV$ a mais na luz horizontal, se comparada com a vertical (pois é onde ocorre mais sombra e sobreposição). Este processo de ajuste da iluminação deve levar em conta a forma de cada objeto e suas peculiaridades relativas a sombras e oclusão.

Objetos mistos

Para objetos que possuam ao mesmo tempo componentes opacos e reflexivos/espelhados, o processo de configuração da iluminação deve prever ambas as situações. Uma vez que a iluminação ideal para objetos extremamente opacos é diferente da iluminação para objetos reflexivos, pode não ser possível encontrar uma situação ideal de iluminação que contemple os dois cenários. A Figura 83 apresenta um grampeador com componentes opacos e reflexivos.

Na Figura 83 são apresentadas duas condições de iluminação: uma adequada para visualizar os componentes opacos (Figura 83(a)) do grampeador e outra para visualizar os componentes reflexivos (Figura 83(b)). É possível ver na ampliação da Figura 83(c) que detalhes de componentes opacos estão bem visíveis. Na ampliação da Figura 83(e) percebe-se que o brilho das partes metálicas atrapalha a visualização de detalhes. Por outro lado, a Figura 83(b) apresenta uma condição de iluminação adequada para visualizar os componentes reflexivos do grampeador. Na Figura 83(f) é possível notar que detalhes de componentes reflexivos estão bem visíveis, porém componentes opacos perderam detalhes(Figura 83(d)).

Configuração da câmera

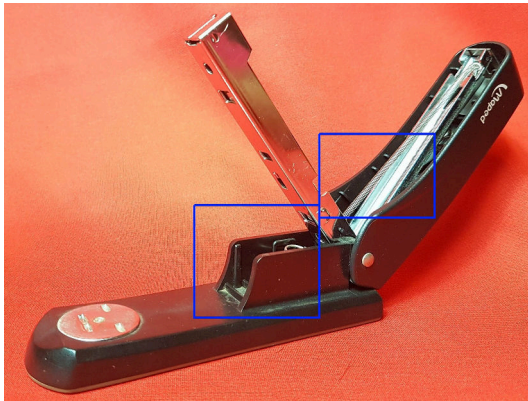
O processo de configuração da câmera deve ser totalmente manual. Qualquer parâmetro que esteja configurado no modo automático pode ser alterado durante o processo de aquisição, fazendo com que esta alteração seja perceptível durante a visualização. Considere que ao visualizar um objeto ele não deve mudar de cor (alteração na exposição) ou tamanho (alteração no foco ou distância focal).

Exposição

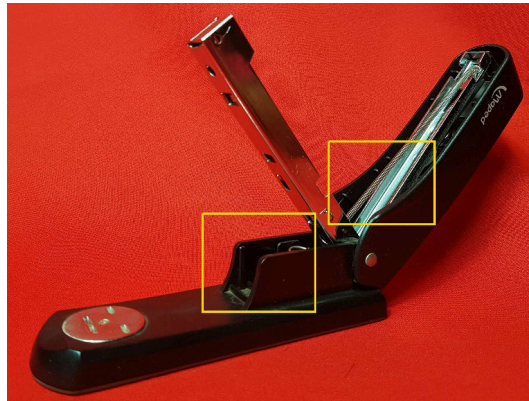
Além de controlar a iluminação, é necessário controlar a exposição nas fotografias. O processo de configuração da iluminação controla brilhos e sombras dependendo do objeto que está sendo digitalizado. A exposição pode ser usada para maximizar a quantidade de detalhes

Figura 83 – Objetos mistos - luz adequada para partes opacas e reflexivas.

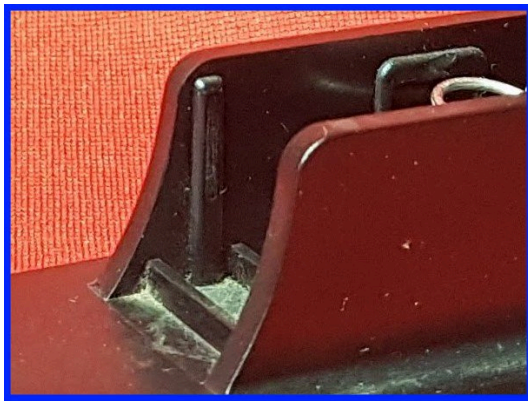
(a) Imagem completa - Mais luz.



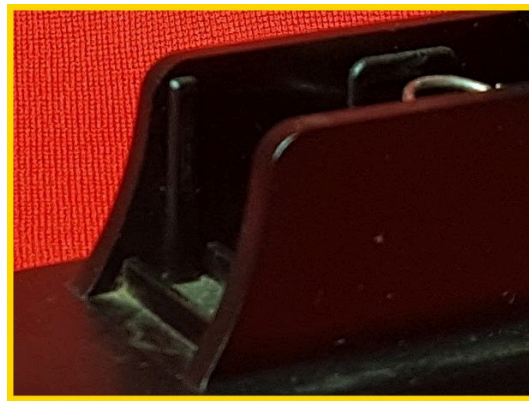
(b) Imagem completa - Menos luz.



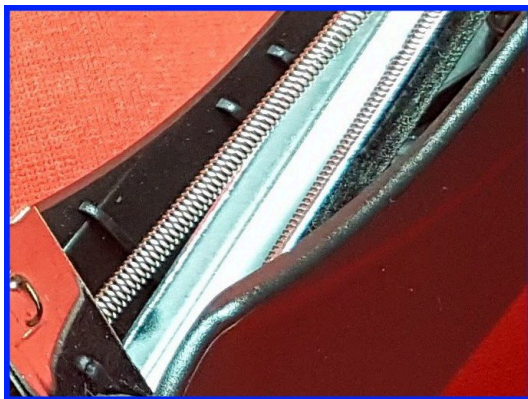
(c) Imagem ampliada 1 - Mais luz.



(d) Imagem ampliada 1 - Menos luz.



(e) Imagem ampliada 2 - Mais luz.



(f) Imagem ampliada 2 - Menos luz.



Fonte: Autoria própria.

que serão capturados em cada imagem. A exposição controla a quantidade de luz que o sensor da câmera receberá. Trata-se de uma configuração fundamental para o processo de fotografia. Para controlar a exposição manualmente três mecanismos devem ser configurados:

- Sensibilidade ISO: controla quão sensível à luz o sensor será. Para este parâmetro, considerando a capacidade de iluminação do F2S2 e que quanto menor este valor, menos ruído

estará presente na imagem, o valor padrão é 100;

- Abertura do diafragma (f): diâmetro que restringe a passagem de luz pelo conjunto óptico. Quanto menor a abertura, maior a profundidade de campo, assim deve ser usado o maior valor de f disponível para o conjunto óptico.
- Velocidade do obturador: controla por quanto tempo a luz atinge o sensor. Considerando que os outros dois mecanismos possuem um valor padrão para uso no F2S2, este é o mecanismo que efetivamente controlará a exposição no F2S2.

O fotômetro oferece um parâmetro de configuração da exposição da câmera para determinada iluminação. Porém a exposição também deve ser configurada de forma dependente do objeto. A unidade de medida da exposição também é EV , de forma que para uma iluminação com $10EV$, caso a câmera esteja configurada para também para $10EV$ a exposição é apropriada. Porém, a exposição pode ser ajustada para receber mais ou menos luz, de forma que neste caso, se a câmera for configurada para $9EV$ a foto é dita subexposta e se for configurada para $11EV$ a foto é dita superexposta.

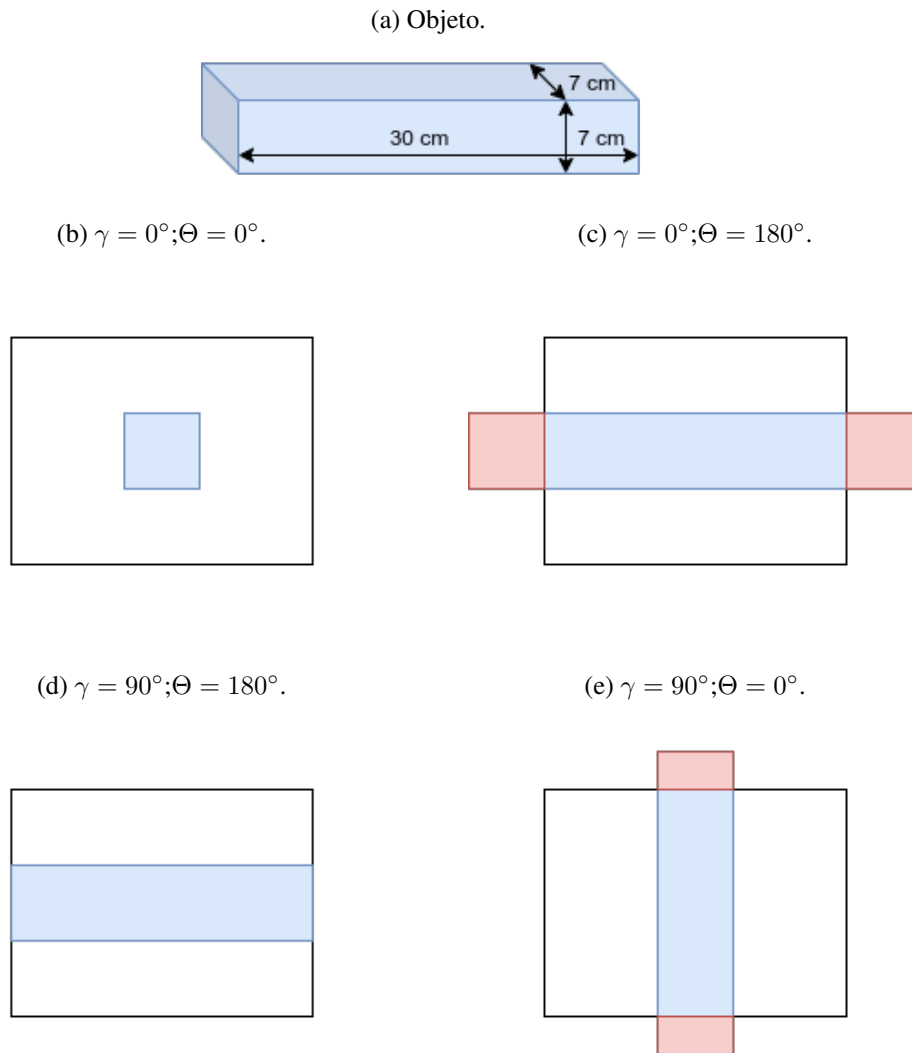
De forma geral, para objetos opacos e escuros deve-se superexpor a fotografia em até $\frac{2}{3}EV$ e para objetos claros e reflexivos deve-se subexpor a foto em até $\frac{2}{3}EV$. Este processo ajuda a preservar mais detalhes do objeto na fotografia, uma vez que o fotômetro de luz incidente não leva em consideração a luz refletida pelo objeto. Como é comum na fotografia, este processo deve ser realizado através de testes para cada objeto. Ou seja, algumas fotografias devem ser adquiridas e analisadas para definir qual a melhor configuração para a aquisição.

Enquadramento

Novamente a diferença entre a técnica de fotografia comum, que enquadra o objeto para apenas uma fotografia, e o F2S2, é que o enquadramento deve considerar todo o conjunto de fotografias da aquisição. Neste caso o enquadramento deve garantir que o objeto ocupe a maior parte dos pixels do sensor sem que em nenhuma das fotografias qualquer parte do objeto seja excluído da fotografia.

De forma similar ao que acontece no ajuste de foco, apresentado na seção 4.3.1, ao mover o objeto ou a câmera, o enquadramento é alterado. Com a alteração do enquadramento duas situações podem acontecer. A Figura 84 apresenta estas situações. Com objetos desproporcionais (Figura 84(a)), com o enquadramento para um determinado ângulo do objeto (Figura 84(b)), ao mover o objeto, este, devido as suas proporções pode sair da área fotografada (Figura 84(c)). Além disso, principalmente em ângulos próximos de $\gamma = 90^\circ$, considerando que a fotografia possui um formato 3:4 e não quadrado, para determinado ângulo o objeto está adequadamente enquadrado (Figura 84(d)), visto que o comprimento da fotografia é maior que a altura; ao girar o objeto a fotografia pode excluir partes do objeto do campo de visão (Figura 84(e)).

Figura 84 – Problemas no enquadramento.



Fonte: Autoria própria.

A forma de resolver o problema do enquadramento que foi mais eficiente durante as aquisições foi realizar uma aquisição de teste com poucos frames. Inicialmente o usuário do scanner tinha a sensação de que ao ajustar o enquadramento todas as fotografias estariam satisfatórias. Porém, após o processo de aquisição algumas fotos estavam incorretas, obrigando a realização de uma nova aquisição. Assim, estabeleceu-se que dada a dificuldade em prever o enquadramento de todas as fotografias da aquisição, sempre seria realizada uma aquisição de teste com algo em torno de 20 fotos para analisar o enquadramento.

Foco

Conforme apresentado no início da Seção, o foco deve levar em consideração a forma e proporções do objeto. Deve ser configurado de forma que o plano de foco fique um pouco atrás do início do objeto, visando respeitar a profundidade de campo e não apenas o plano de

foco, como ocorre em fotografias comuns. Este processo deve ser realizado manualmente, pois o ajuste automático levaria em conta apenas o início do objeto e não seu volume como um todo.

Além disso, como mencionado nesta subseção, a abertura do diafragma deve ser a menor que a câmera/conjunto óptico conseguir, de forma a também maximizar a profundidade de campo. Por último, outro fator capaz de aumentar a profundidade de campo é a distância da câmera ao objeto. Esta distância deve ser a maior possível para os limites das lentes e do scanner, visando uma maior profundidade de campo.

Após realizar estas análises, o foco pode ser configurado manualmente, de forma que seu plano esta um pouco depois do início do objeto. Para realizar esta operação, configura-se o foco normalmente, no início do objeto, então ajusta-se o mesmo para que fique um pouco atrás deste ponto, cuidando para que o início do objeto esteja nítido na imagem (ou seja, ainda dentro da profundidade de campo).

Tamanho das fotos

O tamanho das fotografias varia de acordo com a câmera usada. Dentre as possibilidades oferecidas pelo equipamento, o tamanho das imagens vai depender do objetivo da aquisição. Por exemplo: um animal digitalizado para identificação de espécies necessitaria de tamanhos de imagens maiores, que apresentam mais detalhes e mais nitidez, do que objetos sendo digitalizados para publicidade na internet. Além disso, o tamanho das imagens representa também mais recursos de armazenamento e rede, de forma que imagens muito grandes podem ser inconvenientes para transmissão pela internet.

O tamanho dos arquivos influencia na capacidade de aplicar zoom; tamanho da tela de projeção; nitidez; aquisição de detalhes; necessidade de espaço em disco; dificuldade de transmissão pela rede; tempo de processamento das fotografias; recursos computacionais para exibição das aquisições; entre outros fatores.

Aquisição de Teste

Considerando a grande quantidade de parâmetros que devem ser ajustados para uma aquisição e o fato de que nem todos podem ser automatizados, como por exemplo os processos manuais de posicionamento (e fixação) do objeto ou a análise do enquadramento. Existe a possibilidade de iniciar a aquisição enquanto alguns destes parâmetros não estejam satisfatoriamente ajustados. Como uma aquisição pode levar seis, oito, dez ou mais horas, com o objetivo de minimizar possíveis erros, estabeleceu-se que só se inicia uma aquisição após realizar uma aquisição de teste.

O objetivo desta aquisição é adquirir uma pequena quantidade de fotografias para que seja analisar possíveis erros que não podem ser verificados sem um conjunto de fotografias conseguidas sobre a influência do movimento do scanner. Dentre as características analisa-

das podem ser citadas: imagens tremidas por instabilidade/inércia do objeto; problemas com enquadramento durante a aquisição; problemas de foco/profundidade de campo ao longo da aquisição; excentricidade do objeto; fixação insatisfatória do objeto; iluminação inadequada para determinadas posições; configuração incorreta da câmera; além de outros problemas que venham a ser detectados.

Estabeleceu-se então uma aquisição de teste composta de cinco Streams com quatro Frames cada. Esta aquisição é rápida, assim como a análise das fotografias, que são feitas em menos de dez minutos. Diversos erros já foram corrigidos nesta etapa e acabaram por economizar diversas horas de trabalho e processamento das imagens. Em uma aquisição pequena os erros citados são facilmente percebidos, pois é fácil analisar fotos de todos os pontos da esfera. Porém, em uma aquisição de quatro mil fotografias, algumas vezes estes erros só são percebidos posteriormente ao trabalho de processamento.

Uma aquisição de cinco Streams com quatro Frames cada possui $\Delta\gamma = 22.5^\circ$ e $\Delta\Theta = 90^\circ$. Assim é possível observar o objeto girar cinco vezes sobre seu eixo, testando características mecânicas como inercia, estabilidade, excentricidade, fixação. Como são adquiridas fotos de toda a esfera formada pelo F2S2, é possível perceber problemas de enquadramento, foco/profundidade de campo, iluminação e configuração da câmera.

A técnica de testar configurações ou componentes antes de iniciar um processo é aplicada em diversas áreas, para citar uma pode ser usado como exemplo a BIOS da placa mãe que testa componentes como placa de vídeo e memória antes da inicialização. Atualmente o F2S2 não conta com uma análise automática deste processo, porém, usando técnicas de processamento de imagens este processo pode ser automatizado.

Colorchecker

O colorchecker é um cartão com cores impressas com precisão de cores, e através de um software específico é possível calibrar as cores de um conjunto de fotografia para que estas sejam as mais próximas as cores reais. No F2S2 seu uso é opcional e depende do objetivo da aquisição. O processamento é realizado pelo software que acompanha o cartão de cores, o qual pode realizar processamento em lotes de imagens de forma automática.

O uso do colorchecker para ajuste das cores da aquisição precisa que as fotografias estejam com a mesma iluminação e a câmera com a mesma configuração para todas as fotografias que serão tratadas. Isto não é um problema para o F2S2 uma vez que estas são as mesmas condições necessárias para a aquisição. Apenas uma fotografia do colorchecker nestas circunstâncias é o suficiente para tratar todo o *pack* de imagens.

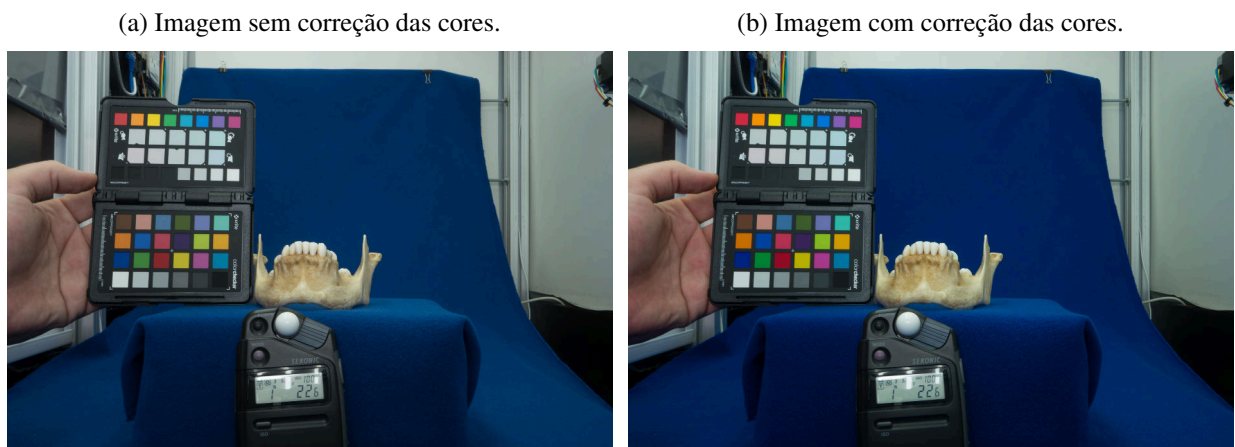
Para garantir que as condições de iluminação e a configuração da câmera durante a fotografia do colorchecker e durante a aquisição sejam as mesmas, é importante que a aquisição do colorchecker seja realizada após a aquisição de teste. Assim todos estes parâmetros já foram

testados e não serão alterados. Inclusive, o Software de controle do scanner possui uma etapa específica para tirar a foto do colorchecker antes da aquisição.

Assim, após a aquisição de teste, o cartão de cores é colocado junto ao objeto que será escaneado; uma fotografia é adquirida e separada do *pack* de imagens, pois ela só será usada para a calibração das cores; após a aquisição e antes da etapa de processamento o software de calibração associado ao cartão é executado e corrige as cores do *pack* de imagens; então as fotografias corrigidas podem passar pela etapa de processamento.

A Figura 85 apresenta o colorchecker fotografado antes de uma aquisição. A Figura 85(a) apresenta a imagem original, ou seja sem correção das cores, enquanto que a Figura 85(b) apresenta a mesma imagem com as cores corrigidas pelo colorchecker. O colorchecker usado no laboratório é o modelo *ColorChecker Passport Photo 2* ⁶ e o software para criação do perfil de cores corrigidas é o *ColorChecker Camera Calibration v2.0*. ⁷

Figura 85 – Colorchecker.



Fonte: Autoria própria.

Aquisição

Quando todas as etapas necessárias para a aquisição foram realizadas, o processo de aquisição pode ser iniciado. É necessário ainda estabelecer algumas condições para o processo de aquisição, como tempo de duração do processo, estabilidade da energia elétrica e espaço necessário para armazenamento das fotografias.

O tempo de duração de uma aquisição depende principalmente do tempo entre cada fotografia e da quantidade de fotografias. Apenas para fins de exemplificação, uma aquisição com

⁶ <<https://www.xrite.com/categories/calibration-profiling/colorchecker-targets/colorchecker-passport-photo-2>> Acessado em 16 de Abril de 2020.

⁷ <https://xritephoto.com/ph_product_overview.aspx?ID=938&Action=Support&SoftwareID=2030> Acessado em 16 de Abril de 2020.

$\Delta\Theta = 3^\circ$ e $\Delta\gamma = 3^\circ$ (ou seja 31 Streams com 120 Frames cada, totalizando 3720 fotografias) e tempo entre fotos de 3 segundos leva aproximadamente 6h.

Estes valores também irão variar de acordo com a distribuição das imagens entre frames e Streams e raio da aquisição, de forma que quanto maior o raio, maiores os movimentos, maior o tempo de aquisição. Além disso, o processo de ajuste do scanner e da peça, leva aproximadamente 1h. Este tempo depende do usuário do scanner e da complexidade da peça.

Como já mencionado, durante o uso do scanner é necessário garantir que o sistema de energia esteja estável para não interferir na movimentação do scanner e na iluminação.

Também é necessário garantir espaço de armazenamento das fotografias no cartão de memória da câmera. A aquisição com $\Delta\Theta = 3^\circ$ e $\Delta\gamma = 3^\circ$ usa aproximadamente 100Gb no formato *RAW*. O cartão usado no laboratório possui 128Gb, de forma que isto limitava a aquisição aos parâmetros acima. Além disso o cartão precisava estar vazio para iniciar a aquisição.

O Software de controle do scanner permite pausar a aquisição, porém a câmera não pode ser desligada, pois esta perde as configurações, principalmente a configuração de foco, que é manual e não pode ser ajustada por parâmetros numéricos, apenas girando manualmente o anel de foco do conjunto óptico. Assim, ao tentar retornar ao foco anterior com exatidão, pode acarretar em erro que seria perceptível na visualização.

Resultado do processo

Assim, como resultado do processo é gerado o conjunto de fotografias do objeto segundo os parâmetros estabelecidos durante a configuração do scanner e da câmera. Estas imagens podem então ser processadas na etapa de processamento para que então possam ser visualizadas.

5 Resultados

Para que seja possível aplicar o protocolo de escaneamento para digitalizar os objetos com o F2S2; para processar as imagens adquiridas construindo o Stream2D; viabilizando a visualização das imagens, foram desenvolvidos dois softwares. O *F2S2-App* é responsável por controlar o F2S2 durante o processo de escaneamento e o *F2S2-BipApp* é responsável pelo processamento do *pack* de imagens e geração do Stream2D.

Este Capítulo inicia apresentando os dois softwares desenvolvidos. Posteriormente é apresentada uma validação do método desenvolvido e por fim uma análise dos resultados conseguidos.

5.1 Software de Aquisição - F2S2-App

Esta Seção apresenta o software de controle do F2S2. As principais funções deste aplicativo são: coletar as configurações da aquisição; gerar as instruções em GMCodes; enviar as instruções ao scanner controlando sua execução. Para isso foi desenvolvido um software que é executado em um computador externo ao F2S2 e que se comunica através de uma conexão USB com o scanner, coordenando o envio das instruções previamente calculadas e analisando as mensagens de retorno do F2S2 para sistematizar os eventos da aquisição.

O software foi chamado de *F2S2-App* e foi desenvolvido em *Python v3.7* usando as prerrogativas de boas práticas de programação em python - PEP8¹, visando maior manutenibilidade do código para futuras atualizações. Para a interface gráfica foi usada a biblioteca Kivy², pois além de oferecer a estrutura de construção de interfaces gráficas como botões, entradas de texto, esta biblioteca permite que na tela sejam desenhados objetos em Canvas, de forma que foi possível apresentar gráficos em tempo real do estado do scanner.

Outras bibliotecas também foram usadas, como: NumPy³, para processamento de cálculos numéricos; JSON⁴, para salvar arquivos de configuração do programa; threading⁵, para uso de threads e semáforos; PySerial⁶, para comunicação serial com o F2S2. Além disto, para oferecer maior facilidade de instalação, manutenção e estabilidade do sistema, o *F2S2-App* foi empacotado dentro de um *container* da plataforma Conda⁷.

O *F2S2-App* foi desenvolvido usando padrões de projetos e a arquitetura MVC (*Model-*

¹ <<https://www.python.org/dev/peps/pep-0008/>> Acessado em 17 de Abril de 2020.

² <<https://kivy.org>> Acessado em 17 de Abril de 2020.

³ <<https://numpy.org/>> Acessado em 17 de Abril de 2020.

⁴ <<https://docs.python.org/3/library/json.html>> Acessado em 17 de Abril de 2020.

⁵ <<https://docs.python.org/3/library/threading.html>> Acessado em 17 de Abril de 2020.

⁶ <<https://pythonhosted.org/pyserial/>> Acessado em 17 de Abril de 2020.

⁷ <<https://www.anaconda.com/distribution/>> Acessado em 24 de Abril de 2020.

View-Controller - em português: Modelo-Visão-Controle) (GIRIDHAR, 2016) que separa o software em três camadas (Modelo - Visão - Controle) de forma a reduzir o acoplamento do código e oferecer maior manutenibilidade. Os padrões de projeto (GAMMA, 2009) usados foram *Singleton*: para garantir que apenas um objeto controlasse o F2S2, controlando assim o acesso ao scanner e ao mesmo tempo garantindo acesso global ao objeto responsável pela aquisição; *Observer*: para notificar a interface gráfica e o sistema de *log*. A interface gráfica desenha o estado atual do scanner e o *log* registra os eventos da aquisição.

A atual versão do programa possui 5.746 linhas (22.722 inserções e 16.976 remoções), 54 commits no sistema de controle de versão onde foi desenvolvido e consome aproximadamente 50Mb de memória para executar. Durante sua execução são criados arquivos com todas as instruções GMC que serão executadas (arquivo “aquisicao.GMC”), um arquivo de *log* da execução (arquivo “log.txt”) e um arquivo com todas as informações da aquisição e instruções (arquivo “aquisicao.json”). Além disso podem ser salvos arquivos de configuração de aquisições (que são salvos com um nome escolhido pelo usuário e a extensão “.json”).

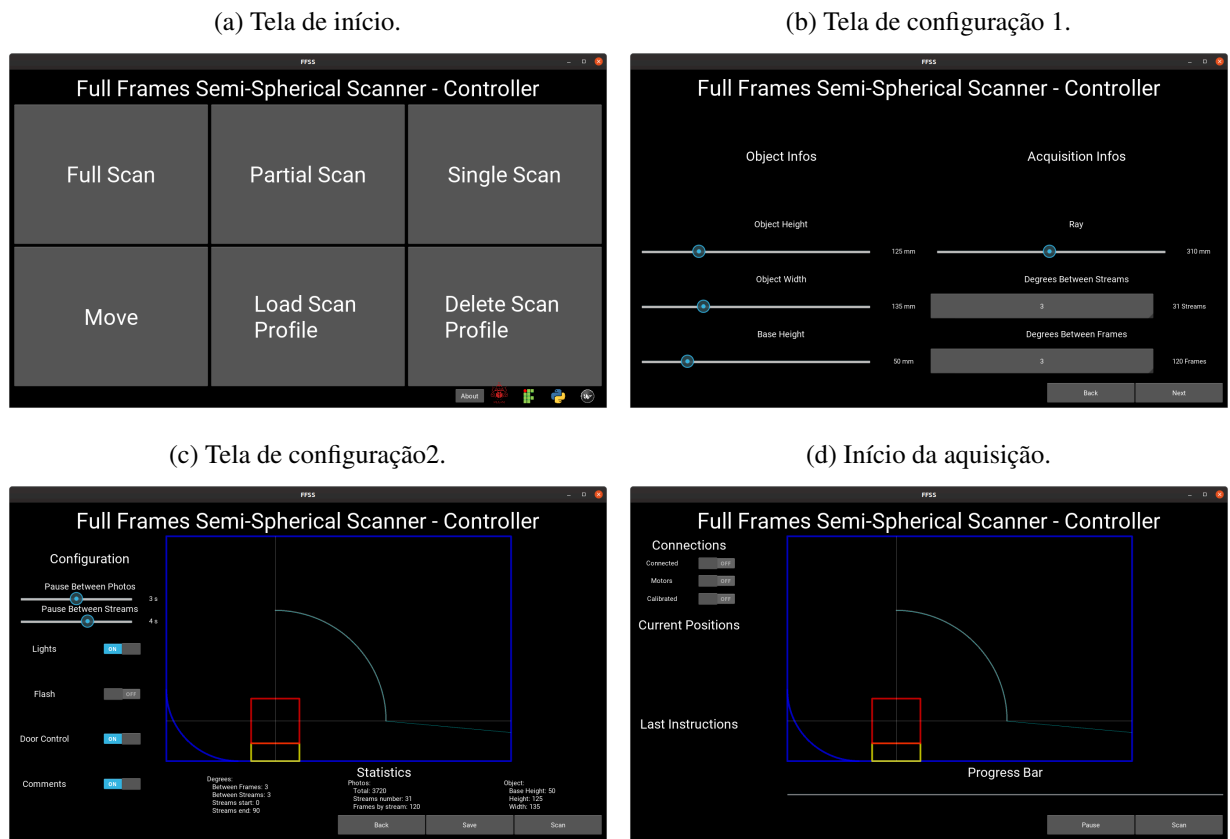
A Figura 86 apresenta as telas principais do *F2S2-App*, sendo a Figura 86(a) a tela inicial do aplicativo, na qual é possível escolher os modos de operação do software, conforme discutido na seção 5.1.1; a Figura 86(b) apresenta a configuração de uma aquisição para um objeto de 125mm de altura, 135mm de comprimento, base de suporte de 50mm, raio de aquisição de 310mm, $\Delta\theta = 3^\circ$ e $\Delta\gamma = 3^\circ$; a Figura 86(c) apresenta o final da configuração com o tempo de espera entre fotos de 3 segundos e entre streams de 4 segundos, luzes ligadas, flash desligado, porta controlada automaticamente e a inserção de comentários (como hora de criação e créditos do *F2S2-App*) no arquivo gerado com as instruções GMC.

Ainda na Figura 86(c), é apresentado um gráfico contendo: em azul o scanner e o fundo infinito, em vermelho o volume do objeto, em amarelo o volume da base de suporte, em branco o centro do objeto, em verde o conjunto de movimentos da câmera durante a aquisição. A Figura 86(d) apresenta o início do processo de aquisição.

A Figura 87 apresenta a tela durante o processo de aquisição. Nela são exibidas as coordenadas da câmera, últimas instruções executadas, a linha verde que representa o caminho da câmera é pintada de rosa para instruções já realizadas e um círculo vermelho apresenta as posições onde a câmera adquiriu fotografias. Nesta imagem também é possível ver um semi-círculo branco, que representa a esfera formada durante a aquisição e uma barra de progresso da aquisição.

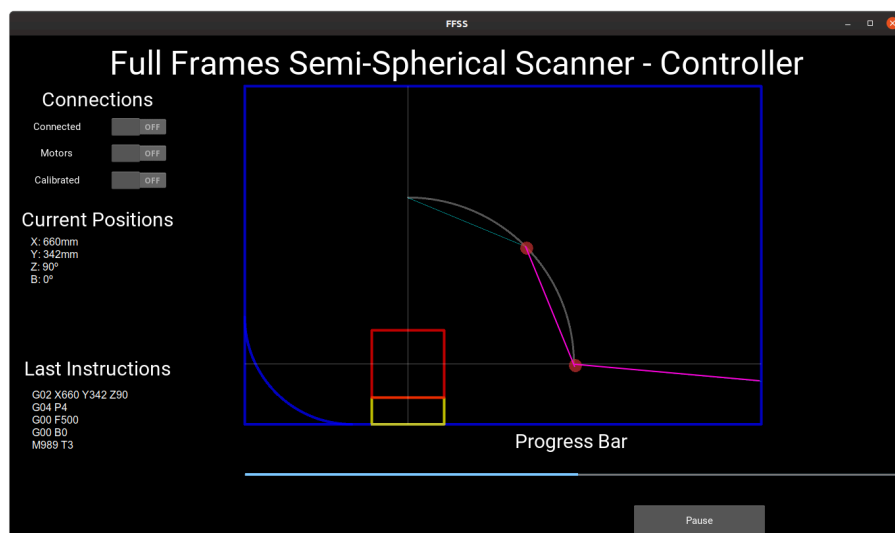
Com os gráficos gerados durante a aquisição é possível acompanhar o processo de aquisição em tempo real, conhecendo cada instrução que foi enviada para o scanner, a posição de cada um dos quatro eixos, o estado da conexão do computador com o F2S2, o estado dos motores e se o scanner já foi calibrado. Estas informações são importantes para auditar em tempo real quaisquer problemas que possam acontecer. Adicionalmente, todas as instruções enviadas são armazenadas em um *log* com a hora de envio de cada instrução, permitindo que uma auditoria da

Figura 86 – F2S2-App.



Fonte: Autoria própria.

Figura 87 – Tela de aquisição do F2S2-App.



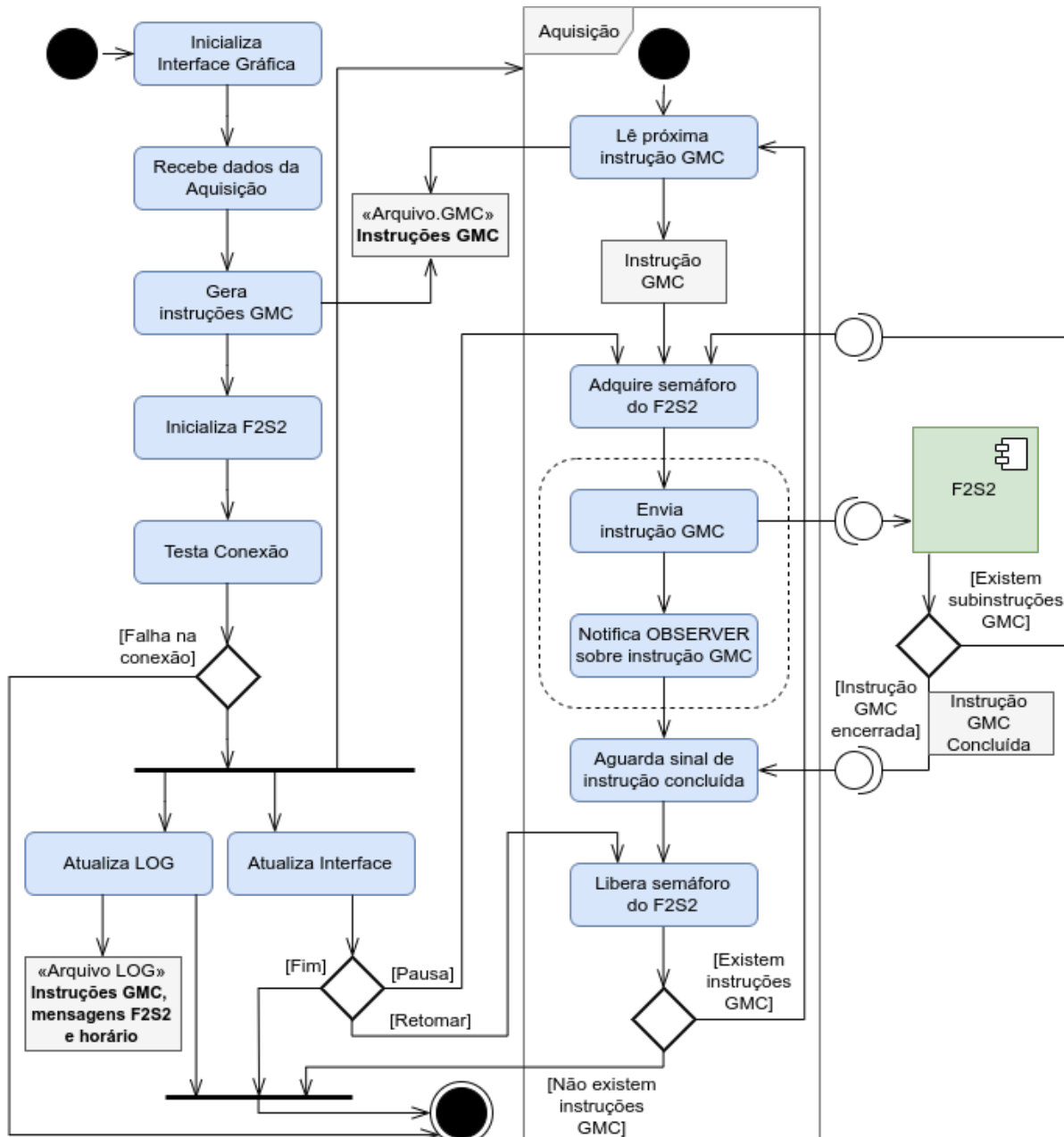
Fonte: Autoria própria.

aquisição seja executada após seu término.

A Figura 88 apresenta o diagrama de atividades do *F2S2-App*. Ao iniciar o software

são solicitadas as informações da aquisição (como na Figura 86b) então é gerado um arquivo “aquisicao.GMC” com os comandos GMC da aquisição solicitada. Ao iniciar a aquisição o aplicativo inicializa o scanner e testa a conexão, se estiver funcionando são lançadas duas threads, uma responsável pelo LOG, outra pelo processo de aquisição (usando o padrão de projetos singleton) e a thread principal fica responsável pela interface gráfica.

Figura 88 – Diagrama de atividade do F2S2-App.



Fonte: Autoria própria.

O processamento da aquisição consiste em ler as instruções do arquivo “aquisicao.GMC”, adquirir um semáforo que garante acesso exclusivo e coordenado de envio de instruções ao scanner. Este semáforo possui três funções principais: pausar a aquisição, sincronizar as instruções e

permitir que o F2S2 execute subinstruções.

Inicialmente é necessário estabelecer que caso diversas instruções sejam enviadas simultaneamente ao F2S2, este pode realizar movimentos que danifiquem a peça ou o próprio scanner. Pode ser feita uma analogia com apertar diversas teclas de uma máquina de escrever que acabam por travar seu mecanismo.

Para que seja possível pausar uma aquisição, é necessário aguardar a instrução, que está sendo executada, finalize. Então não se enviam novas instruções enquanto o usuário desejar o sistema pausado. Para isso, ao clicar no botão de pausa o *F2S2-App* adquire o semáforo que só será liberado quando o usuário comandar a retomada da aquisição.

O sincronismo das instruções é necessário pois o tempo de execução de cada instrução é desconhecido. Considere por exemplo uma instrução “G00 X600” que move a câmera para a posição 600 do eixo X' . O tempo de execução desta instrução depende do valor anterior do eixo X' e da velocidade do movimento, que é configurado pelo comando “G00 F”. Assim, como o tempo de execução de cada instrução não é conhecido e para evitar enviar instruções antes que o scanner tenha terminado as instruções anteriormente enviadas, é usado o semáforo, que só é liberado após a confirmação, feita pelo scanner, de que a instrução anterior foi concluída.

Outro uso do semáforo é permitir que o F2S2 execute instruções compostas. Por exemplo a instrução “M998” calibra o scanner, ou seja, coloca todos os eixos na posição zero. Esta instrução, além de não possuir tempo conhecido, executa movimentos nos 4 eixos do scanner e portanto lança quatro instruções para executar. Para garantir que nenhuma instrução seja enviada antes da conclusão da instrução “M998”, cada subinstrução que o scanner lança, adquire o semáforo, de forma que este só liberará o fluxo de execução após todas as instruções terem sido concluídas e cada uma delas liberar o semáforo que adquiriu. Somente neste momento o *F2S2-App* conseguirá enviar uma nova instrução.

Enfim, após adquirir o semáforo e este liberar o fluxo de execução, o *F2S2-App* envia a instrução ao scanner usando a comunicação serial (cabo USB) e em seguida, usando o padrão de projetos observer notifica o LOG e a interface gráfica sobre a instrução enviada. O LOG armazena a instrução em um arquivo de texto e a interface gráfica desenha a instrução na tela atualizando o usuário do novo estado do F2S2, conforme apresentado na Figura 88.

Neste momento o *F2S2-App* tenta novamente adquirir o semáforo. Como o F2S2 ainda não concluiu a execução da instrução anterior, o semáforo está bloqueado e o software aguarda a execução da instrução pelo scanner. Ao concluir a execução da instrução o F2S2, usando a comunicação serial, comunica a finalização e libera o semáforo. Assim o fluxo de execução é retomado e o *F2S2-App* lê uma nova instrução GMC do arquivo e continua a execução enquanto existirem novas instruções no arquivo.

Ao final da aquisição o arquivo de LOG é salvo e a interface atualizada com uma mensagem de execução finalizada. O software pode então ser fechado e as imagens adquiridas

podem ser baixadas da câmera.

5.1.1 Modos de operação

O *F2S2-App* possui diversos modos de operação, assim é possível realizar aquisições com liberdade para variar o angulo inicial (γ^0) e final (γ^p), o intervalo angular entre de Streams ($\Delta\gamma$) e Frames ($\Delta\Theta$). Para isso foram desenvolvidos 3 modos de escaneamento (*Full Scan*, *Partial Scan* e *Single Scan*), um modo de movimento livre (usado para calibração, manutenção e aquisições livres) e a opção de salvar e carregar perfis de aquisição. Todos os modos de escaneamento possuem em comum a entrada de dados sobre a altura e largura do objeto, altura da base de suporte, raio da aquisição. Estes dados são usados para garantir que a câmera não colida com o objeto ou com o próprio scanner. Assim, para todas as aquisições são conferidos:

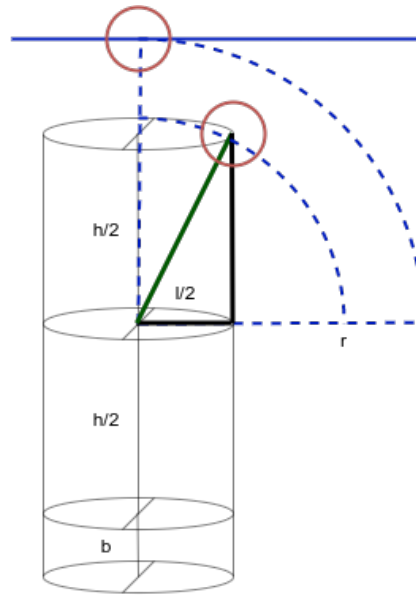
Para evitar colisão da câmera com o objeto : Como o processo de aquisição rotaciona o objeto, os objetos ao longo da aquisição podem ser entendidos como um volume cilíndrico. Considerando que o centro da esfera é alinhado ao centro geométrico do objeto, ou no caso, do volume cilíndrico gerado, o raio da aquisição não pode ser menor que a diagonal de um cilindro com metade da altura do objeto. Portanto, se: $r \leq \sqrt{\frac{l^2}{2} + \frac{h^2}{2}}$ haverá colisão com o objeto. Sendo, r : raio da esfera; l : largura do objeto; h : altura do objeto.

Para evitar colisão da câmera com o scanner : Considerando que a altura do scanner é de $650mm$ e que a esfera fica alinhada com o centro geométrico do objeto, caso: $b + r + \frac{h}{2} > 650$ haverá colisão com o limite superior do scanner. sendo b a altura da base de suporte. A Figura 89 apresenta os dois casos de possíveis colisões do sistema da câmera com o scanner e com o objeto.

Para todas as aquisições o intervalo angular horizontal é de 360° . Ou seja, o objeto sempre faz a rotação completa, não sendo contempladas aquisições horizontais parciais, como, por exemplo, usando apenas 180° na horizontal para a aquisição. Isso posto, o intervalo horizontal $\Delta\Theta$ deve gerar uma quantidade de Frames que contenham todos os 360° . Ou seja, $\Delta\Theta$ deve compreender uma divisão exata pelos 360° . Em outras palavras, não é possível realizar uma aquisição com $\Delta\Theta = 190^\circ$, pois os frames gerados seriam: $\Theta^0 = 0^\circ$, $\Theta^1 = 190^\circ$ e $\Theta^2 = 380^\circ$. Como este último valor não é representado, esta aquisição não é permitida.

Para todos os modos de operação são conferidos os limites de segurança contra colisão com o objeto, o limite superior do scanner, respeitados os limites de colisão com o fundo infinito (ver Figura 56c), e os possíveis valores para $\Delta\Theta$. Os modos de operação do scanner oferecidos pelo *F2S2-App* são:

Figura 89 – Possíveis colisões.



Fonte: Autoria própria.

Full Scan

Este é o modo de aquisição completo, usando como ângulo vertical inicial ($\gamma^0 = 0^\circ$) e final ($\gamma^p = 90^\circ$), com estes valores fixados são escolhidos os intervalos angulares entre Streams ($\Delta\gamma$) e Frames ($\Delta\Theta$). Para ambos os casos os limites angulares devem compreender uma divisão exata pelos intervalos angulares. Ou seja, não é possível realizar uma aquisição com $\gamma^0 = 0^\circ$, $\gamma^p = 90^\circ$ e $\Delta\gamma = 47^\circ$.

Assim, o *F2S2-App* realiza automaticamente o cálculo dos limites de segurança comuns a todos os modos de aquisição e, para este modo calcula os possíveis valores para $\Delta\gamma$, e os oferece para que o usuário escolha o valor, não permitindo a entrada de valores inconsistentes. A Figura 90 apresenta as telas de configuração do modo *Full Scan*.

Partial Scan

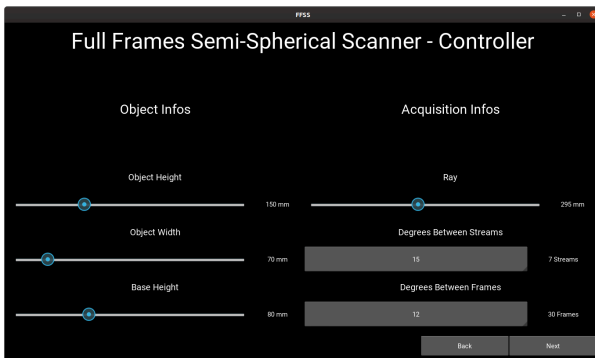
Neste modo de escaneamento é possível selecionar um intervalo angular vertical para a aquisição, variando γ^0 e γ^p . Este processo torna mais complexo a escolha dos possíveis intervalos angulares verticais ($\Delta\gamma$). O processo de escolha dos possíveis valores para $\Delta\gamma$ é definido por: $\{\forall \gamma^0 \leq \Delta\gamma \leq \gamma^p / \Delta\gamma \% (\gamma^0 - \gamma^p) = 0\}$, sendo: % o resto da divisão; $\Delta\gamma \in \mathbb{N}$. A Figura 91 apresenta as telas de configuração do modo de aquisição *Partial Scan*.

Single Scan

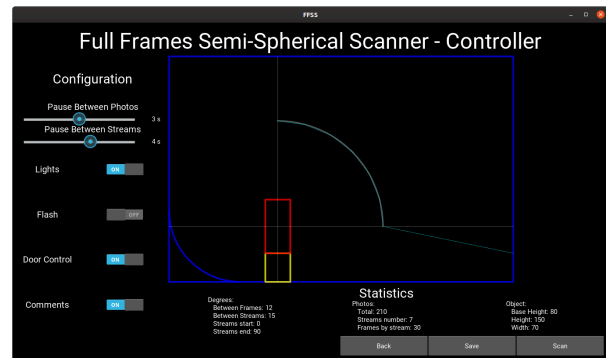
Neste modo de aquisição é realizada uma aquisição de apenas um Stream, de forma que é necessário informar o valor angular da posição desejada. A Figura 92 apresenta as telas de

Figura 90 – F2S2-App - Full Scan.

(a) Tela de configuração 1.



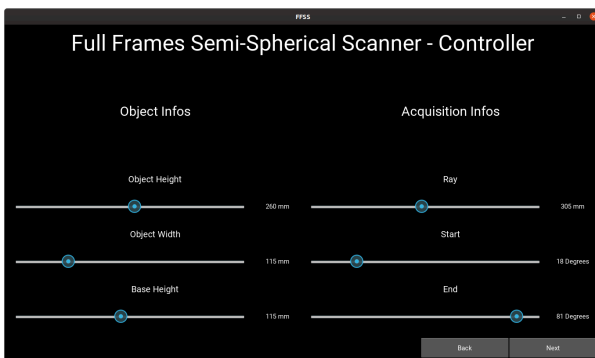
(b) Tela de configuração 2.



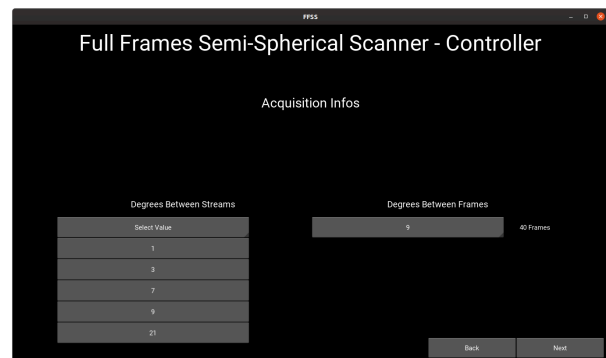
Fonte: Autoria própria.

Figura 91 – F2S2-App - Partial Scan.

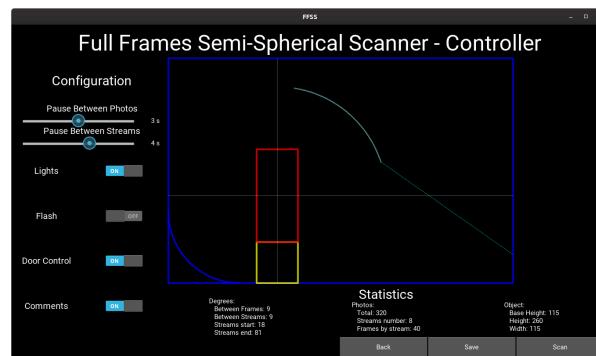
(a) Tela de configuração 1.



(b) Tela de configuração 2.



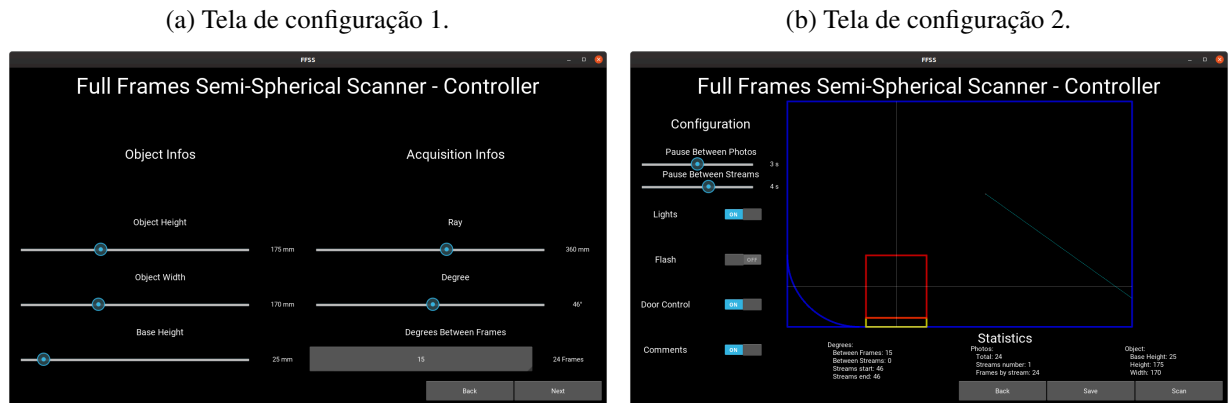
(c) Tela de configuração 3.



Fonte: Autoria própria.

configuração deste modo de aquisição. O Quadro 2 apresenta um resumo dos valores de entrada, checagens de segurança e cálculos automáticos realizados pelo *F2S2-App*.

Figura 92 – F2S2-App - Single Scan.



Fonte: Autoria própria.

Quadro 2 – Quadro de valores de entrada, checagens e cálculos realizados para cada modo de aquisição do *F2S2-App*

Modos de Aquisição		Full Scan	Partial Scan	Single Scan
Entradas	Altura da base	X	X	X
	Altura do objeto	X	X	X
	Largura do objeto	X	X	X
	Raio da Aquisição	X	X	X
	Intervalo angular entre Frames	X	X	X
	Intervalo angular entre Streams	X	X	X
	Intervalo temporal entre Frames	X	X	X
	Intervalo temporal entre Streams	X	X	X
	γ^0	–	X	X
γ^1	–	X	–	
Checagens	Colisão com o objeto	X	X	X
	Colisão com iluminação	X	X	X
	Colisão com o fundo infinito	X	X	X
Cálculos	Possíveis valores para $\Delta\Theta$	X	X	X
	Possíveis valores para $\Delta\gamma$	X	X	–

Fonte: Autoria própria.

Free Move

Neste modo de operação do scanner é possível selecionar os valores de cada um dos eixos (X' , Y' , $Z^a e B^a$) e ao clicar em “move” os eixos se movem para a posição selecionada. Também é possível clicar em “snapshot” para adquirir uma fotografia.

Load/Delete Scan Profile

O *F2S2-App* também permite salvar perfis de escaneamento, este processo pode ser realizado com qualquer modo de escaneamento de forma que, após configurar o escaneamento é

possível clicar em “save” para salvar determinada configuração. Assim, na tela inicial existe a possibilidade de carregar ou excluir determinado perfil.

O salvamento ocorre com a compilação de todas as configuração em um arquivo “JSON” que pode ser carregado pelo software. O Código 5.1 apresenta o exemplo de um arquivo de perfil salvo pelo aplicativo.

código 5.1 – Exemplo de configurações salvas em JSON.

```

1 {"type": "full", "ray": 400, "streams_interval": 45,
2 "frames_interval": 180, "height": 230, "width": 170,
3 "base": 30, "frames_number": 2, "streams_number": 3,
4 "photos_number": 6, "start": 0, "end": 90,
5 "pause_frames": 2, "pause_streams": 4, "lights": true,
6 "flash": false, "door": true, "comment": true }

```

Fonte: Autoria Própria.

5.1.2 Geração das instruções GMC

O processo de construção dos arquivos de instruções GMC iniciam com os comandos de ligar os motores, controle da luz, da porta, do flash e da velocidade do movimento, de acordo com as opções configuradas no *F2S2-App*. Estes comandos são chamados de inicialização do Scanner. Em seguida é realizada a calibração do scanner para posicionar todos os eixos na sua respectiva origem. Após este processo de inicialização e calibração são inseridas as intruções de movimentação.

Criação da esfera/arco

Para calcular os movimentos é necessário primeiro encontrar o centro geométrico do objeto. Este processo para o eixo Y' é realizado com a soma da altura da base de suporte mais a metade da altura do objeto, para o eixo X' o valor é $650mm$, pois esta é a posição do centro do eixo B^a . Centro do objeto = $(X' = 650; Y' = base\ de\ suporte + \frac{altura\ do\ objeto}{2})$. Para exemplificar este cálculo, considere um objeto de $100mm$ de altura sobre uma base de $50mm$. Assim o centro geométrico do objeto está posicionado na coordenada $(X' = 650; Y' = 100)$, pois $Y' = altura\ da\ base + \frac{altura\ do\ objeto}{2} = 50 + \frac{100}{2} = 50 + 50 = 100$

Em seguida é gerada uma lista de todas as posições do eixo B^a de acordo com o valor de $\Delta\Theta$ informado pelo usuário. Por exemplo, caso $\Delta\Theta = 90^\circ$ os valores gerados seriam $\{0^\circ; 90^\circ; 180^\circ; 270^\circ\}$. Note que como $0^\circ = 360^\circ$ esta posição só é adquirida uma vez. O próximo passo é gerar uma lista de todos os valores que γ irá assumir. Caso $\gamma^0 = 0^\circ$, $\gamma^p = 90^\circ$ e $\Delta\gamma = 45^\circ$, os valores gerados seriam $\{0^\circ; 45^\circ; 90^\circ\}$

Com a combinação destes dois conjuntos de valores é possível saber a posição angular de cada foto do Stream2D. A Tabela 4 apresenta a relação de fotos, valores angulares, Streams

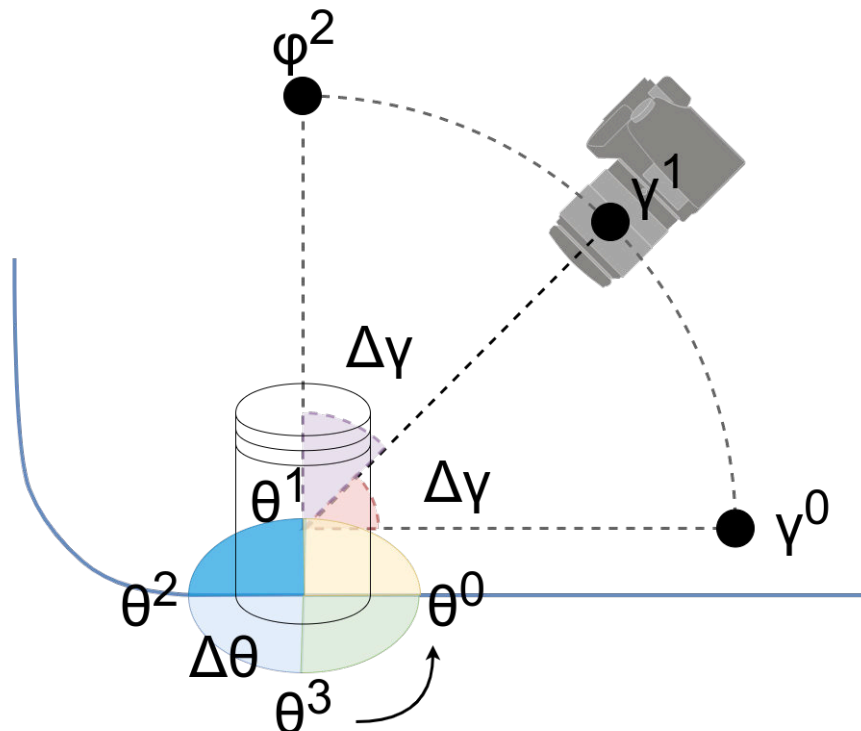
e Frames para o exemplo citado. Neste caso a aquisição teria 12 fotografias organizadas em 3 Streams de 4 Frames cada. A Figura 93 representa graficamente as posições de γ e Θ para $\gamma^0 = 0^\circ$, $\gamma^p = 90^\circ$, $\Delta\gamma = 45^\circ$ e $\Delta\Theta = 90^\circ$.

Tabela 4 – Relação de Frames e suas posições para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$

Foto #	Θ^{id}	Ângulo Θ	γ^{id}	Ângulo γ	Stream	Frame
# 1	Θ^0	0°	γ^0	0°	1	1
# 2	Θ^1	90°	γ^0	0°	1	2
# 3	Θ^2	180°	γ^0	0°	1	3
# 4	Θ^3	270°	γ^0	0°	1	4
# 5	Θ^0	0°	γ^1	45°	2	1
# 6	Θ^1	90°	γ^1	45°	2	2
# 7	Θ^2	180°	γ^1	45°	2	3
# 8	Θ^3	270°	γ^1	45°	2	4
# 9	Θ^0	0°	γ^2	90°	3	1
# 10	Θ^1	90°	γ^2	90°	3	2
# 11	Θ^2	180°	γ^2	90°	3	3
# 12	Θ^3	270°	γ^2	90°	3	4

Fonte: Autoria própria.

Figura 93 – Representação das posições angulares dos Frames para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$.



Fonte: Autoria própria.

Definidas as posições angulares de cada frame, é necessário calcular as coordenadas GMC para cada posição da câmera nos eixos X' , Y' e Z^a . Este processo pode ser feito usando trigonometria, a lista de ângulos de γ , o raio da aquisição e a posição do centro geométrico do objeto, o qual coincide com o centro da esfera formada durante a aquisição do F2S2. O processo de cálculo para cada valor de γ é definido por:

$$X' \approx 650 - (r \times \text{Cos}(\gamma)) \quad (5.1)$$

$$Y' \approx (b + \frac{h}{2}) + (r \times \text{Sen}(\gamma)) \quad (5.2)$$

$$Z^a \approx \gamma \quad (5.3)$$

Sendo:

- r : raio da aquisição.
- b : altura da base de suporte.
- h : altura do objeto.

Assim, os valores das coordenadas GMC para o exemplo citado, usando as Equações 5.1, 5.2 e 5.3 para calcular as coordenadas GMC para os eixos X' , Y' e Z^a , respectivamente, são apresentados na Tabela 5. Os valores são arredondados para um valor inteiro, uma vez que o F2S2 não realiza movimento de fração de milímetro. A Figura 94 apresenta as posições de cada Stream da aquisição com os valores angulares e as coordenadas GMC para cada Stream.

Tabela 5 – Coordenadas GMC para uma aquisição de $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$.

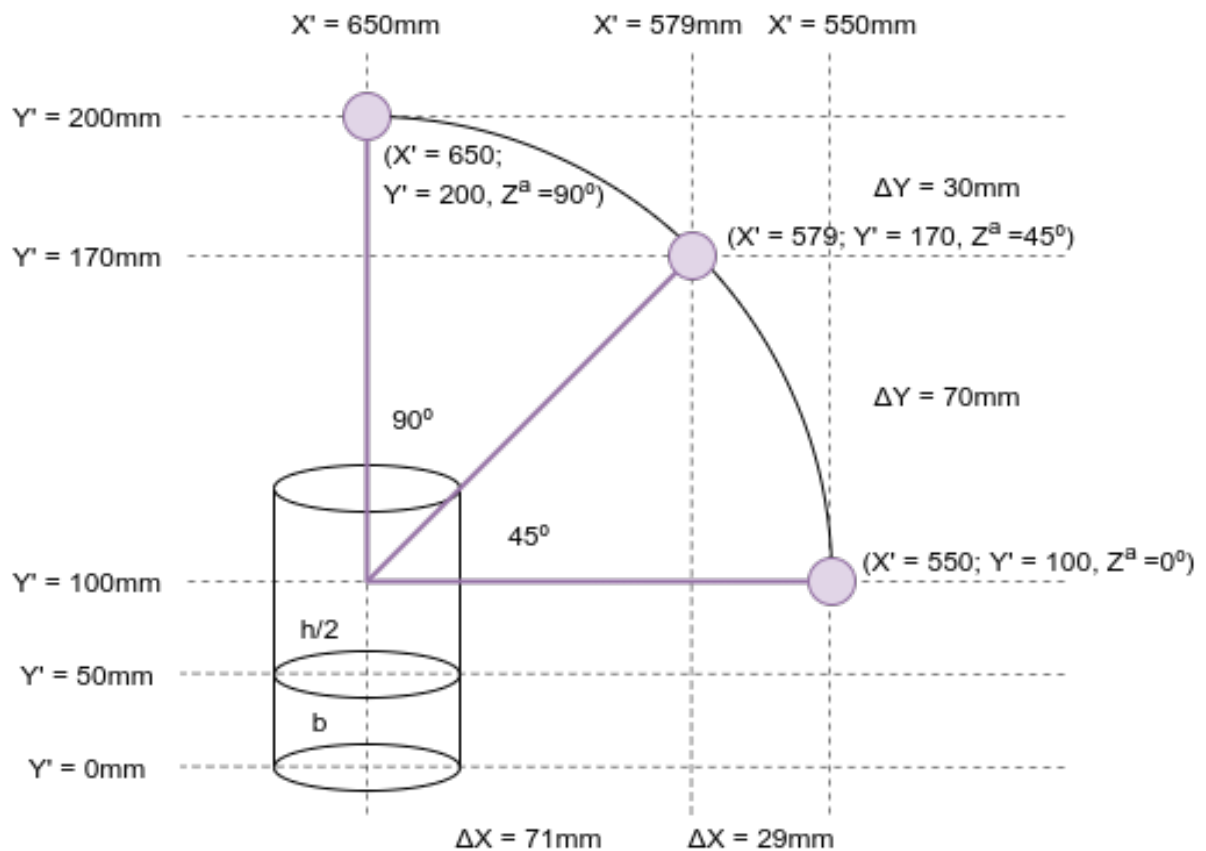
γ		X'	Y'	Z^a
γ^0	0°	550mm	100mm	0°
γ^1	45°	579mm	170mm	45°
γ^2	90°	650mm	200mm	90°

Fonte: Autoria própria.

Rotação do objeto

Após o mapeamento de todos os pontos do arco, é necessário definir as instruções de movimentação do eixo B^a e das fotografias entre os movimentos dos eixos X' e Y' . Para isso é usada a lista de posições do ângulo Θ . Então, a criação do arquivo com as instruções GMC baseia-se em posicionar a câmera no arco, inserir o comando de pausa entre streams para estabilizar a câmera e em seguida o conjunto de instruções de movimentação do eixo B^a ; pausa entre fotos; e aquisição da fotografia.

Figura 94 – Representação das posições angulares dos Streams para uma aquisição de $\Delta\Theta = 90^\circ$ e $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$.



Fonte: Autoria própria.

Código fonte da geração de instruções GMC

O Código 5.2 apresenta uma simplificação do código usado para a geração das instruções GMC. Neste exemplo o código está configurado para calcular as instruções para a aquisição usada como exemplo nesta Subseção, com: $\Delta\gamma = 45^\circ$; $\gamma^0 = 0^\circ$; $\gamma^p = 90^\circ$; $b = 50\text{mm}$; $h = 100\text{mm}$. Ou seja, 3 Streams com 4 Frames cada, totalizando 12 fotos. Estes valores estão em conformidade com as Tabelas 4 e 5; com as Equações 5.1, 5.2 e 5.3; e com as Figuras 93 e 94.

código 5.2 – Geração das instruções GMC.

```

1 from math import sin, cos, radians
2 import numpy as np
3
4 file = open('Aquisicao.GMC', 'w')
5
6 pause_stream = 4
7 pause_photo = 3
8 start_angle = 0
9 end_angle = 90
10 streams_number = 3

```

```

11 frames_number = 4
12 center_x = 650
13 center_y = 100
14 ray = 100
15 feedrate = 4000
16
17
18 def move():
19     x_axis = []
20     y_axis = []
21     z_axis = []
22
23     calc_arch(x_axis, y_axis, z_axis)
24
25     for x, y, z in zip(x_axis, y_axis, z_axis):
26         file.write('G00_X' + str(int(x))
27                 + '_Y' + str(int(y))
28                 + '_Z' + str(int(z)) + '\n')
29
30     calc_frames()
31
32
33 def calc_arch(x_axis, y_axis, z_axis):
34
35     z_axis += list(np.linspace(start_angle,
36                               end_angle,
37                               streams_number,
38                               endpoint=True))
39
40     for angle in z_axis:
41         x_axis.append(center_x - (cos(radians(angle)) * ray))
42         y_axis.append(center_y + (sin(radians(angle)) * ray))
43
44
45 def calc_frames():
46
47     file.write('G00_F' + str(int(feedrate/8)) + '\n')
48     file.write('M999\n')
49     file.write('G04_P' + str(pause_stream) + '\n')
50
51     for b in np.linspace(0.0,
52                          360.0,
53                          frames_number,
54                          endpoint=False):
55
56         if b != 0:
57             file.write('G00_B' + str(int(b)) + '\n')

```

```

58     file.write('G04_P' + str(pause_photo) + '\n')
59     file.write('M989_T' + str(pause_photo) + '\n')
60
61     file.write('G00_F' + str(feedrate) + '\n')
62
63
64 move()

```

Fonte: Autoria Própria.

No código as linhas 1 e 2 são importações das bibliotecas Math e NumPy com funções matemáticas. A Linha 4 realiza a abertura do arquivo onde serão salvas as instruções GMC. As linhas 6 a 15 apresentam as variáveis necessárias para a execução do código. No programa real estas variáveis são obtidas pela interface gráfica e verificadas para que não existam colisões durante a aquisição.

Seus significados são respectivamente: pausa entre Streams, pausa entre fotos, γ^0 , γ^p , número de Streams, número de Frames, centro do objeto no eixo X' , centro do objeto no eixo Y' , raio da aquisição e velocidade do movimento. Quanto maior o valor de *feedrade*, mais rápido é o movimento. Normalmente usa-se valores mais baixos para a movimentação do eixo B^a , pois este eixo move o objeto e para aumentar a estabilidade e portanto deve ser mais lento.

A Função da linha 18 calcula todo o movimento usando as outras Funções do código. As listas instanciadas nas linhas 19, 20 e 21 serão populadas com o conjunto de valores dos eixos X' , Y' e Z^a , respectivamente. Assim, o quarto elemento destas linhas indica a posição de cada eixo durante o quarto movimento da aquisição. Na linha 23 a função “move” executa a função “*calc_arch*” para calcular os movimentos para geração do arco para estes três eixos e popular estas listas.

A Função “*calc_arch*” recebe as três listas como argumento e na linha 35 popula a lista com os valores do eixo Z^a . Para isso é usada a função da biblioteca NumPy linspace, que divide um intervalo numérico linearmente em uma determinada quantidade de espaços iguais. Seus parâmetros são: γ^0 - início do intervalo e primeiro valor da sequência; γ^p - fim do intervalo e último valor da sequência (Caso *endpoint = True*); quantidade de Streams - ou quantidade de espaços para dividir o intervalo; *endpoint* - caso este valor seja verdadeiro o ultimo valor do intervalo, no caso γ^p , será incluído na sequência. No caso da geração do arco este parâmetro é verdadeiro. No caso os valores gerados são: 0, 45, 90.

Na linha 40, para cada valor angular de Z_{axis} é calculado: coordenada do eixo X' em conformidade com a Equação 5.1 - linha 41; e coordenadas do eixo Y' de acordo com a função 5.2 - linha 41. Neste exemplo os valores calculados são os apresentados na Tabela 5. Com as listas de valores dos eixos X' , Y' e Z^a calculados, a função “*calc_arch*” é finalizada e o fluxo de execução volta para a linha 25.

As linhas 25 a 28 geram as instruções de movimento para deslocar a câmera ao longo do arco calculado. No laço de repetição da linha 25, para cada valor “x”, “y” e “z” (ou seja cada posição no arco), é escrito no arquivo a instrução de movimentação “G00” com as coordenadas “x”, “y” e “z” do movimento.

Em seguida, para cada instrução escrita é chamada a função “*calc_frames*”, que calcula os movimentos do eixo B^a , ou seja os Frames. Esta função inicialmente diminui a velocidade do movimento para $\frac{1}{8}$ da velocidade dos movimentos dos outros eixos (instrução “G00 F” - linha 47); calibra o eixo B^a usando o sensor óptico (instrução “M999” - linha 48); aguarda o sistema estabilizar após o movimentos dos 4 eixos (instrução “G00 P” - linha 49).

Então, usando novamente a função *linspace* da biblioteca NumPy ele calcula a posição de cada Frame dentro dos 360° do eixo B^a , iniciando o intervalo em $\Theta = 0^\circ$ (linha 51); finalizando em $\Theta = 360^\circ$ (linha 52); dividindo pela quantidade de Frames desejados (linha 53); e não inserindo a última posição (*endpoint = False* - linha 54) na lista, pois $\Theta = 0^\circ$ é a mesma posição de $\Theta = 360^\circ$.

Então, para cada posição do eixo B^a são inseridas as seguintes instruções GMC: Caso não seja $\Theta = 0^\circ$ (linha 56) é insrido o comando de movimentação do eixo B^a (instrução “G00 B” - linha 57). Esta condição é necessária pois na linha 48 o eixo foi calibrado pelo sensor óptico e portanto o eixo já está na posição correta; aguarda a peça estabilizar após o movimento (instrução “G04 P” - linha 58); adquire a fotografia (instrução “M989 T” - linha 59). Por último, fora do laço de repetição, a velocidade do movimento é acelerada (instrução “G00 F” - linha 61). Assim finaliza o código de geração das intruções GMC. O Código 5.3 apresenta as intruções GMC geradas pelo Código 5.2.

código 5.3 – Códigos GMC gerados.

```

1 G00 X550 Y100 Z0
2 G00 F500
3 M999
4 G04 P4
5 G04 P3
6 M989 T3
7 G00 B90
8 G04 P3
9 M989 T3
10 G00 B180
11 G04 P3
12 M989 T3
13 G00 B270
14 G04 P3
15 M989 T3
16 G00 F4000
17 G00 X579 Y170 Z45
18 G00 F500

```


19	M999
20	G04 P4
21	G04 P3
22	M989 T3
23	G00 B90
24	G04 P3
25	M989 T3
26	G00 B180
27	G04 P3
28	M989 T3
29	G00 B270
30	G04 P3
31	M989 T3
32	G00 F4000
33	G00 X650 Y200 Z90
34	G00 F500
35	M999
36	G04 P4
37	G04 P3
38	M989 T3
39	G00 B90
40	G04 P3
41	M989 T3
42	G00 B180
43	G04 P3
44	M989 T3
45	G00 B270
46	G04 P3
47	M989 T3
48	G00 F4000

Fonte: Autoria Própria.

Dificuldades na execução das instruções

Considerando que o F2S2 não realiza movimentos de fração de milímetro, para aquisições com raios muito pequenos como $20mm$, o arredondamento dos valores pode ser significativo. No exemplo desta Subseção os valores calculados para o movimento foram: $P1 = (X' = 550mm, Y' = 100mm)$, $P2 = (X' = 579.2893218813452mm, Y' = 170.71067811865476mm)$ e $P3 = (X' = 650mm, Y' = 200mm)$. Em P2, os dois valores foram arredondados para $P2 = (X' = 579mm, Y' = 170mm)$. A diferença causada pelo arredondamento pode não ser significativa e perceptível em aquisições com $10cm$ de raio, porém se tornam críticas para aquisições com $20mm$ de raio.

Além disto, o arredondamento, seja excluindo a parte decimal, seja arredondando a cada $0,5mm$, podem causar saltos durante as aquisições. Considere uma aquisição com $\gamma^0 = 0^\circ$ e

$\gamma^p = 90^\circ$ com 61 Streams. Isso resulta em 1.5° entre cada Stream. Assim, os valores angulares calculados, arredondados e o salto entre Streams seriam os vistos na Tabela 6:

Tabela 6 – Valores arredondados para o eixo Z^a .

Valor integral calculado - Z^a	Valor arredondado da instrução - Z^a	$\Delta\gamma$ executado
0.0°	0°	—
1.5°	1°	1°
3.0°	3°	2°
4.5°	4°	1°
6.0°	6°	2°

Fonte: Autoria própria.

Neste caso, o erro acumulado pelo arredondamento causa saltos intercalados de 1° e 2° . Esta diferença pode ser perceptível durante a visualização. A solução encontrada para esta característica foi permitir apenas a seleção de valores para $\Delta\gamma$ e $\Delta\Theta$ nos quais a divisão pelos respectivos intervalos angulares ($\gamma^p - \gamma^0$ no caso de $\Delta\gamma$; 360° no caso de $\Delta\Theta$) seja inteira.

5.2 Software de Processamento - F2S2-BipApp

Esta Seção apresenta o software de processamento das imagens capturadas na etapa de aquisição. Considerando que todo o processamento deve acontecer sobre lote de imagens, o nome escolhido para o software foi *F2S2-BipApp* (*Batch Image Processing App* - Aplicativo de Processamento de Imagens em Lote). Os dados de entrada deste software são as imagens originadas na aquisição, em formato *RAW* ou *JPG*, corrigidas ou não; ao final do processamento é gerado o Stream2D, pronto para que seja feita a visualização. A Figura 95 apresenta a tela inicial do software.

O Software foi desenvolvido em *Python v3.7*, usando as boas práticas de programação em Python - PEP8⁸, visando melhor manutenibilidade. Também foram usados os princípios do livro *Código Limpo* (MARTIN, 2009) para melhorar a organização e legibilidade do código, minimizando a ocorrência de problemas. As dependências do código foram empacotadas dentro de *container* da plataforma Conda⁹, para evitar conflitos de bibliotecas e oferecer uma instalação simples. A interface gráfica foi desenvolvida na plataforma QT¹⁰.

Para manipulação de arquivos de imagem foi usado o OpenCV¹¹ para Python e para algoritmos de matemática, classificação e clusterização foram usadas as bibliotecas: NumPy¹²,

⁸ <<https://www.python.org/dev/peps/pep-0008/>> Acessado em 17 de Abril de 2020.

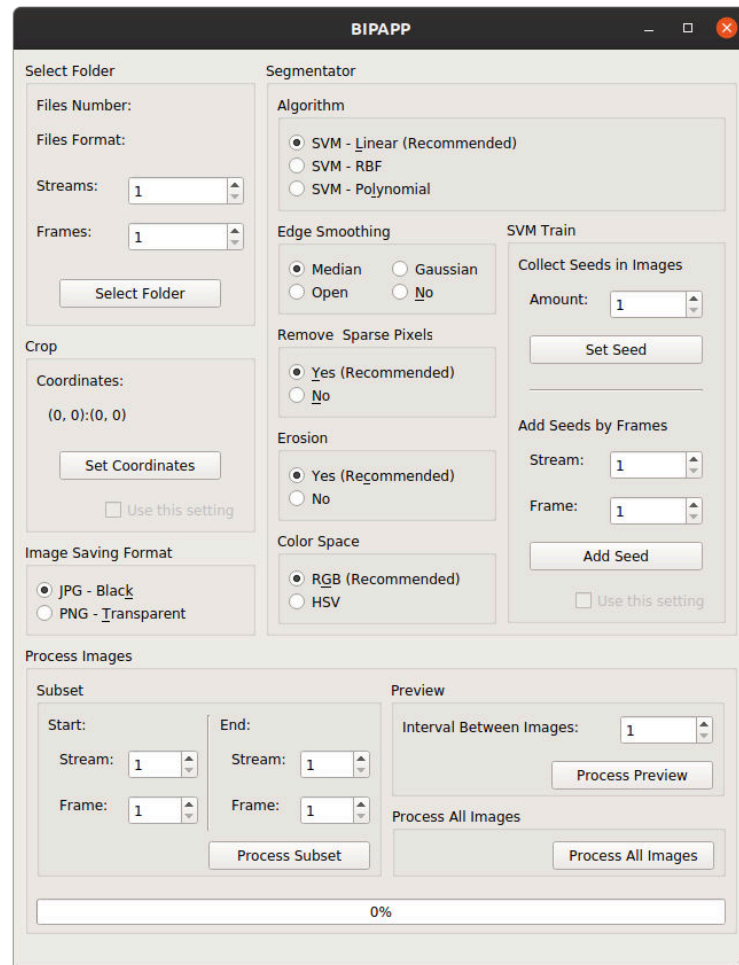
⁹ <<https://www.anaconda.com/distribution/>> Acessado em 24 de Abril de 2020.

¹⁰ <<https://www.qt.io/>> Acessado em 13 de Maio de 2020.

¹¹ <<https://opencv.org/>> Acessado em 13 de Maio de 2020.

¹² <<https://numpy.org/>> Acessado em 13 de Maio de 2020.

Figura 95 – Tela inicial do F2S2-BipApp.



Fonte: Autoria própria.

SciPy¹³ e scikit-learn¹⁴. Um ponto crítico no desenvolvimento foi a manipulação de arquivos *RAW*, principalmente arquivos da câmera Samsung NX500, que por não ser muito popular não podia ser manipulado na maioria das bibliotecas de manipulação de imagens, como o OpenCV. O problema foi solucionado com as bibliotecas LibRaw¹⁵ e rawpy¹⁶. Para o desenvolvimento do software foram usados diversos padrões de projetos (GAMMA, 2009):

Command : Usado para encapsular todas as informações da interface gráfica de forma que ao passar o comando para cada módulo do aplicativo, o módulo conhecesse todo o conjunto de operações que deveriam ser feitas.

Chain of responsibility : Como a ordem das operações de processamento sobre cada imagem é importante, foi criada uma cadeia de responsabilidades, de forma que independentemente

¹³ <<https://www.scipy.org/>> Acessado em 13 de Maio de 2020.

¹⁴ <<https://scikit-learn.org/stable/>> Acessado em 13 de Maio de 2020.

¹⁵ <<https://www.libraw.org/>> Acessado em 13 de Maio de 2020.

¹⁶ <<https://github.com/letmaik/rawpy>> Acessado em 13 de Maio de 2020.

do lançamento de Threads ou ordem de leitura das imagens, todas seriam processadas pelos algoritmos ordenados da mesma forma. O comando gerado a partir dos dados da interface gráfica é enviado para o primeiro nó da cadeia e então cada nó da cadeia ordenada de operações decide as operações que devem ser realizadas.

Factory : Considerando que a entrada de dados é um conjunto de imagens que podem ser de diversos formatos (*JPG* ou *RAW*: “.srw” para Samsung, “.nef” para Nikon, “.cr2” para Canon, “.dng” para Adobe, ...), foi implementada uma fábrica de imagens que decide quais bibliotecas devem ser usadas na abertura e manipulação dos arquivos.

Builder : Ao final do processamento, para salvar cada Frame do Stream2D, foi usado um construtor que recebe informações como: nome do Frame no Stream2D; imagem processada (ainda em memória); nome do arquivo original no disco; informações do Stream2D e então o construtor salva cada Frame em disco, construindo o Stream2D.

A última versão do software possui 4.244 linhas (7.911 inserções e 3.667 remoções), 18 *commits* segundo o gerenciador de versões usado. No que se refere ao consumo de recursos, o software durante o processamento usa aproximadamente 500Mb de memória, dependendo do tamanho e formato das imagens de entrada. O tempo de execução flutua bastante, dependendo da quantidade, tamanho, formato e opções de processamento.

Durante o desenvolvimento do software foi definido informalmente um requisito de que o tempo de processamento das imagens não deveria ultrapassar o tempo de aquisição destas. Isto permitiria o processamento das imagens durante o processo de aquisição, de forma que ao terminar a aquisição o Stream2D estaria pronto. Porém, como não foi possível implementar um sistema de acesso às fotografias durante a aquisição (a câmera usada não consegue disponibilizar as imagens durante a aquisição), não foi possível realizar o processamento simultaneamente à aquisição.

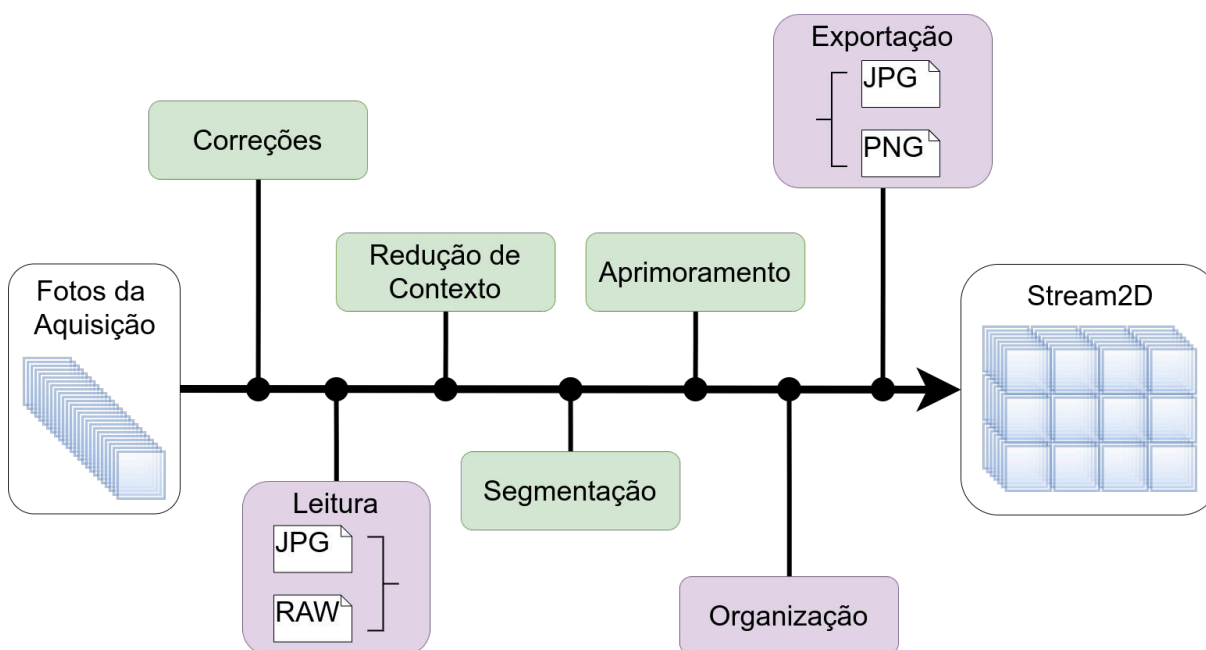
Uma aquisição com 31 Streams e 120 Frames (3720 Fotos) que no formato *RAW* ocupa aproximadamente 100Gb, gasta em torno de 7 horas para a aquisição e 6 horas de processamento, considerando um computador com processador *i7*, 16Gb de RAM e disco rígido sata.

Além disto, analisando o comportamento do software no sistema *\proc* do linux, principalmente o número de trocas de contexto voluntárias (aguardando algum recurso) e involuntárias (durante um processamento) no processador, é possível deduzir que grande parte do tempo de processamento é gasto com leituras do disco. Assim, pode-se supor que o tempo de acesso à disco é crítico para esta aplicação e o uso de um sistema de armazenamento com menor tempo de leitura, como um *SSD-M2* reduziria significativamente o tempo de processamento.

5.2.1 Etapas

Como já mencionado, o processamento ocorre após as etapas de aquisição. Então é iniciado o processamento com as correções e o *F2S2-BipApp*. O processamento pode ser dividido em etapas que precisam seguir a ordem apresentada na Figura 96, sendo que as etapas em verde são opcionais e em roxo obrigatórias para gerar o Stream2D.

Figura 96 – Diagrama das etapas de processamento.



Fonte: Autoria própria.

Correções

A etapa de correções, início do processamento, recebe como entrada o conjunto de fotografias tiradas na aquisição e, após sua execução, retorna o as imagens corrigidas para que possa ser realizado o processamento das imagens pelo *F2S2-BipApp*. Esta etapa usa softwares de terceiros, disponibilizados pelo fabricante dos hardwares usados na etapa de aquisição.

O processo de fotografia, digital ou analógica, possui características intrínsecas do equipamento usado para a captura das imagens; as cores adquiridas são similares às cores verdadeiras da cena, porém não são idênticas; o uso de lentes acaba por deformar as imagens obtidas e gerar aberrações cromáticas; o formato cilíndrico do conjunto óptico escurece a extremidade das imagens.

Considerando que estas características são originadas pelo hardware das câmeras, alguns fabricantes oferecem softwares capazes de minimizar estes efeitos. Considerando a preexistência destes softwares, optou-se pelo seu uso no processo de correção das imagens, em vez do desenvolvimento de software próprio.

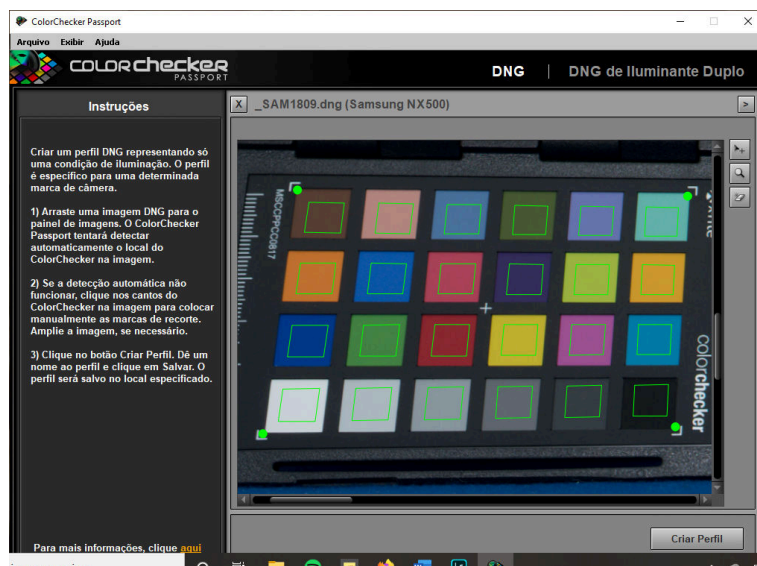
Estes softwares de terceiros são plugins ou tabelas de dados que podem ser usados por softwares de edição de fotografias, como o Lightroom¹⁷, Darktable¹⁸ ou o RawTherapee¹⁹. Nestes softwares é possível realizar as correções aqui citadas em lotes de imagens de forma automática, deixando os arquivos prontos para continuar o processamento no software de processamento do F2S2.

Estas etapas são opcionais e dependem do objetivo da visualização. Aquisições para publicidade ou para uso no ensino médio possivelmente não necessitam destas correções, porém aquisições para uso científico, como identificação de espécies ou doenças, podem necessitar das correções aqui citadas.

Cor – ColorChecker

A aquisição de cores pelos sensores das câmeras não geram representações perfeitas das cores. Este processo é sensível à iluminação da fotografia, à configuração da câmera e ao sensor. Para melhorar a representação das cores é usado o ColorChecker e o software do *ColorChecker Passport Photo 2*²⁰ Este software (Fotografia 97) reconhece o colorchecker em uma fotografia carregada; com base nas cores representadas na imagem do ColorChecker cria uma matriz de correção das cores da fotografia. Este padrão de correção pode ser usado para outras fotografias da mesma máquina, com a mesma configuração e mesmo padrão de iluminação.

Figura 97 – Software ColorChecker Passport.



Fonte: Autoria própria.

¹⁷ <<https://lightroom.adobe.com/>> Acessado em 10 de Maio de 2020.

¹⁸ <<https://www.darktable.org/>> Acessado em 10 de Maio de 2020.

¹⁹ <<http://rawtherapee.com/>> Acessado em 10 de Maio de 2020.

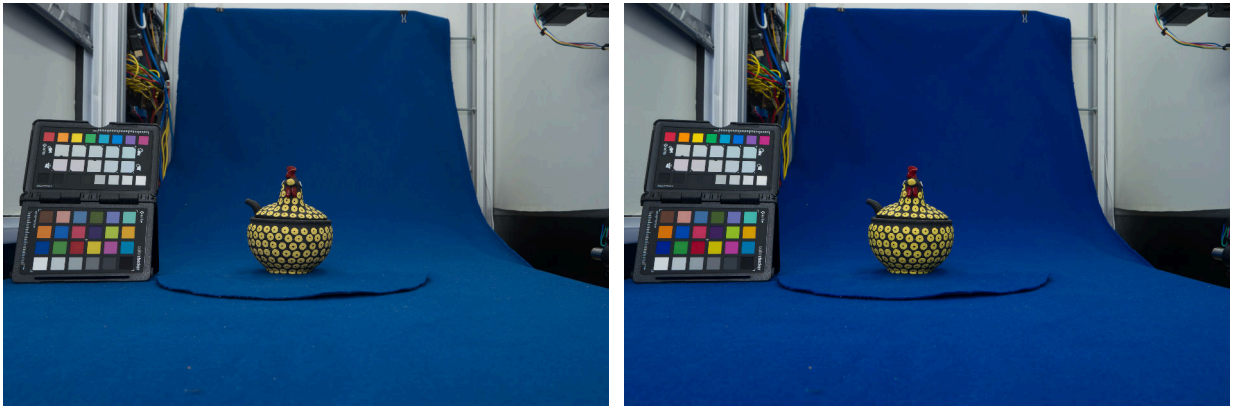
²⁰ <<https://www.xrite.com/categories/calibration-profiling/colorchecker-targets/colorchecker-passport-photo-2/>> Acessado em 16 de Abril de 2020.

A Figura 98 apresenta a fotografia sem a correção de cores (Figura 98(a)) e com a correção de cores criada pelo colorchecker passport (Figura 98(b)). Dependendo da qualidade da impressão/tela em que este trabalho é lido, a diferença de cores pode não ser perceptível - recomenda-se visualizar em uma tela também calibrada.

Figura 98 – Fotografia com correção de cores.

(a) Padrão de cores sem correção.

(b) Padrão de cores corrigido pelo ColorChecker.



Fonte: Autoria própria.

Distorção causada pelo conjunto óptico

O conjunto ótico das câmeras deforma as imagens adquiridas, de forma que o centro das fotografias tende a ser representado em um tamanho menor que as bordas da imagem. Para minimizar este problema os fabricantes de lentes e câmeras oferecem tabelas de correção para cada modelo de lente. A Figura 99 apresenta uma fotografia que apresenta a distorção e uma ampliação da mesma para evidenciar o problema. A Figura 100 apresenta o software DarkTable corrigindo uma imagem com a tabela de correções oferecida pela Samsung.

Aberração cromática

A aberração cromática ocorre pela diferença de refração de luzes de cores diferentes ao atravessar o conjunto ótico. A Figura 101(a) apresenta uma fotografia com o problema e a Figura 101(b) apresenta sua ampliação para evidenciar o problema. A solução também é oferecida pelos fabricantes de câmeras e pode ser feita nos softwares de processamento de fotografias, como visto na Figura 100, onde a aberração cromática aparece na opção TCA (TCA - *Transversal Chromatic Aberrations* - Aberração Cromática Transeversal).

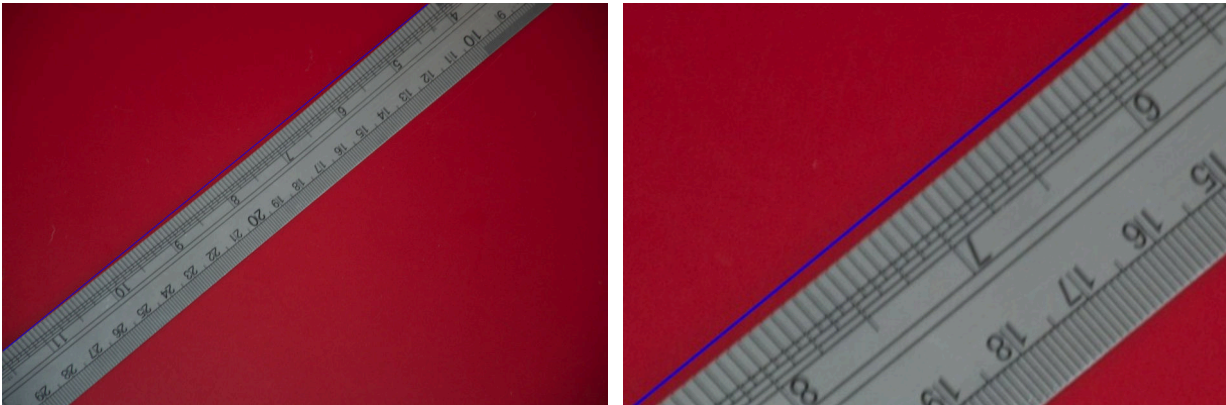
Vinheta

Nas bordas das imagens formam-se uma vinheta, pois o sensor não recebe luz da mesma forma no centro e nas bordas. Esta característica também afeta a fidelidade de cor e textura das

Figura 99 – Distorção causada pelo conjunto óptico.

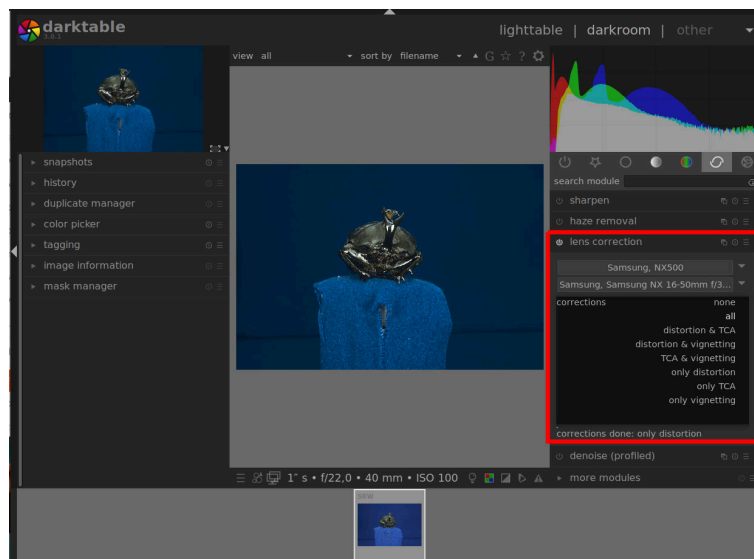
(a) Figura com uma linha reta em azul.

(b) Ampliação.



Fonte: Autoria própria.

Figura 100 – Correção da distorção, vinheta e aberração cromática usando a tabela de correções oferecida pela Samsung no software DarkTable.



Fonte: Autoria própria.

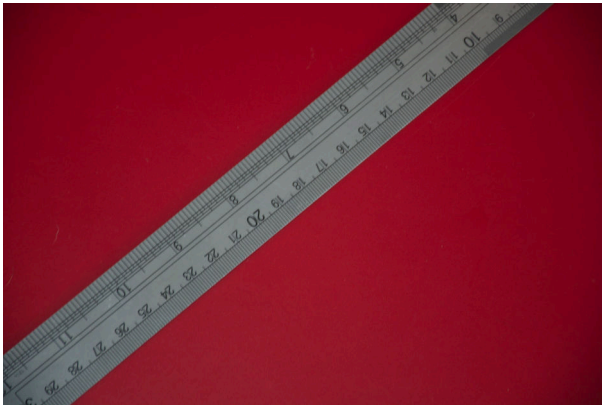
imagens adquiridas. O fenômeno da vinheta pode ser visto na Figura 102, na qual a imagem 102(a) é a original e a imagem 102(b) a corrigida. O processo de correção também consiste em tabelas fornecidas pelos fabricantes.

Considerações

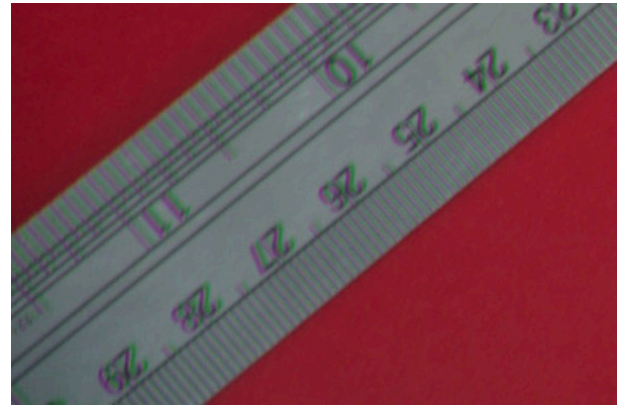
As correções citadas usam softwares dos fabricantes do hardware que necessitam de informações referentes ao modelo da câmera, lentes usadas e a configuração da câmera durante a aquisição. Imagens *JPG* perdem estas informações. Apesar de existirem alternativas para corrigir imagens *JPG*, recomenda-se o uso de imagens *RAW* nesta etapa, além de usar os softwares

Figura 101 – Aberração cromática causada pelo conjunto óptico.

(a) Imagem original.



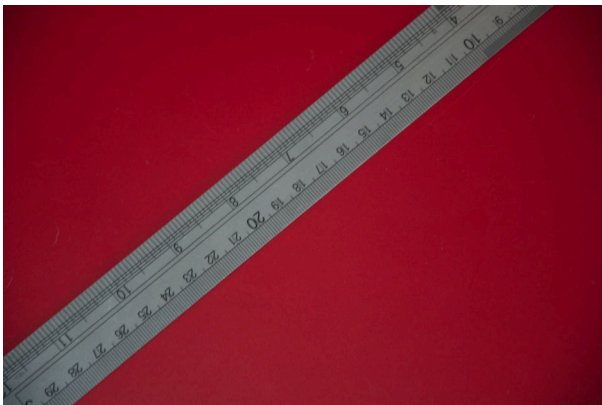
(b) Ampliação.



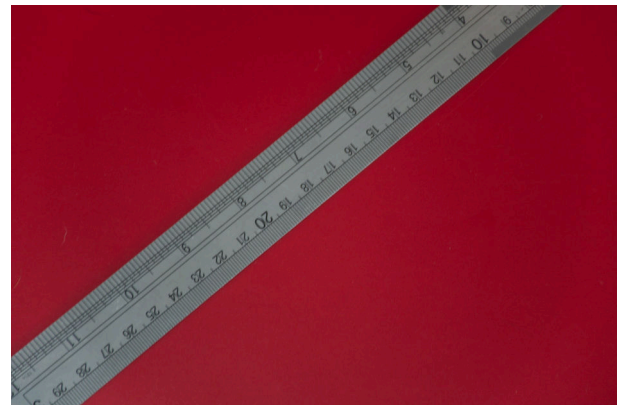
Fonte: Autoria própria.

Figura 102 – Vinheta causada pelo conjunto óptico.

(a) Foto original.



(b) Foto corrigida.



Fonte: Autoria própria.

oficiais dos fabricantes do hardware.

Modos de operação

Antes de apresentar os algoritmos de processamento do *F2S2-BipApp* é necessário apresentar os modos de operação do software. Isto é necessário pois as etapas de processamento usam os modos de operação para verificar se os parâmetros informados na interface gráfica geram um processamento satisfatório ou não, permitindo reconfigurar o software para parâmetros melhores.

Considerando que o *BipApp* processa conjuntos de imagens e que o processamento pode levar horas. Foi necessário desenvolver uma estratégia que evitasse esperar horas de processamento para descobrir que qualquer parâmetro configurado de forma incorreta inviabilizasse o uso das imagens. A estratégia consiste em oferecer ao usuário 3 formas de processamento das

imagens:

Preview : Processa uma imagens para um subconjunto de imagens. Por exemplo, caso o usuário configure o intervalo entre imagens para 100, significa que as imagens 1, 101, 201, ... serão processadas. Este modo de processamento verifica como o *BipApp* se comportara ao longo de todo o conjunto de imagens, pois as imagens processadas estão distribuídas linearmente no *pack* de imagens.

Subset : Processa um subconjunto de imagens, por exemplo: do Stream 2, Frame 1 ao Stream 4 Frame 10. Este modo de operação serve para verificar problemas locais. Considere que o usuário digitalizou um objeto com o topo cromado e não sabe como a segmentação irá se comportar por causa do brilho da peça. Neste caso o usuário pode processar apenas os últimos Streams e verificar se a configuração do algoritmo de segmentação é aceitável ou se é necessário reconfigura-lo.

Full : Neste caso todo o *pack* de imagens é processado.

Assim o usuário do *BipApp* pode testar as configurações nos modos *preview* ou *subset* e ajusta-las até obter uma configuração considerada aceitável. Então o usuário usa o modo *full* para processar o lote completo de imagens. Esta estratégia é importante, pois é difícil para o usuário prever como irão se comportar os algoritmos perante todo o conjunto de configurações que são necessárias para o processamento das imagens. As opções de modos de operação estão no quadro “*Process Images*”, no qual também existe uma barra de progresso que informa a porcentagem de imagens já processadas (Figura 95).

Leitura das imagens

A complexidade da leitura das imagens se dá pelo fato de existirem diversos formatos de arquivos *RAW* que precisam ser processados. Os formatos de imagens baseados em mapas de pixels como *JPG* e *PNG* são facilmente manipulados pela maioria das bibliotecas de processamento de imagens. Porém arquivos *RAW* possuem diversas especificidades, dependendo do fabricante da câmera ou software que gerou o arquivo. Além disto, para poder processar todas as imagens pelo mesmo algoritmo é necessário converter todos os arquivos de entrada do *F2S2-BipApp* para um formato único.

Para realizar a leitura destes diversos tipos de arquivos foi usado o padrão de projetos *Simple Factory*, que cria objetos da classe *imagem*. A super classe *imagem* é o padrão de imagem que será processado pelo *F2S2-BipApp* e usa um objeto da biblioteca OpenCV para armazenar as imagens lidas. Então, ao solicitar à fabrica a leitura de um arquivo (*RAW* ou *JPG*) a fábrica analisa o arquivo e cria objetos das sub classes *imagem_map* ou *imagem_raw*.

Caso o arquivo a ser lido seja um mapa de pixels, usa-se classe *imagem_map* com a biblioteca OpenCV para abrir o arquivo que já está pronto para ser usado. Caso o arquivo a ser

lido seja *RAW*, instancia-se um objeto da classe *imagem_raw*, o qual usa a biblioteca *rawpy* para abrir o arquivo; extrair as informações como abertura, ISO e distância focal da imagem; usa a biblioteca *NumPy* para converter o arquivo *RAW* para um vetor *NumPy* que depois é convertido para o formato de imagens usado pelo *OpenCV*. A partir deste momento todas as funcionalidades da classe *imagem* podem ser usadas para qualquer que seja o arquivo que tenha sido lido.

No *F2S2-BipApp* a abertura dos arquivos é realizada no quadro *Select Folder*, de forma que ao clicar no botão *Select Folder* abre-se uma caixa de diálogo na qual o usuário seleciona um diretório com as imagens, o software então verifica quantos arquivos de imagens existem no diretório e o formato dos arquivos. O Usuário deve então informar o número de Frames e Streams que irão compor o Stream2D, estes valores são conferidos com o total de imagens lidas e caso o valor não esteja correto, o aplicativo não libera as outras funcionalidades e exibe uma mensagem de erro.

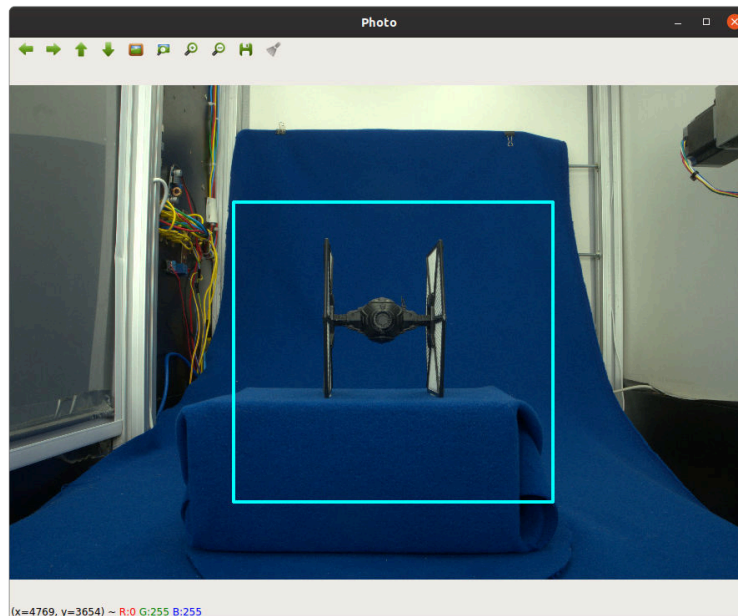
Redução de Contexto

Como já mencionado, a redução de contexto remove partes das imagens que não representam o objeto de interesse. Este processo possui as mesmas dificuldades do processo de enquadramento (Seção 4.3.2), uma vez que é definida uma região que será removida de todas as imagens. Assim, é importante selecionar uma região que ao longo de todo o *pack* de imagens, o objeto de interesse seja preservado integralmente.

Para usar a redução de contexto no *F2S2-BipApp* usa-se o quadro *Crop* (Figura 95), quando pressionado o botão *Set Coordinates* abre-se uma caixa com a primeira imagem da sequência e o usuário usa o mouse para desenhar um quadrado nesta imagem. A região selecionada pelo usuário, será a região que permanecerá em todas as imagens, de forma que as áreas fora deste quadrado serão suprimidas de todas as imagens. Caso deseje, o usuário pode desenhar quadrado, ao finalizar usa-se a tecla “F” para setar as coordenadas da região que será usada dos Frames. A figura 103 apresenta a tela de definição da região.

Esta etapa é opcional, de forma que se não for selecionada nenhuma região, a imagem toda será processada. Como mencionado na Subseção 5.2.1, pode-se avaliar a região selecionada processando apenas subconjunto de imagens do *pack* da aquisição. Isto evita retrabalho, uma vez que o processamento pode levar horas. Para verificar se a região selecionada não está cortando partes do objeto em nenhum Frame, processa-se apenas um subconjunto de imagens e então verifica-se se o objeto permaneceu íntegro, para então prosseguir com o processamento.

Cabe ressaltar que as mesmas dificuldades encontradas no enquadramento da fotografia, como visto na Subseção 4.3.2, se aplicam à redução de contexto, por este motivo recomenda-se sempre testar se a região selecionada para a redução de contexto é adequada.

Figura 103 – *F2S2-BipApp* - Redução de contexto.

Fonte: Autoria própria.

Segmentação

O algoritmo selecionado para fazer a segmentação foi o SVM, o qual é treinado para prever se determinado pixel é de uma das duas classes: Fundo; Objeto. Se for classificado como Objeto o pixel permanece inalterado na imagem, se for classificado como Fundo, o pixel possui sua cor alterada para preto ou transparente. A Figura 104 apresenta a fotografia adquirida e segmentada de um cacto.

Figura 104 – Processo de segmentação.

(a) Figura original.



(b) Figura segmentada e com contexto reduzido.



Fonte: Autoria própria.

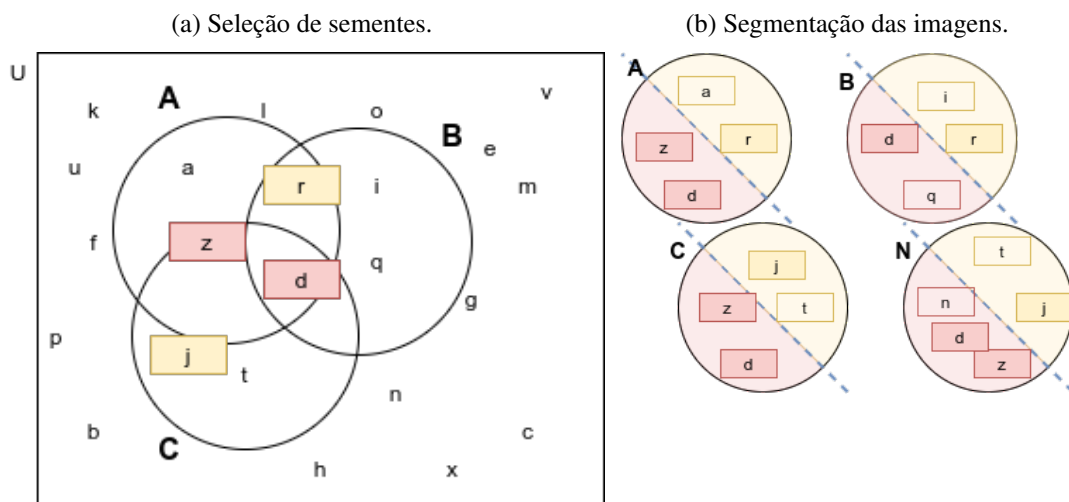
No *F2S2-BipApp* o usuário deve escolher um kernel para o SVM, que pode ser: linear, gaussiano ou polinomial. Além disto as imagens podem ser processadas nos espaços de cores

RGB ou *HSV*. O processo de segmentação então se inicia ao receber dois conjuntos de pixels das imagens a serem processadas: um conjunto representa pontos do fundo infinito do scanner (classe Fundo) e o outro conjunto representa os pixels do objeto (classe Objeto). Com estes dois conjuntos selecionados o SVM é treinado.

Considerando que o conjunto de todas as cores presentes em um pack de imagens, ao selecionar alguns Frames para coletar sementes e algumas sementes destes frames; o SVM deve ser treinado de forma que para todos os frames, o hiperplano gerado pelo SVM deve separar cada cor de pixel como sendo da classe Fundo ou Objeto.

Este processo pode ser representado pela Figura 105, onde o conjunto universo representa todas as cores presentes em pixels do *pack* de imagens, as cores são representados pelos elementos nomeados com letras (Figura 105(a)). Os conjuntos “A”, “B” e “C” representam os Frames escolhidos para se coletar as sementes “r” e “j” (classe Objeto) e “d” e “z” (classe Fundo). O treinamento do SVM determina um hiperplano (representado pela linha azul - Figura 105(b)) que deve separar, para cada frame do Stream2D, os pixels do Fundo e do Objeto.

Figura 105 – Processo de Segmentação.



Fonte: Autoria própria.

A segmentação das imagens ocorre lendo cada pixel de cada imagem do *pack* de imagens e então predizendo (com o SVM) se este pixel é parte do fundo infinito ou do objeto. Para otimizar o processamento e o uso de memória é criada uma máscara binária que informa se determinado pixel é fundo ou objeto. Esta máscara também será usada nas etapas de aprimoramento das imagens, evitando ter que trabalhar com a imagem completa, as quais necessitam de mais recursos computacionais para serem processadas. Também para otimizar a segmentação, caso o usuário tenha optado por usar a redução de contexto, só são processados pixels internos à região selecionada pelo usuário.

Ao final do processamento a máscara que define se cada pixel é fundo ou objeto é usada

para construir a imagem definitiva do Stream2D. Caso o pixels seja classificado como sendo parte do objeto (classe Objeto) seu pixel permanece inalterado na imagem final; caso o pixel seja classificado como sendo parte do fundo infinito (classe Fundo), seu valor é alterado para preto (caso o formato de salvamento seja *JPG*) ou transparente (caso o formato de salvamento seja *PNG*).

Considerando que o *BipApp* realiza o processamento de um conjunto de imagens e que pode levar horas, foi necessário estabelecer uma estratégia na interface gráfica para evitar que ao final do processamento perceba-se que a seleção das sementes das classes Fundo e Objeto não foi satisfatória e será necessário uma nova segmentação. A estratégia consiste na seleção interativa de sementes e no processamento parcial de imagens.

Seleção interativa de sementes

A seleção interativa de sementes se baseia no processo de: inicialmente selecionar um conjunto de sementes para as classes Objeto e Fundo - testar a segmentação em um subconjunto de Frames - analisar se o resultado está satisfatório; Caso o resultado não esteja adequado é possível adicionar novas sementes de quaisquer Frames que se deseje, então testar e analisar novamente o resultado, até que se obtenha um resultado satisfatório.

Este processo não é um simples processo de tentativa e erro. É necessário ajustar as sementes pois os objetos podem apresentar determinadas características apenas em algumas regiões e portanto é necessário que sementes destas cores sejam adicionadas ao conjunto de sementes do Objeto. O mesmo processo pode ser necessário para o fundo, uma vez que como a câmera se move durante a aquisição, pode fotografar o fundo com tonalidades diferentes.

Para exemplificar este processo considere um cubo onde cada face possui uma cor diferente. Ao selecionar sementes de apenas algumas faces, a segmentação pode não ser satisfatória. Assim é necessário selecionar outros Frames, que representem as outras faces, para conseguir sementes que representem todo o conjunto de cores do objeto.

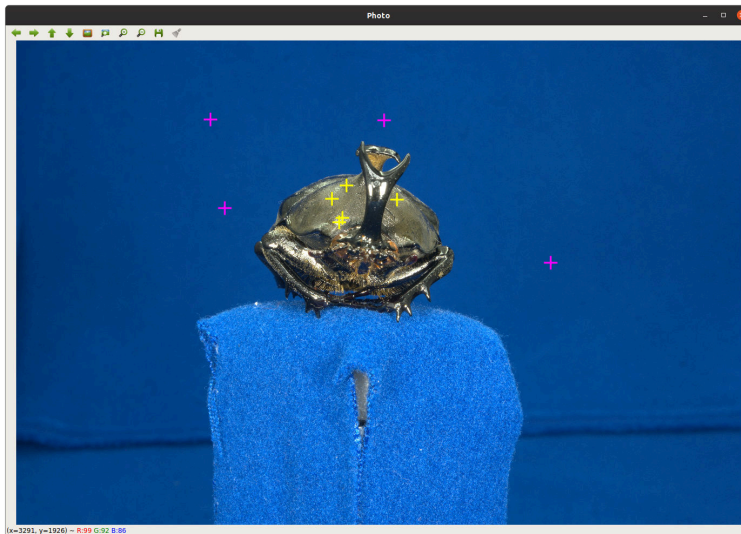
Isto significa distribuir os Frames e cores selecionadas como sementes de forma que estes representem adequadamente o objeto e fundo. Considerando a Figura 105, se o elemento “v” não foi selecionado como semente, pois não está presente em nenhum dos frames “A”, “B” e “C”, pode ser que o hiperplano gerado pelo SVM classifique incorretamente a cor.

A Figura 106 apresenta o processo de seleção interativa de sementes no *BipApp*. A Figura 106(a) apresenta um Frame com 9 sementes selecionadas (4 da classe Fundo em rosa e 5 da classe Objeto em amarelo). A Figura 106(b) apresenta o resultado da segmentação com o SVM treinado com estas 9 sementes. Considerando que a segmentação não foi adequada, adicionou-se 131 sementes (Figura 106(c)), de forma que o resultado desta segmentação foi suficiente (Figura 106(d)).

Seleção de frames para escolha de sementes ocorre no quadro *SVM Train* e pode ser

Figura 106 – Seleção interativa de sementes.

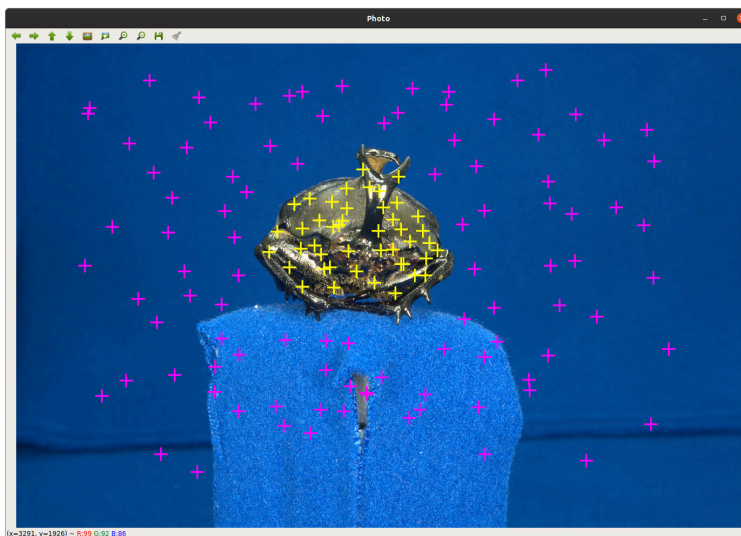
(a) Seleção inicial - 9 sementes.



(b) Resultado da segmentação com 9 sementes.



(c) Seleção final - 142 sementes.



(d) Resultado da segmentação com 142 sementes.



Fonte: Autoria própria.

feita escolhendo uma quantidade de Frames linearmente distribuídas no *pack* de imagens ou indicando especificamente o Frame e Stream que se deseja utilizar.

A quantidade de sementes necessárias para segmentar adequadamente um *pack* de imagens vai depender de diversos fatores, como: complexidade do objeto, profundidade de campo das fotografias, reflexos e transparências. Considerando a dificuldade de determinar a quantidade de sementes adequadas para a segmentação, a seleção interativa permite ao usuário ajustar a quantidade de sementes e principalmente distribuí-las em regiões em que a segmentação não esteja acontecendo corretamente, de forma que a segmentação ocorra adequadamente.

Dificuldades na segmentação

As principais dificuldades de segmentação se originam de duas fontes: fotográfica: quando a profundidade de campo não permite a nitidez nas bordas das imagens; do objeto: quando existem reflexos, transparências e pelos.

A segmentação tem dificuldades para remover uma região de pixels que não representem o fundo ou objeto com o mesmo padrão. Por isso quando a peça apresenta brilho ou transparência, altera a cor do objeto ou fundo, burlando o segmentador. Este processo pode ser visto na Figura 107. A Figura 107(c) apresenta regiões do fundo que permaneceram na imagem pois o brilho alterou a cor da região e a Figura 107(d) apresenta a segmentação de uma região com transparência, na qual existem brilhos, cores do fundo e do objeto.

Figura 107 – Dificuldades na segmentação - transparência e brilho.

(a) Foto original com transparência e brilho.



(b) Foto segmentada.



(c) Foto segmentada e ampliada - brilho.



(d) Foto segmentada e ampliada - transparência.



Fonte: Autoria própria.

A Figura 108 apresenta regiões sem foco e com pelos. Nestes casos, por falta de nitidez, uma região de pixels não apresenta a cor do fundo ou do objeto, fazendo com que o segmentador não classifique corretamente estes pixels. A Figura 108(c) apresenta uma borda azulada, pois aqueles pixels não possuem a cor da abelha, nem do fundo.

Figura 108 – Dificuldades na segmentação - nitidez.



Fonte: Autoria própria.

Aprimoramento

Para melhorar as imagens para a visualização foram selecionadas e implementadas algumas técnicas de processamento de imagens para aprimorar as fotografias segmentadas. Estas técnicas só podem ser aplicadas em imagens segmentadas, uma vez que usam a máscara de pixels gerada durante a segmentação para serem executadas. Além disto, seria incoerente aplicar estas técnicas sem a segmentação, uma vez que elas trabalham com as regiões de fronteira entre objeto e fundo.

Suavização de borda

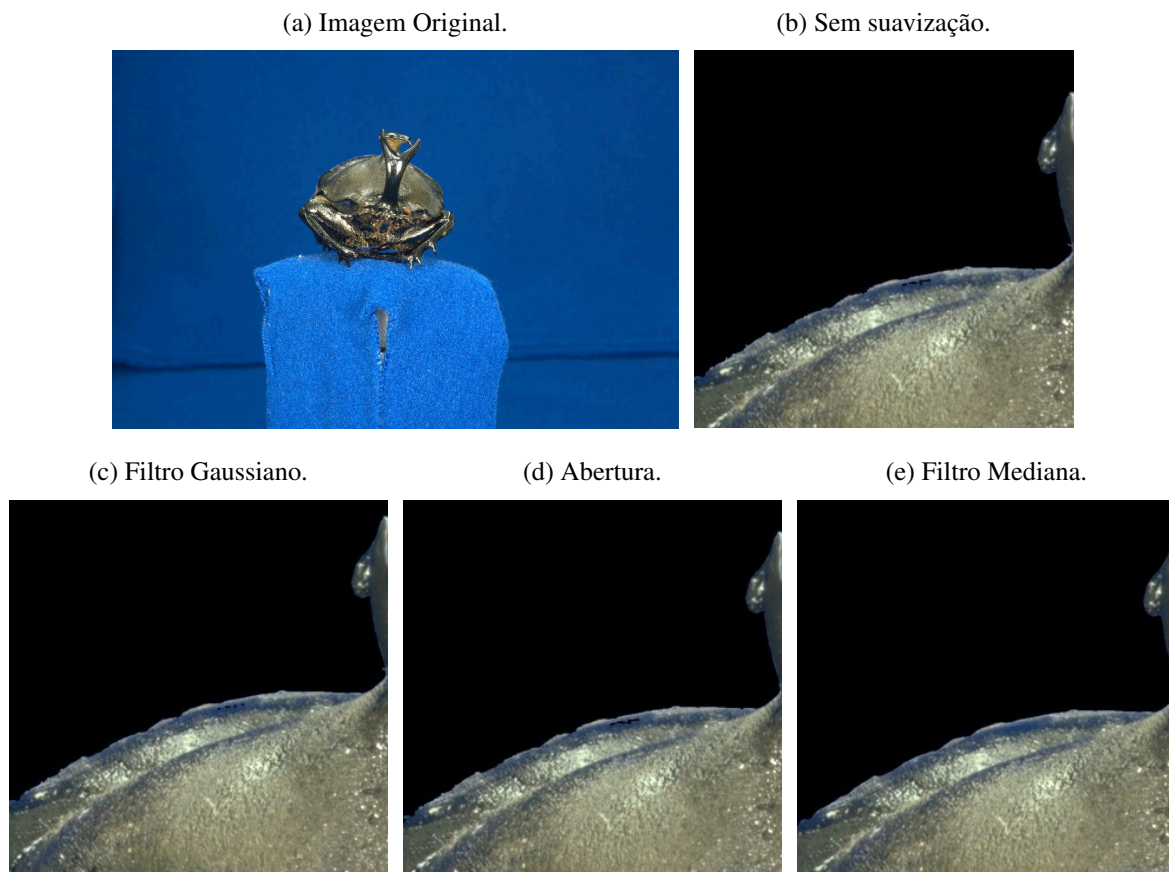
A suavização de borda remove a sensação de serrilhado que pode ser causado pela etapa de segmentação. Foram selecionados 3 técnicas que são aplicados sobre a máscara originada na segmentação:

- Filtro passa-baixa com função Gaussiana e elemento estruturante 7x7. Após a aplicação do filtro é aplicada uma limiarização para que a máscara continue sendo binária (Figura

109(c));

- Operação de abertura da morfologia matemática (erosão seguida de dilatação) com elemento estruturante quadrado de 5×5 (Figura 109(d));
- Filtro passa-baixa mediana e elemento estruturante 11×11 (Figura 109(e)).

Figura 109 – Suavização de borda.



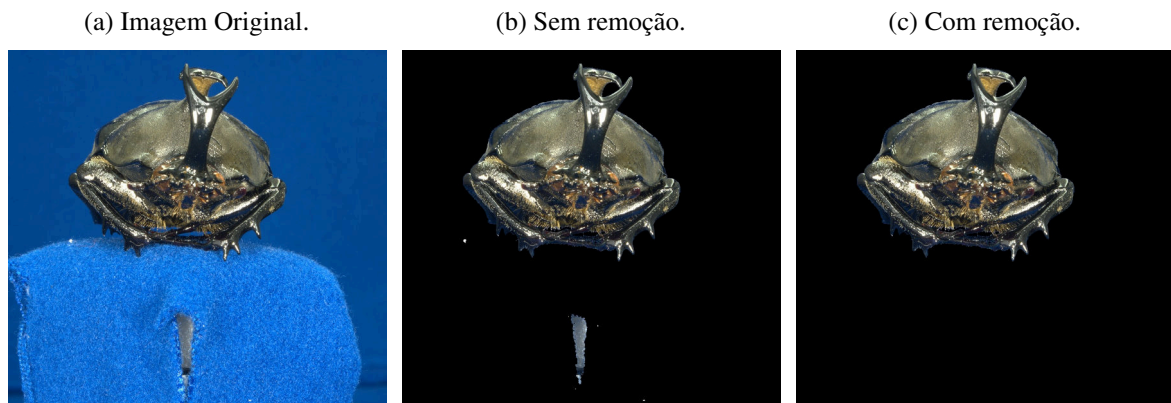
Fonte: Autoria própria.

Remoção de pixels esparsos

Dependendo do tecido usado para recobrir o fundo infinito, de sujeiras sobre o fundo, ou mesmo de regiões de material/cor diferente no tecido que recobre o fundo infinito, após a segmentação, podem permanecer pontos ou regiões que não foram corretamente segmentados. A Figura 110 apresenta uma fotografia sem segmentação 110(a); com segmentação e sem a remoção de pixels esparsos 110(b); e com a remoção de pixels 110(c)

Percebe-se na Figura 110(b) que uma sujeira branca e uma região incorretamente recoberta pelo tecido do fundo infinito (na base de apoio da peça) permaneceram na imagem após a segmentação. Estes problemas poderiam ser resolvidos inserindo pontos destas regiões como

Figura 110 – Remoção de pixels esparsos.



Fonte: Autoria própria.

sementes da classe Fundo para treinar o segmentador. Porém, considerando a grande quantidade de fotografias que podem compor um Stream2D este processo poderia demorar muito para o usuário.

Optou-se então por usar uma abordagem que removesse automaticamente estes componentes da imagem. Foram testadas abordagens que removem regiões menores que determinado tamanho, porém, como no caso do defeito da base de apoio, estas abordagens poderiam perder efetividade por causa do tamanho destas regiões.

Então, manipulando a máscara binária gerada durante a segmentação, identifica-se todos os contornos da imagem; calcula-se a área de cada contorno; e preserva-se apenas o maior contorno, excluindo os outros da imagem. Este processo visa garantir que apenas o objeto de interesse permaneça na imagem, excluindo outras regiões menores.

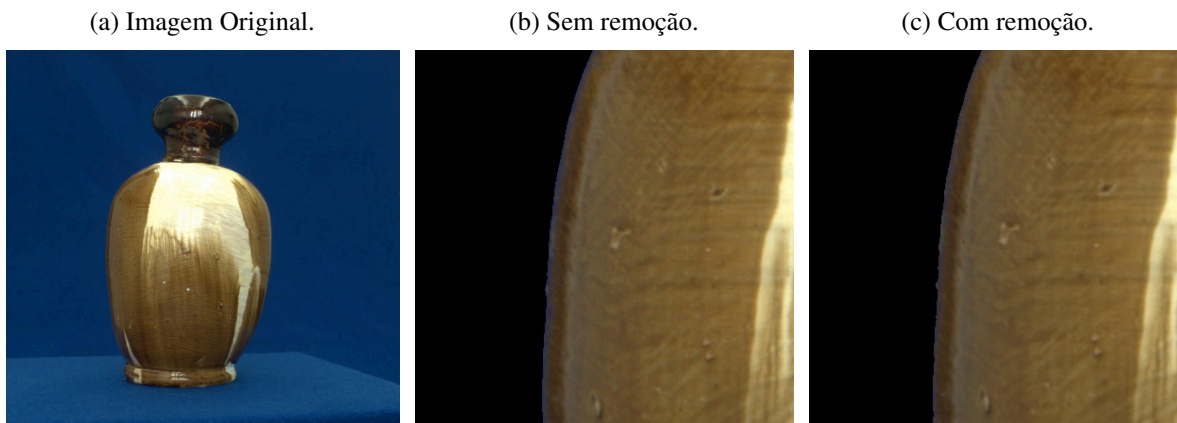
Caso a segmentação tenha gerado regiões mal segmentadas maiores que o objeto, estas regiões seriam preservadas em vez do objeto, mas neste caso recomenda-se que a seleção de sementes seja refeita para gerar uma segmentação mais adequada. Além disto, esta abordagem impede o escaneamento de mais de uma peça, pois a peça menor seria eliminada da imagem.

Remoção da borda

A segmentação pode gerar, na borda entre objeto e fundo, uma linha de pixels da cor do fundo infinito. Esta borda é indiretamente proporcional à nitidez desta borda, de forma que quanto menor a nitidez, maior a borda. A Figura 111 apresenta uma fotografia processada sem a remoção de borda (Figura 111(b)) e com a remoção de borda (Figura 111(c)).

O algoritmo usado para remover a borda foi a erosão (morfologia matemática) com um elemento estruturante de 3×3 . Este processamento é o último da etapa de aprimoramento, pois ela pode gerar descontinuidade e fazer com que a remoção de pixels esparsos remova partes do objeto. Também não faz sentido executar a remoção de borda e em seguida a suavização de

Figura 111 – Remoção de borda.



Fonte: Autoria própria.

borda, que pode acabar reinserindo a borda na imagem.

Organização e Persistência

Para salvar os arquivos com os nomes do Stream2D, inicialmente é necessário estabelecer o nome de determinada fotografia dentro do Stream2D. As câmeras digitais normalmente salvam as fotografias com um identificador numérico que tem seu valor incrementado a cada fotografia. Desta forma, para obter o nome do arquivo dentro do Stream2D, os arquivos de fotografias podem ser lidos do disco e ordenados conforme seus nomes.

Em seguida, para obter o número do Stream a que a foto pertence pode ser usado o índice da fotografia ordenada e a fórmula $Stream^{id} = (\text{Índice} \% \text{número de Frames}) + 1$. O número do Frame pode ser obtido por: $Frame^{id} = (\text{Índice} - ((Stream - 1) \times \text{número de Frames})) + 1$. Com estes dois valores basta construir a String do nome com o formato $S < Stream^{id} > - F < Frame^{id} > . < extensão >$. Por exemplo: $S12 - F34.jpg$.

Com o nome do arquivo já calculado, é necessário verificar na interface gráfica a opção de formato de salvamento selecionada pelo usuário, que pode ser: *jpg*, com fundo preto ou *png*, com fundo transparente. O formato *jpg* ocupa menos espaço em disco, porém o formato *png* considerando que possui fundo transparente, pode ser aplicado em ambientes de simulação sem causar oclusão, pois o fundo da imagem já está transparente.

Após este processo, usa-se a máscara binária que foi gerada na segmentação e processada pelo aprimoramento, junto com a imagem original, de forma que cada pixel marcado na máscara como sendo da classe Objeto, permanece na imagem final, enquanto que os pixels marcados como Fundo são transformados em preto, no caso de *jpg* e transparente, no caso de *png*. Com isso a imagem é salva e está pronta para ser usada na visualização.

Execução

O Tempo de processamento do *BipApp* varia de acordo com a máquina, sendo especialmente sensível ao desempenho do disco, uma vez que a quantidade de dados que precisam ser lidos pode ser grande (Como já mencionado, 100Gb para 31 Stream com 120 frames no formato Raw). O tempo de configuração do software pelo usuário, desde que o usuário esteja familiarizado com o sistema, leva de um a dois minutos, um tempo irrelevante se comparado com as 6 horas de processamento. O mesmo acontece com o treino do SVM, que é irrelevante no tempo total de processamento.

A aquisição de fotografias no formato *jpg* reduz consideravelmente o volume de dados, e consequentemente reduz também o tempo de processamento. Porém câmeras comuns costumam aplicar diversos processamentos em fotografias *jpg* para torna-las mais atrativas, o que pode reduzir a fidelidade das imagens com a peça real. Além de inviabilizar algumas correções como as apresentadas na Seção 5.2.1.

5.3 Validação

Durante o desenvolvimento deste trabalho diversas peças foram escaneadas, tanto para testes do scanner e dos softwares, como para aplicações em outros projetos e estudos. Estes estudos foram desenvolvidos a partir de modelos gerados com as técnicas e softwares apresentados. Para ratificar o funcionamento do F2S2, do processo de aquisição, das técnicas e dos softwares *F2S2-App* e *F2S2-BipApp* serão apresentados os trabalhos e estudos que usaram os modelos gerados.

Cabe ressaltar que estas não foram as únicas peças que foram adquiridas, testadas e avaliadas durante o desenvolvimento deste trabalho. Além disto, nem todas estas peças foram adquiridas e processadas com a última versão do hardware, procedimento de aquisição e software. Além disto, os trabalhos aqui apresentados possuem outros autores e coautores, porém todos envolveram o escopo deste trabalho.

A. Artigo apresentado no CSEDU 2018 (OLSEN et al., 2018)

Este trabalho apresenta objetos de aprendizagem estereoscópicos, interativos e com acuidade visual gerados a partir de escaneamentos feitos pelo F2S2. O trabalho apresenta aquisições de peças de decoração, obras de arte, réplicas de carros, bonecos, componentes de motor de carro.

O artigo apresenta a técnica de aquisição e processamento; discute a interatividade, a acuidade visual e a geração de visualização estereoscópica a partir de modelos construídos com o hardware, software e método apresentados neste trabalho.

Foram desenvolvidas visualizações para telas, projetores, óculos de realidade virtual,

pirâmide holográfica e teatro imersivo (Arena Digital *Full Dome* - PUCPR, Curitiba.) com os modelos gerados.

B. Artigo apresentado no CSEDU 2019 (SILVA. et al., 2019)

O trabalho apresenta os processos de aquisição com o F2S2, processamento, construção do Stream2D, visualização e estereoscopia. Também são apresentadas as equações necessárias para mapear uma fotografia do *pack* de fotografias em um ambiente de computação gráfica tridimensional, viabilizando a aplicação das imagens em sistemas de simulação.

Foi escaneado um crânio humano, com o qual foi desenvolvido um visualizador estereoscópico. Este visualizador é interativo, permitindo navegação e transformações geométricas; com acuidade visual; além de permitir que determinadas partes da peça fossem evidenciadas. O Objetivo deste trabalho foi criar um objeto de aprendizagem com este crânio para ser usado em aulas de anatomia humana.

O objeto de aprendizagem foi apresentado à um grupo de professores e outro grupo de estudantes, os quais responderam a questionários para avaliar a aplicabilidade de materiais deste tipo na educação. A conclusão do estudo foi que o objeto de aprendizagem gerado representa adequadamente o objeto real e que o material pode auxiliar no processo de ensino-aprendizagem.

C. Tese de doutorado de Flávio de Almeida e Silva (SILVA, 2020)

Tese de doutorado defendida por Flávio de Almeida e Silva, com o título “Projeção interativa estereoscópica: visualização de modelos de objetos reais em ambiente virtual”. A Tese trata da geração de visualizações interativas e estereoscópicas usando modelos gerados com as técnicas descritas neste trabalho.

A Geração da estereoscopia baseia-se em identificar um par de Frames, dentro do Stream2D, que ao serem visualizados com as técnicas desenvolvidas oferecem uma sensação de volume e profundidade do objeto real no modelo digital. Os objetos digitalizados pelo Flávio, ou em conjunto, estão descritos no trabalho, bem como as técnicas usadas na aquisição e processamento.

D. Trabalho de Conclusão de Curso de Medicina Veterinária defendido por Daniela Góes Turchenski - PUCPR - Curitiba - 2019

O Trabalho Trabalho de Conclusão de Curso de Medicina Veterinária intitulado “Construção de base de dados digital para estudo da anatomia do cavalo” foi defendido em 2019 por Daniela G. Turchenski sob o orientação do Professor Dr. Pedro Vicente Michelotto Junior.

O Trabalho usou digitalização de uma pata de cavalo que foi seccionada em quatro partes, com um corte longitudinal e outro transversal exibindo a anatomia interna e externa da pata de um cavalo. Cada seção foi digitalizada com 31 Streams de 120 Frames.

E. Trabalhos em andamento

Foram realizados escaneamentos e testes para criação de conteúdo para outro projeto intitulado *The Visible Human Table* (VHT). Este projeto é associado ao Centro de Inovação em Imagens Médicas (CIIM) e foram digitalizados crânios e partes de crânios para serem adicionados como conteúdo da VHT.

Trabalho de Mestrado qualificado em 2020 no PPGIa - PUCPR por Marina de Lara Muller intitulado “Ambientes Imersivos em Caves com Estereoscopia: Uma aplicação para ensino da Entomologia” também usou as técnicas de aquisição, softwares e procedimentos apresentados nesta Tese.

Considerações

A Interatividade pretendida e definida como a possibilidade de navegação e transformações geométricas foi viabilizada com a etapa de aquisição (Seção 4.3), com o software *F2S2-App* (Seção 5.1) e com a definição de uso dos Frames do Stream2D em ambientes de Simulação (Seção 4.1).

A acuidade visual, definida como a possibilidade de ver no modelo virtual as mesmas características do objeto real, foi alcançada com a definição do protocolo de aquisição apresentado na Seção 4.3, com a construção do software *F2S2-BipApp* e definição do protocolo de correção das imagens, apresentados na Seção 5.2.

Considerando que as técnicas e softwares descritos nesta tese conseguiram gerar modelos interativos e com acuidade visual para serem usados neste em outros estudos e projetos, que usam modelos interativos e com acuidade visual, pode-se inferir que o Objetivo Geral deste trabalho foi alcançado. Adicionalmente, foi gerado um registro de software do firmware de controle do F2S2 (JUSTINO et al., 2019).

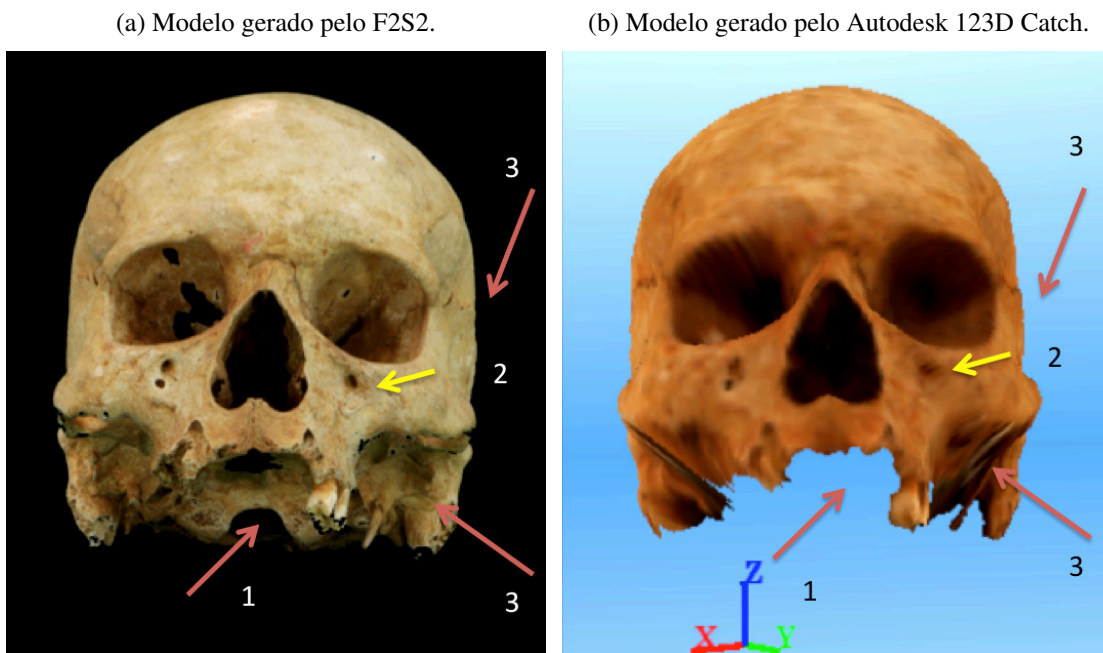
5.4 Análise

No que se refere ao processo de digitalização de objetos reais o método apresentado se difere das técnicas de escaneamento ópticos ativos como o escaneamento a laser e por luz estruturada, pois estas técnicas se baseiam em triangulação para calcular ponto a ponto a superfície dos objetos. O mesmo processo ocorre com as técnicas de digitalização de objetos passivos que usam imagens para realizar a reconstrução tridimensional do objeto.

Ambas as técnicas tradicionais (ópticas passivas e ativas) usam malhas ou nuvens de pontos para representar o volume do objeto. O Método apresentado nesta Tese digitaliza a peça adquirindo cada perspectiva completa, com as cores, iluminação e sombras reais. Não sendo necessário acrescentar estes efeitos posteriormente no Stream2D.

A Figura 112 apresenta uma comparação realizada entre uma aquisição feita pelo F2S2 (Figura 112(a)) e uma reconstrução tridimensional realizada com as fotos desta aquisição no software 123D (Figura 112(b)) Catch da Autodesk²¹. Percebe-se diversas diferenças entre ambas as imagens, como textura, cor e capacidade de visualizar detalhes.

Figura 112 – Comparação entre modelo gerado pelo F2S2 e reconstrução 3D.



Fonte: Laboratório de visão e robótica - PPGIa - PUCPR.

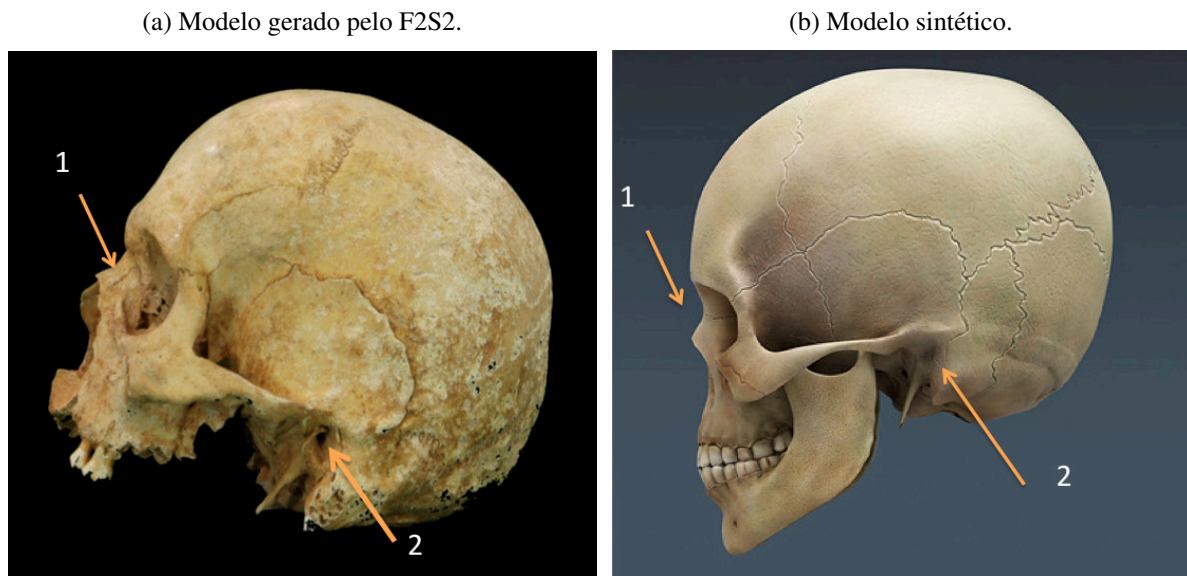
É possível notar na reconstrução (Figura 112(b)) que a parte de baixo do crânio não foi reconstruída (indicado pelo número 1), que o orifício presente no crânio foi coberto (indicado pelo número 2) e que o arco zigomático (responsável por dar forma à maçã do rosto - indicado pelo número 3) foi reconstruído como um volume disforme e não apresenta os mesmos detalhes da Figura 112(a).

Se comparado com desenhos sintéticos, estes dependem do talento em desenhar e da capacidade do artista de perceber detalhes, de forma um desenho dificilmente representará um objeto com grande acuidade visual. Isto pode ser percebido na Figura 113. A digitalização do F2S2 (Figura 113(a)) preserva muito mais detalhes dos ossos do nariz (indicado pelo número 1) e preserva o orifício do ouvido (indicado pelo número 2) se comparado com o modelo sintético (Figura 113(b)).

Tanto o Stream2D quanto as representações tradicionais permitem a interatividade, dependendo do software que gera a visualização. Porém a complexidade computacional para reconstruir uma nuvem/malha de pontos é muito maior do que a complexidade de encontrar o Frame que melhor representa um ponto de vista. Além disto, nas técnicas tradicionais é

²¹ <<http://www.123dapp.com/>> Acessado em: Março de 2018.

Figura 113 – Comparação entre modelo gerado pelo F2S2 e modelo 3D desenhado.



Fonte: Laboratório de visão e robótica - PPGIa - PUCPR.

necessário, após a reconstrução, inserir iluminação e sombras para que o modelo fique realístico, aumentando ainda mais a complexidade.

O Stream2D por sua vez pode ocupar uma quantidade muito grande de espaço em disco para ser armazenada em comparação com os modelos tradicionais. O mesmo pode acontecer com o consumo de memória, que pode ser grande, dependendo do tamanho e quantidade de Frames que sejam necessários carregar simultaneamente para a visualização.

Os modelos tradicionais também podem permitir o cálculo de dimensões e de volumes, dependendo de como estes dados são construídos. O Stream2D por sua vez não permite o cálculo destas características.

Outra possibilidade de uso do Stream2D é a geração de visualizações estereoscópicas. Para gerar visualizações estereoscópicas podem ser usadas técnicas passivas como: perspectiva, iluminação, oclusão e gradiente de textura. Porém, diferente de modelos computacionais que simulam estas técnicas, o Stream2D já garante perspectiva, iluminação, oclusão e gradiente de textura naturais, portanto não é necessário inserir estas características artificialmente nas imagens.

Trata-se de um ganho em complexidade computacional e acuidade visual, uma vez que não são necessários algoritmos e processamentos para varrer a imagem pixel por pixel, identificando determinadas características e alterando a imagem para evidenciar outras características que geram estereoscopia. O Quadro 3 apresenta uma comparação entre os modelos tradicionais (baseados em malhas e núvens de pontos) e o Stream2D.

Em comparação com o trabalho (FARIA et al., 2016), o processo apresentado nesta Tese

Quadro 3 – Quadro comparativo entre o Stream2D e modelos 3D baseados em nuvens e malhas de pontos.

Característica	Stream2D	Tradicionais
Interatividade	Sim	Sim
Acuidade Visual	Mais realístico	Mais sintético
Iluminação	Real	Sintético
Sombra	Real	Sintético
Estereoscopia	Natural	Sintético
Cálculo de Dimensões	Não	Sim
Complexidade Computacional	Menor	Maior
Uso de Disco	Maior	Menor
Uso de Memória	Maior	Menor

Fonte: Autoria Própria.

possui diversas vantagens. Considerando que a aquisição pelo F2S2 é automática, isto permite uma quantidade muito maior de imagens e a precisão no posicionamento da câmera é maior. Isto permite uma aquisição com uma interatividade mais natural, visto que mais perspectivas são adquiridas.

A definição de como aplicar o Stream2D em ambientes de simulação permite uma aplicabilidade maior do Stream2D em outros sistemas, enquanto no trabalho (FARIA et al., 2016), é usado apenas o QuickTime framework. (FARIA et al., 2016) também não explica os processos: fotográfico, de correção ou aprimoramento das imagens. Com isto pode-se inferir que a acuidade visual oferecida neste trabalho é melhor.

6 Considerações Finais

Esta pesquisa objetivou a criação de um procedimento capaz de digitalizar peças reais gerando um modelo interativo e com acuidade visual. O processo se baseia em adquirir um conjunto de fotografias, com acuidade visual, de todos os ângulos do objeto, respeitando um intervalo angular predeterminado. Para visualizar a peça é exibida a fotografia que melhor representa o ponto de vista desejado pelo observador. A interatividade resulta da troca de imagens conforme muda o ponto de vista do observador.

Foi abordado o processo de aquisição fotográfica com o F2S2, que considerou os objetos a serem escaneados, as características da fotografia e as especificidades do scanner usado no processo. A aquisição é realizada de modo controlado para gerar imagens com acuidade visual. Diferentemente do processo de fotografia em que as condições da fotografia são construídas para determinado ponto de vista, na aquisição estas condições devem considerar todo o conjunto de imagens que serão adquiridos.

Discutiu-se o processamento destas imagens para gerar o Stream2D, estrutura composta pelo conjunto de perspectivas que representam o objeto virtual. A construção desta estrutura passa por etapas de correção de problemas intrínsecos da fotografia; redução de contexto e segmentação para manter na imagem apenas o objeto de interesse; aprimoramento das imagens para gerar uma visualização com mais qualidade visual.

Por último, o processamento constrói o Stream2D ao atribuir um identificador único capaz de mapear a posição da câmera em relação ao objeto para cada fotografia. Foram apresentadas também um conjunto de equações necessárias para realizar o mapeamento da posição da câmera, considerando desde as instruções de coordenadas internas do F2S2, até a seleção da fotografia que deve ser visualizada em um ambiente de simulação baseado em computação gráfica tridimensional.

A contribuição deste trabalho foi a elaboração deste procedimento, desde a análise do objeto, até a viabilização da visualização do modelo digital do objeto. É necessário considerar que esta técnica é original, desde o processo de aquisição, que adquire cada ponto de vista do objeto integralmente, com luzes e sombras reais; na representação do objeto virtual, que estrutura perspectivas prontas para exibição em comparação com representar o objeto por um conjunto de pontos; até o procedimento de visualização, que não precisa reconstruir o objeto para exibi-lo.

Para isso os objetivos geral e específicos foram alcançados e a questão de pesquisa “É possível desenvolver um processo de digitalização de objetos reais, baseado em um conjunto de fotografias que representem todos os ângulos de visualização do objeto, gerando um modelo interativo e com acuidade visual?” foi respondida afirmativamente.

Além disto, outras pesquisas foram viabilizadas ou usaram parte deste trabalho em seu desenvolvimento. Contudo, foram identificados outros pontos que podem ser estudados e são enumerados como trabalhos futuros.

Trabalhos futuros

- O Stream2D não permite a medição de distâncias e volumes nos objetos digitalizados. Portanto, pode ser usada uma técnica de reconstrução tridimensional, que associada ao Stream2D viabiliza uma visualização interativa, com acuidade visual e possibilidade de medição de distâncias e volumes. Isto permite diversas aplicações como a impressão tridimensional dos modelos digitalizados.
- Criação de um visualizador capaz de apresentar ao mesmo tempo cenários tridimensionais baseados em câmeras com 360° e modelos Stream2D. Com isto seria possível unir cenários e objetos com interatividade e acuidade visual.
- Investigar outras técnicas de segmentação, principalmente algoritmos específicos para objetos com transparências e reflexos. Atualmente estudos baseados em DeepLearning tem evoluído nesta área e podem ser promissores para este objetivo.
- Criação de repositórios de objetos digitalizados para aplicação em diversas áreas da educação, como: estudo de anatomia, biologia e arte, entre outros.

Referências

ACKERMAN, M. J. The visible human project. *Proceedings of the IEEE*, v. 86, n. 3, p. 504–511, Mar 1998. ISSN 0018-9219. Citado 3 vezes nas páginas 27, 37 e 50.

ANIRUDTH, N. et al. Robust photometric alignment for asymmetric camera system. In: *2018 IEEE International Conference on Consumer Electronics (ICCE)*. [S.l.: s.n.], 2018. p. 1–4. Citado na página 49.

BALMANT, O. *Falta de cadáveres nas universidades aumenta chance de erros médicos - Brasil - iG*. 2015. <<http://ultimosegundo.ig.com.br/brasil/2015-06-10/falta-de-cadaveres-nas-universidades-aumenta-chance-de-erros-medicos.html>>. (Accessed on 06/07/2018). Citado na página 27.

BATTEN, C. F. *Autofocusing and Astigmatism Correction in the Scanning Electron Microscope*. Dissertação (Mestrado) — University of Cambridge, Cambridge, Reino Unido, 01 2000. Citado na página 45.

CAVIEDES, J.; OBERTI, F. A new sharpness metric based on local kurtosis, edge and energy information. *Signal Processing: Image Communication*, v. 19, n. 2, p. 147 – 161, 2004. ISSN 0923-5965. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0923596503000912>>. Citado na página 45.

CHENG, Y.; PEDERSEN, M.; CHEN, G. Evaluation of image quality metrics for sharpness enhancement. In: *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*. [S.l.: s.n.], 2017. p. 115–120. Citado na página 44.

CHERN, N. N. K.; NEOW, P. A.; ANG, M. H. Practical issues in pixel-based autofocusing for machine vision. In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*. [S.l.: s.n.], 2001. v. 3, p. 2791–2796 vol.3. ISSN 1050-4729. Citado na página 45.

CHHATKULI, A. et al. Inextensible non-rigid shape-from-motion by second-order cone programming. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 1719–1727. ISSN 1063-6919. Citado na página 40.

CHOI, D. Y. et al. Sharpness enhancement and super-resolution of around-view monitor images. *IEEE Transactions on Intelligent Transportation Systems*, p. 1–13, 2017. ISSN 1524-9050. Citado na página 42.

CHOI, K.; SUK, H. J. The optimal color temperature of smartphone display under various illuminant conditions. In: *2014 IEEE International Conference on Consumer Electronics (ICCE)*. [S.l.: s.n.], 2014. p. 506–507. ISSN 2158-3994. Citado na página 30.

CHUNG, Y.-C. et al. A non-parametric blur measure based on edge analysis for image processing applications. In: *IEEE Conference on Cybernetics and Intelligent Systems, 2004*. [S.l.: s.n.], 2004. v. 1, p. 356–360 vol.1. Citado na página 46.

DE, K.; MASILAMANI, V. Image sharpness measure for blurred images in frequency domain. *Procedia Engineering*, v. 64, p. 149 – 158, 2013. ISSN 1877-7058. International

- Conference on Design and Manufacturing (IconDM2013). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877705813016007>>. Citado na página 45.
- ERASMUS, S. J.; SMITH, K. C. A. An automatic focusing and astigmatism correction system for the sem and ctem. *Journal of Microscopy*, v. 127, n. 2, p. 185–199, 1982. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2818.1982.tb00412.x>>. Citado na página 44.
- FANG, X. et al. A consistent pixel-wise blur measure for partially blurred images. In: *2014 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2014. p. 496–500. ISSN 1522-4880. Citado na página 45.
- FARIA, J. W. V. de et al. Virtual and stereoscopic anatomy: when virtual reality meets medical education. *Journal of Neurosurgery*, v. 125, n. 5, p. 1105–1111, 2016. PMID: 26871375. Disponível em: <<https://doi.org/10.3171/2015.8.JNS141563>>. Citado 4 vezes nas páginas 40, 41, 195 e 196.
- FERREIRA, R. M. F. *Reconstrução de Modelos CAD 3D Baseada em Imagem Fotográfica Digital*. Dissertação (Mestrado) — Universidade de Aveiro, 2009. Citado 4 vezes nas páginas 37, 38, 39 e 40.
- FERZLI, R.; KARAM, L. J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *IEEE Transactions on Image Processing*, v. 18, n. 4, p. 717–728, April 2009. ISSN 1057-7149. Citado na página 44.
- FIGUEIREDO, E.; OLIVEIRA, L. Simulation training methods in neurological surgery. *Asian Journal of Neurosurgery*, Medknow, v. 14, n. 2, p. 364, 2019. Disponível em: <https://doi.org/10.4103/ajns.ajns_269_18>. Citado na página 41.
- GAMMA, E. *Padrões de Projetos: Soluções Reutilizáveis*. [S.l.]: Bookman, 2009. ISBN 9788577800469. Citado 2 vezes nas páginas 156 e 173.
- GHOSH, S. *Fundamentals of Computational Photogrammetry*. Concept Publishing Company, 2005. ISBN 9788180691874. Disponível em: <<https://books.google.com.br/books?id=HLZ-abpzGCoC>>. Citado na página 28.
- GIL, A. *Métodos e técnicas de pesquisa social*. Atlas, 2008. ISBN 9788522451425. Disponível em: <<https://books.google.com.br/books?id=T3BwPgAACAAJ>>. Citado na página 55.
- GIRIDHAR, C. *Aprendendo Padrões de Projeto em Python: Tire proveito da eficácia dos padrões de projeto (design patterns) em Python para resolver problemas do mundo real em arquitetura e design de software*. [S.l.]: Novatec Editora, 2016. ISBN 9786586057157. Citado na página 156.
- GURBUZ, S.; KAWAKITA, M.; ANDO, H. Color calibration for multi-camera imaging systems. In: *2010 4th International Universal Communication Symposium*. [S.l.: s.n.], 2010. p. 201–206. Citado na página 49.
- HE, X. et al. Adaptive measuring algorithm for cmm based on 5-th bézier curve. In: *IEEE ICCA 2010*. [S.l.: s.n.], 2010. p. 1806–1811. ISSN 1948-3457. Citado na página 37.
- HINCAPIÉ-RAMOS, J. D. et al. Smartcolor: Real-time color and contrast correction for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, v. 21, n. 12, p. 1336–1348, Dec 2015. ISSN 1077-2626. Citado na página 46.

HOSAM, O.; SUN, X. Three-dimensional reconstruction using enhanced shape from stereo technique. In: *2009 Eighth IEEE/ACIS International Conference on Computer and Information Science*. [S.l.: s.n.], 2009. p. 627–632. Citado na página 40.

HUANG, R.; SMITH, W. A. P. Shape-from-shading under complex natural illumination. In: *2011 18th IEEE International Conference on Image Processing*. [S.l.: s.n.], 2011. p. 13–16. ISSN 2381-8549. Citado na página 40.

INTWALA, A. M.; MAGIKAR, A. A review on process of 3d model reconstruction. In: *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. [S.l.: s.n.], 2016. p. 2851–2855. Citado 5 vezes nas páginas 35, 36, 37, 39 e 40.

JACQUESSON, T. et al. Stereoscopic three-dimensional visualization: interest for neuroanatomy teaching in medical school. *Surgical and Radiologic Anatomy*, Springer Science and Business Media LLC, v. 42, n. 6, p. 719–727, fev. 2020. Disponível em: <<https://doi.org/10.1007/s00276-020-02442-6>>. Citado na página 41.

JUNG, J. Y. et al. Camera-based color calibration method for multiple flat-panel displays using smartphone. *Journal of Display Technology*, v. 12, n. 12, p. 1777–1784, Dec 2016. ISSN 1551-319X. Citado na página 46.

JUNG, J.-Y. et al. Camera-based color calibration method for multiple flat-panel display system. In: *2015 IEEE International Conference on Consumer Electronics (ICCE)*. [S.l.: s.n.], 2015. p. 90–91. ISSN 2158-3994. Citado na página 46.

JUSTINO, E. J. R. et al. *Gerenciador do Scanner Semiesférico de Campo Completo*. [S.l.]: Registro de Software - INPI - Instituto Nacional da Propriedade Industrial, 2019. Registro: 512018052451-6. Citado na página 193.

KIM, Y.-T. et al. Color look-up table design for gamut mapping and color space conversion. In: . [S.l.: s.n.], 2004. p. 28–29. Citado na página 48.

KOBUS, B.; BRIAN, F. Camera characterization for color research. *Color Research & Application*, v. 27, n. 3, p. 152–163, 2002. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/col.10050>>. Citado na página 46.

KOIK, B. T.; IBRAHIM, H. A literature survey on blur detection algorithms for digital imaging. In: *2013 1st International Conference on Artificial Intelligence, Modelling and Simulation*. [S.l.: s.n.], 2013. p. 272–277. Citado 2 vezes nas páginas 42 e 43.

KRASULA, L. et al. Quality assessment of sharpened images: Challenges, methodology, and objective metrics. *IEEE Transactions on Image Processing*, v. 26, n. 3, p. 1496–1508, March 2017. ISSN 1057-7149. Citado na página 46.

LAU, C.; HEIDRICH, W.; MANTIUK, R. Cluster-based color space optimizations. In: *2011 International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 1172–1179. ISSN 1550-5499. Citado na página 46.

LAVRINOV, D. S.; KHORKIN, A. I. Problems of internal calibration of precision laser triangulation 2d scanners. In: *2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. [S.l.: s.n.], 2016. p. 1–4. Citado na página 39.

LAWRENCE, F. et al. Comparison of autofocus methods for automated microscopy. *Cytometry*, v. 12, n. 3, p. 195–206, 1991. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/cyto.990120302>>. Citado na página 45.

LI, K.; DAI, Q.; XU, W. High quality color calibration for multi-camera systems with an omnidirectional color checker. In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. [S.l.: s.n.], 2010. p. 1026–1029. ISSN 1520-6149. Citado na página 49.

LI, L. et al. No-reference image blur assessment based on discrete orthogonal moments. *IEEE Transactions on Cybernetics*, v. 46, n. 1, p. 39–50, Jan 2016. ISSN 2168-2267. Citado na página 44.

LIMA, J. D. *Corpos reais para entender anatomia*. 2012. <<https://www.gazetadopovo.com.br/educacao/corpos-reais-para-entender-anatomia-1d2ydx72l3loudjii8ots1rv2>>. (Accessed on 06/07/2018). Citado na página 27.

LIN, W.; KUO, C.-C. J. Perceptual visual quality metrics: A survey. *Journal of Visual Communication and Image Representation*, v. 22, n. 4, p. 297 – 312, 2011. ISSN 1047-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1047320311000204>>. Citado na página 44.

LISSNER, I.; URBAN, P. Toward a unified color space for perception-based image processing. *IEEE Transactions on Image Processing*, v. 21, n. 3, p. 1153–1168, March 2012. ISSN 1057-7149. Citado na página 46.

LIU, H. x.; XIE, M.; HUANG, M. Image color-difference evaluation based on color-difference formula. In: *2011 4th International Congress on Image and Signal Processing*. [S.l.: s.n.], 2011. v. 4, p. 1771–1774. Citado na página 48.

MAHMOOD, M. T. Shape from focus by total variation. In: *IVMSP 2013*. [S.l.: s.n.], 2013. p. 1–4. Citado na página 40.

MALIK, A. S.; CHOI, T. Comparison of polymers: A new application of shape from focus. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, v. 39, n. 2, p. 246–250, March 2009. ISSN 1094-6977. Citado na página 40.

MANTIUK, R. K.; TOMASZEWSKA, A.; MANTIUK, R. Comparison of four subjective methods for image quality assessment. *Computer Graphics Forum*, Wiley/Blackwell (10.1111), v. 31, n. 8, p. 2478–2491, dec 2012. ISSN 14678659. Disponível em: <<http://doi.wiley.com/10.1111/j.1467-8659.2012.03188.x>>. Citado na página 46.

MARTIN, R. *Código Limpo: Habilidades Práticas do Agile Software*. [S.l.]: Alta Books, 2009. ISBN 9788550811482. Citado na página 172.

MARZILIANO, P. et al. Perceptual blur and ringing metrics: application to jpeg2000. *Signal Processing: Image Communication*, v. 19, n. 2, p. 163 – 172, 2004. ISSN 0923-5965. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0923596503000924>>. Citado na página 45.

MENDES, L.; CARVALHO, P. Adaptive polynomial regression for colorimetric scanner calibration using genetic algorithms. In: *IEEE International Workshop on Intelligent Signal Processing, 2005*. [S.l.: s.n.], 2005. p. 22–27. Citado na página 49.

NARVEKAR, N. D.; KARAM, L. J. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *IEEE Transactions on Image Processing*, v. 20, n. 9, p. 2678–2683, Sept 2011. ISSN 1057-7149. Citado na página 45.

OH, T. et al. No-reference sharpness assessment of camera-shaken images by analysis of spectral structure. *IEEE Transactions on Image Processing*, v. 23, n. 12, p. 5428–5439, 2014. ISSN 10577149. Citado na página 46.

OLSEN, D. R. et al. Generation of stereoscopic interactive learning objects true to the original object. In: INSTICC. *Proceedings of the 10th International Conference on Computer Supported Education - Volume 2: CSEDU*, [S.l.]: SciTePress, 2018. p. 259–266. ISBN 978-989-758-291-2. Citado 3 vezes nas páginas 32, 51 e 191.

OPENCLIPART. *Clipart vetorial de câmera fotográfica com um display dobrável*. 2015. Acessado em: 15-01-2020. Disponível em: <<https://publicdomainvectors.org/pt/vetorial-gratis/Clipart-vetorial-de-c%C3%A2mera-fotogr%C3%A1fica-com-um-display-dobr%C3%A1vel/20454.html>>. Citado na página 63.

PANT, D. R.; FARUP, I. Cie uniform chromaticity scale diagram for measuring performance of osa-ucs ee and ciede00 formulas. In: *3rd European Workshop on Visual Information Processing*. [S.l.: s.n.], 2011. p. 18–23. Citado na página 48.

PARASHAR, S.; PIZARRO, D.; BARTOLI, A. Isometric non-rigid shape-from-motion in linear time. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 4679–4687. ISSN 1063-6919. Citado na página 40.

PAYET, N.; TODOROVIC, S. Scene shape from texture of objects. In: *CVPR 2011*. [S.l.: s.n.], 2011. p. 2017–2024. ISSN 1063-6919. Citado na página 40.

PEDERSEN, M. et al. Attributes of image quality for color prints. v. 19, p. 011016, 01 2010. Citado na página 44.

RAFIEE, G.; DLAY, S. S.; WOO, W. L. Automatic segmentation of interest regions in low depth of field images using ensemble clustering and graph cut optimization approaches. In: *2012 IEEE International Symposium on Multimedia*. [S.l.: s.n.], 2012. p. 161–164. Citado na página 46.

ROBINSON, E. M. *Crime Scene Photography (Third Edition)*. Third edition. San Diego: Academic Press, 2016. ISBN 978-0-12-802764-6. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780128027646030014>>. Citado 2 vezes nas páginas 42 e 43.

ROHITH, M. V. et al. Shape from stereo and shading by gradient constrained interpolation. In: *2013 IEEE International Conference on Image Processing*. [S.l.: s.n.], 2013. p. 2232–2236. ISSN 1522-4880. Citado na página 40.

SATTAR, M. U. et al. Effects of virtual reality training on medical students' learning motivation and competency. *Pakistan Journal of Medical Sciences*, Pakistan Journal of Medical Sciences, v. 35, n. 3, maio 2019. Disponível em: <<https://doi.org/10.12669/pjms.35.3.44>>. Citado na página 41.

SHAO, X. et al. Virtual reality technology for teaching neurosurgery of skull base tumor. *BMC Medical Education*, Springer Science and Business Media LLC, v. 20, n. 1, jan. 2020. Disponível em: <<https://doi.org/10.1186/s12909-019-1911-5>>. Citado na página 41.

SILVA, F. de Almeida e. *Projeção interativa estereoscópica: visualização de modelos de objetos reais em ambiente virtual*. Tese (Doutorado) — Pontifícia Universidade Católica do Paraná - Programa de Pós-Graduação em Informática - PPGIa, 2020. Citado 2 vezes nas páginas 95 e 192.

SILVA., F. de Almeida e et al. Stereoscopic interactive objects: Acquisition, generation and evaluation. In: INSTICC. *Proceedings of the 11th International Conference on Computer Supported Education - Volume 2: CSEDU*,. [S.l.]: SciTePress, 2019. p. 165–176. ISBN 978-989-758-367-4. Citado 2 vezes nas páginas 53 e 192.

SMITH, L. N. *Estimating an Image's Blur Kernel from Edge Intensity Profiles*. [S.l.], 2012. Citado na página 46.

SOBOTTA, J. et al. *Sobotta atlas of human anatomy*. 13. ed. [S.l.]: Lippincott Williams and Wilkins, 2001. v. 1-2. ISBN 9780781731737 (v.1) - 9780781731744 (v.2). Citado 2 vezes nas páginas 30 e 36.

SPIELMAN, D. A.; TENG, S.-H. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, ACM, New York, NY, USA, v. 51, n. 3, p. 385–463, maio 2004. ISSN 0004-5411. Disponível em: <<http://doi.acm.org/10.1145/990308.990310>>. Citado na página 35.

STONE, M. C. Representing colors as three numbers [color graphics]. *IEEE Computer Graphics and Applications*, v. 25, n. 4, p. 78–85, July 2005. ISSN 0272-1716. Citado na página 47.

SULISTYO, S. B.; WOO, W. L.; DLAY, S. S. Regularized neural networks fusion and genetic algorithm based on-field nitrogen status estimation of wheat plants. *IEEE Transactions on Industrial Informatics*, v. 13, n. 1, p. 103–114, Feb 2017. ISSN 1551-3203. Citado na página 48.

TANG, L. et al. Camera image quality assessment without reference information. In: *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. [S.l.: s.n.], 2016. p. 666–670. Citado na página 46.

TRIEPELS, C. P. R. et al. Does three-dimensional anatomy improve student understanding? *Clinical Anatomy*, Wiley, v. 33, n. 1, p. 25–33, maio 2019. Disponível em: <<https://doi.org/10.1002/ca.23405>>. Citado na página 42.

TRUCCO, E.; FISHER, R. B. Computing surface-based representations from range images. In: *Proceedings of the 1992 IEEE International Symposium on Intelligent Control*. [S.l.: s.n.], 1992. p. 275–280. ISSN 2158-9860. Citado na página 40.

US, N. *How many pictures has my camera taken? How many will it take?* 2018. Acessado em: 15-01-2020. Disponível em: <https://www.nikonimgsupport.com/ni/Ni_article?articleNo=000003332&configured=1&lang=en_US>. Citado na página 66.

WANG, S. et al. Shape from shading based on maximum entropy. In: *2008 International Conference on Computer Science and Information Technology*. [S.l.: s.n.], 2008. p. 287–289. Citado na página 40.

WEN, S. Color management for future video systems. *Proceedings of the IEEE*, v. 101, n. 1, p. 31–44, Jan 2013. ISSN 0018-9219. Citado na página 30.

- XIA, M. et al. End-to-end color printer calibration by total least squares regression. *IEEE Transactions on Image Processing*, v. 8, n. 5, p. 700–716, May 1999. ISSN 1057-7149. Citado na página 49.
- YU, H. et al. A method for color calibration based on simulated annealing optimization. In: *2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*. [S.l.: s.n.], 2016. p. 54–58. Citado 2 vezes nas páginas 47 e 48.
- YU, J. J.; PARK, C. J. Color correction for stereoscopic images. In: *2015 International Conference on Information and Communication Technology Convergence (ICTC)*. [S.l.: s.n.], 2015. p. 1226–1228. Citado na página 49.
- YUAN, Y. et al. Contour approximation depth image coding for virtual view synthesis. In: *2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP)*. [S.l.: s.n.], 2015. p. 1–6. Citado na página 45.
- ZHANG, B.; ALLEBACH, J. P. Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE Transactions on Image Processing*, v. 17, n. 5, p. 664–678, May 2008. ISSN 1057-7149. Citado na página 42.
- ZHAO, J. et al. Automatic blur region segmentation approach using image matting. *Signal, Image and Video Processing*, Springer, v. 7, n. 6, p. 1173–1181, 2013. Citado na página 46.
- ZHAO, J. et al. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. *BMC Medical Education*, Springer Science and Business Media LLC, v. 20, n. 1, abr. 2020. Disponível em: <<https://doi.org/10.1186/s12909-020-1994-z>>. Citado na página 41.
- ZHAO, Z. et al. Research on color correction based on cmos image sensor. In: *2018 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS)*. [S.l.: s.n.], 2018. p. 699–702. Citado na página 49.

Apêndices

APÊNDICE A – Artigo - Segmentação em lote de imagens com fundo de cor única

Segmentação em lote de imagens com fundo de cor única

Diogo R. Olsen, Flávio de Almeida e Silva,
José Geraldo Noronha Filho, Conrado Feltrin, Edson J. R. Justino

Abstract. *The images created by the “Full Frame Semi-spherical Scanner” have as specificity for use, the single color background, then the background from this object of interest needs to be extracted. For this, seven image databases from different objects were defined and eleven segmentation algorithms were selected and analyzed, searching for the best one that adapts to the problem. The selection of the algorithm was based on computational cost, qualitative and quantitative analysis. Therefore, only one algorithm was validated on a image database of four thousand images.*

Resumo. *As imagens geradas pelo Scanner Semiesférico de Campo Completo têm como especificidade o fundo de cor única e para sua utilização necessita-se extrair o objeto de interesse deste fundo. Para isto, foram definidas sete bases de imagens de objetos diferentes e onze algoritmos de segmentação foram selecionados e analisados, verificando-se qual melhor se adapta ao problema. A seleção do algoritmo foi baseada no custo computacional, análise qualitativa e quantitativa. Assim, apenas um algoritmo foi validado em uma base de quatro mil imagens.*

1. Introdução

Alguns dispositivos de projeção de imagens em Ultra Alta Definição (UHD), níveis 1 (4k) e 2 (8K), permitem visualização em três dimensões (3D). Uma característica do modelo 3D é a manipulação do objeto alterando as três componentes necessárias: rotação, translação e escala. Diferentemente do 3D, a estereoscopia nos dá a noção de profundidade em imagens bidimensionais (2D) ou 3D.

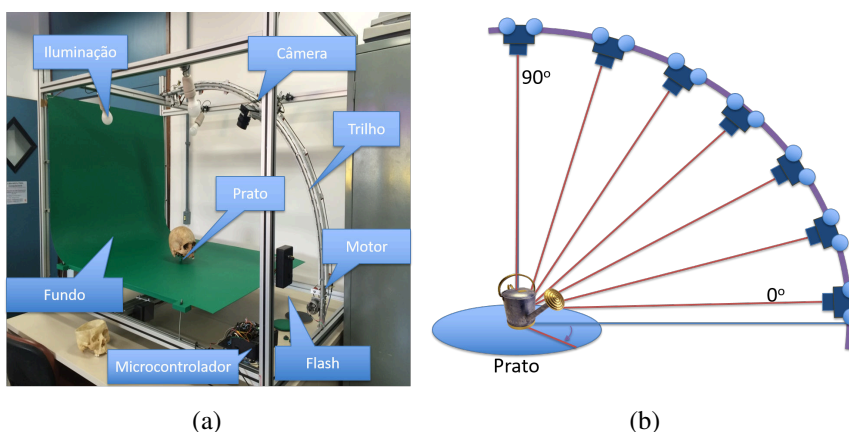
Foi desenvolvido um Scanner Semiesférico de Campo Completo (SSCC). As propriedades apresentadas pelo scanner são: possuir resoluções até UHD; garantir a fidelidade na reprodução de textura e cores do objeto; garantir a visualização em 3D e a possibilidade de visualização estereoscópica.

O SSCC é composto de cinco componentes essenciais (Figura 2(a)): um plano de fundo infinito, que serve de contraste para os objetos digitalizados; um sistema de iluminação; um prato giratório onde os objetos são posicionados; um cabeçote composto por um trilho em arco semiesférico onde a câmera é acoplada; um microcontrolador que controla motores e *flash*.

O processo de aquisição das imagens pelo SSCC consiste em posicionar o objeto centralizado no prato, mover a câmera para a posição de início e configurar a iluminação. Um lote de imagens é formado pela aquisição do mesmo objeto em vários ângulos de movimentação da câmera no trilho, formando a semiesfera. Com a câmera na posição inicial, tira-se as fotos do objeto, com o mesmo fazendo giros em ângulos pré-definidos, a partir do zero até trezentos e sessenta graus.

Para cada posição da câmera, pode-se tirar até quatrocentas fotos, sendo isto determinado pelos ângulos de giro do prato. A câmera pode ser posicionada em até cem posições a partir do zero até noventa graus, formando assim o chamado *Stream-2D*, o qual pode conter até quarenta mil imagens, como mostrado na Figura 2(b).

Figura 1. (a) Componentes do Scanner, (b) Processo de Aquisição.



Fonte: Autoria própria

As imagens que compõem o Stream-2D possuem características similares quanto à cor do fundo, resolução e tamanho, uma vez que o ambiente de aquisição é controlado. O fundo de cor única auxilia no processo de segmentação. Pelo grande volume de imagens que podem ser obtidas, até quarenta mil, percebe-se que a segmentação manual demandaria muitos recursos computacionais e de pessoas.

Realizar o processamento em lote das imagens do Stream-2D de forma automatizada, considerando recursos computacionais, tempo de execução e qualidade do resultado, exige a seleção de um algoritmo que viabilize o processo seguindo as métricas abordadas na seção 2.

Este trabalho aborda o problema de segmentação automatizada em lote de uma grande quantidade de imagens UHD 4K adquiridas pelo SSCC visando a criação de conteúdos para diferentes dispositivos como: TV 4K 3D, pirâmide holográfica, Fulldome[Lantz 2007], óculos de realidade virtual, etc. Para isso serão analisados os algoritmos de segmentação considerando as especificidades do SSCC.

Desta forma, levantou-se as questões de pesquisa que pretendem verificar qual algoritmo melhor se adapta ao problema descrito, considerando: o tempo/custo computacional; e a qualidade da segmentação com relação à taxa de pixels segmentados corretamente e incorretamente.

O restante deste artigo está organizado da seguinte forma: a seção 2 trata da organização dos experimentos realizados; a seção 3 descreve os algoritmos analisados; a seção 4 mostra os resultados obtidos na experimentação; a seção 5 apresenta a conclusão.

2. Método

Os algoritmos selecionados visam segmentar as imagens removendo os pixels do fundo, para isto foram levantadas questões como: quais destes se adaptaram às especificidades das

imagens, quais são os custos computacionais e a qualidade quantitativa e qualitativa dos resultados. Estas questões levam às seguintes hipóteses:

H01 Não há diferença entre algoritmos de segmentação aplicados às imagens UHD com fundo de cor única quanto aos pixels falso positivo/negativo

H02 Não há diferença entre algoritmos de segmentação aplicados às imagens UHD com fundo de cor única quanto ao custo computacional

HA1 Os algoritmos de segmentação não geram falsos positivos/negativos em comparação à imagem gabarito

HA2 O tempo de segmentação dos algoritmos não ultrapassa o tempo de aquisição do lote de imagens

Para responder as questões e as hipóteses, seguiu-se as etapas: definir o equipamento e software utilizados; selecionar os algoritmos de segmentação; avaliar quantitativamente o resultado da segmentação; analisar qualitativamente a segmentação das imagens; definir as bases das imagens; e o experimento.

2.1. Equipamento

O equipamento utilizado foi um notebook com processador Intel Core I7-4910MQ com 2.9GHz, 32GB de RAM, placa de vídeo NVIDIA Quadro K3100M e Intel HD Graphics 4600, sistema operacional Windows 10 PRO 64bits. Nenhum outro processo concorreu pelos recursos computacionais durante os testes. Os algoritmos selecionados, discutidos na seção 3, estão presentes como Toolbox do Matlab, versão R2015b, e no OpenCV-python 3.0 com a versão 3.6.2 do python.

2.2. Seleção dos Algoritmos

A partir das bibliografias encontradas e que abordam algoritmos de segmentação [Pedro et al. 2013][Chauhan et al. 2014], tentou-se mapear aqueles que são amplamente utilizados. Os resultados, as bases e as *strings* de busca podem ser vistas na Tabela 1.

Tabela 1. Resultado da Busca nas Bases

Base	Quantidade	String
IEEE - journals	2857	("image segmentation"NEAR ("review"OR "survey"))
ACM - journals	281	content.ftsec:("image segmentation") AND content.ftsec:("reviewsurvey")
Springer - Image Processing and Computer vision	4317	("image segmentation"AND ("review"OR "survey"))
Science Direct - journals	1163	"image segmentation"AND ("review"OR "survey")

Fonte: Autoria Própria

Com os algoritmos selecionados, deu-se início ao processo de utilização, verificando o tempo, principalmente, pois não pode ser elevado. O tempo é considerado elevado quando o resultado passa para a casa de minutos, com exceção daqueles que levam um determinado tempo para treinamento, mas ao executar o teste esse tempo cai drasticamente.

Um algoritmo que leva aproximadamente quarenta segundos para treinar e menos de dez em um teste, pode ser selecionado. Já, um algoritmo que não tem como diferenciar um tempo de treinamento com o de teste e este leva 1(um) minuto ou mais, pode ser

descartado, pois ao multiplicar o tempo pela quantidade de imagens do Lote, pode acarretar um tempo superior a uma semana de processamento. Inviável para os propósitos das aplicações.

Visando testar as técnicas de segmentação com abordagens diferentes, segundo encontrado em [Chauhan et al. 2014], selecionou-se algoritmos de: Detecção de Bordas (Filtros Canny, Sobel X, Sobel Y, e Laplaciano); Orientados à região (Watershed e Graph Cut); Classificadores (K-means e SVM); Estatístico (Bayes); e Texturas (Texture Segmentation).

Dentre os algoritmos selecionados existem aqueles que necessitam da etapa de treinamento, outros exigem interação manual. Estas especificidades são consideradas, pois todo o tempo, desde o início do processo (abertura das imagens pelo usuário até o resultado final) será levado em conta na escolha dos algoritmos.

2.3. Avaliação quantitativa das imagens

Nas segmentações, a taxa de pixels mantidos ou removidos das imagens é um dos pontos avaliados. Um pixel pode ser segmentado corretamente quando este é mantido na imagem e faz parte do objeto ou quando é removido e não faz parte do objeto, ou seja, é um componente da cor do fundo.

Por outro lado, ele pode ser considerado como segmentado incorretamente quando este for um pixel mantido, mas deveria ser removido (falso positivo - FP), pois faz parte da cor do fundo; ou era um pixel que deveria ser mantido e foi removido da imagem, ou seja, ele fazia parte do objeto, mas foi removido (falso negativo - FN).

A contabilização dos falsos positivos e negativos se dará por meio da comparação entre a imagem segmentada manualmente (*ground truth*) e a segmentada automaticamente. Considerando a inviabilidade de segmentar manualmente todas as imagens, optou-se por fazer a análise quantitativa por amostragem. Desta forma a análise quantitativa foi composta por quatro índices: corretos, incorretos, FP e FN.

2.4. Avaliação qualitativa das imagens

Além de levar em consideração as taxas definidas na avaliação quantitativa, uma vez que para os mesmos índices o impacto visual pode ser diferente, faz-se a verificação de como estes pixels foram segmentados.

Uma distribuição Esparsa (DE) de pixels incorretos tende a ser menos agressiva visualmente do que uma concentração destes. Assim, definiu-se cinco situações de distribuição de pixels incorretos, sendo elas:

DE: Distribuição Esparsa dos pixels incorretos - estes podem ser FP, quando alguns pontos dispersos do fundo são mantidos, ou FN, quando alguns pixels esparsos do objeto são removidos;

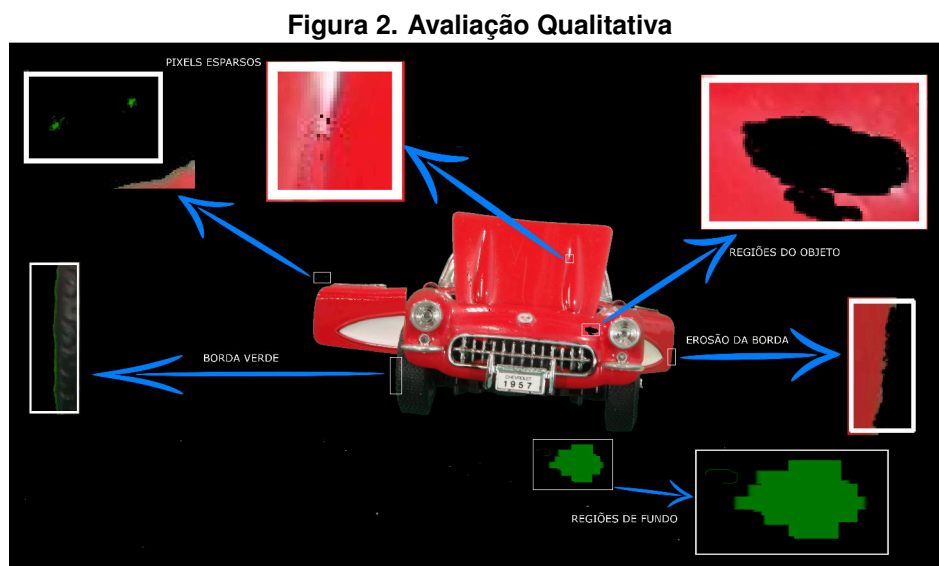
BV: Borda Verde - acontece quando pixels da cor do fundo são mantidos nas bordas do objeto, provocando uma visualização similar a uma sobra verde, tratando-se de um FP;

EB: Erosão da Borda - acontece quando pixels do objeto são removidos indevidamente dos seus limites, causando um serrilhamento, trata-se de um FN;

RF: Regiões do Fundo - acontece quando pixels da cor do fundo são mantidos em regiões de proximidade, trata-se de um FP;

RO: Regiões do Objeto - acontece quando pixels do objeto são removidos em regiões de proximidade, trata-se de um FN.

Essas métricas dão base para a análise qualitativa e tentam não deixá-la como um parâmetro subjetivo. A Figura 2 apresenta as situações descritas por estas métricas.



Fonte: Autoria Própria

2.5. Definir as bases das imagens

Usou-se seis bases de imagens de objetos, quantidades e tamanhos diferentes, como apresentado na Tabela 2. Optou-se por esta abordagem, pois os algoritmos devem ser capazes de resolver o problema de segmentação independentemente do tipo de objeto, suas cores e formas.

Tabela 2. Bases de imagens

Nome	Quantidade	Tamanho da Imagem	Tamanho da base
Cesta de Frutas	19	1632 x 1224	17.1 MB
Corvette	17	4608 x 2592	64.4 MB
Crânio - A	18	2395 x 1985	118.2 MB
Crânio - B	4000	2395 x 1985	2644 MB
Estátua Feminina	20	525 x 635	8.2 MB
Galinha	19	1632 x 1224	16.8MB
Shelby	21	1320 x 800	23.4 MB

Fonte: Autoria própria

A quantidade de imagens para as etapas 1 e 2 foi considerada adequada para o propósito de determinar quais algoritmos não passariam para a terceira. Mesmo que o trabalho tenha sido executado para uma ou até vinte imagens, deve-se considerar que elas são compostas por uma quantidade de pixels que torna o trabalho custoso. Assim como, o lote podendo conter até as quatro mil, como citado, faz com que o custo computacional aumente.

2.6. Experimento

O experimento está dividido em três etapas. Na primeira, deseja-se eliminar os algoritmos que apresentem pior desempenho em tempo de execução e qualidade do resultado, esta baseada nas métricas apresentadas.

A segunda etapa objetiva selecionar o algoritmo para a fase de validação, ela abrangerá todos os parâmetros de custo computacional, interação com o usuário e análise quantitativa e qualitativa dos resultados. A terceira etapa será a validação do algoritmo selecionado.

Para a primeira etapa, cada algoritmo será aplicado sobre uma imagem de cada base, realizando a coleta de dados para as variáveis: tempo de treinamento (para aqueles com a etapa de treinamento), tempo de execução, necessidade de interação, análise qualitativa.

Os algoritmos que forem selecionados para a segunda etapa, serão aplicados para os lote que variam de dezessete a vinte e uma imagens, como apresentado na subseção 2.5. Nesta etapa serão utilizadas as variáveis aplicadas na primeira etapa, além da análise quantitativa e consumo de memória.

Na etapa de validação, analisar-se-á o comportamento do algoritmo para um lote gerado pelo SSCC de quarto mil imagens. Uma vez que a qualidade já foi avaliada, nesta etapa apenas o tempo e consumo de memória serão levados em conta.

3. Algoritmos Testados

Como citado, os algoritmos para segmentação estão classificados em detectores de bordas, orientados à região, classificadores e baseados em textura. Escolheu-se ao menos um algoritmo de cada classe e alguns algoritmos foram testados em configurações diferentes, como apresentado na Tabela 3.

Tabela 3. Algoritmos Testados

Nome	Fonte	Interação	Treinamento	Classe
Canny	OpenCV	Não	Não	Detecção de Borda
Sobel X	OpenCV	Não	Não	Detecção de Borda
Sobel Y	OpenCV	Não	Não	Detecção de Borda
Laplaciano	OpenCV	Não	Não	Detecção de Borda
Watershed	OpenCV	Sim	Não	Orientado a Região
Graph Cut	OpenCV	Sim	Não	Orientado a Região
K-Means Supervisionado	Matlab	Sim	Sim	Classificador
K-Means Não Supervisionado	OpenCV	Não	Sim	Classificador
K-Means Não Supervisionado	Matlab	Não	Sim	Classificador
SVM	Matlab	Sim	Sim	Classificador
Textura	Matlab	Não	Não	Textura

Fonte: Autoria própria

Algoritmos baseados em detecção de bordas identificam alterações nas cores dos pixels de uma região da imagem de forma que sejam detectadas suas bordas [Wu and Leahy 1993]. Espera-se com estes algoritmos que seja possível definir os limites que separam o objeto do fundo. Dentre os algoritmos desta classe foram selecionados: filtro Canny [Canny 1986, Chen et al. 2014], Sobel X, Sobel Y [Jo and Lee 2012] e Laplaciano, todos disponíveis na biblioteca OpenCV.

Algoritmos orientados à regiões partem do princípio que a similaridade entre os pixels podem definir regiões [Ahuja 1996]. Sendo assim, as imagens obtidas pelo SSCC possuem fundo de cor única e portanto estes algoritmos podem separar o objeto do fundo, mesmo que partes internas do objeto também sejam separadas entre si. Estes algoritmos precisam de interação com o usuário para determinar as regiões [Tian and Yu 2016]. Foram selecionados os algoritmos Watershed [Shafarenko et al. 1997] e Graph Cut [Boykov and Jolly 2001].

Classificadores podem ser aplicados sobre pixels de imagens separando-os em classes. Estes podem ser usados de forma supervisionada, quando o usuário escolhe um conjunto de pixels para cada classe, ou não supervisionada, quando são escolhidos as classes de maneira automática. No OpenCV foi testado o algoritmo K-Means [Xu and Wunsch 2005, Xu et al. 2017] não supervisionado e no Matlab foram utilizados o K-means supervisionado e não supervisionado além do SVM [Huang et al. 2012, Tsai et al. 2006] que necessita da interação para escolha das classes, ou seja de forma supervisionada.

Como visto em [NUNES and CONCI 2007], a textura é um característica utilizada pela visão humana no processo de reconhecimento e classificação de imagens. Texturas referem-se a homogeneidade/repetição de padrões visuais sobre uma região ou impressão de rugosidade e contraste, não analisando apenas cor ou intensidade. Usou-se o filtro de entropia do Matlab, que calcula a organização de pixels, para separar as texturas. Esperando-se que este filtro reconheça todo o fundo como uma única textura.

4. Resultados

Como apresentado na seção 2 o experimento está dividido em três etapas. O objetivo da primeira etapa foi desconsiderar aqueles algoritmos que apresentaram pior desempenho em tempo de execução e qualidade do resultado. Os algoritmos que não foram excluídos na etapa anterior foram testados com um conjunto maior de imagens, selecionando-se um algoritmo para a validação.

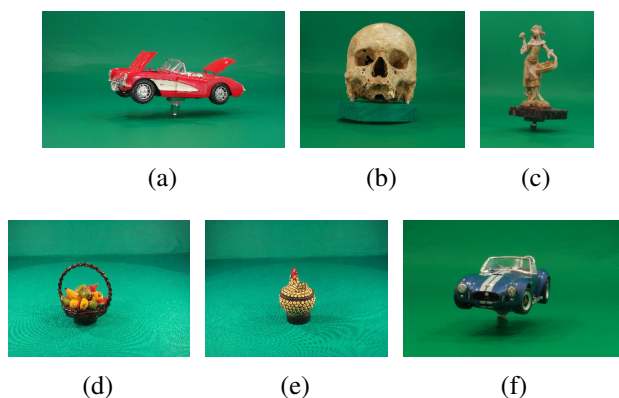
4.1. Primeira etapa

Sobre as imagens (Figura 3) foram aplicados todos os algoritmos e coletados os dados de tempo de treinamento, execução, necessidade de interação e análise qualitativa.

Os dados coletados para o algoritmo Canny variaram entre um tempo de 0.03 e 0.44 segundos, este algoritmo não necessita de interação, quanto análise qualitativa apresentou DE, EB e RF. Ele não foi capaz de separar alguns objetos do fundo e em alguns casos detectou bordas na textura da cor do fundo, além disto, necessita-se outro algoritmo para efetivar a segmentação, inviabilizando sua utilização. Uma imagem segmentada é apresenta na Figura 4(a).

O algoritmo Sobel X e Y tiveram seus tempos variaram entre 0.04 e 1.52 segundos, não necessitam de interação manual, em relação à análise qualitativa, detetou RF, BV e não houve definição da borda do objeto, pois manteve RF e EB na região de borda, além disto, necessita-se outro algoritmo para efetivar a segmentação. Estes algoritmos também foram eliminados para segunda etapa. Uma imagem segmentada é apresenta na Figura 4(b).

Figura 3. (a) Corvette, (b) Crânio, (c) Estátua Feminina, (d) Cesta de Frutas, (e) Galinha e (f) Shelby.



Fonte: Autoria própria

O algoritmo Laplaciano obteve tempo entre 0.02 e 0.65 segundos, não necessita de interação manual, sua análise qualitativa apresentou EB e RF de forma a não ser possível uma segmentação adequada do objeto, além disto, necessita-se outro algoritmo para efetivar a segmentação. Logo este algoritmo também foi excluído da segunda etapa. Uma imagem segmentada é apresenta na Figura 4(c).

O algoritmo Watershed apresentou tempo entre 0.03 e 0.89 segundos, não necessita de interação manual, sua análise qualitativa em relação aos parâmetros definidos (DE, BV, EB, RF e RO) não se enquadrou, no entanto, como visto na Figura 4(d), o fundo foi separado em várias classes, o objeto, em alguns casos também houve separação em mais de uma classe, em outros juntou o fundo com o objeto. Portanto, o algoritmo foi excluído da segunda etapa.

O algoritmo Graph Cut apresentou tempo entre 4.44 e 1071.56 segundos, necessita de interação manual para definir uma região onde está o objeto de interesse, na análise qualitativa ele apresentou BV e RF, contudo o resultado é dependente da região selecionada pelo usuário, o que dificulta a automatização do processo. Considerando o tempo e qualidade do resultado, optou-se por excluí-lo da segunda etapa. Uma imagem segmentada é apresenta na Figura 4(e).

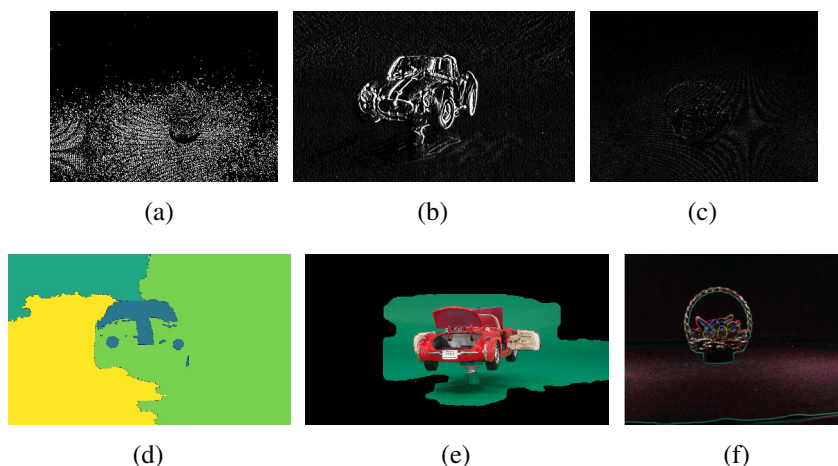
Outro algoritmo excluído da segunda etapa foi o de Textura. Seu tempo foi entre 6.00 e 89.21 segundos, não há necessidade de interação manual, o resultado da análise qualitativa apresentou BV e não houve definição da borda do objeto, pois manteve RF e EB na região de borda. Uma imagem segmentada é apresenta na Figura 4(f).

4.2. Segunda etapa

Os algoritmos K-means e SVM foram aqueles classificados para esta etapa. Além das variáveis analisadas na etapa anterior, foram coletados dados da análise quantitativa e consumo de memória. A partir deste momento foram usadas mais de uma imagem por objeto, formando um lote.

As implementações do algoritmo K-means obtiveram performance diferenciadas entre si, o supervisionado apresentou, baseado nos parâmetros definidos, resultado melhor do que os não supervisionados. A supervisão ocorre em apenas uma imagem para obter

Figura 4. (a) Canny, (b) Sobel, (c) Laplace, (d) Watershed, (e) Graph Cut e (f) Textura.



Fonte: Autoria própria

os pixels referentes a cada classe, determinando os centróides que serão aplicados às outras imagens do lote. Este processo viabiliza a automatização. O SVM parte do mesmo princípio de supervisão descrita.

O K-means não supervisionado implementado no OpenCV apresentou tempo de execução entre 3.5 e 111.88 segundos para o treinamento, para determinar os centróides. O tempo total para o lote, ou seja, aplicar os centróides a todas as imagens, adicionado ao tempo de treinamento, para a base que levou mais tempo, que foi a do Corvette: 111.88 (treinamento) + 665.68 = 777.56 segundos. Com sementes aleatórias, este algoritmo apresentou resultados irregulares para diferentes imagens.

Na análise quantitativa foi realizada por amostragem com três imagens de cada base, para a base do crânio-A cada imagem tem um total de pixels de 4.754.075, em média as imagens apresentaram 1.906.784 (13.368%) pixels incorretos, sendo 2.596 (0.054%) FP e 632.999 (13.314%) FN. Portanto, a porcentagem de pixels corretamente segmentados foi de 86.632%. A análise qualitativa resultou em DE, BV, EB, RF e RO dependendo das sementes. Uma imagem segmentada é apresentada na Figura 5(a).

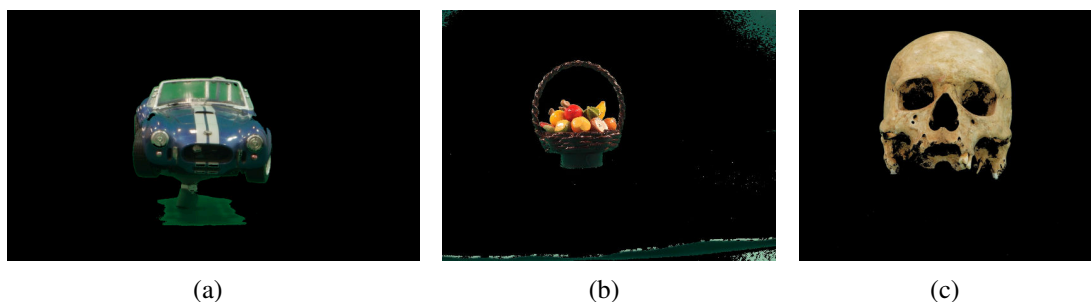
O K-means não supervisionado implementado no Matlab apresentou tempo de execução entre 7.2 e 80.00 segundos para o treinamento, para determinar os centróides. O tempo total para o lote adicionado ao tempo de treino para a base do Corvette: 80.00 (treino) + 544.00 = 624 segundos. Assim como os resultados do K-means implementado no OpenCv, apesar do tempo um pouco menor, foram irregulares para diferentes imagens. Uma imagem segmentada é apresentada na Figura 5(b).

Diferentemente, o K-means supervisionado apresentou resultados mais consistentes. O tempo de interação varia conforme o usuário, mas para fim de cálculo este foi considerado de 40 segundos para seleção de 50 pixels por classe. O tempo total de processamento do lote: 40.00 (interação) + 80.00 (treino) + 547.00 = 667.00 segundos.

Na análise quantitativa foi realizada por amostragem com três imagens de cada base, para a base do crânio-A cada imagem tem um total de pixels de 4.754.075, em

média as imagens apresentaram 172.775 (3.633%) pixels incorretos, sendo 111 (0.002%) FP e 172.665 (3.631%) FN. Portanto, a porcentagem de pixels corretamente segmentados foi de 96.367%. Quanto a análise qualitativa, apresentou DE, BV em alguns pontos e RO. Uma imagem segmentada é apresentada na Figura 5(c).

Figura 5. (a) K-means Não Supervisionado no OpenCV, (b) K-means Não Supervisionado no Matlab e (c) K-means Supervisionado no Matlab.



Fonte: Autoria própria

4.3. Terceira etapa

Dentre todos os algoritmos, o SVM foi o que obteve melhores resultados nas etapas anteriores. Ele foi validado na base do crânio-B que contém quatro mil imagens. O tempo de interação varia conforme o usuário, mas para fim de cálculo este foi considerado de 40 segundos para seleção de 50 pixels por classe.

O tempo de treino para 100 pixels, sendo 50 para cada classe, foi de 0.04 segundos e o tempo total de processamento para as quatro mil imagens foi de: 40 (interação) + 0.04 (treino) + 4429.93 = 4469.97 segundos. Em minutos isso perfaz 74.50 minutos, ou 1h e 15 minutos.

Quanto ao consumo de memória, o importante é que o algoritmo seja suportado pela quantidade disponível na máquina. Durante a medição foi constatado que o Matlab, sozinho, consumiu 366 MB e ao se executar o programa de segmentação o consumo variou entre 544 e 1298 MB (matlab + algoritmo + dados).

Na análise quantitativa foi realizada por amostragem com dez imagens da base crânio-B cada imagem tem um total de pixels de 4.754.075, em média as imagens apresentaram 6979 (0.14%) pixels incorretos, sendo 5004 (0.10%) FP e 1974 (0.04%) FN. Portanto, a porcentagem de pixels corretamente segmentados foi de 99.853%.

A análise qualitativa apresentou DE apenas de FP em algumas imagens. A Figura 6 apresenta o resultado do algoritmo SVM. A Figura 6(a) é a original e a Figura 6(b) é a segmentada, onde os retângulos em vermelho mostram o problema dos pixels esparsos que não foram segmentados.

5. Conclusão

O presente trabalho abordou o problema de segmentação de imagens geradas pelo SSCC em UHD com cor de fundo única para diferentes tipo de objetos. A estas imagens foram aplicados algoritmos que buscou-se na literatura por meio de pesquisa em bases de trabalhos científicos. Os algoritmos selecionados envolviam diferentes técnicas e durante o experimento foram aplicados a sete bases de imagens.

Figura 6. (a) Imagem Original e (b) Imagem Segmentada



(a)

(b)

Fonte: Autoria própria

Foram definidas três etapas para a escolha e validação de um algoritmo para ser usado no SSCC. Na primeira etapa foram excluídos aqueles que não apresentaram viabilidade de uso, já na segunda, foi segmentado um lote de imagens para comparação dos algoritmos. Dentro todos os algoritmos, o SVM se mostrou melhor do que os outros em qualidade e consumo de recursos computacionais. Na terceira etapa, o algoritmo SVM foi validado com uma base de quatro mil imagens.

No que se refere às variáveis definidas para o trabalho, o SVM se mostrou a alternativa mais adequada na resolução do problema proposto. Pois, ele foi capaz de segmentar um lote de imagens em um tempo inferior ao necessário para a aquisição da mesma quantidade de imagens pelo SSCC, de forma que o processo de segmentação pode ser realizada ao mesmo tempo em que ocorre a aquisição.

Neste trabalho percebeu-se que a cor de fundo em verde interferiu no processo de segmentação, pois causou reflexo nos objetos digitalizados, portanto estuda-se a alteração desta cor. Sendo assim, com a possibilidade de uso de cores diferentes, dependendo inclusive das cores presentes nos objetos, o trabalho de segmentação deverá ser reavaliado para essa nova configuração. Além disso, outros algoritmos de segmentação podem ser testados.

Portanto, refuta-se a H01, pois houve diferença entre os algoritmos quanto aos FP e FN. Refuta-se a H02, pois há diferença entre algoritmos quanto ao custo computacional. Confirma-se as hipóteses HA1 e HA2, uma vez que o SVM atingiu a segmentação desejada em algumas imagens, não apresentando os FP e FN, além do tempo estar dentro do esperado para um lote, ou seja, não ultrapassando o tempo de aquisição para o mesmo lote.

Referências

Ahuja, N. (1996). A transform for multiscale image segmentation by integrated edge and region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12):1211–1235.

- Boykov, Y. Y. and Jolly, M. P. (2001). Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 105–112 vol.1.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698.
- Chauhan, A. S., Silakari, S., and Dixit, M. (2014). Image segmentation methods: A survey approach. In *2014 Fourth International Conference on Communication Systems and Network Technologies*, pages 929–933.
- Chen, H., Ding, H., He, X., and Zhuang, H. (2014). Color image segmentation based on seeded region growing with canny edge detection. In *2014 12th International Conference on Signal Processing (ICSP)*, pages 683–686.
- Huang, G. B., Zhou, H., Ding, X., and Zhang, R. (2012). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(2):513–529.
- Jo, J. H. and Lee, S. G. (2012). Sobel mask operations using shared memory in cuda environment. In *2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012)*, pages 289–292.
- Lantz, E. (2007). A survey of large-scale immersive displays. In *Proceedings of the 2007 Workshop on Emerging Displays Technologies: Images and Beyond: The Future of Displays and Interacton*, EDT '07, New York, NY, USA. ACM.
- NUNES, E. O. and CONCI, A. (2007). Texture segmentation and edge location of regions in multiband images. *IEEE Latin America Transactions*, 5(3):184–191.
- Pedro, R. W. D., Nunes, F. L. S., and Machado-Lima, A. (2013). Using grammars for pattern recognition in images: A systematic review. *ACM Comput. Surv.*, 46(2):26:1–26:34.
- Shafarenko, L., Petrou, M., and Kittler, J. (1997). Automatic watershed segmentation of randomly textured color images. *IEEE Transactions on Image Processing*, 6(11):1530–1544.
- Tian, X. and Yu, W. (2016). Color image segmentation based on watershed transform and feature clustering. In *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pages 1830–1833.
- Tsai, C.-F., McGarry, K., and Tait, J. (2006). Claire: A modular support vector image indexing and classification system. *ACM Trans. Inf. Syst.*, 24(3):353–379.
- Wu, Z. and Leahy, R. (1993). An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1101–1113.
- Xu, J., Han, J., Nie, F., and Li, X. (2017). Re-weighted discriminatively embedded k -means for multi-view clustering. *IEEE Transactions on Image Processing*, 26(6):3016–3027.
- Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678.

APÊNDICE B – Fotografia Digital

Fotografia Digital

Diogo Roberto Olsen¹

¹Pontifícia Universidade Católica do Paraná - PUCPR
Programa de Pós-Graduação em Informática - PPGIa

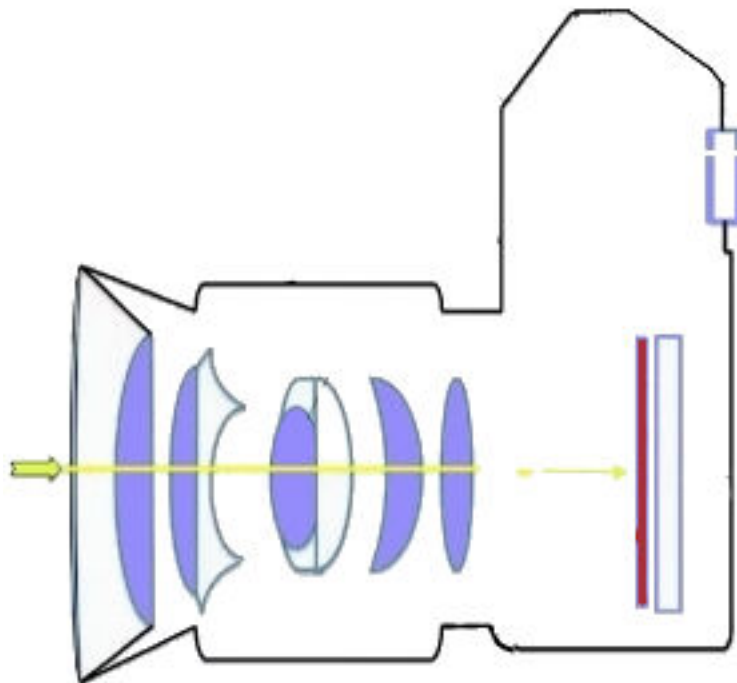
diogoolsen@gmail.com

1. Introdução

Uma opção para aquisição digitalização de objeto são as fotografias digitais. Estas fotografias podem ser adquiridas e processadas de forma a garantirem acuidade visual, porém, por serem bidimensionais, não são interativas por natureza.

O funcionamento básico da fotografia digital baseia-se em dois componentes principais das câmeras: a luz é capturada com um conjunto de lentes (chamado de conjunto óptico) e estes feixes de luz são direcionados para um sensor que faz a digitalização das informações. A Figura 1 apresenta este processo, na qual os raios de luz são representados em amarelo, o conjunto óptico em lilás e o sensor em vermelho.

Figura 1. Aquisição fotográfica.



Fonte: Adaptado de: [Hashen 2017].

Para entendimento de fotografia digital é importante conhecer alguns conceitos que vão desde o entendimento do hardware da câmera, até o processo de configuração desta. Para isso serão usados como base os seguintes livros [Hedgecoe 2013, Freeman 2001, Kelby 2011a, Kelby 2011b, Kelby 2011c, Kelby 2011d].

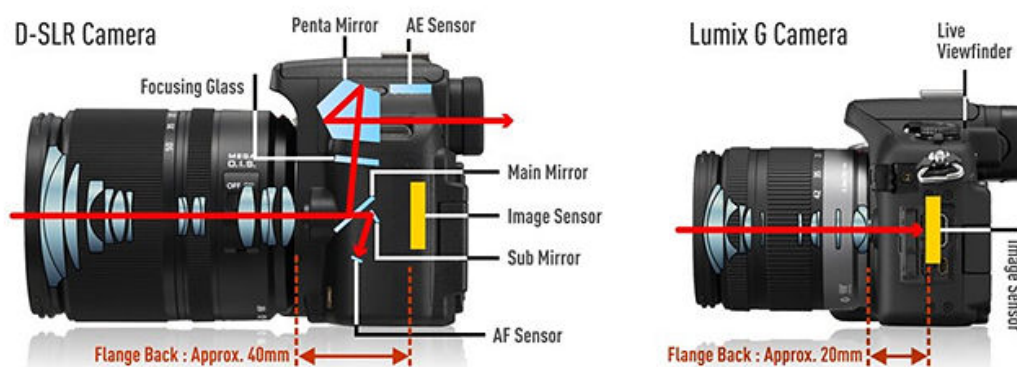
Hardware - Tipos de Câmeras

As câmeras possuem diversos outros componentes, dependendo de marca e modelo, como: sensores de foco, processador, memória de armazenamento, tela, motores, espelhos, estabilizadores de imagens, flash, entre diversos outros. Além disso, atualmente, existem duas classes de câmeras digitais: D-SLR e Mirrorless. A Figura 2 apresenta as duas câmeras.

D-SLR (Reflex Monobjetiva Digital - *Digital Single-Lens Reflex*) : Nestas câmeras o sensor da imagem fica atrás de um espelho que se move, expondo o sensor, apenas no momento de adquirir a imagem. Devido ao movimento mecânico, a vida útil de uma câmera D-SLR gira entre cem mil e duzentos mil cliques. Por exemplo, segundo o manual, a câmera D-SLR Nikon D7100 possui uma vida útil de duzentos mil fotos¹;

Mirrorless (Sem espelho): Nestas câmeras o sensor de imagens fica exposto o tempo todo e, como não existe movimento mecânico durante a aquisição, são menos suscetíveis a problemas causados pelo número de fotos.

Figura 2. Câmeras DSLR e Mirrorless.



Fonte: Divulgação Panasonic <https://blogs.panasonic.com.au/consumer/10th-anniversary-of-lumix-g-series-cameras/>.

Hardware - Objetivas

Alguns modelos de câmeras permitem a troca do conjunto óptico, através da troca de objetivas - ou lentes - que são o conjunto de todos os componentes ópticos e componentes mecânicos/elétricos que permitem sua regulação para uma determinada fotografia. A troca de objetivas faz com que a câmera possua um comportamento diferente, uma vez que altera o ângulo de visão da câmera é alterado.

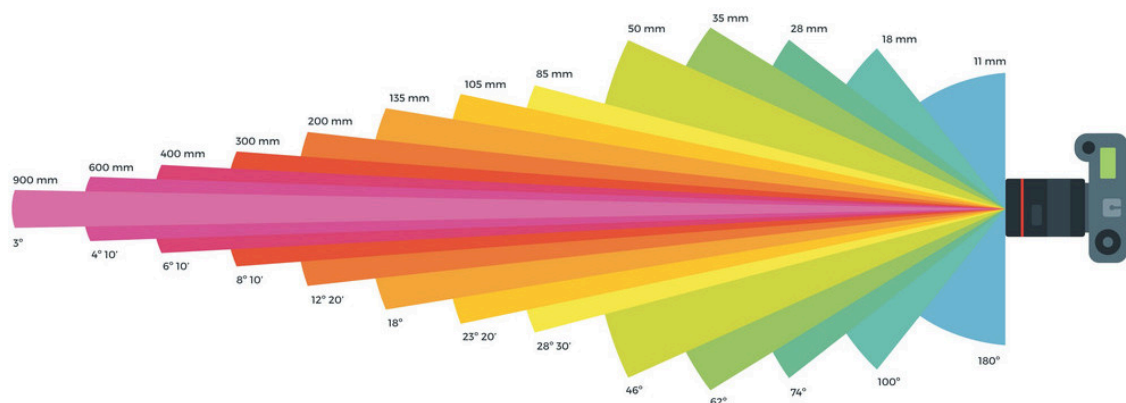
Algumas objetivas também possuem regulação de distância focal/ângulo de visão, alterando a amplitude de visualização da lente. A Figura 3 apresenta a relação entre distância focal e ângulo de visão de objetivas. A Figura 4 apresenta o efeito que a alteração da distância focal/ângulo de visão faz em fotografias de uma mesma posição.

¹<https://downloadcenter.nikonimglib.com/pt-br/products/27/D7100.html>

Distância focal : Distância, normalmente calculada em milímetros do ponto onde os feixes de luz convergem a um único ponto, até o sensor da câmera;

Ângulo de visão : Medida em graus, a distância focal altera o ângulo de visão da lente, quanto maior a distância focal, menor o ângulo de visão. Aumentar a distância focal, faz com que a cena fotografada pareça mais próxima, o que é popularmente conhecido como *zoom* da lente.

Figura 3. Objetivas - Distância focal e ângulo de visão.



Fonte: adaptada de Vector Stock

<https://www.vectorstock.com/royalty-free-vector/lens-focal-length-and-camera-zoom-lens-vector-7326276>.

Figura 4. Objetivas - Efeito de aproximação de objetivas.

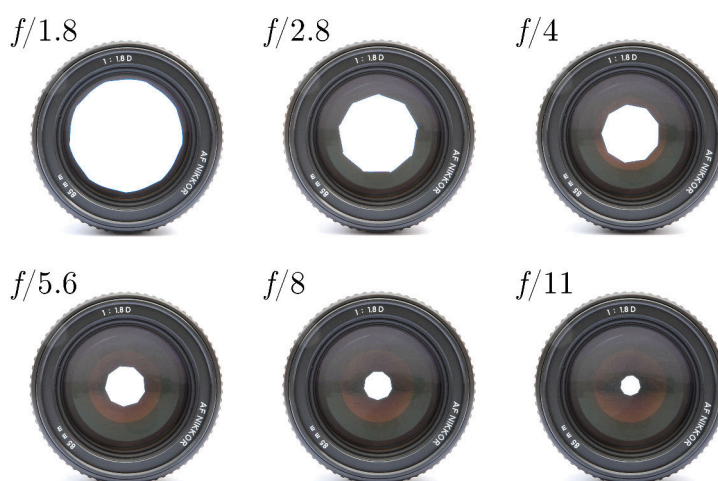


Fonte: Divulgação Nikon <https://www.nikon.com.br/learn-and-explore/a/tips-and-techniques/introduo-entendendo-a-distncia-focal.html>.

Hardware - Diafragma

O diafragma, também conhecido como abertura, controla a quantidade de luz que entra na câmera e atinge o sensor. Trata-se de um orifício com o tamanho da abertura controlado mecanicamente pela inserção/remoção de lâminas semicirculares que fazem com que mais/menos luz passe através dele. A Figura 5 apresenta uma objetiva com diferentes aberturas do diafragma.

Figura 5. Diafragma.



Fonte: Wikipédia - Alemanha https://de.wikipedia.org/wiki/Datei:Lenses_with_different_apetures.jpg.

Seu valor é calculado em pontos e denotado pela letra f seguido por “/”, “.” ou “:” e um valor numérico, como em $f/1.8$. O valor de f representa o número pela qual se divide a distância focal da objetiva para obter o diâmetro da abertura. Este processo garante que para quaisquer objetivas, determinado valor de f vai permitir a entrada da mesma quantidade de luz. Note-se que quanto maior o valor de f , menor a quantidade de luz que passará.

Nem todas as objetivas possuem os mesmos limites de abertura. Objetivas com aberturas máximas de $f/1.2$ a $f/2$ são consideradas lentes “claras”, pois deixam mais luz passar. Lentes com abertura máximas com valores maiores que dois são consideradas objetivas escuras, pois a maior abertura que a lente suporta não permite tanta passagem de luz.

A abertura do diafragma interfere diretamente em duas características da fotografia adquirida: fotometria e profundidade de campo. Estes fatores serão explicados no decorrer desta seção. Além disso a abertura interfere na nitidez da fotografia.

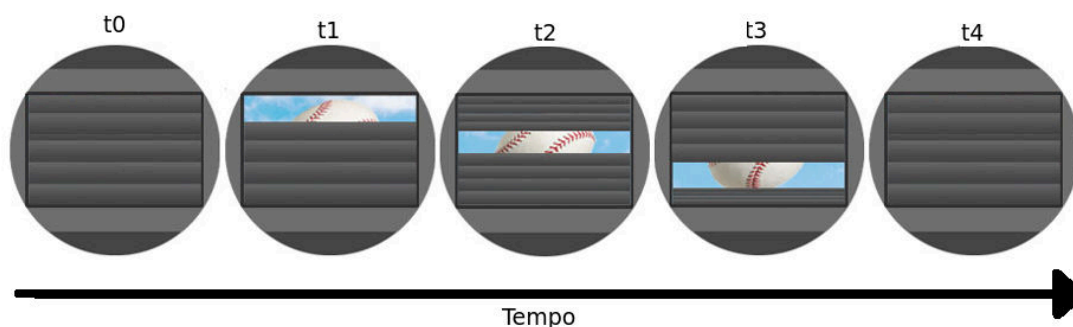
Alguns autores afirmam que quanto menor a abertura - ou seja, maior o valor de f - mais nítida é a fotografia [Freeman 2001]. Outros por sua vez defendem que cada lente possui um valor de f que oferece melhor nitidez e que este valor é “normalmente” (segundo o autor) dois passos abaixo da abertura máxima da objetiva [Kelby 2011a]. Será necessário fazer experimentos para determinar qual abertura das objetivas do projeto oferecem melhor nitidez.

Hardware - Obturador

O principal objetivo do obturador é controlar o tempo de exposição do sensor da câmera. Existem diversos formatos de obturador, como os centrais e os de janela (mais comuns atualmente). Estes equipamentos são barreiras físicas para a luz, eles se abrem e fecham com grande precisão de tempo e de forma que cada parte do sensor seja exposta pelo mesmo intervalo de tempo.

A Figura 6 apresenta um obturador com formato de janela. Nestes obturadores existem duas carreiras independentes, chamadas de cortina, dispostas uma na parte superior e outra a parte inferior. Inicialmente uma das cortinas cobre todo o sensor e no momento do disparo da foto esta cortina se recolhe expondo o sensor. Após um intervalo de tempo configurado na câmera, a outra cortina se move paralelamente à primeira cobrindo o sensor. Desta forma cada parte do sensor fica exposto pelo mesmo intervalo de tempo, garantindo uma iluminação homogênea.

Figura 6. Obturador.



Fonte: Adaptado de

<https://www.dpreview.com/forums/thread/4079862>.

O tempo que o obturador fica aberto é, normalmente, medido em segundos ou fração de segundos. Este tempo é chamado de velocidade do obturador. Além de controlar a quantidade de luz que atinge o sensor, esta velocidade é importante em fotos com movimento. Ao fotografar um carro em alta velocidade, se o obturador ficar aberto por um segundo, todo o deslocamento do carro será percebido pelo sensor e a foto ficará borrada e sem nitidez.

Hardware - Sensor

Atualmente existem duas tecnologias mais comuns de sensores para câmeras: CCD (*Charge Coupled Device*) e CMOS (*Complementary Metal Oxide Semiconductor*). A grande diferença entre eles é que o CCD envia sinais analógicos para um circuito externo processar enquanto que o CMOS processa os sinais e envia para o restante da câmera os sinais digitais que formarão a imagem. As principais características de um sensor são:

Tamanho físico : Quanto maior o sensor, mais luz irá incidir sobre ele e mais sensível à luz ele será, adquirindo fotos mais rapidamente. Além disso, sensores fisicamente

maiores são menos propensos a gerar ruídos nas imagens que sensores menores. A Tabela 1 apresenta o tamanho e área dos sensores mais comuns nas câmeras atuais com lentes intercambiáveis.

Tabela 1. Tamanho dos sensores (em milímetros)

Tipo do sensor	APS-C	DX	Super35	Full Frame
Largura	22,2mm	23,6mm	24,6mm	36mm
Altura	14,8mm	15,5mm	13,8mm	24mm
Área	328,5mm ²	366mm ²	339,5mm ²	864mm ²

Fonte: Autoria própria.

Quantidade de pixels : Quanto maior o número de pixels adquiridos por um sensor, maior o tamanho da imagem gerada e mais detalhes podem ser vistos na foto. O tamanho do sensor é medido em megapixel, que significam em torno de um milhão de pontos, dependendo do formato/proporção da imagem gerada;

ISO ²: Está associado à sensibilidade do sensor à luz, seus valores iniciam em 100 e são multiplicados por dois, de forma que a cada multiplicação a sensibilidade do sensor é dobrada. Cada sensor possui um limite máximo de ISO, por exemplo a Nikon D5³ possui iso máximo de 102400. Valores altos de ISO auxiliam em situações de pouca luz e fazem a aquisição ser mais rápida, porém geram mais ruído nas imagens se comparado à valores de ISOs mais baixos. Para fotos mais nítidas é recomendável um tempo de exposição maior e valores de ISO menores [Kelby 2011a].

Hardware - Fotômetro

Exposição é a intensidade de luz que chega ao sensor pelo tempo que o sensor fica exposto [Freeman 2001]. A exposição pode fazer fotos ficarem muito claras ou escuras, omitindo detalhes importantes. Alterações na exposição são muito usadas em fotos artísticas, porém neste trabalho o objetivo é adquirir a maior quantidade de detalhes possíveis, portanto a exposição da foto precisa privilegiar este aspecto. A Figura 7 apresenta uma mesma cena fotografada com diferentes exposições para ilustrar como a exposição altera uma fotografia.

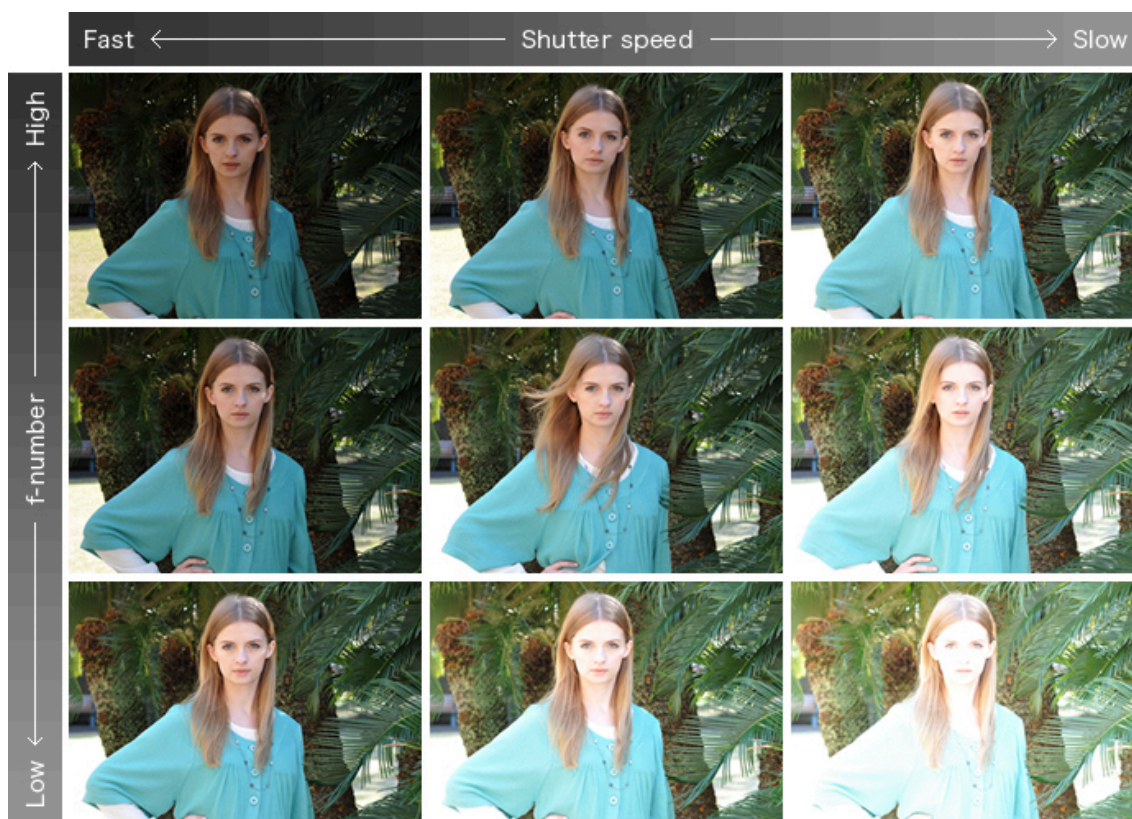
Além disso, fotografias não possuem a mesma intensidade luminosa em todos os pontos. Isso ocorre, por exemplo, ao fotografar um carro em uma estrada. Um pixel da estrada poderá ter uma intensidade luminosa muito menor - por ser escura e fosca - do que um pixel do farol ou cromados - por emitirem/refletirem muito mais luz do que o asfalto.

Para que seja possível o ajuste da exposição de uma fotografia, podem ser usados fotômetros, que medem a luz de uma cena. A unidade usada nos fotômetros é o ponto, e sua escala varia entre valores negativos, zero e positivos. Normalmente os pontos são divididos em terços. Uma fotografia considerada ideal está com o fotômetro marcando

²A sigla ISO faz referência à International Standards Organization <https://www.iso.org>, responsável pela padronização destes valores, de forma que sua seu nome não está associado diretamente à sensibilidade do sensor.

³<https://imaging.nikon.com/lineup/dslr/d5/spec.htm>

Figura 7. Exposição.

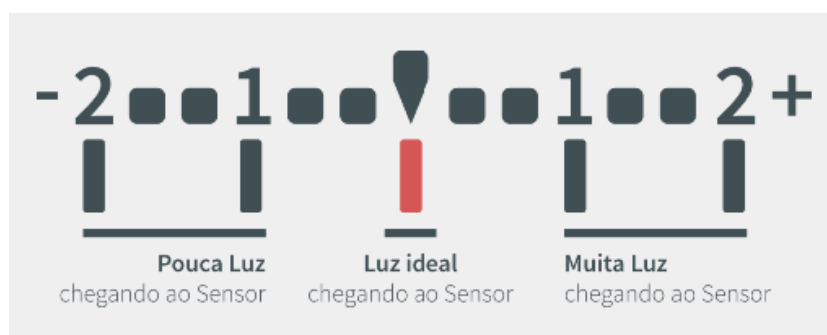


Fonte: Divulgação Nikon -

<http://imaging.nikon.com/lineup/dslr/basics/04/02.htm>.

zero. Valores negativos são considerados subexposição - fotos mais escuras - e positivos superexposição - fotos claras. A Figura 8 representa um fotômetro.

Figura 8. Fotômetro.



Fonte: <https://falandodefoto.com.br/o-que-e-fotometria-fotografia/>.

As câmeras atuais possuem fotômetros integrados - do tipo reflexivo. Estes equipamentos fazem a leitura da luz refletida pela cena fotografada e, por este motivo, já considera as cores, tons e texturas dos objetos fotografados. O inconveniente desta medição

é que para realizar a medição o fotômetro tentará normalizar a fotografia a um valor de cinza 18%, fazendo superfícies claras parecerem mais escuras e superfícies escuras mais claras. Estes fotômetros possuem 3 modos de medição:

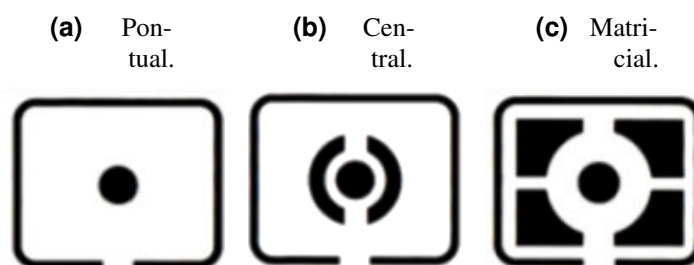
Pontual : Significa que apenas a região mais central da fotografia será considerada no cálculo realizado pelo fotômetro. Este modo é ideal para garantir que determinada parte da cena esteja com a fotometria correta, mesmo que outras partes não estejam com a exposição adequada. Ao tirar uma foto de um carro, pode-se fazer a fotometria na lataria do carro (objeto mais importante nesta foto) sabendo que o asfalto pode ficar muito escuro ou o céu muito claro;

Central : Considera a fotometria média da região central da foto, uma área maior que a pontual, de forma que pode ser enquadrado um objeto maior nesta região. Seu uso é similar ao da fotometria central;

Matricial : Considera uma média da imagem como um todo. Tenta adequar a quantidade de luz de toda a foto, de forma que a média da luminosidade seja considerada. Para fotografias com áreas muito diferentes, pode tender a não conseguir uma exposição adequada para nenhuma região da foto, mas para fotografias com intensidade luminosa homogênea, pode ser adequado.

A Figura 9 apresenta os modos de medição dos fotômetros das câmeras digitais. Os ícones podem ser diferentes para diferentes fabricantes, nas imagens é apresentado o padrão da Nikon. As regiões escuras dentro do quadro representam as regiões que serão levadas em consideração na fotometria.

Figura 9. Modos de medição da fotometria.



Fonte: Adaptado de:

<https://fotodicasbrasil.com.br/modos-de-medicao/>.

Para fotômetros reflexivos, existe ainda uma técnica chamada de leitura substitutiva. Nesta técnica não se usa o objeto a ser fotografado para o cálculo da fotometria, e sim um cartão conhecido como “cartão cinza médio”, que nada mais é do que um cartão com cinza 18%. Como o fotômetro usa este cinza médio como referência para o cálculo, ao medir a luz refletida por este cartão a fotografia tende a ficar mais natural, deixando regiões claras mais expostas e escuras subexpostas. A foto tende a ficar mais próxima da visão humana.

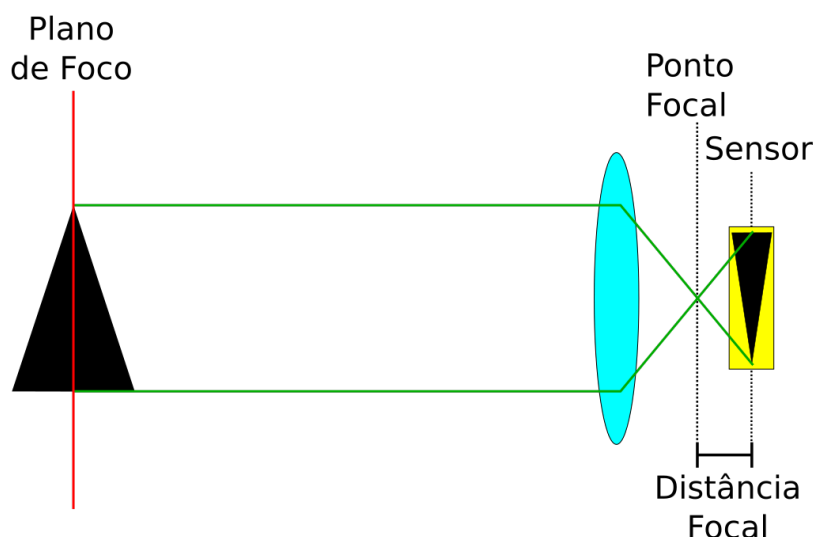
Além destes fotômetros internos/reflexivos, que medem a luz refletida, podem ser usados fotômetros externos, que medem a luz incidente em um determinado local a ser fotografado. Estes fotômetros, por medirem a luz do ambiente e não a refletida, tendem a sugerir a subexposição de objetos escuros e foscos; e a superexposição de objetos claros, brilhantes ou que emitem luz. Estes equipamentos só fazem a medição pontual, não

sendo possível uma medição da cena como um todo - exemplo da medição matricial. Estes fotômetros informam os parâmetros de configuração da câmera para determinada iluminação.

Configuração - Foco

Para aquisição da fotografia, a luz oriunda da cena atravessa os elementos ópticos da objetiva, converge em um ponto único entre as lentes e o sensor - ponto focal - e diverge até atingir o sensor, formando a imagem. Devido a este processo, apenas um determinado plano, chamado de plano de foco, forma perfeitamente a imagem no sensor. O que estiver antes ou depois do plano de foco forma imagens sem nitidez. A Figura 10 apresenta este processo.

Figura 10. Foco.



Fonte: Autoria Própria.

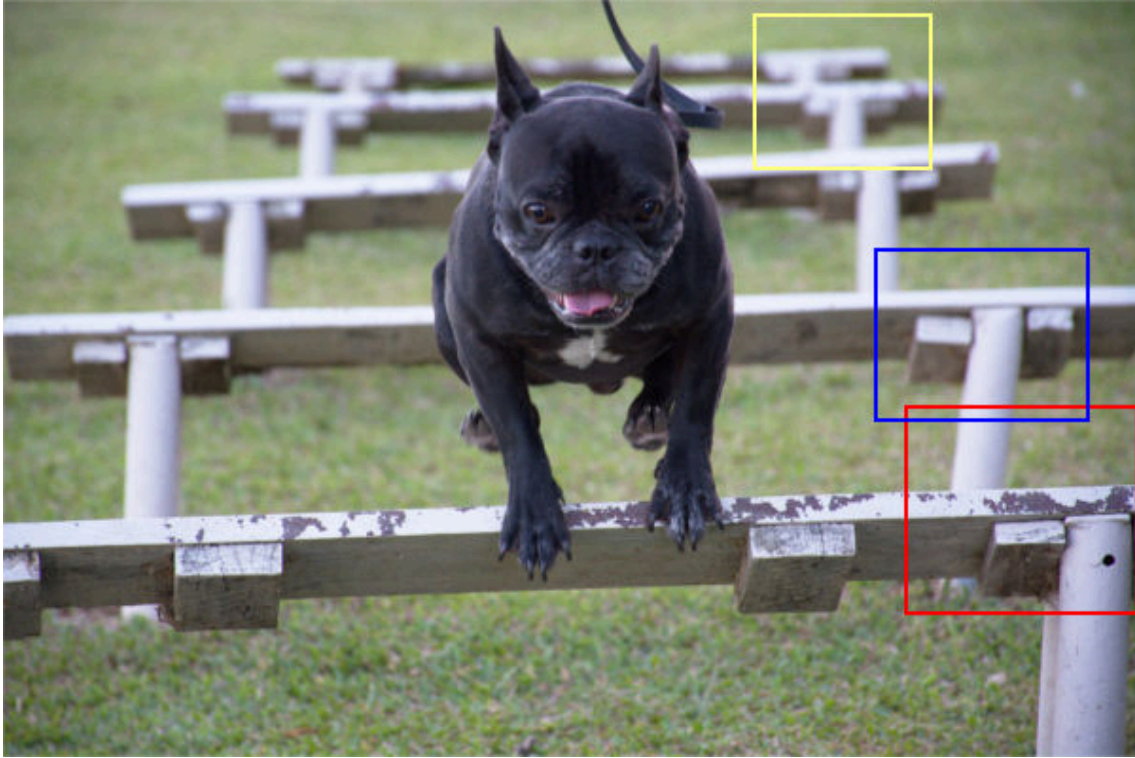
A falta de foco ocorre pela dispersão dos feixes de luz, de forma que desordena os feixes, transformando pontos em círculos de confusão, e conseqüentemente a imagem aparece borrada, embaralhando cor e forma da imagem gerada. Imagens adquiridas sem foco podem ser aprimoradas por software, mas o problema da falta de foco dificilmente pode ser corrigido. Quanto mais afastado do plano de foco, menos definido aparecerá na imagem.

Este efeito pode ser visto na Figura 11, na qual o plano focal está alinhado com o plano do primeiro obstáculo. É possível observar pelo recorte vermelho - Figura 11b - que é possível observar detalhes do obstáculo (cor, forma e textura). No recorte azul - Figura 11c - que está um pouco mais afastado do plano focal, não é possível observar detalhes como a textura. Na figura 11d, a mais afastada de todos os recortes, a imagem está mais degradada.

Nas câmeras digitais atuais existe a possibilidade de foco automático ou manual. Na focagem manual o usuário opera a objetiva e ajusta o foco conforme sua necessidade.

Figura 11. Alteração da nitidez gerada pelo afastamento do plano focal.

(a) Imagem original.



(b) Recorte vermelho.



(c) Recorte azul.



(d) Recorte amarelo.



Fonte: Autoria própria.

Na focagem automática os dois métodos mais comuns de *autofocus* usados nas câmeras atuais são:

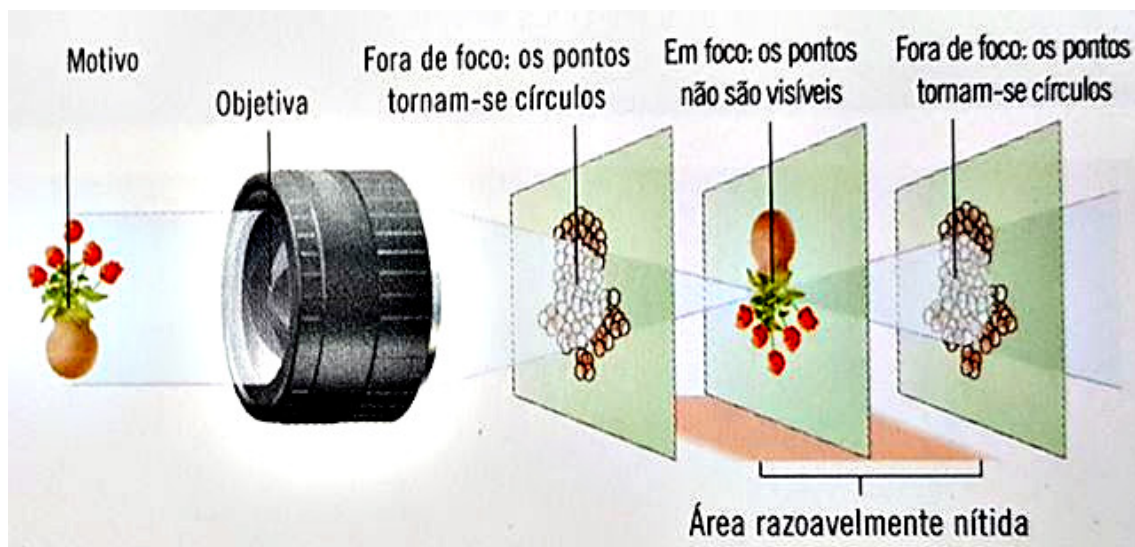
Autofoco ativo : utiliza um hardware específico que compara o contraste da imagem para verificar se a imagem está focada ou não e ajusta o elemento óptico se necessário, aproximando ou afastando o plano de foco. Este método não funciona em ambientes com pouca luz, ou emite uma luz específica para medir o foco na imagem;

Autofoco passivo : Utiliza infravermelho para medir a distância de um ponto até a câmera e assim ajusta o foco da objetiva com base na distância. É mais rápido que o método ativo e funciona mesmo em ambientes escuros, porém seu alcance é limitado.

Configuração - Profundidade de campo

Apesar de que objetivas só conseguem focalizar com precisão o plano focal, existe um limite no qual os círculos de confusão ainda são pequenos o suficiente para gerar imagens consideradas nítidas. Quanto mais próximo do plano focal, menores os círculos de confusão e mais aceitável é a nitidez da imagem gerada.

Figura 12. Profundidade de campo.



Fonte: [Hedgecoe 2013]

É possível entender a profundidade de campo usando a Figura 13. Ao analisar a Figura 13a percebe-se que o plano focal está situado próximo ao primeiro obstáculo. É possível verificar com o recorte vermelho (Figura 13b) que toda a taboa do primeiro obstáculo está com nitidez, é possível identificar a cor, forma e textura do objeto, assim como enxergar a sujeira presa na parte de traz da taboa.

Já no recorte azul (Figura 13c) não é possível identificar a textura do objeto. Ou seja a profundidade de campo desta fotografia conseguiu capturar toda a primeira taboa com nitidez, mas as outras estavam muito longe do plano focal e ficaram sem nitidez.

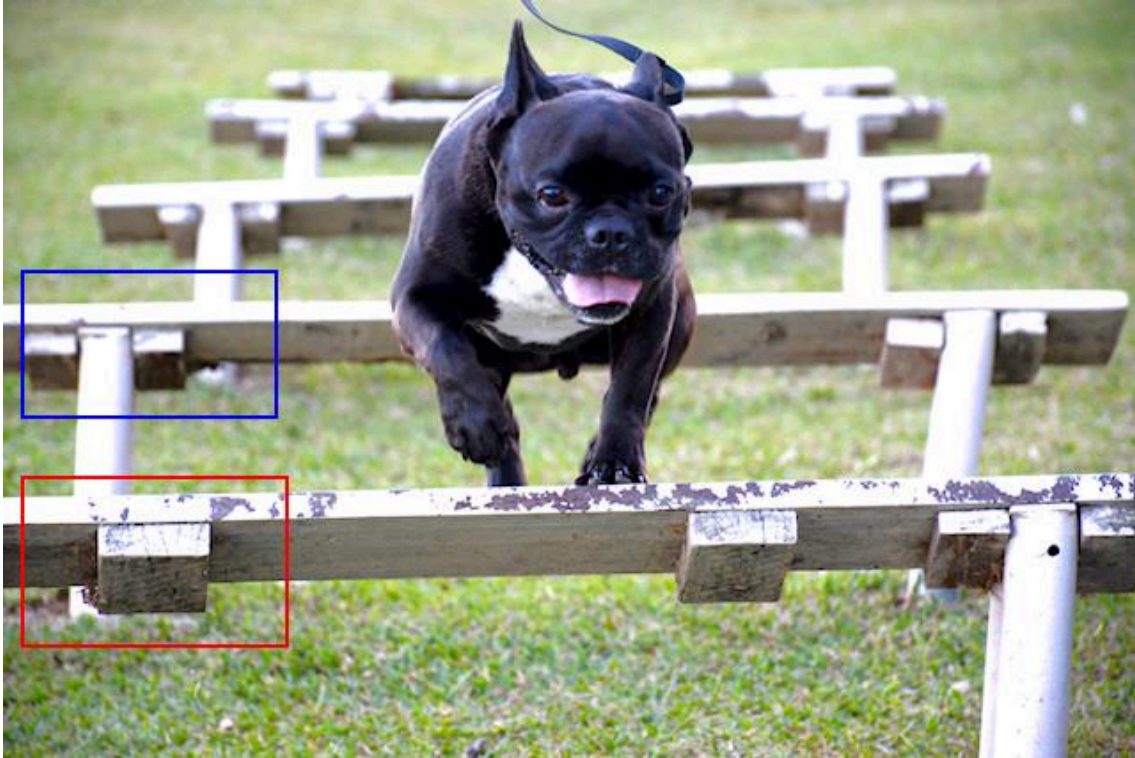
A profundidade de campo normalmente é maior do plano focal para traz do que do plano focal para frente, dependendo da construção do conjunto óptico, da câmera e da configuração usada. Existem três maneiras de alterar a profundidade de campo: abertura; distância focalizada; e distância focal.

A abertura do diafragma altera a profundidade de campo. Quanto mais aberto o diafragma, menor a profundidade de campo da foto. Esta alteração pode ser realizada sem alterar a objetiva nem a distância da câmera e o objeto/plano focal. A figura 14 apresenta como a profundidade de campo muda com a mudança na abertura da lente.

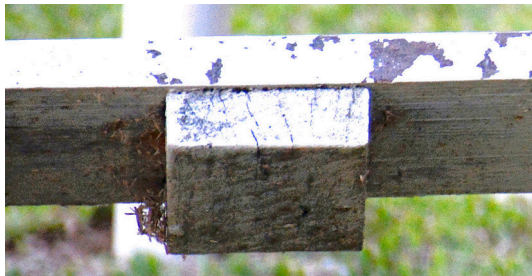
A profundidade de campo também é alterada pela distância focalizada - distância entre a câmera e o plano de foco. Quanto maior a distância focalizada, maior a profundidade de campo, porém para usar esta técnica também é necessário alterar a posição da

Figura 13. Profundidade de Campo.

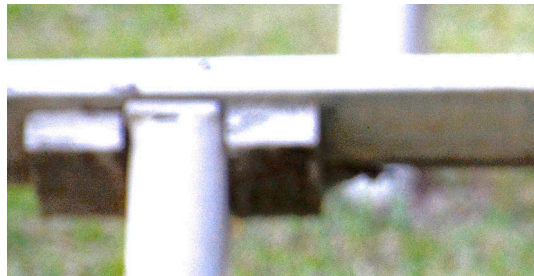
(a) Imagem original.



(b) Recorte vermelho.



(c) Recorte azul.



Fonte: Autoria própria.

Figura 14. Alterar a profundidade de campo - abertura.



Fonte: [Hedgecoe 2013]

câmera e/ou objeto. Isto irá alterar também o enquadramento da fotografia. A figura 15 apresenta este processo.

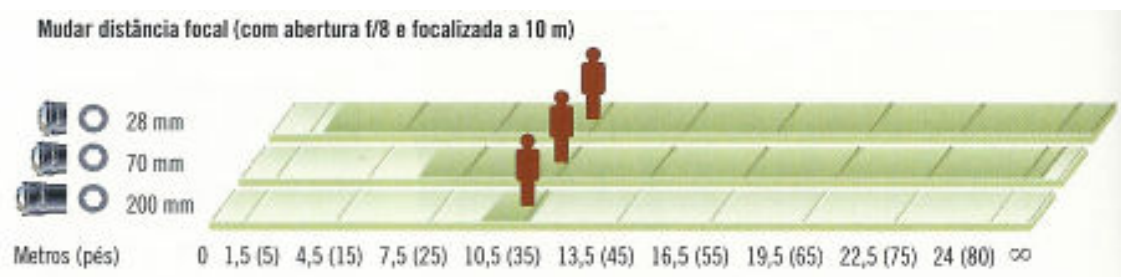
Figura 15. Alterar a profundidade de campo - distância focalizada.



Fonte: [Hedgecoe 2013]

A outra forma de alterar a profundidade de campo é alterando a distância focal. Isso pode ser feito alterando a objetiva ou, em objetivas com distância focal variável, configurando o objetivo. Este processo porém altera o enquadramento da fotografia. Quanto maior a distância focal, menor a profundidade de campo. A Figura 16 apresenta este processo.

Figura 16. Alterar a profundidade de campo - distância focal.



Fonte: [Hedgecoe 2013]

As três técnicas podem ser usadas separadas ou em conjunto para conseguir alcançar a profundidade de campo desejada. Porém é necessário considerar que ao alterar a distância focal ou a distância focalizada, o enquadramento (e conseqüentemente a fotometria) da fotografia também será alterado. A única forma de alterar a profundidade de campo sem alterar o enquadramento é alterando a abertura, a qual altera a fotometria (apresentada na seção 1).

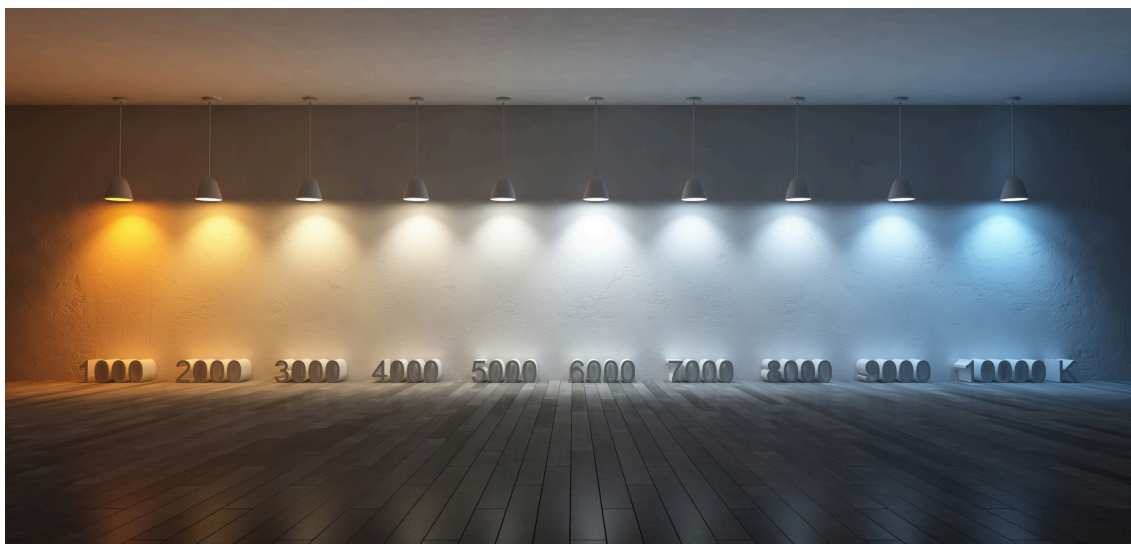
Configuração - Balanço de branco

A cor de qualquer fonte luminosa depende do comprimento de onda da luz que a fonte emite, porém a maioria das fontes luminosas não emite luz em apenas uma frequência, e sim em um espectro contínuo de frequências/cores. Caso exista, dentro deste espectro, uma frequência com mais energia, a luz possuirá uma cor definida [Freeman 2001].

Iluminar um objeto com determinada luz, altera a percepção (humana e da câmera) da cor do objeto. O olho humano, por sua vez, possui a capacidade de entender a cor natural de um objeto com mais facilidade do que as câmeras, que são mais sensíveis às diferenças da iluminação no processo de aquisição de fotografias. Ao fotografar um objeto branco, uma câmera pode adquiri-lo como sendo da cor creme, ou até alaranjado/avermelhado (diz-se cor quente), ou azulado (cor fria).

Considera-se como cor real de um objeto, a cor do objeto iluminado por luz natural do sol ao meio dia. Esta é a fonte luminosa predominante e mais natural, portanto assume-se que a cor natural de determinado objeto, é a cor refletida pelo objeto nesta iluminação. A Figura 17 apresenta uma parede branca sendo iluminada por fontes luminosas diferentes e mostra como esta alteração é capturada por câmeras.

Figura 17. Efeito da alteração da iluminação em fotografias.



Fonte: <https://medium.com/@Dropality/matching-lights-color>.

Para mapear a luz predominante em uma fotografia, usa-se a temperatura de cor da luz. Esta escala assume um corpo hipotético, metálico e negro (que absorve toda a luz) que é aquecido e com isso passa a emitir luz própria. Inicialmente a luz emitida tende ao vermelho e quanto mais é aquecido o corpo, a luz passa pelo branco e chega ao azul.

A escala assume o valor da temperatura do corpo em *Kelvin* (K) e a associa à luz emitida. A luz considerada natural, do sol ao meio dia, possui um valor em torno de $5.500K$. Objetos brancos só ficam brancos se iluminados com esta temperatura de cor, como pode ser visto na Figura 17. A Figura 18 apresenta as cores da escala *Kelvin* e suas respectivas medidas.

Para que seja possível adquirir fotografias com iluminações de temperatura de cor diferentes de $5.500K$ e mesmo assim exibir cores mais naturais, é possível ajustar a temperatura de cor das fotografias. Isso pode ser feito na própria câmera, ou por software na edição posterior. Esta configuração é chamada de balanço de branco e ajusta a fotografia para que o branco seja representado como branco, mesmo que a iluminação com outras temperaturas de cor o façam parecer de outra cor.

Figura 18. Escala Kelvin das temperaturas de cor.



Fonte: <http://www.newline.ind.br/voce-sabe-o-que-e-temperatura-de-cor/>

Este processo funciona sobre a cor branca, pois o branco é a junção de todas as cores. Assim ao corrigir o branco é possível corrigir todas as outras cores também. Nas câmeras atuais este processo pode ser feito de modo manual - escolhendo a temperatura de cor da iluminação da cena manualmente - ou de forma automática - quando a câmera determina a temperatura de cor da fotografia.

Porém este processo possui limitações, por exemplo, em cenas com várias fontes luminosas não é possível mapear a influencia de cada fonte luminosa em cada pixel da imagem. Isto pode ser observado na Figura 17.

Em uma cena interna, com luz natural entrando por uma janela, uma luz incandescente acesa e uso de flash (o qual normalmente não atua por todo o tempo de exposição) não será possível mapear para cada pixel o branco correto. Será necessário optar por uma correção e os pontos com outras iluminações não ficarão representados com a cor mais natural.

Portanto, para conseguir um balanço de branco é importante controlar a iluminação, de forma que todas as fontes luminosas usadas na fotografia possuam a mesma temperatura de cor. Isso permite que a correção de balanço de branco seja feita de forma efetiva.

Configuração - Exposição e fotometria

Exposição é a quantidade de luz necessária para registrar uma imagem no sensor de uma câmera digital. Para qualquer configuração da câmera, se pouca luz atingir o sensor a imagem formada será escura e ocultará detalhes da cena. Se muita luz atingir o sensor, a imagem ficará muito clara, e detalhes também serão omitidos.

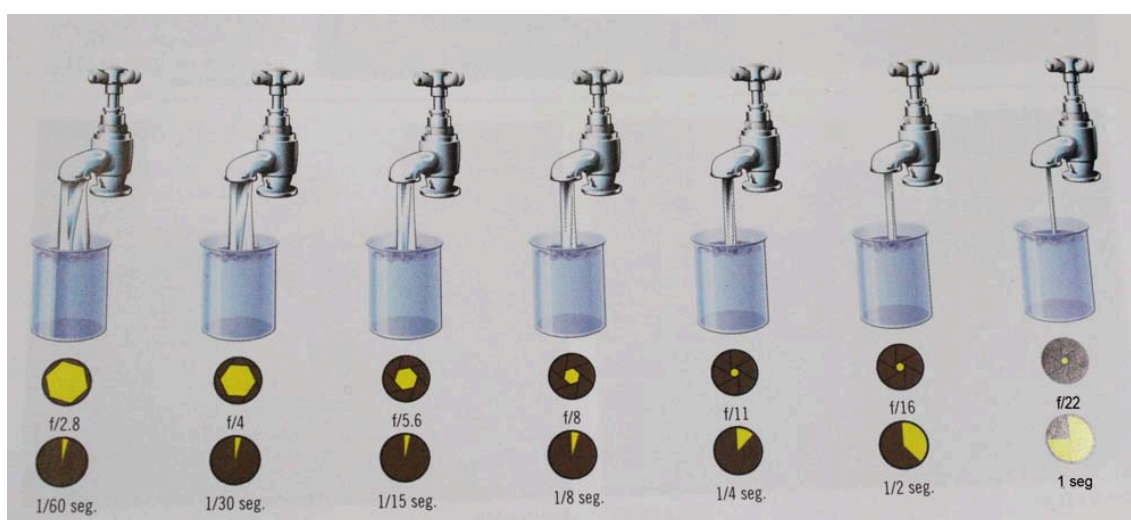
Para controlar a exposição de uma fotografia podem ser usadas duas abordagens: a primeira é controlar - se possível - a iluminação da fotografia; a outra abordagem é a configuração da câmera. Em ambientes com iluminação controlada, como estúdios fotográficos, o planejamento da iluminação deve ser feita de forma individual e depende muito do objeto a ser fotografado. Deve-se considerar regiões de sombreamento, reflexo, transparência, entre outros.

A configuração da câmera para atingir uma exposição adequada depende de três fatores relacionados entre si: Abertura do diafragma (representada por f); velocidade do

obturador; e ISO. Uma analogia que pode ser feita para entender melhor a exposição é o processo de encher um copo com água na torneira [Hedgecoe 2013].

Considerando que uma foto precisa de uma determinada quantidade de luz, na analogia a água simboliza a luz e o volume de água do copo a quantidade correta de luz. A sensibilidade ISO faz com que seja necessário mais ou menos luz, então pode ser entendido como o volume do recipiente a ser preenchido. A quantidade de luz que atinge o sensor pode ser controlada pela abertura do diafragma, na analogia isso pode ser entendido como a torneira mais/menos aberta. A velocidade do obturador é o tempo que a torneira fica aberta, permitindo a passagem de água.

Figura 19. Configuração da exposição.



Fonte: [Hedgecoe 2013].

Para se configurar a exposição é preciso levar em consideração a relação entre as três variáveis. Quanto maior a sensibilidade ISO, menor precisam ser a abertura do diafragma - maior será o valor de f - e/ou menor o tempo que o obturador permite a passagem de luz - maior a velocidade do obturador.

Quanto maior a abertura do diafragma - menor valor de f - menor precisa ser o valor do ISO e menor o tempo de exposição. O mesmo acontece com a velocidade do obturador: quanto mais lento o obturador - maior o tempo de exposição - menor o ISO e abertura necessários.

Além de equalizar estas três variáveis para atingir uma exposição correta, é preciso considerar o efeito destas variáveis na fotografia adquirida. Ao aumentar a sensibilidade ISO do sensor, a imagem tende a apresentar ruído. Quanto maior a sensibilidade, e o valor de ISO, mais ruído será inserido na imagem. Dependendo da quantidade de ruído, este pode ser removido por software - na própria câmera ou externamente - porém se a quantidade for muito grande isso não poderá ser feito com qualidade.

A velocidade do obturador está relacionada com a percepção de movimento. Em cenas completamente estáticas - objeto parado e câmera parada e estabilizada - o tempo de exposição pode ser longo. Mas caso exista algum movimento na cena ou câmera, este

movimento será captado na fotografia, diminuindo a nitidez da imagem. Ver subseção 1.1.

A abertura do diafragma interfere na nitidez da imagem capturada de duas formas principais: Quanto maior a abertura, menor a profundidade de campo, podendo deixar partes da cena sem foco e, conseqüentemente, sem nitidez; as objetivas possuem mais nitidez com menores aberturas.

As Tabelas 2, 3 e 4 apresentam um resumo das opções de configuração da exposição e como estas alteram a fotografia. Como as configurações são relacionadas - ao alterar ISO será necessário alterar também f e/ou velocidade do obturador - as tabelas foram divididas da seguinte forma: cada item da configuração (ISO, f e velocidade de obturação) possui uma tabela própria, definindo a alteração do seu valor para mais ou menos (primeira coluna - valor).

Cada tabela informa o efeito direto da alteração daquela configuração na fotografia. Por exemplo, alterar ISO influencia diretamente no nível de ruído da imagem (colunas: Efeito Direto). A tabela informa também o que será necessário fazer com cada uma das outras configurações (colunas: Ajustes Necessários). Por exemplo, ao diminuir o ISO será necessário (para equilibrar a fotometria) aumentar a abertura e/ou diminuir a velocidade do obturador.

Considerando a necessidade de alterar as outras variáveis que também influenciam na exposição, estas alterações também causam efeitos na imagem. Nas tabelas estes efeitos foram chamados de efeitos indiretos. Por exemplo, ao diminuir o ISO, é necessário aumentar a abertura, o que causa um efeito indireto de diminuir a profundidade de campo e a nitidez da imagem. O significado das colunas são:

ISO : valor da sensibilidade ISO;

f : abertura do diafragma, quanto maior o valor de f , menor a abertura;

Vel. Obturador : velocidade do obturador, quanto maior a velocidade, menor o tempo;

Valor : valor da variável principal da tabela, pode assumir valores maiores ou menores;

Ruído : nível de ruído inserido na imagem pelo ISO;

Prof. Campo : profundidade de campo;

Nitidez : nível de nitidez da imagem causada pela abertura da lente;

Estática : Quanto maior este valor, menor a possibilidade de movimento aparecer na fotografia, diminuindo a nitidez;

Aquisição : tempo necessário para adquirir uma fotografia.

Tabela 2. Resumo de configuração da exposição - ISO.

Sensibilidade ISO							
Valor	Efeito Direto Ruído	Ajustes Necessários		Efeitos Indiretos			
		f	Vel. Obturador	Prof. Campo	Nitidez	Estática	Aquisição
Maior	Mais	Maior	Maior	Maior	Maior	Mais	Rápida
Menor	Menos	Menor	Menor	Menor	Menor	Menos	Lenta

Fonte: Autoria própria.

No momento de configurar a câmera é necessário entender que não existe uma configuração padrão, ou que seja adequada para todas as situações. O objeto que está

Tabela 3. Resumo de configuração da exposição - Velocidade do Obturador.

Velocidade do Obturador							
Valor	Efeitos Diretos		Ajustes Necessários		Efeitos Indiretos		
	Aquisição	Estática	f	ISO	Prof. Campo	Nitidez	Ruído
Maior	Rápida	Mais	Menor	Maior	Menor	Menor	Mais
Menor	Lenta	Menos	Maior	Menor	Maior	Maior	Menos

Fonte: Autoria própria.

Tabela 4. Resumo de configuração da exposição - Abertura (Valor de f).

Abertura (Valor de f)							
Valor	Efeitos Diretos		Ajustes Necessários		Efeitos Indiretos		
	Prof. Campo	Nitidez	ISO	Vel. Obturador	Ruído	Estática	Aquisição
Maior	Maior	Maior	Maior	Menor	Mais	Menos	Lenta
Menor	Menor	Menor	Menor	Maior	Menos	Mais	Rápida

Fonte: Autoria própria.

sendo fotografado, sua cor, forma e textura, precisam ser considerados. Alguns objetos refletem muita luz, outros produzem muita sombra, e isso altera a exposição. Alguns objetos possuem forma irregular, como por exemplo um cano, o que vai exigir maior profundidade de campo, se comparado com fotografar uma bola de tênis de mesa.

1.1. Problemas intrínsecos da fotografia digital.

As câmeras digitais apresentam alguns problemas inerentes ao processo de fotografia, sendo alguns deles considerados usos incorretos da câmera, como falta de foco ou fotos tremidas. Deformações causados pelas lentes, aberrações cromáticas e a presença de vinhetas não podem ser corrigidos por configuração da câmera e precisam de softwares para serem corrigidos.

Aberração Cromática.

Para exemplificar a aberração cromática, pode ser usada a Figura 20. A imagem é 20a é a imagem original e aparentemente nenhuma distorção pode ser vista. Porém ao aumentar a escala da imagem podem ser notadas faixas verdes e vermelhas causadas pela lente 20b. A imagem 20c mostra a imagem com a mesma escala, porém com a aberração cromática corrigida.

A aberração cromática é causada pelo fato de cores diferentes possuírem índices de refração diferentes. Isso faz com que as cores tomem caminhos diferentes dentro das objetivas e se separem. O processo de correção consiste em aplicar tabelas de correção de imagens e cores oferecidos pelos fabricantes das câmeras.

Vinheta.

A mesma foto pode ser usada para mostrar outra aberração criada pelo fato da lente ser curva e o sensor da câmera plano. Nas bordas das imagens formam-se uma vinheta, pois o

Figura 20. Aberração cromática.

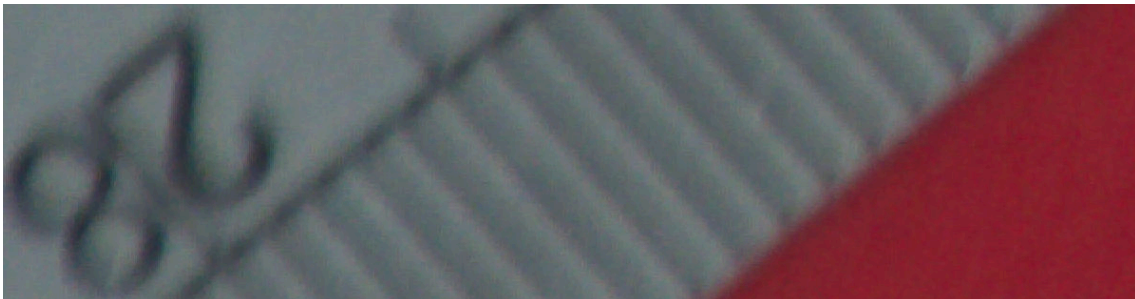
(a) Imagem Original



(b) Aberração cromática



(c) Imagem corrigida



Fonte: Autoria própria.

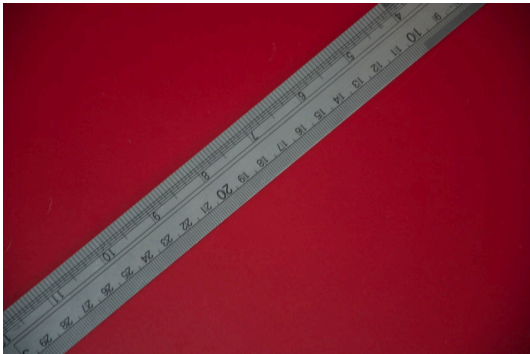
sensor não recebe luz da mesma forma no centro e nas bordas. Esta característica também afeta a fidelidade de cor e textura das imagens adquiridas. O fenômeno da vinheta pode ser visto na Figura 21, na qual a imagem 21a é a original e a imagem 21b a corrigida. O processo de correção também consiste em tabelas fornecidas pelos fabricantes.

Deformidades causadas pelas lentes.

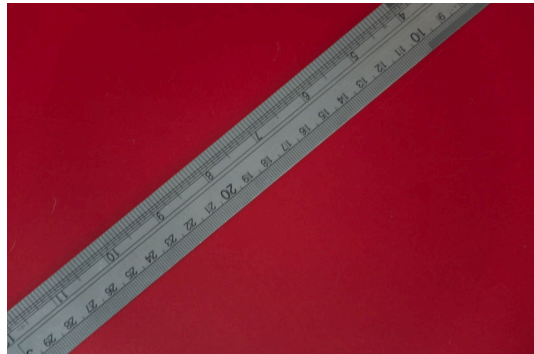
As lentes podem causar deformidades nas imagens pelo fato de serem curvas. Um exemplo de deformidade causado pelas lentes pode ser visto na imagem 22. A Figura 22a apresenta a foto original; a Figura 22b mostra a foto original com uma reta traçada do início ao final da régua. A imagem 22c apresenta a imagem 22b aproximada para evidenciar o problema, que pode ser resolvido por software.

Figura 21. Vinheta.

(a) Imagem original



(b) Imagem corrigida



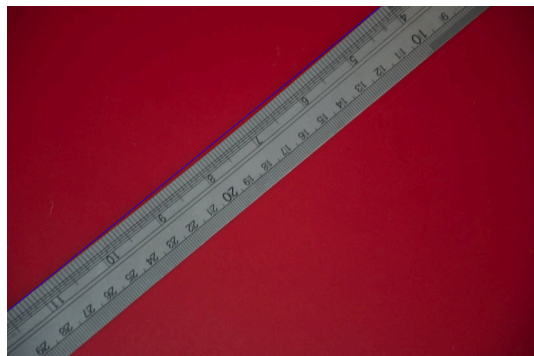
Fonte: Autoria própria.

Figura 22. Distorção.

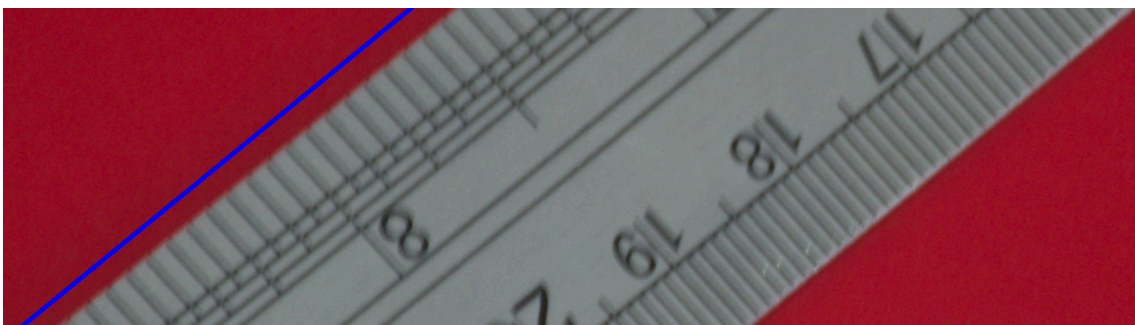
(a) Imagem original



(b) Imagem editada



(c) Imagem aproximada para evidenciar a distorção



Fonte: Autoria própria.

Aquisição de movimento na foto.

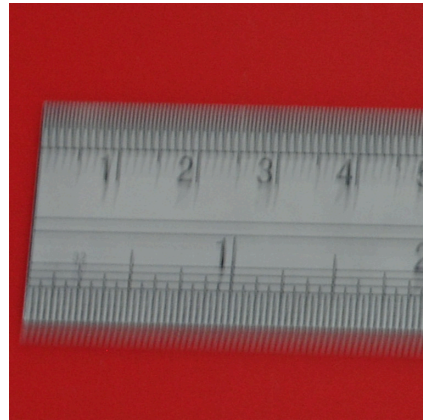
No contexto deste trabalho um problema importante que pode alterar a nitidez das imagens é o movimento da câmera ou do objeto durante o processo de aquisição, estas fotos são chamadas popularmente de “fotos tremidas”. Um exemplo destas fotos pode ser vista na imagem 23

Figura 23. Fotografia sem nitidez causada pelo movimento da câmera.

(a) Imagem original



(b) Imagem ampliada



Fonte: Autoria própria.

Referências

- Freeman, M. (2001). *Grande manual da fotografia*. Dinalivro, Lisboa, 2 edition.
- Hashen, A. (2017). Difference between DSLR and Mirrorless Cameras. acessado em: 17-07-2018.
- Hedgecoe, J. (2013). *O novo manual de fotografia: guia completo para todos os formatos*, volume volume. SENAC, São Paulo, 4 edition.
- Kelby, S. (2008-2011a). *Fotografia digital na prática*, volume 1. Pearson, São Paulo, 1 edition.
- Kelby, S. (2008-2011b). *Fotografia digital na prática*, volume 2. Pearson, São Paulo, 1 edition.
- Kelby, S. (2008-2011c). *Fotografia digital na prática*, volume 3. Pearson, São Paulo, 1 edition.
- Kelby, S. (2008-2011d). *Fotografia digital na prática*, volume 4. Pearson, São Paulo, 1 edition.