

Flávio de Almeida e Silva

**Projeção interativa estereoscópica:
visualização de modelos de objetos reais em
ambiente virtual**

Curitiba - PR, Brasil

2020

Flávio de Almeida e Silva

Projeção interativa estereoscópica: visualização de modelos de objetos reais em ambiente virtual

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 José Rodrigues Justino
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
Pamela Travassos de Freitas – CRB 9/1960

S586p
2020

Silva, Flávio de Almeida e
Projeção interativa estereoscópica : visualização de modelos de objetos em ambiente virtual / Flávio de Almeida e Silva ; orientador: Edson José Rodrigues Justino ; coorientador: Flávio Bortolozzi. – 2020.
196 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2020
Bibliografia: f. 177-187

1. Simulação (Computadores digitais). 2. Polarização (Luz). 3. Programação orientada a objetos (Computação). 4. Sistemas de Indicação visual tridimensional. 5. Visão estereoscópica. I. Justino, Edson José Rodrigues. II. Bortolozzi, Flávio. III. Pontifícia Universidade Católica do Paraná. Pós-Graduação em Informática. IV. Título.

CDD 22. ed. – 003.353

DECLARAÇÃO

Declaro para os devidos fins que o aluno **FLÁVIO DE ALMEIDA E SILVA**, defendeu sua tese de doutorado intitulada “**Projeção interativa estereoscópica: visualização de modelos de objetos reais em ambiente virtual**”, na área de concentração Ciência da Computação, no dia 31 de março 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, 23 de abril de 2020.



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

Dedicatória

Dedico este trabalho à minha mãe, Abigail de Carvalho e Silva (In Memoriam), por tudo que me ensinou e pelo exemplo de pessoa que foi. Ao meu pai, João de Almeida e Silva, por tudo que foi e ainda é em minha vida, pelos seus ensinamentos, pela sua dignidade e caráter. Aos meus irmãos, sobrinhos e amigos por toda a força e ajuda. Dedico também ao meu orientador, Prof. Dr. Edson José Rodrigues Justino, e co-orientador Prof. Dr. Flávio Bortolozzi, por toda paciência e conhecimentos passados, antes e durante a realização deste trabalho. Uma dedicação especial a Deus, pela minha saúde e auxílio nas horas difíceis.

Agradecimentos

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

Agradeço ao meu pai, irmãos e sobrinhos pela força.

Ao meu Orientador e Coorientador.

Aos meus amigos pelo incentivo.

Aos que colaboraram de forma direta, em especial:

Diogo Roberto Olsen, Lucas Murbach Pierin, Marina de Lara,

Andrei Rafael Brongel, Felipe Teixeira de Almeida e Silva, Guilherme Bertoldo,

William John Pereira Brobouski e Prof. Manoel de Campos Almeida.

A Nvidia GPU Grant Program pela doação da placa TITAN XP GPU.

Aos que contribuíram de forma indireta.

Em especial a Deus, por não me deixar abater diante das dificuldades.

Epígrafe

*“Quando tiver algum problema,
faça alguma coisa! Se não puder passar por cima,
passe por baixo, passe através, dê a volta,
vá pela direita, vá pela esquerda.
Se não puder obter o material certo, vá procurá-lo.
Se não puder encontrá-lo, substitua-o.
Se não puder substituí-lo, improvise.
Se não puder improvisar, inove.
Mas acima de tudo, faça alguma coisa!!
Há dois gêneros de pessoas que nunca
chegam a lugar nenhum: as que não
querem fazer nada e as que só
inventam desculpas.”*

In: A Essência da Verdade, coleção Pensamentos da Sabedoria

*“O propósito do aprendizado é crescer, e nossas mentes,
diferentes de nossos corpos, podem continuar
crescendo enquanto continuamos a viver ”*

Mortimer Adler

Resumo

As simulações almejam proporcionar realismo em observações de objetos reais digitalizados, no entanto o padrão é a utilização de objetos tridimensionais reconstruídos a partir de malhas ou nuvens de pontos, com renderização de cores e texturas. Por outro lado, tem-se a visualização tridimensional, que não faz a reconstrução, mas utiliza técnicas que buscam a projeção de objetos com simulação de como os olhos dos seres humanos captam as imagens. Os olhos capturam duas imagens distintas, uma para o olho esquerdo e outra para o direito, fazendo com que o cérebro as interprete, resultando na noção de profundidade (estereopsia), o que é percebido como tridimensional no mundo real. Isso é denominado como visão binocular, também chamada de visão estéreo. Para se ter uma visualização que permita observar o objeto por completo e também com noção de proporção e profundidade, as duas técnicas, visualização tridimensional e estereoscopia, devem ser empregadas. Assim como, dar ao observador a liberdade de manipular a visualização por todas as perspectivas adquiridas, ou seja, manipular a visualização tridimensional estereoscópica, é prover interatividade. Este trabalho apresenta os resultados de uma pesquisa aplicada, com objetivos exploratórios e procedimentos experimentais, que teve a finalidade de desenvolver uma visualização estereoscópica que permita aos observadores manipulá-la, ou seja, interagir com a visualização para observá-la por todos os ângulos, simulando a observação de um objeto real. Por meio das aquisições do “*Full Frames Semi-Spherical Scanner*” é possível adquirir um objeto que permite a geração da visualização interativa estereoscópica, sem a necessidade de tratamento ou ajustes da geometria epipolar no par estéreo. O volume de dados adquiridos, que é a quantidade de imagens, pode ultrapassar trinta mil, assim, um algoritmo de carregamento e gerenciamento das imagens em memória foi elaborado para garantir que a visualização seja sem interrupções ou lentidão. Os objetos adquiridos são de diferentes tamanhos, texturas, cores e materiais, para a geração da estereoscopia, além de permitir a verificação do desempenho dos algoritmos. Essa visualização possibilita o desenvolvimento dos objetos de aprendizagem, que hoje são debatidos e apontados pela comunidade científica como auxiliares no processo ensino/aprendizagem. Assim, fez-se a projeção via projetor *Digital Light Processing* (DLP) com óculos polarizado ativo e a interação via teclado, para todos esses objetos.

Palavras-chave: Simulação, par estéreo, polarização da luz, visualização tridimensional, estereoscopia interativa.

Abstract

Simulations aim to provide realism into real digitalized objects observation, although the standard method used is the tridimensional objects reconstruction based on mesh points or cloud points, with color and texture rendering. Conversely, there is the tridimensional visualization, that makes no reconstruction, but instead uses techniques to search objects projection simulating how human beings' eyes capture images. Eyes capture two distinct images, one for the left eye and one for the right eye, forcing the brain to interpret them, resulting into depth notion (stereoscopy), what is understood as tridimensional in real life. This is called as binocular vision, also called stereo vision. In order to obtain a visualization that allows completely observing an object with depth and proportion notion, both techniques, tridimensional and stereoscopy must be applied. As well giving to the observer freedom to manipulate all acquired perspectives, by means, manipulate the stereoscopy tridimensional visualization is the same to provide interactivity. This work presents the results of an applied research, which has exploratory objectives and experimental proceedings that had the purpose of developing a stereoscopy visualization which enables observers to manipulate it, in other words, interact with a visualization to observe it by all possible angles, simulating observation of a real object. Using Full Frames Semi-Spherical Scanner acquisitions, it is possible to acquire an object which allows stereoscopy interactive visualization generation of it, without needing treatment or epipolar geometry adjustments into the stereo pair. The amount of acquired data can overpass third thousand images, that said, an images memory loading and management algorithm was elaborated to assure visualization with no interruptions or delays. The acquired objects to stereoscopy generation have different sizes, textures, colors and materials in order to enable algorithms performance verification. This visualization enables learning objects development, that nowadays are debated and pointed by scientific community as supporting tools into teach/learning process. This way, it was made the projection through Digital Light Processing (DLP) projector using active polarized eyeglasses and keyboard interaction, to all these objects.

Keywords: Simulation, stereo pair, light polarization, anaglyph, interactive stereoscopy.

Lista de ilustrações

1	Temas relacionados à aquisição e reconstrução 3D	45
2	Face da estátua Davi de Michelangelo: a) Foto e b) Varredura e colorização	49
3	Temas abordados relacionados ao trabalho	55
4	Fotografia com efeitos passivos: (a) Estacionamento, (b) Calçada.	56
5	(a) Paralaxe Zero, (b) Paralaxe Positiva e (c) Paralaxe Negativa.	57
6	Ângulos, DI e distâncias da paralaxe	58
7	Relações entre DI e paralaxe: (a) $Pa < DI$, (b) $Pa \approx DI$ e (c) $Pa > DI$	59
8	Horóptero	59
9	Anáglifos: (a) cinza, (b) colorido, (c) vermelho-verde e (d) vermelho-azul.	61
10	Estereoscópio compacto de Davi Brewster.	62
11	Estrutura para aquisição de objetos - Light Field	64
12	Versões anteriores do F2S2: (a) estrutura da versão 1; (b) plano de fundo; e (c) versão 2 do scanner.	78
13	Versão 3 do F2S2: (a) estrutura e (b) plano de fundo e iluminação.	79
14	Aquisição de dois crânios para geração de OA.	80
15	Estrutura do F2S2 terceira versão.	82
16	Posicionamentos da bandeja.	83
17	Prato ou base giratória e motor.	83
18	Sistema de iluminação do F2S2.	84
19	Contraste entre o fundo infinito e os objetos	85
20	Distância mínima de configuração do raio: (a) sem BE e (b) com BE.	85
21	Mínimos e máximos dos eixos do F2S2	86
22	Planos do F2S2: (a) eixos; (b) coordenadas cartesianas; e (c) coordenadas esféricas.	86
23	Fotometria do objeto; (a) superexposta; (b) subexposta e (c) exposição considerada ideal.	87
24	Calibração da bandeja e câmera: (a) iniciando o posicionamento; (b) calibrando; e (c) retornando para posição de início.	88
25	Tratamento das imagens: (a) imagem completa e (b) redução de contexto.	89
26	Tratamento das imagens: (a) ângulo 90 graus completa e (b) ângulo 90 graus com redução de contexto.	89
27	Tratamento das imagens: (a) com redução de contexto e (b) segmentada.	89
28	Composição do Stream2D.	90
29	Formalização do Stream2D	91
30	Esquema do Stream2D no plano esférico	92
31	Perspectiva na aquisição do F2S2	94
32	Iluminação insidente apresenta diferença no objeto: (a) saliência pouco visível e (b)	

	saliência mais aparente	94
33	Oclusão do talo da maçã	95
34	Objeto sem oclusões de suas partes	95
35	Fluxograma do modelo proposto	97
36	Imagens de Dom Pedro II: (a) imagem esquerda; (b) imagem direita; e (c) Estereoscopia anáglifo	110
37	Preparação para aquisição em paralelo: (a) posições dos lasers, câmera e objeto; (b) régua de distância entre a esquerda e direita; e (c) régua de distância entre objeto e câmera	111
38	Câmera com os pontos: seta vermelha é o centro da lente e seta amarela é a posição do sensor	111
39	Aquisição paralela: (a) imagem Esquerda e (b) imagem direita	112
40	Algoritmo de Matching SURF: (a) matriz fundamental e (b) imagem da direita após ajustes da geometria - Aquisição Paralela	112
41	Imagem resultante da diferença entre as imagens paralelas	113
42	preparação para aquisição convergente.	114
43	Posicionamento do objeto, lasers e régua para a aquisição convergente	114
44	Imagens convergentes: (a) esquerda e (b) direita.	115
45	Ângulos na aquisição convergente	115
46	Algoritmo de Matching SURF: (a) matriz fundamental e (b) imagem da esquerda após ajustes da geometria - Aquisição Convergente	116
47	116
48	Processo de centralização por laser da parte frontal e traseira	118
49	Ajustes dos lasers da lateral	118
50	Imagem com 15cm na aquisição: (a) frame esquerdo e (b) frame direito	121
51	Submatriz da $Stream2D_{(i,j)}$ com 3° entre streams e frames	122
52	Posição fictícia da câmera durante a aquisição	122
53	Algoritmo SURF em Frames do F2S2: (a) matriz fundamental e (b) imagem original e imagem ajustada	123
54	Cálculo da diferença entre imagens da Figura 53.b	124
55	Diferença entre os frames 0° e 6°	124
56	Diferença entre os frames 0° e 9°	124
57	Fomação de par estéreo no plano esférico do Stream2D	125
58	Stream2D no plano esférico, visão: (a) frontal; (b) direita; (c) esquerda; e (d) traseira .	126
59	Estereoscopia Anáglifo da mesma imagem	127
60	formação da estereoscopia com imagens que distam 2°	128
61	Ligação entre um frame e seus adjacentes	129
62	Processos de carregar e descarregar um subconjunto da $Stream2D_{(i,j)}$	130
63	Janelas do possível visualizador: (a) tela principal com botões e (b) secundária com	

	projecção do par estéreo para polarização da luz.	134
64	Crânio estereoscópico com: (a) etiqueta para informações e (b) conteúdo sobre o osso selecionado	134
65	Divisão das janelas para estereoscopia polarizada lado a lado	135
66	Sobreposição da estereoscopia polarizada com janela dupla sem divisão: (a) escala normal e (b) escala aumentada	136
67	Movimento na janela deslizante: (a) horizontal; (b) vertical e (c) diagonal	139
68	Aquisição do par estéreo paralelo	142
69	Preparação do par estéreo com distância de 25cm entre câmera e objeto - StereoPhoto Maker	143
70	Ajuste horizontal aplicado ao par estéreo	144
71	Ajuste horizontal aplicado ao par estéreo no Gimp	144
72	Imagens estereoscópicas	145
73	Aquisição do par estéreo convergente	145
74	Par estéreo paralelo segmentado - diferença	146
75	Par estéreo convergente - StereoPhoto Maker	146
76	Imagem anáglifo resultante do ajuste do par convergente	147
77	Diferença entre os canais em valores de quatidade de pixels	148
78	Probabilidade da diferença entre os canais	149
79	Anáglifo do par estéreo gerado pelo F2S2	150
80	Imagens do Grilo	151
81	Profundidade de campo em aquisições	151
82	Estereoscopia anáglifo: (a) imagem macro e (b) sem macro	152
83	Centralização do objeto papel	153
84	Aquisição sem centralização do objeto	154
85	Estereoscopia sem centralização do objeto	154
86	Definição dos eixos e variáveis para o movimento circular/elíptico	155
87	Grilo com centralização	157
88	Inclinação da cabeça em 90°	160
89	Estereoscopia com paralaxe positiva e negativa	167
90	Estereoscopia em relação ao plano de projecção	168
91	Inclinação da cabeça na diagonal	168
92	Estereoscopia na diagonal	169
93	Imagens em sentido oposto	170

Lista de tabelas

1	Consumo de Memória	166
---	------------------------------	-----

Lista de quadros

1	Autores citados na revisão de literatura	53
2	Diferença entre olhos	56
3	Autores citados na revisão de literatura	72
4	Autores citados na revisão de literatura	73
5	Base de objetos adquiridos	81
6	Arquivo de configuração para o visualizador	137
7	Diferença entre os canais	147
8	Probabilidades diferença entre os canais	149

Lista de algoritmos

1	Pseudocódigo da janela deslizante	131
2	Pseudocódigo da definição do par estéreo	138
3	Pseudocódigo do Algoritmo Força Bruta	159
4	Pseudocódigo do Algoritmo Força Bruta	161
5	Estrutura na lista dinâmica da janela deslizante	163
6	Pseudocódigo do Algoritmo Janela Deslizante Inicial	164

Lista de abreviaturas e siglas

3D	Tridimensional
3dFDT	Detecção e Rastreamento 3D de Dedos
4D	Quatro Dimensões
AVA	Ambientes Virtuais de Aprendizagem
BTF	Bidirectional Texture Functions
CAD	Computer Aided Design - em português: Desenho assistido por computador
CCD	Charge-coupled Device - em português: dispositivo de carga acoplada
CIIM	Centro de Inovação em Imagens Médicas
CMOS	Complementary Metal Oxide Semiconductor - em português: semicondutor de óxido metálico complementar
CNC	Computer Numeric Control
CNN	Convolutional Neural Network
CPU	Central Process Unit
DI	Distância Inter-axial ou Distância Intraocular ou Distância Inter-pupilar
DICOM	Digital Imaging and Communications in Medicine item[DLP] Digital Light Processing
DM	Diferença Média
DTW	Dynamic Time Wrapping
F2S2	Full Frames Semi-spherical Scanner
FPGA	Field Programmable Gate Array
fMRI	Ressonância Magnética funcional
GB	Gygabyte
GLOH	Gradient Location and Orientation Histogram
GPU	Graphics Processing Unit

HD	Hard Disk
IFPR	Instituto Federal do Paraná
ISO	International Standards Organization
KD	k-Dimensional Tree
MB	Megabyte
MicroCT	Microtomografia Computadorizada
NCC	Normalized Cross-Correlation
NGFF	Next Generation Form Factor
NN	Nearest Neighbor
O2O	one-to-one
OA	Objetos de Aprendizagem
OSI	Open System Interconnection
RM	Ressonância Magnética
RNC	Redes Neurais Convolucionais
ROI	Region of Interest - em português: Region de Interesse
rpm	Rotações por Minuto
RV	Realidade Virtual
SB-GFRLS	Selective Binary and Gaussian Filtering Regularized Level Set
SDTW	Structured Dynamic Time Wrapping
SIFT	Scale Invariant Feature Transform
SPECT	Single-Photon Emission Computed Tomography
SSD	Solid-State Drive
Stc3D	Stereoscopy 3D
SURF	Speeded Up Robust Features
SUV	Sport Utility Vehicle - em português: Veículo utilitário esportivo
SVM	Support Vector Machine

TB	Terabyte
TC	Tomografia Computadorizada
TCP	Transmission Control Protocol
TV	Televisão
UHD	Ultra High Definition - em português: Ultra-alta Definição
USP	Universidade de São Paulo
YCbCr	Espaço de cor (componente luminância (Y) e duas componentes de crominância (Cb e Cr))
ZPS	Zero Parallax Setting

Lista de símbolos

D_a	Ponto em que o objeto está no foco
D_b	Ponto em que o objeto está no foco
D_A	Ponto em que o objeto está no foco
D_B	Ponto em que o objeto está no foco
ϕD_a	Ângulo de paralaxe
ϕD_b	Ângulo de paralaxe
cm	Centímetros
D	Olho Direito
E	Olho Esquerdo
ξ	Medida Angular entre a Distância média entre olhos e distância ao plano de projeção
d	Plano de projeção
DI	Distância interpupilar
Dist	Distância até qualquer objeto que se queira olhar
DM	Distância média entre os olhos humanos
Pa	Paralaxe
κ	Intensidade de saída da luz de um analisador
κ_0	Intensidade de entrada
Θ	ângulo formado no eixos de transmissão do polarizador e analisador
m	Metros
mm	Milímetros
X_p	Posição que representa a soma de todos os pixels da imagem
I	Imagem de entrada
xp	Soma das posições do pixel na imagem

y_p	Soma das posições do pixel na imagem
i_p	Varia de 0 até x_p
j_p	Varia de 0 até y_p
σ	Escala na imagem
H_p	Matriz de Hessian
L_{xpxp}	Convolução da derivada gaussiana
L_{xpyy}	Convolução da derivada gaussiana
L_{yppy}	Convolução da derivada gaussiana
O	Big O - complexidade computacional
n	Pontos-chave na complexidade computacional
m	Vizinhos utilizados para encontrar a complexidade
q	Número de pontos detectados
kg	Quilogramas
H	Eixo horizontal no F2S2
V	Eixo vertical no F2S2
C	Ângulo da câmera no F2S2
B	Giro da base no F2S2
λ	Altura do objeto no F2S2
r	Raio da câmera ao objeto no F2S2
BE	Base de elevação do objeto no F2S2
X	Eixo cartesiano no F2S2
Y	Eixo cartesiano no F2S2
Z	Eixo cartesiano no F2S2
P	Ponto no plano cartesiano no F2S2
F	Ponto (ou Frame) no plano esférico no F2S2
θ	Ângulo de giro do Eixo "B" no F2S2

γ	Ângulo de deslocamento da câmera no F2S2
numFrames	Quantidade total de frames para um stream
dbS	Degrees between streams - Graus entre streams no F2S2
dbf	Degrees between frames - Graus entre frames no F2S2
$Stream2D_{(i,j)}$	Matriz composta por streams e frames
i	Índice de uma linha da matriz
j	Índice de uma coluna da matriz
p	Total de streams da aquisição no F2S2
k	Total de frames da aquisição no F2S2
$\delta\gamma$	Diferença angular entre streams no F2S2
$\delta\theta$	Diferença angular entre frames no F2S2
S	Stream no plano esférico no F2S2
τ	Subconjunto de frames do Stream2D
ς	Subconjunto de streams do Stream2D
R	Red (vermelho) no canal de cores
G	Green (verde) no canal de cores
B	Blue (azul) no canal de cores
ImgDif	Imagem diferença
RimgE	Canal vermelho na imagem esquerda
GimgE	Canal verde na imagem esquerda
BimgE	Canal azul na imagem esquerda
RimgD	Canal vermelho na imagem direita
GimgD	Canal verde na imagem direita
BimgD	Canal azul na imagem direita
η	Total de pixels de uma imagem
α	Ângulo do foco na aquisição convergente

β	Ângulo da câmera na aquisição convergente
ρ	Ângulo na visualização dos frames no Stream2D
p1	Posição 1 - aquisição de um frame no Ângulo θ
p2	Posição 2 - aquisição de um frame no Ângulo θ
posxJan1	Posicionamento inicial, em x, na janela da esquerda
posyJan1	Posicionamento inicial, em y, na janela da esquerda
posxJan2	Posicionamento inicial, em x, na janela da direita
posyJan1	Posicionamento inicial, em y, na janela da direita
LarguraImagem	Largura total de uma imagem
AlturaImagem	Altura total de uma imagem
LarguraMaxima	Largura total da resolução de um monitor
AlturaMaxima	Altura total da resolução de um monitor
totalDifR	Total da diferença no canal vermelho
totalDifG	Total da diferença no canal verde
totalDifB	Total da diferença no canal azul
probTotal	Probabilidade total da diferença
PImageE	Pixel na imagem esquerda
PImageD	Pixel na imagem direita
Lin	Número da linha onde está o pixel
Col	Número da coluna onde está o pixel
probDifR	Probabilidade do canal vermelho na diferença
probDifG	Probabilidade do canal verde na diferença
probDifB	Probabilidade do canal azul na diferença
\vec{D}_p	Vetor de distâncias dos pontos na aquisição descentralizada
\vec{D}_c	Vetor de distâncias da câmera na aquisição descentralizada
dp	Distância de um ponto na aquisição descentralizada

dc	Distância da câmera a um ponto na aquisição descentralizada
zp	Distância entre o eixo x e o ponto onde está o objeto
a	Hipotenusa do triângulo retângulo entre a e zp
w	Valor determinado pelo ângulo θ
x	Eixo x no plano cartesiano
y	Eixo y no plano cartesiano
z	Eixo z no plano cartesiano

Sumário

1	INTRODUÇÃO	37
1.1	Contextualização e Problematização	37
1.2	Questão de Pesquisa	40
1.3	Objetivos	40
1.3.1	Objetivo Geral	40
1.3.2	Objetivos Específicos	40
1.4	Justificativa	40
1.5	Contribuições	42
1.6	Organização	43
2	CONCEITOS, PREMISSAS E REVISÃO DA LITERATURA	45
2.1	Introdução	45
2.2	Reconstrução tridimensional a partir de varredura laser e imagens DICOM	46
2.2.1	Reconstrução tridimensional por varredura a laser	46
2.2.2	Outras formas de aquisição e reconstrução tridimensional	50
2.2.3	Citações acerca de aquisições para reconstrução ou representação 3D	53
2.2.4	Considerações sobre aquisição e reconstrução 3D	54
2.3	Conceitos e técnicas de visualização 2D/3D, aquisições por fotos e estereoscopia	55
2.3.1	Estereoscopia	55
2.3.2	Aquisição e visualização por fotos	63
2.3.3	Algoritmo <i>Speeded Up Robust Features</i> para <i>matching</i> entre imagens	67
2.3.4	Algoritmos para carregamento de imagens em memória	70
2.3.5	Citações acerca dos trabalhos correlatos: fotos, estereoscopia e algoritmos	72
2.3.6	Considerações acerca dos conceitos e técnicas de aquisição por fotos e estereoscopia	74
3	METODOLOGIA E INFRAESTRUTURA PARA O DESENVOLVIMENTO DA PESQUISA	77
3.1	Introdução	77
3.2	Infraestrutura para o desenvolvimento da pesquisa	77
3.3	<i>Full Frames Semi-spherical Scanner</i>	82
3.3.1	Arquitetura do F2S2	82
3.3.2	Funcionamento do F2S2	84
3.3.3	Vantagens nas aquisições do F2S2 para estereoscopia	92
3.3.4	Limitações do F2S2	96
3.4	Modelo Proposto	97
3.5	Delineamento da pesquisa	98

3.5.1	Descrição das etapas	98
3.5.2	Definição dos testes	100
3.5.3	Descrição dos testes	101
3.5.4	Processo de realização dos testes	103
3.5.5	Análise de Resultados dos testes	107
4	DETALHAMENTO DO MODELO PROPOSTO	109
4.1	Introdução	109
4.2	Aquisição paralela e convergente para formação da estereoscopia	109
4.3	Centralização do objeto e preparação das imagens	117
4.4	Stream2D na formação da estereoscopia	120
4.4.1	Par estéreo no Stream2D	125
4.5	Janelas deslizantes no Stream2D	128
4.6	Projeção interativa estereoscópica para visualização objetos reais em ambiente virtual	132
5	RESULTADOS EXPERIMENTAIS E DISCUSSÃO	141
5.1	Introdução	141
5.2	Aquisições em paralelo e convergente	141
5.3	Distâncias nas aquisições de objetos e formação da estereoscopia	150
5.4	Centralização do objeto para estereoscopia	152
5.5	Algoritmo de força bruta	157
5.6	Algoritmo par-a-par	159
5.7	Algoritmo de janelas deslizantes na formação da estereoscopia	162
5.8	Movimentação e formação estereoscópica	166
6	CONSIDERAÇÕES FINAIS	173
6.1	Trabalhos Futuros	174
	REFERÊNCIAS	177
	APÊNDICE A – OBJETOS ADQUIRIDOS	189

Capítulo 1

Introdução

1.1 Contextualização e Problematização

A evolução da tecnologia, em especial os computadores e programas, possibilitou a criação de modelos virtuais tridimensionais (3D) de objetos do mundo real. Esses modelos podem ser construídos por meio de malhas ou nuvens de pontos (WANG et al., 2015; ROMANONI et al., 2016), os quais constituem algumas das características físicas de superfície, tais como as silhuetas, cavidades e protuberâncias apresentadas pelos objetos reais.

As finalidades para esse tipo de modelo abrangem áreas como a industrial (JANKOVIC; JOVICIC, 2018), a paleontologia (DARDON et al., 2010; FAU; CORNETTE; HOUSSAYE, 2016) e a educacional (THOMAS et al., 2016; ALAM; ULLAH; ALI, 2017).

No ambiente instrucional, treinamentos e educação, destaca-se a utilização da modelagem 3D, principalmente em design de ambientes (HUI, 2016; ZHANG, 2017). Esse tipo de trabalho está presente na arquitetura, engenharia civil, mas também outras áreas do conhecimento utilizam a reconstrução e visualização 3D, por exemplo, a medicina, onde corpos humanos são reconstruídos para o estudo da anatomia (MARTÍN, 2018; KAZOKA; PILMANE, 2019).

Estas reconstruções e visualizações, quando aplicadas ao ensino, encaixam-se em abordagens como metodologias ativas, tema que vem sendo debatido pela comunidade científica (DIESEL; BALDEZ; MARTINS, 2017; MOYA, 2017; MOTA; ROSA, 2018; BESSA; COSTA, 2019).

Atrelados a essas metodologias, vê-se o crescimento dos chamados ambientes virtuais de aprendizagem (AVA) e sua utilização (NAZARENKO, 2015; ALAM; ULLAH; ALI, 2017; SCOTT; SORIA; CAMPO, 2017; KOH et al., 2017), além dos objetos de aprendizagem (OA) (ALVARENGA; GINESTIÉ; BRANDT-POMARES, 2017; REDMOND et al., 2018; NABIZADEH et al., 2020).

Atualmente discute-se a melhor forma de disponibilizá-los, o que levou à criação dos repositórios (COHEN; REISMAN; SPERLING, 2015; KOUTSOMITROPOULOS; SOLOMOU, 2018; OSORIO-ZULUAGA; DUQUE-MENDEZ, 2018). Estes são vistos como locais de armazenamento de dados, os quais podem ser reutilizados por diferentes áreas do conhecimento para gerarem seus próprios OA.

Tanto para a criação dos OA como para apenas apresentação de conteúdo, a reconstrução e posterior visualização desses objetos virtuais, que representam os reais, partem da digitalização.

Esta pode ser realizada por meio de equipamentos a laser, luz estruturada, sonar, tomografia computadorizada (TC), ressonância magnética (RM), raio-x ou por fotos (BARONE et al., 2018; LUCHT et al., 2018; KIGKA et al., 2018; YEO et al., 2018; WU et al., 2018; RAMOS; PRIETO, 2015; KLECKA; BOSTIK, 2018).

Existem algumas considerações quando se faz a reconstrução 3D, pois, no que se refere à representação fidedigna do objeto real, problemas como as oclusões, perda de detalhes, da textura ou cor, podem ocasionar erros na reprodução, principalmente quando o objeto é desenvolvido em algum programa de criação 3D.

Por outro lado, as visualizações tridimensionais, que não precisam de reconstrução por nuvens ou malhas de pontos, podem apresentar o objeto tridimensional por técnicas com a utilização das chamadas Imagens Integrais (JAVIDI; HONG; MATOBA, 2006). Esse tipo de visualização é gerada a partir de aquisições com, por exemplo, câmeras fotográficas, utilizando ou não filtros (MOON; JAVIDI, 2008). Essas técnicas se diferenciam, pois uma reconstrói o objeto com malhas ou nuvens de pontos e a outra utiliza Imagens Integrais a partir de fotos.

Em objetos desenhados ou reconstruídos ocorre a ocultação de detalhes, principalmente na representação de cavidades. Um exemplo disso ocorre no atlas anatômico do *Visible Body*¹, o canal auditivo (canal do ouvido) e os forames (aberturas para passagem de músculos, vasos, artérias, veias e outras estruturas, através de ossos) não estão representados, a não ser por marcas sem profundidade. No site e vídeo citados, pode-se ver isso (tempo 1:40 min.).

Sobre os atlas anatômicos por fotos, a apresentação de fotos estáticas não se compara ao realismo de manipular o objeto real, pois elas não possibilitam, isoladamente, a troca de perspectiva (ACKERMAN, 1998). Segundo o autor, é indispensável a visualização tridimensional, em estudos anatômicos, para se ter o entendimento das estruturas biológicas, uma vez que elas são tridimensionais. Apenas visualizar os objetos em 3D também não traz o realismo necessário, pois o objeto pode sofrer os problemas das reconstruções supracitados.

Mesmo com esses problemas, não há inviabilidade no seu uso, todos tem sua validade, principalmente para a educação, pois diferentes técnicas de ensino podem levar a um aumento eficaz na retenção de conhecimento. O uso de diferentes técnicas de ensino é entendido como pedagogia multimodal (GOODARZI et al., 2017).

A visualização 3D é a noção de perceber proporção e profundidade, que são inerentes aos objetos reais tridimensionais. Essa noção é atingida quando se insere técnicas como luz e, conseqüentemente, sombra em todas as perspectivas do objeto virtual. No entanto, tem-se que contornar problemas com a luz em objetos virtuais (KASPER; HECKMAN, 2019; KHADKA; REMAGNINO; ARGYRIOU, 2018; WANG et al., 2018). Os trabalhos citados utilizam técnicas para contornar os problemas, contudo permanecem as dificuldades com oclusões e problemas na reconstrução.

¹ <<https://www.visiblebody.com/librarians/overview>>

Essas noções de proporção e profundidade estão relacionadas à capacidade humana de visualização por meio da visão binocular, estereopsia ou visão estéreo (KASPER; HECKMAN, 2019). Que também está relacionada ao número e posição dos olhos dos seres humanos. Um ser humano, normal, tem dois olhos e cada olho recebe uma imagem, sendo diferentes entre si (BAŞGÖZE; MACKEY; COOPER, 2018).

Com a criação da fotografia, cita-se aqui o Daguerreótipo (GUARNIERI, 2018), seção “II.C. - *Photography*”, cenas e objetos são capturados e visualizados de forma bidimensional. No entanto, vislumbrou-se a possibilidade de observá-los com profundidade, ou seja, com estereoscopia, como é da natureza dos seres humanos com visão binocular (visão com dois olhos). Para que isso seja possível, obtém-se duas imagens díspares, além da utilização de um equipamento próprio para a visualização. Esse tipo de equipamento, estereoscópio, nasceu quase que paralelamente ao surgimento das câmeras fotográficas e é creditado a Charles Wheatstone, contudo o aprimoramento do aparelho foi realizado por David Brewster (WHEATSTONE, 1838; KING, 2013).

Visualizar um objeto de forma estereoscópica não garante que seja também 3D. Isso se deve ao fato de que é possível perceber a profundidade utilizando duas perspectivas e não necessariamente todas. A partir do momento em que se visualiza todas as perspectivas com proporção e profundidade, tem-se um objeto estereoscópico 3D.

Diante desse panorama foi criado o grupo de pesquisa “Centro de Inovação em Imagens Médicas - CIIM”, dentro do qual foi desenvolvido o *Full Frames Semi-spherical Scanner* (F2S2). Com esse equipamento é possível adquirir um objeto completo, todas suas perspectivas, resolvendo o problema de oclusão. Além da possibilidade de realizar aquisições em resolução de até 8k (7680x4320 pixels), na atual configuração da câmera, permitindo qualidade de detalhes dos objetos.

Com todas as perspectivas consegue-se criar um objeto que é estereoscópico e que permite a interação do usuário, portanto a visualização se torna interativa estereoscópica. A interatividade nesse processo é a possibilidade de um observador poder manipular a visualização por todos os seus ângulos, assim como aplicar as transformações geométricas de translação, escala e rotação.

Essa visualização interativa estereoscópica permite a criação de um sistema de simulação para os objetos que forem adquiridos, ou seja, independentemente da área de conhecimento, se o objeto real puder ser adquirido, ele poderá fazer parte do simulador e também ser transformado em OA.

1.2 Questão de Pesquisa

Como visualizar projeções interativas estereoscópicas de objetos reais em ambiente virtual?

1.3 Objetivos

1.3.1 Objetivo Geral

Desenvolver um modelo de projeção interativa estereoscópica para visualização objetos reais em ambiente virtual.

1.3.2 Objetivos Específicos

Sendo o objetivo geral o desenvolvimento de um processo que permita a visualização interativa estereoscópica, para alcançá-lo, traça-se os objetivos específicos:

1. analisar a formação do par estéreo a partir do conjunto de imagens gerado pelo F2S2;
2. analisar o processo computacional de seleção das imagens que formam o par estéreo, considerando as quatro dimensões da interatividade;
3. selecionar dentre as imagens do conjunto de perspectivas as quais melhor formam o par estéreo, considerando as quatro dimensões da interatividade;
4. analisar o processo computacional de interação considerando as três transformações geométricas que caracterizam um objeto tridimensional - translação, escala e rotação;
5. desenvolver um processo computacional, que tenha baixo consumo de recursos como memória e acesso a disco, de geração de uma visualização interativa a partir de todas as perspectivas adquiridas de um objeto.

1.4 Justificativa

As projeções tridimensionais são utilizadas em âmbito instrucional, como também empregadas na indústria. Isso pode ter sido alavancado pela crescente evolução das tecnologias de apresentação desse tipo de conteúdo. Entretanto, a noção de proporção e profundidade apresentadas são apenas por meio de inserção de luz e sombra, sem utilização de duas imagens. Apresentar essas duas imagens, para a geração da proporção e profundidade tem relação com a visão binocular humana.

Os processos de colorização e texturização também acabam limitando a representação fiel ao objeto real, contudo existem técnicas que buscam essa proximidade, no entanto exigem

um custo operacional (tempo, pessoas e computacional) elevado. A varredura a laser ou luz estruturada, por exemplo, pode acarretar oclusões com alguma perda de informação.

Além disso, a formação estereoscópica tem que ocorrer com recursos computacionais, onde após a reconstrução do objeto, insere-se elementos virtuais como focos luminosos, para gerarem luz e sombra, duas câmeras para a representação do par estéreo. A cada movimento de mudança de visualização, terão que ocorrer todos os cálculos de iluminação, sombra e posicionamento das câmeras para a nova visualização.

As aquisições realizadas pelo F2S2 contam com um ambiente controlado em iluminação, como também precisão no deslocamento da câmera. Cada perspectiva adquirida é completa, ou seja, tem-se a face toda do objeto, não necessitando uma reconstrução parte a parte para tê-la, como ocorre nas varreduras a laser ou pela luz estruturada. Assim, nenhum artefato (cores ou texturas) é inserido, como também não ocorrem oclusões (ocultação de detalhes), pois todas as perspectivas são capturadas.

Para a educação, a manipulação dos conteúdos interativos estereoscópicos vem como uma motivação aos acadêmicos. Esses conteúdos são auxiliares aos processos pedagógicos tradicionais, assim como é uma forma atrativa de manter o foco e permitir que os acadêmicos gerenciem o próprio desenvolvimento de seus conhecimentos. O que é almejado pela abordagem das metodologias ativas.

Os conteúdos interativos e estereoscópicos buscam minimizar os problemas em diferentes frentes, como a falta de objetos reais ou exposição desnecessária a elementos nocivos e perigosos. Por exemplo, como ocorre na medicina, acadêmicos não dispõem de uma quantidade necessária de objetos reais do corpo humano para seus estudos.

Assim como, determinadas áreas, por exemplo educação física, podem diminuir suas incursões a laboratórios de anatomia, onde entram em contato com produtos químicos nocivos ou equipamentos perigosos, como bisturis.

Esses problemas podem ser contornados com a aquisição de ossos, órgãos, entre outros componentes do corpo, mantendo a fidelidade de cor, textura e forma, provendo um material de estudo, assim como fazem os atlas anatômicos, porém com interatividade e estereoscopia.

Os objetos adquiridos podem ser bases de dados disponibilizadas em repositórios para a geração de OA digitais. A geração desses OA a partir de objetos reais pode contornar algumas dificuldades, como ter apenas um exemplar de um animal raro e este ter que ser compartilhado com outras instituições, para o estudo. O transporte desse exemplar pode danificá-lo ou mesmo destruí-lo. Assim, com o OA criado, pode-se disponibilizá-lo para qualquer instituição, tendo a garantia de poderem estudá-lo sempre que necessitarem, sem a preocupação de danificá-lo ou destruí-lo.

Uma área que também pode se beneficiar é a de preservação do patrimônio cultural, pois a forma de aquisição, visualização e interatividade pode ser empregada na conservação do acervo

histórico, além de permitir uma aproximação entre os frequentadores dos museus e as obras expostas.

A visualização de conteúdos estereoscópicos com tempo prolongado pode ocasionar desconforto visual, um problema que é estudado, mas que ainda se buscam soluções. Os esforços atuais foram direcionados à criação de equipamentos de visualização, deixando de lado a criação dos conteúdos. Entretanto, melhorar o conteúdo é também uma forma de diminuir esses esforços. Assim, a aquisição realizada provê um conjunto de imagens que já traz uma correção do par estéreo, privilegiando a estereoscopia.

Um problema para técnicas de aquisição por fotos é realizar as transformações geométricas (translação e rotação) em imagens adquiridas, para que essas gerem estereoscopia. Isso pode gerar imagens que ao serem visualizadas estereoscopicamente causem o desconforto visual, mesmo em curto tempo de exposição.

Esse tipo de desconforto está ligado às técnicas de aquisição das imagens. Com isso contornado pelo F2S2, a busca é pelas imagens que formam o par estéreo, o qual pode apresentar a estereoscopia e mesmo em exposição por tempo prolongado não causar desconforto.

Assim, este trabalho se justifica pela junção de técnicas (aquisição controlada pelo F2S2 e geração de estereoscopia com recursos computacionais modestos) para apresentar o objeto estereoscópico que permita a aplicação das componentes de tridimensionalidade (translação, escala e rotação), além das transformadas de movimentação, aproximação, afastamento e giro, necessárias para visualizar todas as perspectivas, trazendo a sensação da manipulação do objeto real.

1.5 Contribuições

A contribuição deste trabalho é possibilitar a interação, em tempo de execução, de observadores em visualizações estereoscópicas de objetos do mundo real digitalizados, com acuidade visual, pelo F2S2. Filmes estereoscópicos, chamados de 3D, não permitem que usuários interajam, a não ser pausar, avançar ou recuar as cenas. Este trabalho, permite que os observadores interajam, manipulando a visualização, possibilitando observar o objeto completo, ainda aplicar as transformações geométricas. Isso traz uma nova perspectiva de utilização da estereoscopia e visualização tridimensional.

O F2S2 desempenha um papel fundamental, pois suas aquisições são em ambiente controlado. Os objetos são posicionados e digitalizados com um plano de fundo infinito em uma cor que contrasta com os objetos, como também se controla a iluminação. Com isso, consegue-se extrair apenas o objeto de interesse e atende ao requisito de qualidade visual das imagens. O que não se consegue, por exemplo com as varreduras a laser e luz estruturada.

Por meio das aquisições realizadas, não há necessidade de criar malhas e nem nuvens

de pontos, com posterior adição de cores e texturas, para ser ter uma aproximação aos objetos reais. A qualidade conseguida pela aquisição já provê todas as perspectivas em uma resolução até 8k, mas isso depende da câmera, no entanto melhorando sua tecnologia, também reflete na melhoria da qualidade das aquisições. Assim, o F2S2 fornece uma qualidade de aquisição que outros equipamentos ainda têm dificuldades em fornecer.

Essas características discriminadas permitirão a construção de um sistema para a manipulação dos objetos com interatividade estereoscópica, o que significa mover, aproximar, afastar e girar, além das transformações de tridimensionalidade (translação, escala e rotação). Essa interação não é prevista em vídeos, uma vez que esses são apresentados de forma unidimensional e limitados aos controles do sistema de visualização, os quais podem não dispor desse tipo de manipulação. As apresentações estáticas de fotos não atingem o mesmo realismo conseguido por meio da interatividade com o conjunto de fotos adquiridas pelo F2S2.

A visualização parte do princípio de que não se está apresentando uma parte do objeto, mas sim todas as perspectivas e com a técnica de estereoscopia, ou seja, dando as noções de proporção e profundidade, comum a visão humana. Assim, alcança-se um nível diferenciado de visualização e manipulação de conteúdo, seja para a educação, indústria ou mesmo comércio.

Isso é entendido como contribuição, pois há a possibilidade de criação de conteúdos para as áreas do conhecimento, área comercial, preservação do patrimônio cultural, treinamentos em diferentes segmentos, como a indústria 4.0, que tem a realidade virtual e aumentada como uma linha ainda em exploração.

Esses são nichos que apresentam diversas aplicações sociais, por exemplo, a virtualização de um museu, onde as obras que seriam vistas somente presencialmente, estariam disponíveis em diferentes mídias, com a possibilidade de servir também para o estudo, em cursos como o de artes. As aquisições atendem a um requisito que outras, como as aquisições paralelas e convergentes, não atendem, pois o F2S2 consegue a similaridade com a visão humana (capturar imagens e ter noção de proporção e profundidade). Inicialmente, o foco é a geração de conteúdos instrucionais, os OA, para diferentes áreas do conhecimento, auxiliando o processo de ensino/aprendizagem.

1.6 Organização

O trabalho está organizado em 6 (seis) capítulos e um apêndice. No capítulo 1, a Introdução, foi descrita a contextualização, a questão de pesquisa, objetivos, a justificativa e contribuições que norteiam esta pesquisa.

No segundo capítulo aborda-se os temas relativos à revisão da bibliografia, assim como os trabalhos correlatos a este. Dando ênfase àqueles que, de alguma forma, deixaram lacunas a serem resolvidas.

No terceiro é abordado a metodologia e infraestrutura para o desenvolvimento. Onde

são definidos o Full Frames Semi-Spherical Scanner, apresenta-se o modelo proposto e faz-se o delineamento da pesquisa.

Na sequência, apresenta-se, no quarto capítulo, o detalhamento do modelo proposto. Como se realizou o estudo das formas como geram estereoscopia por meio de imagens adquiridas em paralelo e convergentes, além de abordar como a centralização do objeto influencia na aquisição. Apresenta-se a formação da estereoscopia a partir do conjunto de imagens adquiridas e o gerenciamento destas em memória. Por fim, apresenta-se a projeção interativa estereoscópica.

Existem várias alternativas para gerar objetos tridimensionais, contudo a geração da estereoscopia ainda está ligada aos métodos de sobreposição e ajustes das imagens. Por isto, no capítulo 5, apresenta-se os resultados obtidos com as aquisições paralelas e convergentes, assim como a necessidade de se ajustar as imagens para a estereoscopia. Qual a relação da distância da câmera com o objeto para a visualização dos objetos estereoscópicos. Como um objeto descentralizado é observado em estereoscopia. Os resultados de consumo de memória e acesso a disco dos algoritmos utilizados para o gerenciamento dos processos de carregar e descarregar imagens da memória. Por fim, a interação de movimentação e como ocorreu a formação estereoscópica.

O capítulo 6 é destinado às considerações finais e trabalhos futuros. Em seguida são apresentadas as bibliografias utilizadas nesta pesquisa. No apêndice são apresentados os objetos digitalizados.

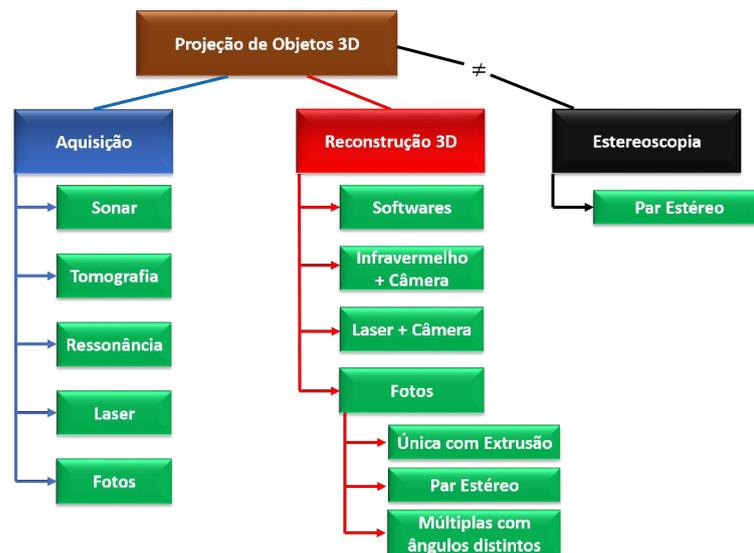
Capítulo 2

Conceitos, Premissas e Revisão da Literatura

2.1 Introdução

Este capítulo aborda os temas que estão relacionados ao trabalho. Para tanto, foca-se nas definições, apresentações do que pesquisadores vêm desenvolvendo, afim de fundamentar esta pesquisa. Na Figura 1 são apresentados esses temas.

Figura 1 – Temas relacionados à aquisição e reconstrução 3D



Fonte: Autoria Própria

Assim, são apresentados os conceitos e premissas sobre: a reconstrução tridimensional a partir de varredura laser e imagens DICOM; e outras formas de aquisição e reconstrução tridimensional. Também aborda: as considerações sobre aquisição e reconstrução 3D; os conceitos e técnicas de aquisição por fotos e estereoscopia; a estereoscopia; a aquisição e visualização por fotos; o algoritmo *Speeded Up Robust Features* para *matching* entre imagens; o algoritmos para carregamento de imagens em memória; e as considerações acerca dos conceitos e técnicas de aquisição por fotos e estereoscopia.

2.2 Reconstrução tridimensional a partir de varredura laser e imagens DICOM

A reconstrução 3D utilizando varredura a laser, para a construção de malhas de pontos, será abordada pela propriedade encontrada nessa técnica, que é a aquisição dos pontos da superfície dos objetos, os quais passam por um pós-processamento para ajustes das malhas, além de, algumas vezes, realizar-se a colocação de texturas e cores via fotografia.

Em abordagens por reconstrução a laser, ressonância magnética (RM), tomografia computadorizada (TC) e ultrassonografia, mesmo os propósitos sendo similares, cada uma pode diferenciar da outra, ou seja, cada técnica pode produzir dados de formas diferentes, necessitando técnicas diferentes de reconstrução. Faz-se uma apresentação desses tipos de reconstruções para apontar algumas lacunas e problemas decorrentes dessas aquisições, os quais podem gerar dificuldades na formação da estereoscopia.

2.2.1 Reconstrução tridimensional por varredura a laser

O objetivo desta seção é apresentar a reconstrução 3D a partir da varredura a laser. A ideia é mostrar como um objeto é gerado e suas similaridades ao objeto real, além de verificar se é viável para a geração da visualização interativa estereoscópica. Para esses objetos, necessita-se que a câmera virtual seja duplicada, com a alteração angular de uma delas para prover a estereoscopia e isso, por padrão, é realizado em pós-processamento.

A digitalização de objeto, utilizando equipamentos a laser, teve seu início nos anos 60 e utilizavam luzes, câmeras e projetores. Suas limitações, como tempo e esforço para digitalizar objetos, não permitiram maiores avanços nessa tarefa. Nos meados dos anos 80 surgiram equipamentos que utilizavam luz branca, lasers e sombreamento para a captura de superfícies. Alguns trabalhos já se preocupavam com a digitalização e reconstrução de objetos por meio de laser e outras técnicas (BESL; JAIN, 1985).

Scanners 3D são capazes de capturar formas e aparências de objetos para construir modelos tridimensionais a partir dos dados coletados. São variados os propósitos dessas aquisições, podendo ser utilizados para reprodução e análise da forma do objeto, engenharia reversa, até mesmo no desenvolvimento de réplicas, por meio de impressão 3D.

O avanço na construção dos lasers, nos sensores CCD (em português: dispositivo de carga acoplada) e CMOS (em português: semicondutor de óxido metálico complementar), além de unidades de aquisição em altas velocidades, permitem que se faça rapidamente a captura de milhares de pontos, por exemplo, 10 mil amostragens por segundo. Isso pode gerar milhares ou milhões de pontos dependendo do objeto.

Essa nuvem de pontos, normalmente não organizada, faz com que o trabalho computacional seja custoso, não apenas pela grande quantidade de pontos que se tem, mas também

pelo fato de ocorrerem problemas como descontinuidade, dados incompletos, pontos esparsos e ruídos (REMONDINO, 2003).

Para criar objetos 3D, existem duas classes (Sistemas baseados em medições de objetos e sistemas não baseados em medições) nas quais os métodos se enquadram. A classe a qual pertencem os métodos que não são baseados em medições não será abordado neste trabalho. Dentro dessa classe estão softwares comerciais de desenvolvimento de objetos 3D, por exemplo, 3D Studio Max¹ e Lightwave².

A fotogrametria se enquadra na primeira classe e trata-se de uma ciência que faz medições a partir de fotografias, especialmente para recuperar as posições exatas dos pontos de superfície (MARKIEWICZ; ZAWIESKA; PODLASIAK, 2017; OWDA; BALSABARREIRO; FRITSCH, 2018; DO; NGUYEN, 2019).

Quando se fala em reconstrução, vê-se que existem restrições quanto a objetos adquiridos por meio de scanners laser, uma vez que não se pode garantir que o objeto criado seja um reflexo do original, em características como textura e cor, podem ocorrer casos de também não se conseguir reproduzir a forma.

Esses problemas decorrem de um conjunto finito de pontos de amostragem, também por não satisfizerem certas propriedades exigidas por um algoritmo de reconstrução, como boa distribuição, densidade e poucos ruídos. Os pontos medidos podem ser desorganizados e ruidosos, além disso o objeto pode conter uma topologia com elementos agudos, entre outras, fazendo com que o método de reconstrução tenha que inferir a geometria e topologia a partir do conjunto finito de pontos de amostra.

Existem algoritmos nos quais a topologia é conhecida a priori, algo como ter uma topologia fixa. Outros, exploram a estrutura ou informação de orientação, ou seja, exploram a estrutura dos dados para a reconstrução da superfície. Por exemplo, no caso de múltiplas varreduras, eles podem usar a relação de adjacência dos dados dentro de cada imagem do intervalo. Existem ainda, aqueles que usam o conhecimento da orientação da superfície que é fornecida com os dados.

Para a visualização desses modelos, verifica-se que o interesse é obter um objeto 3D que se assemelhe ao objeto real, tanto em realismo quanto em precisão. Assim, no início da fotogrametria, as primeiras tentativas na visualização de modelos 3D foram realizadas no início dos anos 90. Os objetos como carros eram exibidos usando, principalmente, pacotes de CAD.

Com a evolução dos softwares de reconstrução, elementos como sombra e textura puderam ser adicionados a todos os modelos, contudo a visualização necessita de grandes conjuntos de dados, porém as informações contidas em modelos fotogramétricos são frequentemente reduzidas.

¹ <<https://www.autodesk.com.br/products/3ds-max/overview>>

² <<https://www.lightwave3d.com/>>

Como consequência dessa redução, sabe-se que a precisão dos dados é perdida e as texturas de alta resolução são deixadas de lado. Por outro lado, a baixa precisão na visualização não atrai os usuários finais e não justifica o alto custo de produção de um modelo 3D fotogramétrico.

Abordados problemas que podem surgir no decorrer do processo de aquisição até a visualização 3D do objeto, discorre-se então sobre algumas aplicações utilizando a digitalização a laser.

Uma utilização para essas formas de aquisição é a engenharia reversa, mesmo com a necessidade de processamento intensivo para a reconstrução do objeto, devido aos dados digitalizados serem, inicialmente, pontos dispersos. Assim, uma possibilidade de integração, câmeras e laser, pode melhorar o processo de reconstrução (FISCHER; MANOR, 1999).

Outros métodos que utilizam o scanner a laser, mesmo utilizando também a fotogrametria, o objeto é renderizado e colorizado via computação gráfica, o que ainda resulta em tratamento custoso para ajustar a iluminação para todos os ângulos de visão do objeto 3D. A aquisição por laser e imagens para colocar a textura e cor no objeto sofre com perda de qualidade (ALYAMAN; ABD-RAHEEM; ALDEIRI, 2019).

Os problemas de oclusões, ou seja, ocultação de partes não atingidas pelo laser, também é recorrente nesse tipo de aquisição (IZQUIERDO et al., 1999). No trabalho citado, o autor tentou reduzir o problema com uma plataforma giratória, onde foram fixados o laser e a câmera. Um exemplo para o problema de ocultação é a aquisição de uma mão com dedos abertos. Lateralmente, um dedo oculta detalhes do outro, isso só seria corrigido com outras varreduras por ângulos em que o laser atinja as partes ocultadas. No entanto, na reconstrução haveria a necessidade de unir os pontos de todas as varreduras.

Trabalhos com scanner a laser em alta resolução podem ser onerosos, tanto na construção da malha ou nuvem de pontos, como nos processos de colocação da textura e cor. A geração de bilhões de polígonos pode contribuir para se apresentar o objeto real com fidelidade, contudo o pós-processamento pode ser custoso computacionalmente.

O resultado pode ser aproximado ao real, mas exige um trabalho minucioso. Em objetos grandes, como uma estátua, um método automático de planejamento das varreduras pode fazer a diferença no tempo de aquisição e pós-processamento (LEVOY et al., 2000).

No trabalho citado, foi adquirida a estátua Davi de Michelangelo, como visto na Figura 2, percebe-se que a coloração e alguns detalhes anatômicos não ficaram similares, apesar de muito próximo do real.

As setas em vermelho indicam as diferenças encontradas após a renderização e colorização da estátua, como a coloração do mármore e diferença no olho direito (observador como referencial). O círculo que envolve a pupila também não foi reconstruído como apresentado na foto, pois apresenta um excesso de sombreamento abaixo. Na sobrancelha também houve um escurecimento e perdeu um pouco da forma. A estátua ganhou um ar mais sisudo, pois as

Figura 2 – Face da estátua Davi de Michelangelo: a) Foto e b) Varredura e colorização



Fonte: (LEVOY et al., 2000)

cavidades entre as sobrancelhas e nariz ficaram acentuadas.

A varredura a laser é utilizada em diferentes áreas e resoluções de problemas, por exemplo na área jurídica, em reconstrução de uma cena onde ocorreu um acidente fatal (EYRE et al., 2017). Contudo, os problemas na reconstrução e tempo para se ter o modelo podem ser empecilhos para a utilização de tais recursos.

Alguns tipos de aquisição laser envolvem a luz estruturada e câmeras para fotografar os padrões de luz. Mesmo com a tentativa de adquirir o objeto completo com feixe de laser e câmera, fixados em um braço robótico e giro do objeto, ainda apresenta problemas tanto na aparência quanto na cor (NIOLA; ROSSI; SAVINO, 2010).

Isso pode levar a técnicas com cálculos para distâncias e limiares das cores entre os pontos mais próximos, o que minimizaria os problemas na aquisição (TAKIMOTO et al., 2016). Porém, como em toda aquisição a laser, existem os problemas de estimar e corrigir as cores dos objetos digitalizados. Além do problema de vibração, quando se faz aquisições em objetos com movimentação, pois ela pode causar erros nas reconstruções 3D.

Os scanners a laser apresentam dificuldades em partes de objetos como curvas, reentrâncias, mas principalmente em superfícies com texturas de pelos, como um brinquedo de pelúcia. Estes problemas, principalmente as curvas e reentrâncias, levam a estudos de aquisições por superfícies reflexivas, ou seja, os objetos colocados sobre uma plataforma espelhada (FASANO et al., 2003).

A utilização desse espelho leva a dois objetivos, a automatização do processo de aquisição 3D e os problemas de peças difíceis. Subentende-se que pelos exemplos dados no artigo, sejam peças com curvas e cavidades, como vasos e conchas do mar. Houve também uma tentativa de aquisição de estátuas irregulares, nos quais somente o laser, sem o espelho, não haveria como atingir determinadas partes. Mesmo com este aparato, fica perceptível, nos resultados apresentados, que houve perda de algumas partes, principalmente das estátuas.

Com os problemas apresentados nas aquisições a laser, passa-se a debater outras formas de aquisições, uma vez que, a priori, o F2S2 sobressai sobre essas aquisições, em termos de representação fiel ao objeto real.

2.2.2 Outras formas de aquisição e reconstrução tridimensional

Esta seção tem por objetivo apresentar algumas formas de aquisição e a reconstrução 3D a partir das imagens adquiridas. A intersecção com a pesquisa desenvolvida é no âmbito de como se faz a aquisição, assim como também foi o objetivo na seção anterior com a varredura a laser.

A representação tridimensional de imagens digitais não é nova, já era debatida em 1981 dentro do campo da TC, que era apresentada como uma moderna tecnologia de aquisição. As técnicas para a época eram consideradas promissoras para a representação de imagens 3D discretas (SRIHARI, 1981). Nesse tipo de reconstrução, a imagem digital 3D é um mapeamento que associa cada voxel a um valor numérico, podendo ser real, inteiro ou binário. O voxel é a representação volumétrica do pixel.

Algoritmos para a geração de superfícies 3D a partir de imagens de TC, RM e também de tomografia computadorizada de emissão de fóton único (do inglês: *single-photon emission computed tomography* - SPECT) fazem uso do chamado “cubo de marcha” ou “cubo em movimento” (do inglês: *Marching Cubes*), onde se cria cubos de intersecção entre as fatias para determinar o que será reconstruído e o que será descartado (LORENSEN; CLINE, 1987).

O nome “cubo de marcha” ou “cubo em movimento” vem do fato de que ao terminar de analisar uma parte da fatia, para verificar se os pontos dentro ou na superfície dela faria parte da construção 3D, passava-se a analisar outra parte, assim por diante. Com isto, a reconstrução 3D era realizada com os pontos de intersecção das partes, chamadas de cubos, entre e dentro das fatias.

Ainda na década de 80, foi desenvolvido um trabalho para realizar a reconstrução 3D por meio de imagens que continham materiais diferentes, por exemplo gordura, tecido e ossos (DREBIN; CARPENTER; HANRAHAN, 1988). Buscava-se identificar os materiais dentro de um voxel, assim a reconstrução poderia evitar os chamados artefatos produzidos por *aliasing* (serrilhamento, efeito escada) e quantização (discretização do espaço de cor). A reconstrução era realizada por meio de polígonos, no entanto, segundo os autores, era uma abordagem que se sobressaía pelo fato de montar os polígonos na ordem do tamanho de um pixel.

Em 1991 foi publicado um *survey*, no qual são apresentadas as operações básicas e algoritmos para imagens médicas 3D, apontando as arquiteturas, desempenho, resolução produzida, tempo e grau de paralelismo de oito máquinas utilizadas para a renderização destas imagens (STYTZ; FRIEDER; FRIEDER, 1991). Nesse *survey* os autores colocam que não há diferença entre a geração de imagens médicas 3D e outros aplicativos de processamento gráfico. Que existem dois fatores bem conhecidos e intimamente relacionados, os quais ainda hoje são motivo

de preocupação para os desenvolvedores, que são o volume de dados e o custo computacional.

Em relação ao volume de dados, a TC média gera mais de 2 milhões de voxels por exame. Procedimentos de RM e ultrassom produzem quantidades semelhantes de dados. O segundo fator é que os algoritmos usados para imagens médicas em 3D têm um grande custo computacional mesmo com resolução 3D moderada.

Acerca da reconstrução de um volume 3D, pode-se utilizar duas formas de reconstruir estes volumes a partir da TC, sendo a renderização por fatias, uma de cada vez, ou renderização por matriz de pixels (plano bidimensional) (CABRAL; CAM; FORAN, 1994). A segunda é a mais natural, pois a saída de qualquer método renderizador de volume é uma matriz bidimensional de pixels. Mesmo assim, os autores também tinham a preocupação quanto ao custo computacional impetrado pelos cálculos necessários para a reconstrução 3D.

Em reconstruções 3D a partir de imagens no padrão DICOM, encontraram-se problemas de imprecisão. Por exemplo, na construção de modelos de mandíbulas com e sem dentes, com o intuito de contribuir para as subáreas de ortodontia, implantodontia, ortopedia e traumatologia, isso ocorreu na representação da geometria específica, devido à segmentação manual e imprecisão da tomografia, ausência de propriedades dos materiais e estruturas no osso medular (VASCO et al., 2016).

Ainda na área odontológica, pode utilizar a classificação de diferentes estruturas anatômicas dos dentes por imagens de microtomografia computadorizada (MicroCT), o que fornece informações para a análise de estresse (WANG et al., 2017). Este processo de rotulação é realizada manualmente por médicos clínicos experientes, consumindo um tempo considerável. Os autores propuseram uma segmentação e classificação automáticas, porém, segundo eles, devem vencer alguns desafios, pois as estruturas anatômicas e o entorno do dente em imagens médicas são bastante complexas.

Uma forma de aquisição que ganhou adeptos foi o ultrassom 3D, pois as gestantes podem olhar em monitores ou impressões o bebê em formação (LERTWIRAM et al., 2011). Além disso, há também a junção de outros métodos, como imagens de RM, para verificação de volumes (RUANO; JOUBIN, 2004). Contudo, é uma representação do volume com algumas restrições, já que não apresenta o bebê por completo, apenas a parte virada para a barriga da gestante.

Técnicas para melhorar essa visualização buscam diminuir os ruídos e interferências de ondas eletromagnéticas que sofrem dispersão após a interação com algum meio físico. Estes problemas podem dificultar a reconstrução 3D, principalmente na representação de superfícies contínuas e lisas. Técnicas como Classificação Variacional e *oriented splatting* são utilizadas com o intuito de melhorar a reconstrução (FATTAL; LISCHINSKI, 2001). Assim, os autores apresentam o resultado dessas técnicas, dizendo funcionar bem, particularmente em conjuntos de dados em que o objeto de interesse é grande e delimitado por uma superfície contínua, o que ocorre frequentemente em ultrassonografia fetal.

Vê-se nesses artigos que a reconstrução 3D apresenta a delimitação da forma do que foi adquirido, por exemplo o feto, no entanto, como ocorre neste tipo de aquisição, há perda de detalhes na renderização, porém isso é inerente a essa forma, já que se trata de algo interno ou envolto em diferentes tipos de materiais, como gordura, pele, osso e líquido.

As aquisições em RM e a RM funcional (fMRI) geram informações para a reconstrução 3D, neste contexto existem diversos trabalhos. Por exemplo, utilizar a reconstrução como forma de planejar e avaliar as melhores técnicas cirúrgicas em pacientes (TEKES et al., 2014). Neste trabalho, a reconstrução é realizada para prover uma visualização em uma interface 3D estereoscópica e coordenada com as imagens médicas. Utiliza-se os planos sagital, coronal e axial para a reconstrução.

Os volumes de dados DICOM podem ser visualizados com a percepção de profundidade 3D estereoscópica e isso pode dar suporte à visualização neurocirúrgica tridimensional (HINCKLEY et al., 1998). O foco dos autores foi produzir uma visualização que permitisse a manipulação com as mãos.

A utilização destes recursos 3D e de realidade virtual está crescendo vertiginosamente. Na educação tem-se abordado os repositórios e objetos de aprendizagem, os quais proporcionam a troca de conteúdos e permitem a docentes e discentes terem contato com materiais que muitas vezes não existem em suas instituições.

Para atender a essa demanda, pesquisadores da área médica focam em tecnologias como microscopia tridimensional e endoscopia, realidade virtual, simulação cirúrgica, robótica cirúrgica, cirurgia reconstrutiva e neuroimagem avançada. Todas elas contribuem com a relação cirurgião-computador e com o desenvolvimento de neurocirurgias. Essas ferramentas podem reduzir a curva de aprendizado, melhorar a compreensão conceitual e melhorar as habilidades de novos cirurgões (BERNARDO, 2017; ARDILA et al., 2018).

Uma alegação para o desenvolvimento dessas pesquisas é que mostra-se viável para utilização por diferentes acadêmicos e centros de saúde localizados em diferentes locais, como países em desenvolvimento. No entanto, um ponto desfavorável é que modelos desenhados ou colorizados por computador, como os apresentados nesses estudos, podem omitir partes importantes em relação ao real.

2.2.3 Citações acerca de aquisições para reconstrução ou representação 3D

Quadro 1 – Autores citados na revisão de literatura

Autor(es)	Assunto	Relevância
(BESL; JAIN, 1985)	Reconstrução 3D e aquisição laser	Apresenta a preocupação com recursos computacionais para execução da tarefa
(REMONDINO, 2003)	Reconstrução 3D e aquisição laser	Reconstrução por malhas de pontos de objetos digitalizados
(MARKIEWICZ; ZAWIESKA; PODLASIAK, 2017)	Fotogrametria e malhas de pontos	Reconstrução por malhas de pontos e fotogrametria
(OWDA; BALSÁ-BARREIRO; FRITSCH, 2018)	Fotogrametria e aquisição laser	Reconstrução de patrimônio histórico cultural
(DO; NGUYEN, 2019)	Fotogrametria e reconstrução 3D	Utilizar a fotogrametria estéreo para reconstruir o objeto
(FISCHER; MANOR, 1999)	Fotografia e aquisição laser	Utilizar câmera e aquisição a laser para reconstrução de objetos
(ALYAMAN; ABD-RAHEEM; ALDEIRI, 2019)	Fotogrametria e reconstrução 3D	Utilizar a fotogrametria para geração de um objeto 3D
(IZQUIERDO et al., 1999)	Aquisição por luz estruturada e laser	Fazer a reconstrução 3D utilizando luz estruturada e laser
(LEVOY et al., 2000)	Aquisição por laser e câmera	Reconstrução 3D de estátuas do Museu de Florença
(EYRE et al., 2017)	Aquisição por laser e câmera	Reconstrução 3D de ambientes por aquisições laser e câmeras
(NIOLA; ROSSI; SAVINO, 2010)	Reconstrução 3D por laser	Reconstrução 3D em tempo real
(TAKIMOTO et al., 2016)	Aquisição por laser e sensor RGB-D	Reconstrução 3D usando laser e sensor estilo Kinect
(SRIHARI, 1981)	Representação 3D por imagens	Representação 3D a partir de TC
(LORENSEN; CLINE, 1987)	Representação 3D por imagens	Representação 3D a partir de TC e RMI e SPECT
(DREBIN; CARPENTER; HANRAHAN, 1988)	Reconstrução 3D por imagens	Reconstrução 3D por imagens para diferentes materiais no corpo humano
(STYTZ; FRIEDER; FRIEDER, 1991)	Imagens médicas 3D	Survey apresentando técnicas para imagens médicas 3D
(CABRAL; CAM; FORAN, 1994)	Reconstrução 3D e imagens médicas	Técnicas para reconstrução a partir de TC
(VASCO et al., 2016)	Reconstrução 3D e imagens médicas	Reconstrução de mandíbula por meio de imagens
(LERTWIRAM et al., 2011)	Reconstrução 3D e ultrassom	Reconstrução 3D a partir de ultrassom
(RUANO; JOUBIN, 2004)	Reconstrução 3D e ultrassom	Reconstrução 3D a partir de ultrassom
(FATTAL; LISCHINSKI, 2001)	Reconstrução 3D e ultrassom	Técnicas para reconstrução 3D a partir de ultrassom
(TEKES et al., 2014)	Reconstrução 3D e imagens médicas	Reconstrução 3D por Rm e fRMI
(HINCKLEY et al., 1998)	Reconstrução 3D e imagens médicas	Reconstrução 3D por DICOM
(BERNARDO, 2017)	Reconstrução 3D e imagens médicas	Reconstrução 3D educação médica
(ARDILA et al., 2018)	Reconstrução 3D e imagens médicas	Reconstrução 3D educação médica

Fonte: Autoria Própria

2.2.4 Considerações sobre aquisição e reconstrução 3D

Abordou-se até aqui a aquisição e reconstrução 3D por diferentes tipos de equipamentos, como laser, TC, RM etc. O objetivo principal foi identificar se, mesmo com ótimos resultados derivados do processo, para o propósito de apresentação visual interativa e estereoscópica haveriam barreiras. Essas barreiras estão em dois principais pontos, sendo eles a constante avaliação e reconfiguração do foco de luminosidade e da câmera virtuais a cada movimento, além da avaliação da estereoscopia para todas as perspectivas do objeto. Além disso, os problemas encontrados nas reconstruções a partir de varredura a laser, luz estruturada e imagens DICOM, podem inviabilizar a visualização estereoscópica.

O foco de luminosidade é importante, pois ele permite a sensação de tridimensionalidade neste tipo de objeto. As funções ligadas à manipulação de objetos em 3D envolvem um conjunto de elementos para, ao final, obter-se o objeto e sua movimentação. Estas funções são a modelação, *rigging*, texturização, animação, simulação de fluidos, fumaça e partículas, renderização, composição e detecção de movimentos, assim como edição do vídeo.

O termo *rigging* está relacionado ao processo de animação 3D, o qual consiste na criação de um esqueleto para figuras animadas. Ao se terminar a criação de um objeto, este é estático tal como se fosse uma escultura. Por isso, há a necessidade do *rigging*. Isto não foi discutido ou apresentado no capítulo, pois focou-se na aquisição e reconstrução dos objetos.

Em trabalhos como a aquisição do Davi de Michelangelo em Levoy et al. (2000), apesar do resultado ser um objeto com algumas características do real, levantou-se alguns pontos que dificultariam sua utilização no objetivo final desta pesquisa. Um dos pontos é a aplicação da cor sobre o objeto, o que em muitos casos não condiz com o objeto real, outra é o tempo para se ter o objeto tridimensional, sem contar a aplicação das câmeras e focos luminosos para obtê-lo em estereoscopia.

Quando se faz uma aquisição por equipamentos laser, a oclusão se torna um problema, pois partes que sejam ocultadas ao laser, não serão capturadas e muito provavelmente não serão representadas como deveriam. Um exemplo disso é a parte interna entre dedos, a qual envolve um trabalho para ajustar, pois o laser tem que passar por todas estas partes e posteriormente o casamento delas para que sejam apresentadas.

Assim, para o trabalho com a estátua de Davi, o pós-processamento levou um tempo considerado grande para se ter a reconstrução em todos os detalhes. Além disso, as texturas como veias e outros elementos foram inseridos “artificialmente”, pois não foram capturados pelo laser.

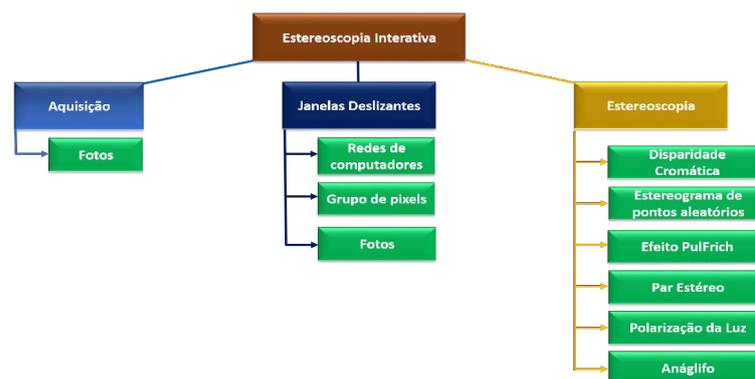
Para os outros tipos de aquisição e reconstrução, também não é promissor o uso para a geração dos objetos interativos estereoscópicos, pois em alguns casos eles não adquirem o objeto em todos os lados e em outros ocorre o mesmo caso da aquisição por laser, há a criação da malha e colorização por aproximação.

Mediante essas constatações, acredita-se que o objetivo de conseguir uma visualização bem mais próxima do que é o objeto real em forma, cor e textura, seria utilizando a fotografia em todos os ângulos dos objetos, como a aquisição realizada pelo F2S2.

2.3 Conceitos e técnicas de visualização 2D/3D, aquisições por fotos e estereoscopia

Nesta seção abordam-se os trabalhos que apresentam técnicas similares ou próximas aos métodos desenvolvidos nesta pesquisa. Assim sendo, serão apresentados os temas: algoritmos para gerenciamento e manipulação das imagens, aquisição por meio de fotografia e estereoscopia. Na Figura 3 são apresentados os temas que serão abordados no capítulo.

Figura 3 – Temas abordados relacionados ao trabalho



Fonte: Autoria Própria

2.3.1 Estereoscopia

As informações de profundidade e proporção podem ser geradas utilizando algum recurso artificial. Um exemplo disso é uma imagem única capturada por uma câmera fotográfica comum. Dependendo do tipo de imagem que se capture, tem-se as noções de profundidade e proporcionalidade (MURRAY, 1994).

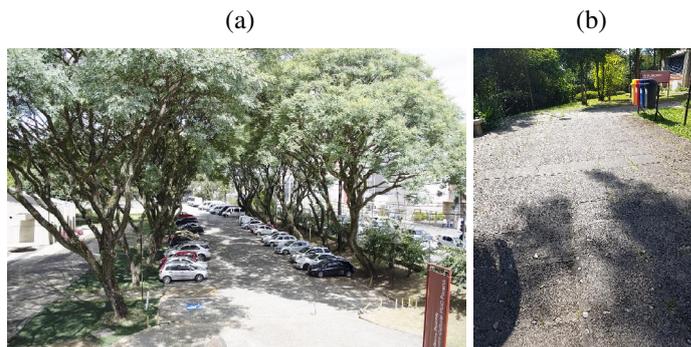
Em imagens estáticas têm-se os chamados efeitos passivos. São inerentes ao mundo, como enxergado, e independem da forma de sua captura. As características presentes nesse tipo de efeito são: perspectiva (KIM et al., 2011), iluminação e sombra (NUR et al., 2012), oclusão (SÁRA, 2010) e gradiente de textura.

Para se ter uma noção de profundidade, pode utilizar-se também o efeito de **gradiente de textura**. Esse efeito é uma repetição dos padrões presentes no ambiente ou objeto virtual. Por exemplo, ao se desenhar linhas onduladas paralelas iniciando na parte inferior da imagem indo

até a parte central ou superior em sentido cônico, causa-se a sensação de que as linhas estão indo no sentido do horizonte, distanciando-se do ponto de visão.

Existe a possibilidade de visualização desses efeitos em fotografias, sem a necessidade de aplicação por software. Uma imagem capturada por uma câmera comum pode apresentar os efeitos desejados, depende apenas de como a cena se apresenta. Por exemplo, nas fotos da Figura 4, as Figuras (a) e (b) contêm praticamente todos os efeitos passivos.

Figura 4 – Fotografia com efeitos passivos: (a) Estacionamento, (b) Calçada.



Fonte: Aatoria Própria

Diferentemente do tipo de efeito passivo citado, a estereoscopia realizada pelos seres humanos é do tipo efeito ativo. Isso significa que são necessárias duas imagens para que o cérebro forme a noção de profundidade e proporção. Uma imagem capturada por apenas um olho é tratada pelo cérebro como plana, perdendo, principalmente, a noção de profundidade.

Portanto, a estereoscopia advém do fato de que a visão humana apresenta uma diferença entre as imagens capturadas pelos olhos. O motivo para isso ocorrer é a posição deles, que em seres humanos são voltados para frente e com distância média de separação em 6.5cm. O Quadro 2 apresenta as medições feitas por pessoas que voluntariamente se disponibilizaram a passar os dados.

Quadro 2 – Diferença entre olhos

Observador	Medidas entre olhos em centímetros
Observador 1	6.5
Observador 2	6.8
Observador 3	6.7
Observador 4	6.4
Observador 5	6.1
Observador 6	6.6
Observador 7	6.8
Observador 9	6.6
Observador 10	6.4

Fonte: Aatoria Própria

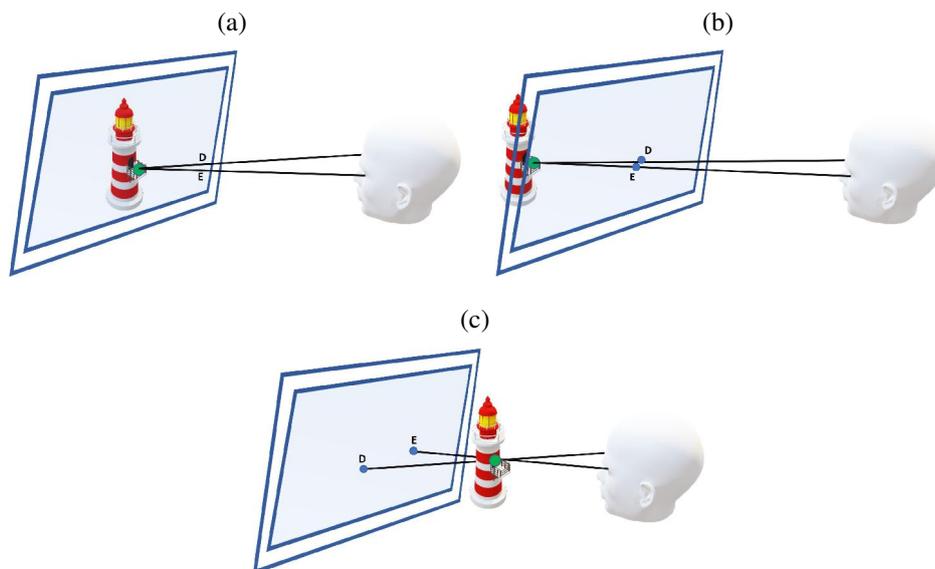
Essa diferença, chamada de **disparidade binocular** (TORI; KIRNER; SISCOOTTO, 2006), permite que sejam fornecidas duas imagens ligeiramente distintas, uma para cada olho, possibilitando as noções citadas, o que é conhecida como visão estereoscópica (do grego antigo “*στερεός*” (stereós) “firme, duro, sólido” + “*σκοπέω*” (skopeō) “ver”: Visão sólida).

Quando uma pessoa direciona sua visão para um objeto a certa distância, o campo de visão de cada olho se sobrepõe. Com isso, as imagens vistas pelos olhos são processadas pelo cérebro e este as funde, formando uma imagem única com as noções supracitadas.

Na formação de estereoscopia em dispositivos como TVs, monitores, celulares ou projetores, também há o conceito de distância. Contudo, este está relacionado com a **paralaxe** (CATELLI et al., 2017). Entende-se paralaxe como sendo as mudanças de posicionamento de um objeto em relação aos pontos diferentes de observação ou o deslocamento aparente de um referencial a partir do deslocamento de um observador. Ela permite que se tenha a sensação de profundidade, em outras palavras, tem-se a visão estereoscópica por meio da diferença de posicionamento dos olhos e do ponto de cruzamento das imagens em cada olho.

Destacam-se três classificações, as quais se baseiam no ponto de cruzamento das imagens para cada olho, como visto na Figura 5. Na Figura 5.a tem-se a Paralaxe Zero, nesta o ponto está exatamente no plano do dispositivo que disponibiliza a imagem. Na Figura 5.b tem-se a Paralaxe Positiva, onde a imagem é vista atrás do plano do dispositivo. Na Figura 5.c tem-se a Paralaxe Negativa, nesta a imagem é vista no plano a frente do dispositivo. Sendo E e D os olhos esquerdo e direito respectivamente.

Figura 5 – (a) Paralaxe Zero, (b) Paralaxe Positiva e (c) Paralaxe Negativa.



Fonte: adaptada de (LACOTTE, 1995; TORI; KIRNER; SISCOOTTO, 2006)

A relação da Paralaxe e da Disparidade, para a formação das imagens estereoscópicas, é que as duas são similares. Como visto, a paralaxe está relacionada ao plano de projeção das

imagens, assim sua medição é realizada nesse plano. Já a disparidade é uma medição relacionada à retina, diferença entre os olhos. O uso de algum aparato para se ver imagens estereoscópicas, como óculos, faz com que a paralaxe se torne a chamada disparidade retiniana. Assim, a paralaxe produzirá a disparidade retiniana e a disparidade será responsável pela produção da visão estereoscópica.

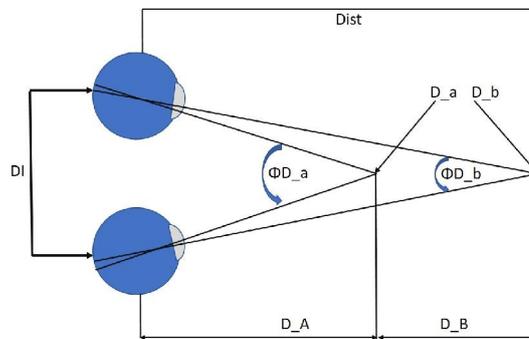
A paralaxe pode ser dada em termos de medida angular, assim relacionando-a com a disparidade ao levar em consideração a distância entre um observador e o plano de projeção. Como citado, existe uma diferença média (DM) entre os olhos, dada em centímetros, tendo também a distância, em centímetros, do observador para o plano (d), pode-se calcular a medida angular (ξ), como visto na Equação 2.1.

$$\xi = 2 \arctan \frac{DM}{2d} \quad (2.1)$$

Caso ξ seja superior a 1.5° poderá causar desconforto na visualização da imagem estereoscópica.

A disparidade binocular também é vista como sendo distância inter-axial, distância intraocular ou distância interpupilar, todas remetendo à sigla DI. Para algumas definições que se seguirão, esta será a sigla adotada. Algumas variáveis como ângulos e distâncias relacionados à paralaxe e DI são apresentadas na Figura 6.

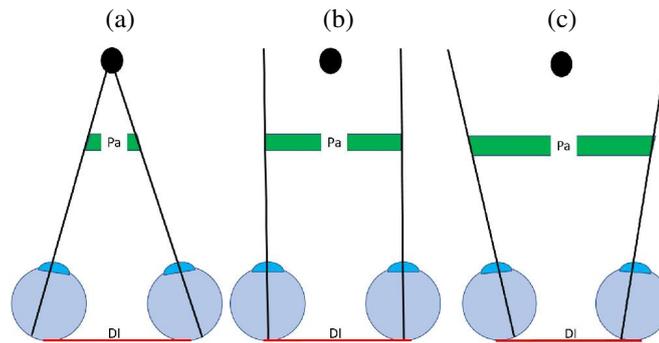
Figura 6 – Ângulos, DI e distâncias da paralaxe



Fonte: adaptada de (LACOTTE, 1995)

Sendo “Dist” a distância até qualquer objeto que se queira olhar. ϕD_a e ϕD_b são ângulos de paralaxe. Os pontos “D_a” e “D_b” são objetos que estão no foco para os ângulos definidos. As distâncias entre os olhos direito (D) e esquerdo (E) são “D_A” para o ponto “D_a” e “D_B” para o ponto “D_b”. Por fim, “DI” que já foi definido.

Com estas variáveis, pode-se encontrar informações como o menor ângulo e distância para visualizar um objeto, além da maior distância, aquela que a partir dela não se tem mais a noção de profundidade. A relação entre a DI e paralaxe, representada por “Pa”, é vista na Figura 7.

Figura 7 – Relações entre DI e paralaxe: (a) $Pa < DI$, (b) $Pa \approx DI$ e (c) $Pa > DI$.

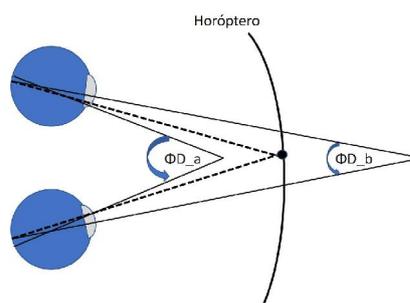
Fonte: Autoria Própria - baseadas em (LACOTTE, 1995; TORI; KIRNER; SISCOOTTO, 2006)

Das três relações, a que permite uma melhor visualização da estereoscopia é a com paralaxe menor do que a DI, visto na Figura 7.a. Existem limites para a visualização e esta é aproximadamente de 25cm para a menor distância.

Quando as duas forem muito próximas, o foco é distante, o que causa perda da noção de profundidade (PONZONI; DISPERATI, 1992). O limite para essa distância maior é de aproximadamente 720m. Se a paralaxe é maior do que DI, há perda de noção da profundidade.

Existe uma barreira imaginária, chamada horóptero, onde a disparidade é zero (do inglês: *zero parallax setting - ZPS*), ou seja, é o foco do observador. O horóptero é mostrado na Figura 8.

Figura 8 – Horóptero



Fonte: adaptada de (GROVE; HARROLD, 2013)

O foco é alterado conforme o observador muda o ponto de visualização entre as paralaxes. A vergência pode ser encontrada por meio da diferença entre os ângulos ϕD_a e ϕD_b , assim diz-se que a vergência é positiva se o observador afastou o foco do ZPS, negativa se ele aproximou o foco de ZPS ou nula se o foco foi mantido em ZPS.

Esses processos podem ser simulados por meio de câmeras, como já citado. Pode-se colocá-las lado a lado obedecendo a DI, a configuração de posicionamento pode ser em paralelo

ou convergente (*toed-in*). Para se conseguir a convergência nas imagens em paralelo, faz-se um deslocamento ou dos sensores da câmera ou por software diretamente na imagem capturada, tendo-se que verificar pontos correspondentes entre elas para se ter a distância ou paralaxe que permita a visão estereoscópica.

Esse deslocamento horizontal, transformação geométrica epipolar, pode não ocasionar uma sobreposição das imagens. Esta sobreposição pode ocorrer com um ajuste para o campo de visão comum, entretanto os planos de convergência podem não ser os mesmos e a estereoscopia causar algum desconforto.

O problema em câmeras convergentes, é que pode ocasionar um desalinhamento vertical ou paralaxe vertical entre os pontos das imagens, novamente causando desconforto. O efeito é uma distorção (distorção *Keystone*), a qual pode ocorrer pelo fato das câmeras estarem em planos diferentes (TORI; KIRNER; SISCOOTTO, 2006).

Não se pode confundir a paralaxe vertical com a formação da visão estereoscópica por meio de duas imagens, o par estéreo. A angulação das imagens do par estéreo, quando a aquisição é realizada por duas câmeras devem apresentar um alinhamento e sobreposição com maior parte de campo de visão comum possível, assim diminui-se a possibilidade de desconforto.

O par estéreo é uma das formas de montagem da estereoscopia mais utilizada, uma vez que ela é a base para três tipos diferentes de construção. O par dá origem ao anáglifo, polarização da luz e o próprio par estéreo.

Anáglifo

A estereoscopia anáglifo é uma técnica que exige a utilização de óculos para a visualização, que podem mudar conforme as cores definidas para cada imagem. Por padrão, são utilizadas duas cores e elas determinam também os filtros nos óculos. As cores vermelho e ciano (azul) são as comumente utilizadas.

A cor vermelha é aplicada sobre a imagem, por exemplo, que será apresentada ao olho esquerdo e a outra cor, ciano, é aplicada sobre a imagem que será apresentada ao olho direito. Então, os filtros das lentes funcionam como barreiras para as cores, sendo que a lente vermelha filtra, deixando passar o vermelho e a lente ciano filtra, deixando passar o azul.

A montagem de um anáglifo depende de duas imagens, sendo elas um par estéreo, apresentando uma pequena diferença de ângulos de captura, no entanto tem trabalho que utiliza apenas uma imagem e faz uma duplicação desta, mas exige cálculos de ajustes da paralaxe (CURTI; SIRTORI; VELLA, 2002), ou seja, uma transformação geométrica para corrigir a rotação de um ou mais eixos (x, y ou z).

Existem diferentes tipos de anáglifo, por exemplo, o cinza, o colorido, o semicolorido e o otimizado. Na Figura 9 estão quatro imagens com os tipos cinza, colorido, vermelho-verde e vermelho-azul.

Figura 9 – Anáglifos: (a) cinza, (b) colorido, (c) vermelho-verde e (d) vermelho-azul.



Fonte: Autoria própria

Polarização da luz

Quando se fala de polarização da luz (HUARD, 1996; WANG et al., 2015; LIVATINO et al., 2015), refere-se a filtrar os raios luminosos da luz natural. Um emissor de luz qualquer, como lâmpada e até mesmo o sol, produz raios luminosos formados por ondas eletromagnéticas que vibram em diversas direções durante a emissão, incidência e reflexão.

Cada onda faz uma vibração em determinada direção, sendo que para uma sempre há outra em um plano perpendicular. Um simples esquema que apresenta como se comporta a luz natural é a Lei de Malus (KUMAR, 2010; LAGES; GIUST; VIGOUREUX, 2008; HUARD, 1996).

Por meio de fenômenos como absorção, espalhamento, reflexão e birrefringência, consegue-se polarizar uma luz não polarizada (luz natural). Existem diversos materiais que permitem a passagem de apenas parte desta luz (onda), como algumas lentes de óculos de sol, polarizadores próprios para lentes de câmeras fotográficas e *insu filmTM*.

O processo de filtragem consiste em deixar passar uma luz por apenas um plano, assim para se fazer a estereoscopia, dois projetores enviam as imagens, sendo um para o plano horizontal e outro para o plano vertical. Apenas para constar, a intensidade é proporcional ao quadrado do cosseno da amplitude, dado pela Equação 2.2 da lei do cientista francês Étienne-Louis Malus. Uma luz normal passa por um cristal polarizador que tem um eixo de transmissão, a onda polarizada chega em um cristal analisador, o qual também tem um eixo de transmissão.

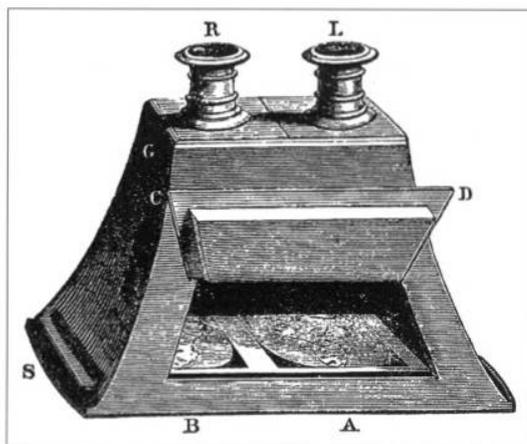
$$\kappa = \kappa_0 \cdot \cos^2 \Theta \quad (2.2)$$

Sendo, a intensidade de saída da luz de um analisador (κ) é dada pelo produto entre a intensidade de entrada (proveniente do polarizador) (κ_0) e o quadrado do cosseno do ângulo formado pela defasagem dos eixos de transmissão do polarizador e do analisador ($\cos^2 \Theta$).

Par Estéreo

A fotografia estereoscópica (HUARD, 1996; MEESTERS; IJSSELSTEIJN; SEUNTIËNS, 2004; SON; JAVIDI; KWACK, 2006), também foi alvo dos experimentos realizados pelo inglês David Brewster. Eram tiradas duas fotos da mesma cena e utilizando um estereoscópio para visualizar, tinha-se a noção de profundidade. O estereoscópio desenvolvido por Brewster é visto na Figura 10.

Figura 10 – Estereoscópio compacto de Davi Brewster.



Fonte: (ZONE, 2007)

Como nota histórica, esse estereoscópio foi apresentado em 1851 para a rainha Vitória do Reino Unido (ZONE, 2007). A Google desenvolveu o seu cardboard³ com duas lentes biconvexas, apenas adaptada para se acoplar um celular. O modelo mais simples utiliza papelão na sua confecção e pode ser encontrado, para que seja construído em casa, no site (notas de rodapé).

Para se criar um par estéreo pode-se fotografar algo obedecendo à distância média entre os olhos, como já citado é de 6.5cm. Então, utilizando essa medida, faz-se uma foto, em seguida, distancia-se aproximadamente os centímetros referentes à distância média dos olhos, no mesmo plano e inclinação, em relação a primeira foto e tira-se outra. No entanto, mesmo utilizando um tripé, pode-se ter imprecisão nas duas fotos, obrigando a realização de ajustes, como já citado.

Desconforto na visualização

Um tema que é debatido entre os pesquisadores e que merece atenção está relacionado com quão confortável uma pessoa se sente ao assistir conteúdos estereoscópicos. Nem sempre há relação com problema de boa sobreposição das imagens, mas sim com o equipamento utilizado. Aqui se faz uma abordagem para mostrar os problemas e o que vem sendo discutido para

³ <<https://vr.google.com/cardboard/get-cardboard/>>

minimizá-los. Contudo, a preocupação abordada neste trabalho é prover o par estéreo que não provoque o desconforto, mas não visa propor soluções para os equipamentos.

A fadiga visual em dispositivos de estereoscopia está relacionada à diminuição do desempenho do sistema visual, o qual pode ser medido por exames oftalmológicos. O desconforto está relacionado a questões subjetivas, decorrentes da fadiga (LAMBOOIJ et al., 2009).

As visualizações podem apresentar problemas de imagens borradas ou duplas, disparidade em telas, não ter relação de convergência-acomodação, associada à visualização de conteúdos estereoscópicos, além de conflitos que ocorrem na correção óptica optométrica/oftálmica. Corrigir esses problemas para uma visualização confortável, demanda estudos da compreensão dos conflitos de convergência e acomodação (SHIBATA et al., 2012). A questão do desconforto na visualização de conteúdos estereoscópicos 3D pode atingir pessoas em uma escala de até 50%, essas pessoas podem sofrer algum tipo de problema durante a visualização (TERZIC; HANSARD, 2016).

A aquisição é considerada um dos fatores que geram o desconforto, devido às configurações das câmeras. A configuração convergente é a mais popular, pelo motivo de reduzir a necessidade de cortes (*crooping*), no entanto causa o problema de incompatibilidade entre as visualizações. As filmagens ocorrem em cenas com movimento, o que pode gerar uma série de ajustes nas imagens, o que nem sempre é possível.

Para evitar tais problemas, acabam utilizando as câmeras em paralelo a uma distância igual à média da distância interpupilar. Contudo, esta forma também exige cuidados para que não haja o problema de paralaxe vertical. Neste tipo de aquisição se faz os cortes em áreas que não estão dentro do campo de visão de ambas as câmeras. Outro problema é a distância focal, pois dificilmente haverá alguma técnica que determine robustamente e rapidamente essa distância focal, uma vez que ela depende da geometria das câmeras e da zona de conforto.

A aquisição e correção dos problemas podem ser em tempo real (HEINZLE et al., 2011), no entanto os ajustes devem ser direcionados para diminuir o desconforto (JUNG; SOHN; RO, 2012). Contudo, as imagens capturadas necessitam de adaptações de acordo com diferentes ambientes de visualização estereoscópica.

Em relação aos problemas encontrados, percebe-se uma preocupação na forma de aquisição, assim como no pós-processamento e nos dispositivos de apresentação. A conclusão a que se chega é que existem problemas que ainda estão sem solução, mas que algumas combinações entre a aquisição e o pós-processamento buscam minimizar o desconforto na visualização estereoscópica.

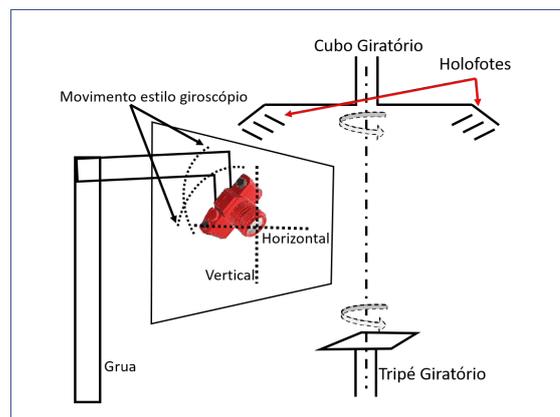
2.3.2 Aquisição e visualização por fotos

Esta seção se destina aos trabalhos que utilizam câmeras para realizar a aquisição e texturização de objetos, assim como, em alguns casos, a reconstrução tridimensional. As aborda-

gens que surgiram antes dos anos 2000 debatem a renderização baseada em imagem. Nessas abordagens, o sistema de renderização baseia-se em um conjunto de imagens pré-adquiridas para gerar diferentes visualizações de um ambiente. Mesmo assim, algumas aquisições apresentam o problema de oclusões e para corrigir isso foram propostas técnicas que apresentavam liberdade para as possíveis visualizações, tendo a representação do *Light Field*, no qual o foco luminoso é uma função de posição e direção nas regiões de espaço livres de oclusores (espaço livre) (LEVOY; HANRAHAN, 1996).

A representação do *Light Field* se assemelha à geometria epipolar (KUTULAKOS; VALLINO, 1998; ILLINGWORTH; HILTON, 1998; SHUM; SZELISKI, 1999). A semelhança reside no fato de que os volumes epipolares são formados por uma matriz de imagens criada por meio da mudança incremental em uma única direção de uma câmera. No trabalho de Levoy e Hanrahan (1996) foi utilizada uma câmera que adquire o objeto em quatro posições com rotação em 90° , feita manualmente. Um esquema da estrutura descrita no artigo citado é visto na Figura 11.

Figura 11 – Estrutura para aquisição de objetos - Light Field



Fonte: (LEVOY; HANRAHAN, 1996)

O *Light Field* também foi utilizado para a construção, compressão, renderização interativa e edição de superfícies de objetos reais (WOOD et al., 2000). A reconstrução dos objetos foi realizada por aquisição de fotos, onde foi utilizada uma grua com movimento esférico da câmera. Para a reconstrução 3D fizeram o registro do conjunto de fotografias na malha reconstruída, selecionando manualmente as correspondências entre os pontos da malha e os pontos de um pequeno subconjunto das fotografias.

A aquisição de objetos pode apresentar problemas, dependendo do tipo de objeto. Assim, algumas técnicas são testadas, afim de minimizar o problema. Uma das técnicas é a aquisição dos contornos do objeto, em alguns casos ele é fotografado juntamente com sua imagem opaca. Nessa técnica utiliza-se algum equipamento para produzir uma reflexão, por exemplo, duas

televisões, sendo uma como base e outra como plano de fundo (MATUSIK et al., 2002).

Esses trabalhos também envolvem a ideia de *light field*, pois apresentam uma abordagem de representar um modelo apenas conseguido sob um esquema de iluminação, o qual adquire o campo de refletância da superfície de múltiplas posições da visualização ao redor do objeto. Essa técnica, chamada de “casca” (do inglês: *hull*, que também pode ser entendida como carcaça ou invólucro) foi aplicada à renderização de superfícies inferiores e côncavas (FENG; SONG; ZHOU, 2008).

O processo de reconstrução 3D por meio de múltiplos pontos de vista, ou visualizações de um objeto, veio tomando corpo e melhorando sua qualidade conforme novas técnicas e equipamentos foram surgindo. Assim, a reconstrução 3D de objetos reais, buscando ter alta resolução tanto de textura quanto de forma, ainda é objeto de estudo. A utilização de computador, câmera digital e mesa giratória de alta precisão, busca aliar aquisições na horizontal (objeto fotografado em um número igual de passos, fazendo a cobertura completa em 360°) com algoritmos, como “octree”, para a reconstrução e visualização virtual do objeto (YEMEZ; SCHMITT, 2004).

A utilização de câmeras comuns para capturar as imagens e posterior reconstrução, também utiliza um esquema de matrizes, podendo gerar duas, uma para as cores e outra para o modelo. O modelo representa as superfícies como frontal, traseira, esquerda, direita, superior e inferior, assim, faz-se o preenchimento sequencial das superfícies no elemento correspondente da matriz de cores, obedecendo ao critério que é ocupar as cores referentes à visualização (TU; LIN; LEE, 2006). Esse processo, estando correto, gera o modelo 3D fazendo a combinação entre as matrizes.

Reconstrução de objetos 3D, a partir de fotos, por meio de diferentes técnicas, ainda é um problema da visão computacional. Grande parte dos trabalhos são voltados para uma reconstrução por malha de pontos, pois o objetivo é alcançar uma representação volumétrica virtual do objeto adquirido. Contudo, existem técnicas que fazem uso de uma malha de voxels, como também um conjunto de fotos, para obter a reconstrução (VOGIATZIS et al., 2007). Neste trabalho, os algoritmos utilizados foram o *Graph-cuts* e a votação *Normalized Cross-Correlation - NCC*. Os objetos foram colocados sobre uma mesa giratória e as fotos foram tiradas de forma automática.

Os resultados mostram que a reconstrução foi possível, porém percebe-se alguns pontos em que detalhes minuciosos não foram reconstruídos. Isso pode ser visto em todos os objetos, como na casa de brinquedo, templo e busto, além disso, a colorização não foi realizada, pois não era um dos objetivos.

A reconstrução completa e detalhada de objetos brilhantes apresenta desafios, principalmente para a textura, pois o reflexo pode interferir na representação do objeto. Já a silhueta, pode ser gerada por meio de múltiplos pontos de vista. A reconstrução pode levar em consideração

a estimativa do posicionamento da iluminação. Isto poderia ser utilizado em desenvolvimento via software de objetos 3D, pois um dos problemas encontrados em objetos dessa natureza é como representar um foco luminoso e os efeitos deste sobre o objeto. Contudo, um ponto a ser observado é que os objetos brilhantes refletem a luz, ou seja, suas superfícies contribuem para estimar o posicionamento e até a intensidade do foco (HERNANDEZ; VOGIATZIS; CIPOLLA, 2008).

Neste segmento de reconstrução 3D com apresentação “realística” de objetos reais em um ambiente virtual, como a internet, duas linhas que despertaram interesse foram as coleções de museus e instituições de Patrimônio Cultural. Essa abordagem busca aliar a reconstrução e apresentação de objetos de museu e patrimônio cultural. Para a reconstrução e visualização são utilizadas técnicas como a informação de refletância, por meio de uma técnica chamada *Bidirectional Texture Functions - BTF*, e do chamado *streaming* progressivo (SCHWARTZ; RUITERS, 2013). Em 2012 os mesmos autores aprearam o equipamento e detalharam a aquisição dos objetos. (SCHWARTZ; KLEIN, 2012).

Com uma proposta diferente da apresentada por Schwartz e Ruiters (2013), mas ainda no segmento de obras de museus, Nicolae et al. (2014) abordam a reconstrução utilizando malha de pontos e a texturização por fotos para objetos que apresentam materiais com propriedades ópticas difíceis (absortividade, refletividade, espalhamento), texturas e formas/geometrias complexas.

As técnicas de fotogrametria e a chamada *Lumigraph* (GORTLER et al., 1996) são empregadas ou dão base para grande parte destes trabalhos, pois segundo Berrier et al. (2015), um ponto chave da fotogrametria é não exigir equipamentos caros, além de necessitar de apenas uma câmera e tripé, o que seria facilmente manipulado por uma pessoa.

O trabalho de Faria et al. (2016) apresenta uma proximidade com este trabalho de doutorado. Os autores desenvolveram um equipamento para aquisição de fotos com uma câmera e um prato giratório. A câmera é presa em um braço com movimento na vertical, sendo que ela pode ser posicionada em diferentes pontos do braço, assim como o braço pode ser posicionado em diferentes ângulos na vertical. Subentende-se que os giros do braço aliados ao posicionamento na vertical gere uma aquisição em semiesfera. Entretanto, o processo foi manual, nos dois movimentos da câmera, giro em volta do objeto e deslocamento vertical.

A aquisição das perspectivas foram em ângulos determinados pelo equipamento que utilizaram, um controlador de giro de câmeras que pode ser acoplado aos tripés comuns para fotografia. Para testarem sua pesquisa, fizeram a aquisição de um cérebro real, criando uma base de entre 7300 imagens. A partir dessas imagens criaram a estereoscopia por meio de um software denominado PhotoMaker e utilizaram a biblioteca QuickTime para gerar as visualizações.

Pelo que foi apresentado no artigo, alguns problemas podem ser citados, por exemplo os ângulos de aquisição estão relacionados ao eixo de convergência (discutido neste trabalho na Seção 2.3.1), os ângulos da formação estereoscópica citados foram de 5° e 15°. Também

identificou-se que a interatividade era por meio de vídeos criados com a movimentação do objeto. Assim, a manipulação ficou restrita aos comandos do programa de execução do vídeo, os quais foram disponibilizados na internet⁴. Um vídeo apresenta a estereoscopia anáglifo, no entanto utilizando um óculos para visualizá-la, não se percebeu a formação estereoscópica, isto pode ser devido aos ângulos que escolheram para as imagens.

Em resumo, apesar das similaridades com este trabalho, elas estão apenas na forma como tentaram criar a estereoscopia anáglifo, no entanto, percebe-se que o trabalho não alcançou o que os autores esperavam, quanto ao objeto estereoscópico, como eles apresentam nos seus resultados. Os problemas em visualizações estereoscópicas também são frequentes em filmes, havendo a necessidade de ajustes, sem estes, é inviável sua visualização (SOUZA, 2009).

2.3.3 Algoritmo *Speeded Up Robust Features* para *matching* entre imagens

O algoritmo *Speeded Up Robust Features* (SURF) é utilizado para extração e descrição de pontos de interesse em imagens (BAY; TUYTELAARS; GOOL, 2006). Ele é considerado mais rápido e robusto do que alguns antecessores, como o algoritmo *Scale Invariant Feature Transform* (SIFT). Uma biblioteca SURF é fornecida pelos autores, no entanto, é de código fechado, contudo existem outras implementações de código aberto que podem garantir a qualidade do algoritmo original. Neste trabalho utiliza-se uma biblioteca para o Matlab⁵ chamada OpenSURF.

Seu funcionamento parte da detecção dos pontos de interesse, da aproximação da matriz de Hessian, além do conceito de imagens integrais, apresentado por Viola e Jones (2001). No artigo onde o algoritmo SURF é apresentado, os autores propõem a utilização de um detector baseado em uma matriz Hessian, contudo utilizam uma aproximação básica, como o detector Laplaciano básico. Assim como nas imagens integrais citadas, para reduzir o tempo computacional, esse processo recebeu o nome de detector “*Fast-Hessian*”.

As imagens integrais permitem uma rápida implementação de filtros de convolução do tipo caixa, onde a entrada dessa imagem é dada por $I_{\Sigma}(Xp)$, em uma posição $Xp = (xp, yp)$, que representa a soma de todos os pixels na imagem de entrada “I” de uma região retangular formada pelo ponto Xp e a origem, $I_{\Sigma}(Xp) = \sum_{ip=0}^{ip \leq xp} \sum_{jp=0}^{jp \leq yp} I(ip, jp)$.

A partir do valor encontrado para I_{Σ} , o algoritmo faz algumas adições na busca pela soma das intensidades sobre qualquer posição vertical, área retangular, independentemente de seu tamanho. O detector *Fast-Hessian* baseado na matriz Hessian é dado por um ponto $Xp = (xp, yp)$ na imagem “I”, sendo a matriz $Hp(Xp, \sigma)$ em Xp para a escala σ definida como na Equação 2.3.

⁴ <<https://vimeo.com/142885660>> e <<https://vimeo.com/142885661>>

⁵ <<https://la.mathworks.com/products/matlab.html>>

$$Hp(Xp, \sigma) = \begin{bmatrix} L_{xpxp}(Xp, \sigma) & L_{xpyy}(Xp, \sigma) \\ L_{yypx}(Xp, \sigma) & L_{yyyp}(Xp, \sigma) \end{bmatrix} \quad (2.3)$$

Sendo $L_{xpxp}(Xp, \sigma)$ a convolução da derivada gaussiana de segunda ordem $\frac{\partial^2}{\partial x^2}gp(\sigma)$ com a imagem “I” no ponto Xp e da mesma forma para $L_{xpyy}(Xp, \sigma)$ e $L_{yyyp}(Xp, \sigma)$. Aos símbolos originais, definidos pelos autores do artigo citado, foram alterados para não serem confundidos com os mesmos símbolos usados na definição do F2S2.

No artigo citado há maiores detalhes sobre o algoritmo, por isso, aqui foca-se em trabalhos que buscam a melhoria em tempo de execução e na utilização para gerar uma matriz fundamental, para, a partir dela, alterar uma imagem possibilitando o *matching*. Em Gossow, Decker e Paulus (2011) tem avaliação de implementações do algoritmo *Open source SURF*. No artigo, os autores concluem que o algoritmo implementado no software Pan-o-matic⁶ produz resultados similares aos da implementação original.

Em Han (2010) é abordado que a implementação original para a atribuição da orientação principal dos pontos importantes não é otimizado, pois utiliza a técnica de janelas deslizantes com deslocamento de 60 graus em torno de uma região circular. Com isso, obtém o cálculo das somas de todas as respostas Haar para produzir um vetor. Isso produz uma repetição em muitas regiões, pois há sobreposição. Consequentemente, há uma queda na velocidade, por conseguinte os autores propuseram uma modificação em que limpam todas as somas de respostas repetidas nas regiões sobrepostas. Com isso, todas as respostas de Haar eram calculadas apenas uma vez, com resultado de diminuição da complexidade e aumento da velocidade.

O trabalho de Dong Hui e Han Dian Yuan (2012) tem como objetivo utilizar o algoritmo SURF para detectar e descrever os pontos de interesse. Após obterem os pontos, estes são lançados em outro algoritmo, a árvore KD (do inglês *k-Dimensional Tree*), para realizar o *matching* nas imagens. Fazer o *matching* entre as imagens não é o que se pretende com este trabalho, mas a ideia é similar, pois o par estéreo deve ter uma compatibilidade entre os pixels para poder gerar a estereoscopia sem causar desconforto visual. Portanto, demonstrar-se-á que nas aquisições paralelas e convergentes necessita-se de ajuste da epipolar para que o *matching* seja realizado.

Um dos problemas que se pode elencar quanto ao uso desses softwares de *matching* é a complexidade computacional, uma vez que se está manipulando imagens. A imagem ocupa um espaço em memória e para analisar algo sobre ela, manipula-se os pixels, ou seja, uma análise pixel a pixel pode gerar um custo computacional. Por isso, alguns trabalhos, como o de Šluzek e Paradowski (2010), debate qual é essa complexidade quando se trabalha com duas imagens.

No trabalho dos autores supracitados, a ideia é analisar duas imagens quase duplicadas, por exemplo, um vídeo onde um frame é diferente do outro, mas o frame posterior mantém

⁶ <<http://aorlinsk2.free.fr/panomatic/>>

quase todos os detalhes do anterior. O objetivo é detectar fragmentos quase duplicados em tempo real. Por isso, uma análise da complexidade computacional se torna primordial para identificar possíveis pontos fracos e/ou gargalos do algoritmo. Para cada par de pontos-chave correspondente, o algoritmo de construção e análise de vizinhança tem a complexidade $O(n + m^3)$, se “m” vizinhos forem usados.

O algoritmo é aplicado a todos os pares de pontos-chave correspondentes, para que a complexidade geral da verificação da consistência topológica seja igual a $O(n^2 + nm^3)$. Assim, dadas duas imagens I e J, assume-se que elas contêm correspondentemente n_I e n_J pontos-chave descritos por vetores de característica de “f” dimensionalidade. Portanto, os métodos de *matching* de ponto-chave O2O (*one-to-one, mutualmente vizinhos próximos*) e NN (*Nearest Neighbor - caso especial de muito-para-muitos não mutualmente vizinhos próximos*) requerem operações fn_In_J . Assumindo que $n = \max(n_I, n_J)$, a complexidade computacional da correspondência de pontos-chave é $O(fn^2)$.

Os autores encontraram a complexidade computacional de $O(\log * n)$ para um conjunto de tamanho “n”. A complexidade total do algoritmo de localização de união é $O(mn \log * n)$. Onde, assumiram que $\log * n < f$ e $\log * n < m$, portanto a complexidade computacional da detecção de fragmentos quase duplicados com base em topologia em um par de imagens não é maior do que $O(fn^2) + O(n^2 + nm^3) + O(fn^2 + nm^2) = O(fn^2 + nm^3)$. Neste caso, pega-se o pior entre fn^2 e nm^3 , aquele que for maior caracteriza a complexidade.

A complexidade dos algoritmos SIFT, *Gradient Location and Orientation Histogram*-GLOH, e SURF é alvo de estudos (DREWS; De Bem; De Melo, 2011). Neste artigo, o resultado da análise do SURF, assim como do SIFT, foi que a complexidade é $O(mn + q)$, sendo “q” o número de pontos detectados. No entanto, concluíram que as constantes do SIFT são maiores e que isso o faz mais lento que o SURF. Contudo, quando se tem imagens, como as geradas pelo F2S2, deve-se analisar qual algoritmo pode ser apropriado, uma vez que as constantes deixam de ter importância.

A avaliação dos detectores e descritores vem sendo abordada para reconhecimento de objetos ou ajustes em imagens, como rotação etc. Em Canclini et al. (2013) é realizada uma avaliação abrangente com alguns algoritmos, para determinar aqueles ou a união de alguns que podem trazer o menor custo computacional. Entre os estudados estão o SIFT e o SURF, onde concluem que para alguns tipos de processos, como detector e descritor, o SURF e SURF com outra técnica, têm bons resultados. Contudo, após a análise de todos os algoritmos utilizados, os autores concluem que a avaliação mostrou que os descritores binários podem alcançar um desempenho semelhante, até mesmo maior, do que os descritores não binários e com complexidade menor. Indicam em trabalhos futuros o uso para a reconstrução 3D e também estereoscopia.

2.3.4 Algoritmos para carregamento de imagens em memória

As janelas deslizantes são utilizadas em diferentes segmentos da computação. A transmissão de dados utiliza os protocolos de janela deslizante na camada de enlace de dados (modelo OSI - Open System Interconnection), e também no TCP (camada de transporte do modelo OSI). Contudo, o interesse para este trabalho são as janelas utilizadas em imagens.

Neste contexto, aborda-se as implementações em diversos trabalhos, buscando focar nas diferentes maneiras de implementação ou propósitos, uma vez que para esta pesquisa será testada uma forma reconfigurável e dinâmica.

Um desses trabalhos diz respeito à identificação de caracteres em documentos produzidos durante o Império Otomano (YALNIZ et al., 2009). Foi utilizada uma união de técnicas como métodos de janela deslizante, segmentação de histograma, abordagens de reconhecimento usando recursos espaciais, redes neurais e um modelo baseado em grafos.

O processo de janela deslizante empregado no artigo é similar ao adotado neste trabalho, pois são de tamanhos variáveis, tendo um mínimo chegando ao máximo do tamanho do componente. Este componente é um bitmap que foi isolado pelos outros processos e corresponde a uma escrita ou palavra em otomano. A variação da janela se deve ao fato de que não há tamanho de fonte fixo nos documentos históricos.

O termo componente corresponde a um único caractere otomano ou a um número de caracteres combinados, além disso, os caracteres na biblioteca não são necessariamente do mesmo tamanho. Então, a janela está relacionada ao processo de obtenção de um segmento candidato do componente isolado para ser comparado com a base de caracteres. Em Fan e Baek (2018) também há a identificação de textos usando janelas e *Random Forest*, contudo em cenas cotidianas, como placas em ruas, fachadas, letreiros etc.

Com propósito diferente ao apresentado no reconhecimento de caracteres otomanos, as janelas foram utilizadas para a concatenação de partes de uma imagem, recompondo-a após ter sido dividida em blocos (SHEN; MIAO, 2010). A ideia principal foi o reconhecimento de ações humanas em vídeos em tempo real, para isto encontravam os contornos de uma pessoa na imagem.

A partir dos contornos, dividiam em blocos e calculavam os gradientes de orientação de cada pixel, incluindo a intensidade. Esses gradientes eram posicionados em oito caixas, calculado o histograma dos gradientes orientados para cada bloco. Assim extraíam a orientação de intensidade máxima do gradiente, o qual dava a característica do bloco. Cada bloco era uma parte da pose humana. Tendo isso, utilizaram as janelas deslizantes para concatenar os blocos, formando a pose completa. Então, utilizaram as poses como entrada para treinar uma Máquina de vetores de suporte (do inglês: *support vector machine-SVM*) e posteriormente aplicar em vídeo de tempo real e identificá-las.

Em Tang et al. (2018), o reconhecimento de gestos contínuos foi abordado para o

reconhecimento da trajetória da mão em escrita contínua, por exemplo, em cartas. O método proposto é *Structured Dynamic time wrapping - SDTW*, onde para validar o processo utilizaram a versão *Dynamic time wrapping - DTW* combinado com SVM e janelas deslizantes. Essas foram para o reconhecimento das trajetórias contínuas de segmentos em momentos de redução de velocidade, onde a abordagem, segmentação automática da trajetória contínua, combina modelos e informações de velocidade para identificar os pontos inicial e final nas trajetórias de gestos das mãos.

Uma proposta que utiliza as janelas para extração da região de interesse no campo da biometria pode ser vista em Yang et al. (2013). Segundo os autores, existem benefícios da técnica biométrica por análise das veias em relação a outros métodos. Neste trabalho são utilizadas as janelas de tamanho estipulados previamente para dois propósitos, no caso, o primeiro é o corte (*crop*) de uma região candidata de veia do dedo, para extração da região de interesse (*Region of Interest - ROI*). O segundo é a identificação da articulação falangeana, pois esta nem sempre é determinada com maior valor de cinza em imagens da veia do dedo. Assim, é realizada uma melhora no método de detecção da articulação com janela deslizante, na qual são calculados a soma dos valores de cinza dentro da janela.

Ainda na linha da biometria, a detecção automática da orelha humana tem despertado o interesse. O reconhecimento de orelhas está sendo visto com uma possibilidade na aplicação de segurança, como um complemento no reconhecimento de face. Em Wang, Du e Huang (2017) encontra-se um estudo utilizando classificadores e CNN para a detecção de orelhas. Além disso, utilizam as características de janela deslizante para uma rede neural totalmente convolucional a fim de melhorar a precisão na detecção de orelhas.

Por fim, uma abordagem interessante desenvolvida na Universidade de São Paulo (USP) por Mendes, Osório e Wolf (2017)). Eles utilizam um esquema de eleger um elemento presente em uma matriz e este determinará a janela deslizante, no entanto ele não faz parte da janela.

Para a formação da janela é utilizado um elemento chamado de Pivô, este é um ponto central na janela quando se utiliza a convolução de imagens. Entretanto, no artigo citado, a janela é construída acima desse pivô. A imagem do intervalo é usada para os propósitos de indexação e a computação em si é executada em domínio de nuvem de pontos.

A principal diferença é que as convoluções de imagem são geralmente executadas com um tamanho fixo de janela, enquanto o método desenvolvido usa um tamanho variável de janela que depende da coordenada do pivô. Este elemento é utilizado para uma busca dos pontos compatíveis e formam um cone invertido, que contém os pontos que formarão a nuvem.

Tanto o método desenvolvido pelo autores, quanto as convoluções em imagens, utilizam as janelas deslizantes guiadas por um elemento, que foi chamado de pivô. No caso das convoluções, esse elemento central faz parte da janela e ela é montada ao seu redor. No caso do trabalho citado, o elemento não faz parte da janela, mas é peça importante na sua construção.

2.3.5 Citações acerca dos trabalhos correlatos: fotos, estereoscopia e algoritmos

Quadro 3 – Autores citados na revisão de literatura - Parte 1

Autor(es)	Assunto	Relevância
(MURRAY, 1994)	Perspectiva e percepção da profundidade	Aborda a visão e a percepção da profundidade
(KIM et al., 2011)	Estereoscopia e Light Field	Framework para geração de conteúdos estereoscópicos
(NUR et al., 2012)	Disparidade Binocular e iluminação	Percepção da profundidade em diferentes condições de iluminação
(SÁRA, 2010)	Disparidade e oclusão	Modelagem da oclusão para correção da disparidade
(TORI; KIRNER; SISCOUTTO, 2006)	Realidade virtual e aumentada	Fundamentos da estereoscopia
(CATELLI et al., 2017)	Paralaxe	Determinar os ângulos da paralaxe
(LACOTTE, 1995)	Aquisições paralelas e convergentes	Problema da distorção trapezoidal - Keystone
(PONZONI; DISPERATI, 1992)	Estereoscopia e paralaxe	Limites nas visualizações estereoscópicas
(GROVE; HARROLD, 2013)	Horóptero	Apresenta a relação do horóptero vertical e horizontal
(CURTI; SIRTORI; VELLA, 2002)	Estereoscopia e visão Monocular	Apresenta cálculos da paralaxe para estereoscopia
(HUARD, 1996)	Estereoscopia e Polarização da Luz	Técnica estereoscópica por polarização da luz
(WANG et al., 2015)	Estereoscopia e Polarização da Luz	Técnica estereoscópica por polarização da luz
(LIVATINO et al., 2015)	Estereoscopia e Polarização da Luz	Técnica estereoscópica por polarização da luz
(KUMAR, 2010)	Polarização da Luz e Lei de Malus	Apresenta a lei de Malus relacionada à polarização da luz
(LAGES; GIUST; VIGOUREUX, 2008)	Polarização da Luz e Lei de Malus	Apresenta a lei de Malus relacionada à polarização da luz
(MEESTERS; JJSSELSTEIJN; SEUNTIENS, 2004)	Estereoscopia e Par Estéreo	Fotografia e a geração da estereoscopia por par estéreo
(SON; JAVIDI; KWACK, 2006)	Estereoscopia e Par Estéreo	Fotografia e a geração da estereoscopia por par estéreo
(ZONE, 2007)	Cinema e estereoscopia	Nota histórica sobre a origem dos filmes 3D
(LAMBOOIJ et al., 2009)	Desconforto visual e estereoscopia	Aborda o desconforto e fadiga visual em estereoscopia
(SHIBATA et al., 2012)	Desconforto visual e estereoscopia	Aborda a zona de conforto visual em estereoscopia
(TERZÍĆ; HANSARD, 2016)	Desconforto visual e estereoscopia	Aborda a zona de conforto visual em estereoscopia
(HEINZLE et al., 2011)	Aquisição e Estereoscopia	Correção de problemas estereoscópicos em tempo real
(JUNG; SOHN; RO, 2012)	Aquisição e Estereoscopia	Correção de problemas estereoscópicos em tempo real
(LEVOY; HANRAHAN, 1996)	Aquisição por fotos e reconstrução	Reconstrução 3D utilizando fotos
(KUTULAKOS; VALLINO, 1998)	Light Field e geometria epipolar	Semelhanças entre Light field e geometria epipolar
(ILLINGWORTH; HILTON, 1998)	Light Field e geometria epipolar	Semelhanças entre Light field e geometria epipolar
(SHUM; SZELISKI, 1999)	Light Field e geometria epipolar	Semelhanças entre Light field e geometria epipolar

Fonte: Autoria Própria

Quadro 4 – Autores citados na revisão de literatura - Parte 2

Autor(es)	Assunto	Relevância
(WOOD et al., 2000)	Light Field e reconstrução 3D	Light field em reconstrução 3D de superfícies
(MATUSIK et al., 2002)	Fotografia e reconstrução 3D	Reconstrução 3D por meio de fotos e bordas (“cascas”)
(FENG; SONG; ZHOU, 2008)	Fotografia e reconstrução 3D	Reconstrução 3D por meio de fotos e bordas (“cascas”)
(TU; LIN; LEE, 2006)	Fotografia e reconstrução 3D	Reconstrução 3D por fotos e matrizes de pixels
(VOGIATZIS et al., 2007)	Fotografia e reconstrução 3D	Utilização de malhas de voxels para a reconstrução 3D
(HERNANDEZ; VOGIATZIS; CIPOLLA, 2008)	Fotografia e reconstrução 3D	Reconstrução 3D por meio de fotos
(SCHWARTZ; KLEIN, 2012)	Fotografia e reconstrução 3D	Equipamento para aquisição de objetos
(SCHWARTZ; RUITERS, 2013)	Fotografia e reconstrução 3D	Reconstrução de patrimônio histórico
(NICOLAE et al., 2014)	Fotografia e reconstrução 3D	Propriedades de materiais e reconstrução 3D
(GORTLER et al., 1996)	Fotogrametria e lumbigraph	Base para reconstruções 3D por malhas e texturização
(BERRIER et al., 2015)	Fotogrametria e lumbigraph	Técnicas para a reconstrução 3D
(FARIA et al., 2016)	Estereoscopia e anáglifo	Apresenta a estereoscopia por ângulos de um objeto
(BAY; TUYTELAARS; GOOL, 2006)	Algoritmo de ajuste da epipolar	Algoritmo SURF para ajuste da geometria epipolar
(Viola; Jones, 2001)	Algoritmo de ajuste da epipolar	Algoritmo SURF para ajuste da geometria epipolar
(GOSSOW; DECKER; PAULUS, 2011)	Algoritmo de ajuste da epipolar	Algoritmo SURF para ajuste da geometria epipolar
(HAN, 2010)	Algoritmo de ajuste da epipolar	Algoritmo SURF para ajuste da geometria epipolar
(Dong Hui; Han Dian Yuan, 2012)	Algoritmo de ajuste da epipolar	Algoritmo SURF para ajuste da geometria epipolar
(ŚLUZEK; PARADOWSKI, 2010)	Algoritmo de ajuste da epipolar	Complexidade computacional em duas imagens
(DREWS; De Bem; De Melo, 2011)	Algoritmo de ajuste da epipolar	Complexidade computacional em imagens
(CANCLINI et al., 2013)	Algoritmo de ajuste da epipolar	Complexidade computacional em imagens
(YALNIZ et al., 2009)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(SHEN; MIAO, 2010)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(TANG et al., 2018)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(YANG et al., 2013)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(WANG; DU; HUANG, 2017)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(MENDES; OSÓRIO; WOLF, 2017)	Algoritmos e memória	Algoritmos para gerenciamento de imagens em memória
(SOUZA, 2009)	Estereoscopia	Problemas com aquisições de imagens para estereoscopia

Fonte: Autoria Própria

2.3.6 Considerações acerca dos conceitos e técnicas de aquisição por fotos e estereoscopia

Nesta seção foram debatidos trabalhos e definições da estereoscopia, janelas deslizantes, algoritmos de *matching*, a aquisição e visualização por meio de fotos. O intuito foi apresentar as correlações entre os trabalhos citados e o que se fez nesta pesquisa.

Sobre os algoritmos de janelas deslizantes, o que se verificou é a possibilidade de utilizá-los de uma forma diferenciada das apresentadas, pois não se trata de trabalhar com subconjuntos de uma imagem, mas com uma janela de imagens completas. Isso terá que ser avaliado via experimentos e em comparação com outras formas de gerenciamento das imagens em memória, como citado são duas outras implementações, uma de força bruta - carregar todas as imagens em memória, antes de gerar a visualização. A outra é carregar um par estéreo de cada vez gerando em tempo de execução a estereoscopia, que tem como ponto crítico, o constante acesso a disco.

Acerca dos algoritmos de *matching* e a complexidade computacional, estudou-se como eles são aplicados para gerarem uma imagem que pode proporcionar a noção de profundidade. A complexidade dos algoritmos foi abordada pelo fato de que, em alguns casos, as imagens não geram estereoscopia. A estas imagens devem ser aplicados os algoritmos detectores e descritores, para que sejam encontrados os pontos ou matriz de pontos que são utilizados para se fazer os ajustes de uma das imagens. Assim, percebe-se que haverá um custo computacional para que um par estéreo seja criado e que a estereoscopia seja gerada.

Quando se fala de complexidade computacional, no caso a notação *Big O*, em imagens, é clássica e corresponde a $O(n^2)$. Sendo quadrática, tende a se tornar muito ruim quando a quantidade de dados é grande. As imagens geradas pelo F2S2 são em 8K, o que pode levar a um custo computacional, pois a análise é em tempo real e o par estéreo, por meio da interação do usuário, pode ser trocado a todo momento.

Sobre a estereoscopia, o que mais se alinha a este trabalho são aquelas que utilizam duas imagens para sua formação, o par estéreo, pois aquelas que utilizam outras formas não apresentam o objeto como é o desejado.

As técnicas que se baseiam no par estéreo compreendem a anáglifo, polarização da luz e o próprio par estéreo. Quando se tem o par, pode-se colocar as imagens lado a lado e olhar via um estereoscópio, o qual pode ser facilmente montado tendo duas lentes biconvexas e seguindo os manuais que se encontram na internet. Contudo, atualmente se faz uso dos óculos RV e de celulares. Para a visualização polarizada, utiliza-se os equipamentos que forneçam esta estereoscopia, como é o caso de algumas televisões. Elas utilizam o par e seus próprios softwares embutidos encarregam-se de criar e apresentar a estereoscopia na tela e nos óculos, os quais podem ser passivos ou ativos.

A única regra para se visualizar ou ter a noção de profundidade, é que estas imagens tenham uma diferença angular entre elas. Isso pode ser conseguido pegando uma câmera e batendo

uma foto de uma paisagem em um determinado ângulo, mudar a câmera para aproximadamente 7° para o lado, qualquer lado, e tirar outra foto da mesma paisagem, utilizando algum programa para colocá-las lado a lado. Fazendo isso, coloca-se o resultado da junção das fotos para ser visualizado no celular e nos óculos RV (Realidade Virtual).

Contudo, apenas visualizar uma parte do objeto, no exemplo utilizou-se o termo paisagem, não é objetivo deste trabalho, o que se quer é poder olhar o objeto por completo, todos os ângulos de forma estereoscópica. Assim, necessita-se tirar as fotos de todas as perspectivas e fazer o casamento daquelas que estiverem na distância angular que propicie a estereoscopia.

Para a formação do anáglifo, o par deverá sofrer uma alteração de canais de cores, para deixar a imagem esquerda apenas com o canal vermelho e a imagem direita com o canal azul e verde, formando o ciano. Também faz-se a alteração estrutural nos pixels. Essa alteração está relacionada com operações do processamento de imagens, como adição e subtração.

O último trabalho apresentado na seção de aquisição por fotos, apesar de similaridades com esta pesquisa, mostrou que alguns erros levaram ao fracasso na apresentação estereoscópica, a ponto de descartarem sua utilização. Acredita-se que o problema esteja na formação do par estéreo, uma vez que os ângulos utilizados não são os ideais. Como visto na seção 2.3.1 e como será mostrado neste trabalho, não basta pegar os ângulos de convergência para montar a estereoscopia, necessita-se dos cálculos da paralaxe ou pelo menos usar a distância interpupilar para a formação do par estéreo. Além disso, uma aquisição com ângulos até 6° pode gerar uma visualização mais contínua do objeto, acima deste ângulo, pode ser que fique uma visualização quebrada, ou seja, não apresentar o objeto como ele é no mundo real.

Neste capítulo foram abordadas as principais premissas e os conceitos utilizados nesta pesquisa. No capítulo seguinte são apresentadas a metodologia e a infraestrutura para o desenvolvimento da pesquisa.

Capítulo 3

Metodologia e infraestrutura para o desenvolvimento da pesquisa

3.1 Introdução

Esta pesquisa é de natureza Aplicada, pois objetiva a geração de conhecimentos para, especialmente, à construção de OA, que é uma aplicação prática dirigida para a área instrucional. Seus objetivos são de cunho Exploratório, uma vez que adentra-se em um campo que, apesar de bem definido, que é a estereoscopia, não há conhecimento suficiente quanto ao comportamento das aquisições realizadas pelo F2S2 na geração da visualização estereoscópica. Assim, classifica-se os procedimentos realizados como Exploratório, já que propõe-se uma nova abordagem para representar e manipular um objeto real em ambiente virtual, sem que haja necessidade de reconstruir tridimensionalmente o objeto ou de ajustar as imagens do par estéreo para visualização estereoscópica.

Então, neste capítulo são apresentados a metodologia e infraestrutura para o desenvolvimento da pesquisa. São abordados: a infraestrutura, o *Full Frames Semi-Spherical Scanner*, o modelo proposto e o delineamento da pesquisa.

3.2 Infraestrutura para o desenvolvimento da pesquisa

O grupo de pesquisa CIIM surgiu em 2014 com o desenvolvimento de dois projetos que correm paralelamente. O primeiro projeto é o “*The Visible Human Table*”, onde há a projeção dos dois corpos humanos do 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, entretanto 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 doutorandos Diogo Roberto Olsen e Flávio de Almeida e Silva (autor deste trabalho).

A idealização do F2S2 foi do professor Me. Manoel Campos de Almeida e o professor Dr. Edson José Rodrigues Justino. Eles, em reunião com professores da medicina, em especial professor Dr. Luiz Roberto Aguiar, foram informados do descontentamento com a forma com que a anatomia é ensinada, além dos problemas encontrados em objetos desenhados, fotografados ou reconstruídos em 3D, os quais acabam ocultando detalhes ou perdendo características como

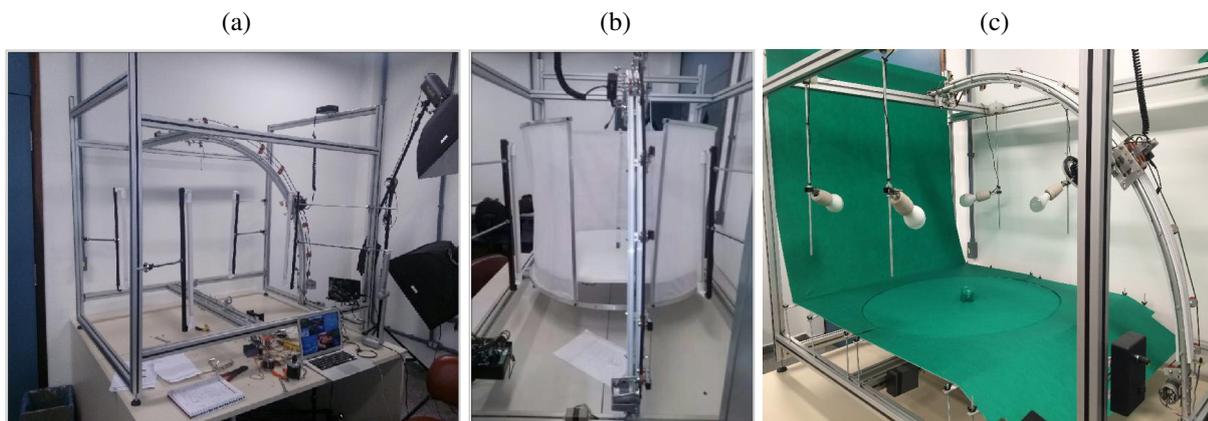
cor e textura, como apresentado no Capítulo 2.

Assim, foi elaborado 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 12.b). 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 12.c).

O sistema anterior utilizava lâmpadas tubulares de Led (Figura 12.a). O problema das lâmpadas foram a temperatura de cor, isso causava diferenças na aquisição. No desenvolvimento de hardware e software de controle do scanner houve a participação do acadêmico citado. Na Figura 12 são apresentas as duas versões do scanner.

Figura 12 – Versões anteriores do F2S2: (a) estrutura da versão 1; (b) plano de fundo; e (c) versão 2 do scanner.



Fonte: CIIM

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 (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 de gravidade.

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 a estereoscopia, uma vez que os ângulos poderiam não privilegiar a visão binocular.

Como será visto na Seção 3.3, página 82, na terceira versão, alguns desses problemas foram contornados, melhorando a aquisição. O scanner foi fechado, deixando de ter a influência 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. Na Figura 13 é apresentado o início da construção da terceira versão.

Figura 13 – Versão 3 do F2S2: (a) estrutura e (b) plano de fundo e iluminação.



Fonte: CIIM

Alguns objetos foram digitalizados ainda na segunda versão do scanner, dando abertura à procura de como encaixar tais aquisições no ambiente instrucional, principalmente na educação. Percebeu-se que as aquisições podem se tornar OA com 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.

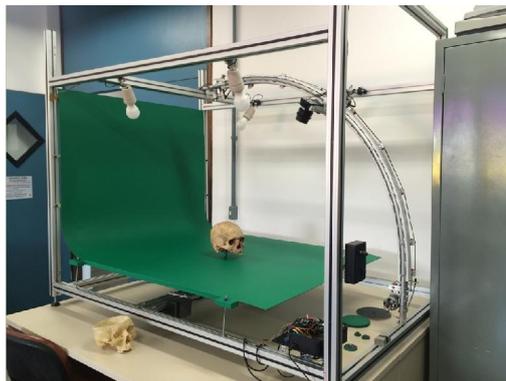
Uma das vantagens das aquisições feitas no scanner é permitir a visualização dos objetos adquiridos em diferentes tipos de dispositivos, como televisões com e sem 3D, óculos de realidade virtual, pirâmides holográficas (técnica conhecida com Ghost Pepper's), ambientes imersivos etc. Isso foi apresentado e debatido por Olsen. et al. (2018) no artigo intitulado “*Generation of Stereoscopic Interactive Learning Objects True to the Original Object*”.

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. Na Figura 14 é apresentado um dos crânios sobre a base giratória e o outro sobre a mesa.

Criou-se uma visualização 3D, a qual foi chamada de “visualização interativa”, pois

Figura 14 – Aquisição de dois crânios para geração de OA.



Fonte: CIIM

permite visualizar todas as perspectivas de um objeto digitalizado. Além dessa visualização, também se criou 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, em sua terceira versão, onde, a partir desta, houve a participação do doutorando Diogo Roberto Olsen e Flávio de Almeida e Silva. Estes contribuíram com ideias para a montagem do plano de fundo e também do posicionamento da iluminação, uma vez que a concepção da estrutura da aquisição, por deslocamento angular da base e da câmera, manteve-se fiel à ideia original, pensada pelos professores Justino e Manoel.

Como será visto na Seção 3.3, página 82, houve melhorias no hardware, buscando precisão nas aquisições e no posicionamento dos objetos. Além do F2S2, o trabalho conta com um notebook da marca DELL, modelo Inspiron 15 Série 7000 (I5-7559-A20) com uma unidade de estado sólido (em inglês: *solid-state drive* - SSD, *Next Generation Form Factor* - NGFF), também conhecida como M2, com 16 gigabytes de memória, processador Intel CORE i7 e duas placas de vídeo, sendo uma Intel(R) HD Graphics 530 e a outra uma NVidia Geforce GTX 960M. Para testar as projeções estereoscópicas foi utilizado um Projetor Optoma HD27HDR 3D Full HD 3400 Lumens 1080p.

Em questão de softwares, foi utilizado o Gimp¹, software de edição de imagens, o StereoPhoto Maker², para produção de algumas estereoscopia, o Matlab com licença trial, para executar o software SURF, que é de distribuição livre, para o ajuste das imagens do par estéreo, assim como alguns algoritmos de teste para redução de contexto e segmentação das imagens.

Também foi utilizado o visual studio para desenvolvimento de um visualizador para

¹ <<https://www.gimp.org/>>

² <<https://stereo.jp.org/eng/stphmkr/index.html>>

prova de conceito, juntamente com a linguagem C++ e a biblioteca SDL2³ para jogos, também de distribuição livre. O ambiente de desenvolvimento foi o Microsoft Visual Studio⁴, licença para acadêmicos da PUCPR.

Os objetos relacionados no Quadro 5 foram adquiridos, no entanto outros objetos não relacionados também foram digitalizados, alguns em quantidade diferente de streams e frames. As imagens de todos os objetos, mesmo os que não estão no quadro, foram colocadas no Apêndice A.

Quadro 5 – Base de objetos adquiridos

Nome do objeto	descrição	Aquisição
Massa de Modelar	Massa moldada em forma irregular de 20cm largura por 12cm altura	30 Streams e 120 Frames
Papel	Papel dobrado e moldado com fita adesiva de 18cm largura por 10cm altura	30 Streams e 120 Frames
Boneco	Boneco de plástico de 15cm largura por 23cm altura	30 Streams e 120 Frames
Concha	Concha do mar de 20cm largura por 10cm altura	30 Streams e 120 Frames
Boneco2	Boneco de madeira com bola de plástico de 7cm largura por 25cm altura	30 Streams e 120 Frames
Maçã	Maçã vermelha de 6.5cm largura por 8.3cm altura	30 Streams e 120 Frames
Peça de carro	Parte interna do motor de arranque de 7cm largura por 12.8cm altura	30 Streams e 120 Frames
Vaso	Vaso de cerâmica de 12cm largura por 21cm altura	30 Streams e 120 Frames
Perna de Cavalo	Metade da parte da perna, corte medial, de um cavalo, 12cm largura por 31cm altura	30 Streams e 120 Frames
Casco de Cavalo	Metade da parte da perna e casco, corte medial, de um cavalo, 16cm largura por 33cm altura	30 Streams e 120 Frames
Carro de brinquedo	Carro metálico de brinquedo de 11cm largura, 22 comprimento e 6cm altura	30 Streams e 120 Frames

³ <<https://www.libsdl.org/download-2.0.php>>

⁴ <<https://visualstudio.microsoft.com/pt-br/>>

3.3 Full Frames Semi-spherical Scanner

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. 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.

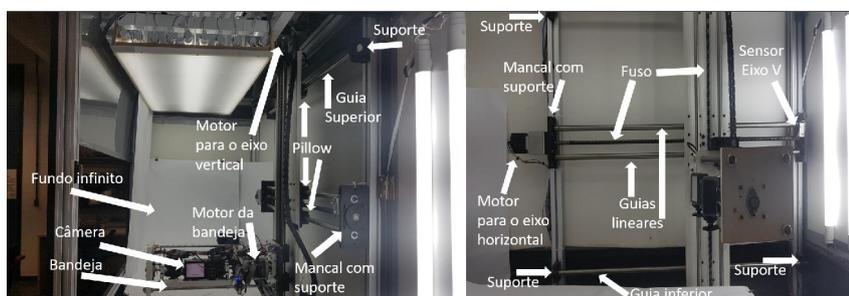
O termo “*Semi-spherical*” se refere à possibilidade de iniciar a aquisição em um ângulo, normalmente 0° , posição frontal ao objeto e paralela à base giratória e plano de fundo, terminando em um ângulo de 90° em relação à base giratória. Assim, criou-se o nome a partir dos dois “F” (*Full Frames*) e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

3.3.1 Arquitetura do F2S2

A atual versão (terceira) do F2S2 conta com quatro graus de liberdade, sendo o movimento na horizontal (Eixo H), na vertical (Eixo V), ângulo da câmera (Eixo C) e giro da base (Eixo B). Os dois eixos de deslocamento H e V são formados por componentes similares a de um Controle Numérico Computadorizado (sigla CNC, do inglês: *Computer Numeric Control*): guias lineares cilíndricas; mancal com suportes para guias; suporte para barras guias; *pillow* (ou *plummer block*); fuso com castanha; case para castanha; e acoplamento.

Em cada extremidade dos fusos estão motores com torque para até dez quilogramas de peso. São motores de passo com precisão de 0.9 graus. Na Figura 15 estão discriminados os eixos, motores, fusos, bandeja, câmera, fundo infinito, guias, barras guias inferior e superior. Alguns componentes não estão visíveis, como a castanha e seu *case*, alguns *pillows* e os acoplamentos. Estes são junções entre os eixos dos motores e os fusos.

Figura 15 – Estrutura do F2S2 terceira versão.



Fonte: autoria própria

O eixo V contém uma madeira, para que possa dar sustentação ao deslocamento no sentido horizontal, nela estão fixados os *pillows* e o *case* com castanha do eixo H. Além das guias lineares e fuso para o deslocamento da estrutura que contém a bandeja e câmera.

A bandeja (Eixo C) tem um movimento angular (de inclinação), o qual é configurado para sempre direcionar o centro da lente da câmera para metade da altura do objeto que será adquirido. Com isso garante-se que a câmera iniciará paralelamente e terminará em um ângulo de 90 graus em relação à base giratória. Na Figura 16 a bandeja é vista em diferentes posições.

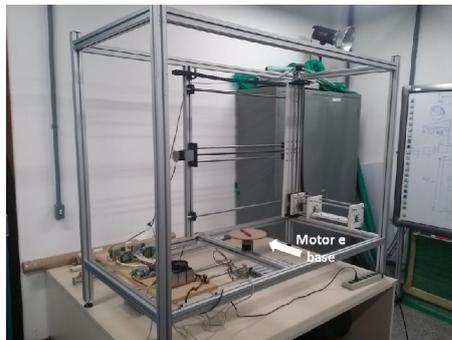
Figura 16 – Posicionamentos da bandeja.



Fonte: autoria própria

A base giratória (Eixo B) é composta de um prato acoplado diretamente ao eixo do motor, como apresentado na Figura 17. Esse motor também é de passo e tem precisão de 0.9 graus. Ela comporta um peso de até dez quilogramas. No entanto, há um problema com determinadas bases e objetos, isso será debatido na Seção 3.3.4 - limitações.

Figura 17 – Prato ou base giratória e motor.



Fonte: autoria própria

No sistema de iluminação (Figura 18), as lâmpadas podem ser ligadas ou desligadas, conforme os parâmetros dos objetos e fundo. Na parte superior o conjunto conta com dez lâmpadas e na porta de acesso à câmera são quatro.

Todas são de *Led* com 6500k de temperatura de cor, para não influenciar ou alterar a cor e textura do objeto, ou seja, se elas forem de configurações diferentes, a intensidade luminosa pode mudar, dando uma tonalidade ou textura diferente onde essa luz incidir no objeto.

Na Figura 18 são apresentadas apenas três configurações possíveis de iluminação da parte superior, mas isso pode ser feito também para a iluminação posicionada na porta.

Figura 18 – Sistema de iluminação do F2S2.



Fonte: autoria própria

Os controles dos motores e disparo automático da câmera são efetuados por um microcontrolador ATMEGA 328, implementado com G-Code. Existem alguns códigos para isso e uma discussão ampla pode ser vista em fóruns, como o encontrado no site do Arduino⁵. Uma apresentação mais completa do G-Code pode ser vista na Wiki da RepRap⁶.

O conjunto de imagens resultante da aquisição é processado para se ter o chamado Stream2D. O processo inicial tem relação com o trabalho do doutorando Diogo Roberto Olsen, que são as correções como balanço de branco, correção de lente, entre outros. Outros processos como redução de contexto, segmentação e organização dos nomes, que também estão relacionados ao trabalho desse doutorando, estão descritos na subseção 3.3.2.

3.3.2 Funcionamento do F2S2

O processo de aquisição parte de um planejamento inicial, levando em consideração as características do objeto. Por exemplo, é desejável que o fundo seja o mais contrastante possível com qualquer parte do objeto.

Contudo, alguns objetos multicoloridos podem levar a um planejamento mais detalhado, considerando-se a iluminação e a cor do fundo, de forma que não dificultem o processo da segmentação. Um exemplo dessas considerações é apresentado na Figura 19.

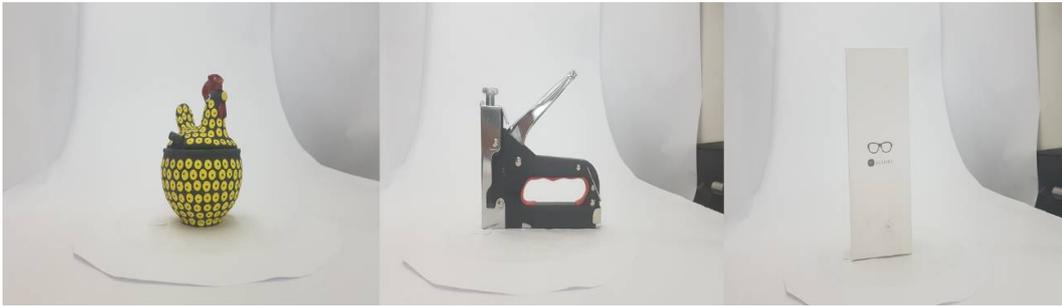
Na Figura 19 percebe-se três diferentes situações, sendo que na primeira o objeto está totalmente contrastante com o fundo branco, na segunda existem as cores preta e vermelha que contrastam, mas a prata pode dificultar o processo de segmentação. Na última tem-se um problema, uma vez que a caixa é branca e não há contraste, a segmentação seria trabalhosa, não sendo indicado fazer a aquisição com esse fundo.

Ao se determinar o fundo, parte-se para a quantidade necessária de imagens para o trabalho ao que se destinará. Se a escolha for uma aquisição de 1° em 1°, serão 90 streams com

⁵ <<https://forum.arduino.cc/index.php?topic=259360.0>> ou <<https://forum.arduino.cc/index.php?topic=318302.0>>

⁶ <<https://reprap.org/wiki/G-code>>

Figura 19 – Contraste entre o fundo infinito e os objetos



Fonte: autoria própria

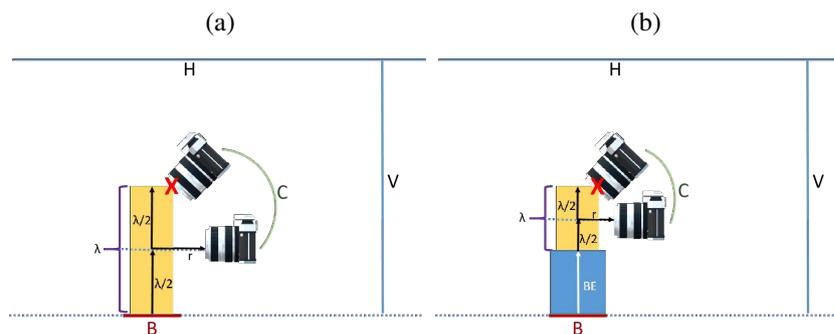
360 frames cada, perfazendo um total de 32.400 (trinta e duas mil e quatrocentas) imagens. Este trabalho mostrará, no Capítulo 4, página 109, quais são os graus entre frames e também streams para se realizar a visualização estereoscópica.

A altura e largura do objeto determina a mínima distância que se pode posicionar a câmera para fazer a aquisição, pois o deslocamento em arco tem que ser calculado para não haver choque entre a estrutura da bandeja ou a lente da câmera e o objeto ou suporte da iluminação superior. Então, o cálculo leva em consideração a altura do objeto, juntamente com uma base de elevação (BE), caso haja necessidade, e o raio de deslocamento.

Assim, se a altura do objeto (λ – lamba, uma vez que “H” está representando o eixo horizontal) for de 18cm, o posicionamento inicial da câmera não poderá ser igual ou inferior a $\frac{\lambda}{2}$. A distância entre a câmera/estrutura de suporte da bandeja e o objeto, com ou sem BE, é determinada pelo raio (r).

O “r” sempre é fixo para uma aquisição e obedece ao critério citado acerca do tamanho do objeto. Por exemplo, sendo $\lambda = 18\text{cm}$, não poderá ter um raio igual ou menor do que 9cm, pois assim a estrutura da bandeja se chocaria com o objeto. Na Figura 20 é apresentada a situação com e sem BE.

Figura 20 – Distância mínima de configuração do raio: (a) sem BE e (b) com BE.

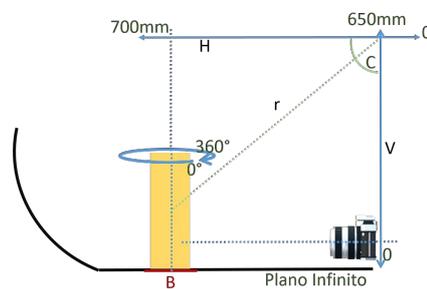


Fonte: autoria própria

Portanto, o “r” terá que obedecer ao seguinte critério: $\frac{\lambda}{2} < r < 65\text{cm}$, onde 65cm (650mm) é a altura do sistema de iluminação superior. Se o “r” for igual ou superior a este tamanho, a estrutura de suporte da bandeja se chocará com esse sistema de iluminação, pois tem a restrição de tamanho de objeto para ser adquirido, o que será debatido nas limitações do F2S2.

Os eixos têm seus inícios e fins determinados de maneira diferente do habitual, ou seja, para o F2S2 a posição 0(zero) do Eixo “V” encontra-se na parte superior direita, pois o deslocamento é da direita para a esquerda. Já o 0(zero) do Eixo “H” encontra-se na parte inferior direita, uma vez que o deslocamento é para cima. Essa configuração é vista na Figura 21.

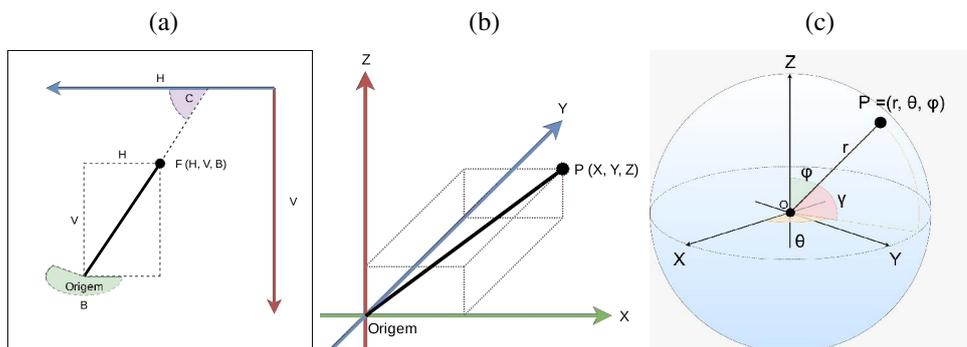
Figura 21 – Mínimos e máximos dos eixos do F2S2



Fonte: autoria própria

Os eixos X, Y, Z, H, V, B e C, com suas representações nos planos cartesiano e esférico, são apresentados na Figura 22, porém suas definições matemáticas e discussões fazem parte do trabalho do doutorando Diogo Roberto Olsen. Cabendo aqui a apresentação para fundamentação da visualização estereoscópica.

Figura 22 – Planos do F2S2: (a) eixos; (b) coordenadas cartesianas; e (c) coordenadas esféricas.



Fonte: Diogo Roberto Olsen

Os eixos X, Y e Z são respectivos ao plano cartesiano, onde um ponto deste plano é $P = (X, Y, Z)$. Os eixos H, V e B são relativos ao F2S2, onde um ponto é representado por $F = (H, V, B)$ e diz respeito a um ângulo ou perspectiva do objeto. O Eixo “C”, refere-se ao posicionamento da câmera que tem sempre um ponto, ou seja, para cada “F” o Eixo “C” será direcionado ao

ponto médio de altura do objeto. Contudo, a cada mudança de “F”, mesmo “C” sendo o ponto médio, a perspectiva capturada é diferente.

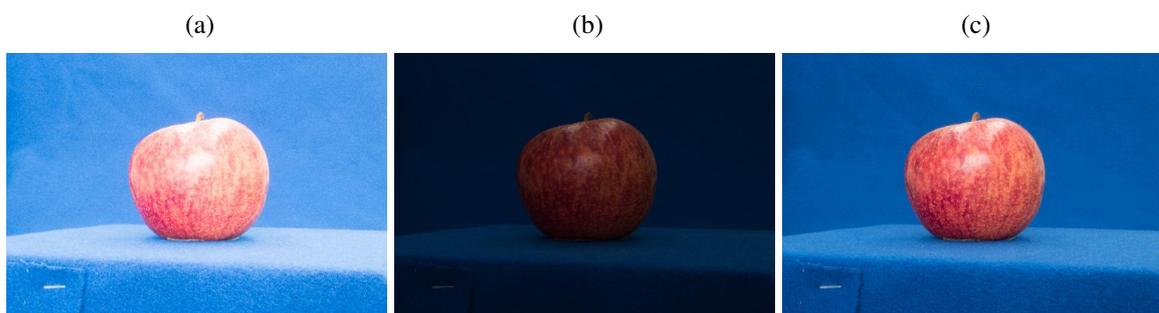
O ponto P, agora no plano esférico (Figura 22.c), é representado por “r”, “ θ ” e “ γ ”, que são respectivamente o raio, ângulo de deslocamento da base giratória (Eixo “B”) e ângulo de deslocamento da câmera (eixos H e V) com o ângulo de inclinação da câmera.

Após o planejamento e levando em consideração os critérios de λ , determina-se a iluminação, a fotometria, configurando-a conforme uma observação nas partes superior e frontal ao objeto. Uma boa fotometria é determinada quando a diferença entre a medição frontal e superior não ultrapassar dois pontos. Por exemplo, se o valor medido na parte frontal for “22.3”, na superior não poderá ser acima de “22.5”. Essa medição determina a luz incidente no objeto, permitindo a configuração da câmera para a aquisição.

Para as aquisições, algumas configurações da câmera são fixas, sendo elas o ISO (*International Standards Organization*), a velocidade e a abertura do diafragma. Assim, a fotometria é usada para verificar se a iluminação superior e na porta não estão com leituras diferentes, superior aos dois pontos citados. Caso esteja, liga-se ou desliga-se alguma ou algumas lâmpadas, até que se tenha uma leitura dentro do parâmetro estipulado.

O fotômetro da câmera também é observado, para verificar se ele fica acima, abaixo ou em zero. Por exemplo, uma regulagem de +1 significa que há um aumento da quantidade de luz em 1 ponto da leitura do fotômetro; -1 significa que há uma diminuição da quantidade de luz em 1 ponto da leitura do fotômetro. Um exemplo para a configuração é apresentado na Figura 23, onde na Figura 23.a é superexposta, na Figura 23.b é subexposta e na Figura 23.c é com o fotômetro da câmera na configuração considerada ideal para o objeto.

Figura 23 – Fotometria do objeto; (a) superexposta; (b) subexposta e (c) exposição considerada ideal.



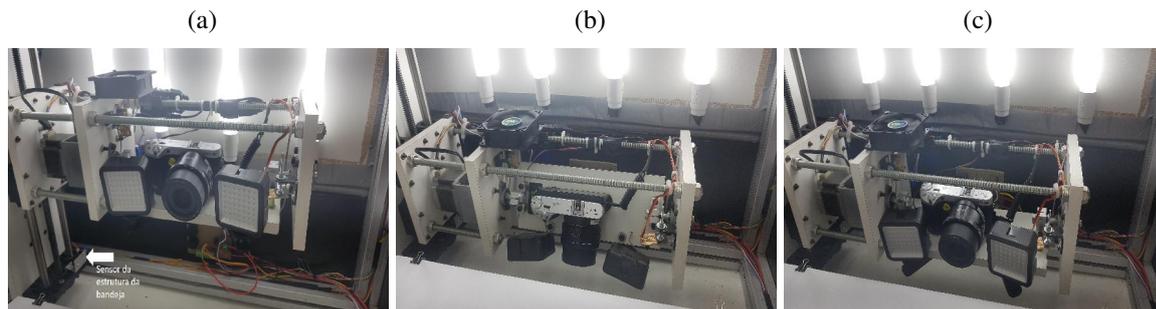
Fonte: autoria própria

Após essas configurações ajustadas, lança-se no sistema de controle do F2S2 os parâmetros de tamanho do objeto, da base de elevação, se houver, da largura do objeto, dos ângulos e raio da aquisição, do tempo de espera entre cada frame e cada stream.

Esse tempo é importante para não haver foto com desfoque ou tremida, pois leva em

consideração a inércia ao terminar um giro ou iniciar um novo *stream*. No início da aquisição há a calibração da estrutura da bandeja e da bandeja, conseqüentemente o posicionamento da câmera. Inicialmente o Eixo “V” é deslocado até um sensor de limite (Figura 15, página 82), que ao ser acionado parará o movimento do eixo. Ao tocar o sensor também fará com que a bandeja seja calibrada, colocando a câmera paralelamente ao eixo “B”. O sensor e a situação de calibração da bandeja (ao ser tocado e retorno à posição), podem ser vistos na Figura 24.

Figura 24 – Calibração da bandeja e câmera: (a) iniciando o posicionamento; (b) calibrando; e (c) retornando para posição de início.



Fonte: autoria própria

Ao final das calibrações, os Eixos “H” e “V” são acionados para posicionarem a câmera na distância estipulada pelo parâmetro “r”. Assim a aquisição iniciará e após iniciada não há como interromper. Se algo for identificado como errado, o processo tem que ser reiniciado, uma condição para que todas as imagens sejam capturadas como estipulado pelas configurações.

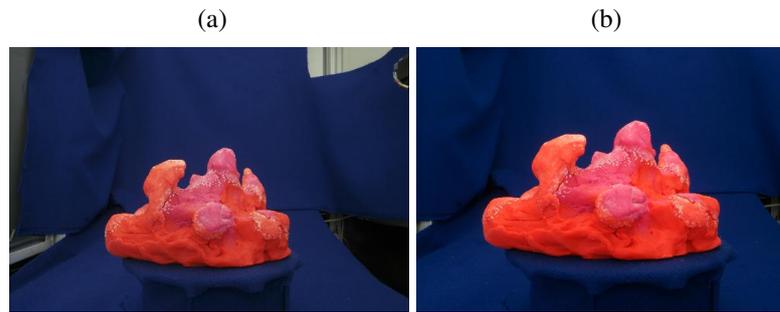
Se por algum motivo o processo for interrompido ou se a câmera desligar, as configurações que estão nela se perdem, por isso deve ser reiniciado. Contudo, existe a opção de “pausar”, para que possa abrir o F2S2 e verificar se não há algo errado, como desligamento da câmera, objeto caído ou deslocado da posição em que foi colocado, lâmpada apagada, entre outros. Se nada for identificado, retoma-se o processo.

Com o término da aquisição, faz-se as correções nas imagens e a redução de contexto, que é extrair o máximo de conteúdo sem importância. Na Figura 25 são apresentadas duas imagens, sendo uma completa (Figura 25.a) e a redução de contexto (Figura 25.b).

Para realizar a redução de contexto, deve-se levar em conta todas as imagens, de forma que o objeto fique enquadrado, sem que perca pedaços. Isso normalmente deixa as imagens nos primeiros ângulos com algum conteúdo sem importância. Contudo, as dos últimos ângulos (p.ex.: 90 graus) ficam com pouco desse tipo de conteúdo. Essa situação pode ser vista na Figura 26.

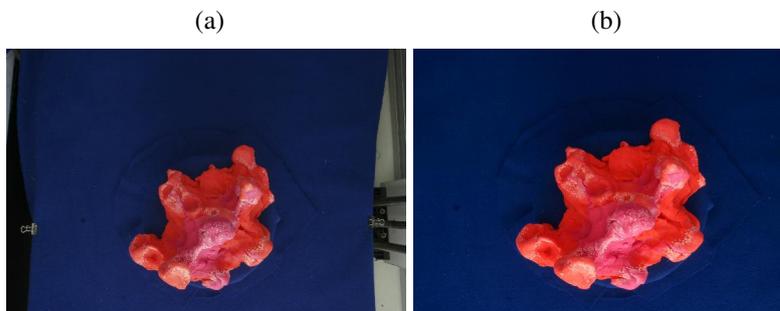
O que deve ser percebido é que todas as imagens do conjunto deverão ter o mesmo tamanho em pixels, ou seja, se a redução for estipulada em 2500x2500 pixels, todas elas seguirão esse padrão, não podendo existir imagens com tamanhos variados, pois isso prejudicaria a visualização.

Figura 25 – Tratamento das imagens: (a) imagem completa e (b) redução de contexto.



Fonte: autoria própria

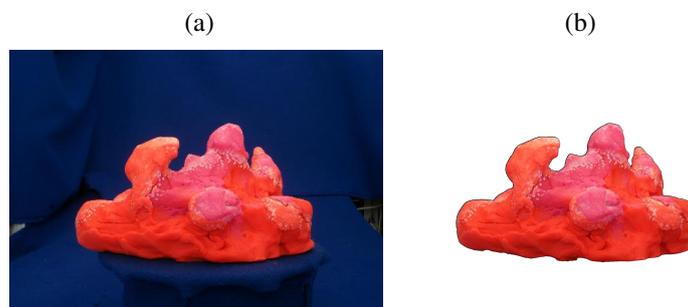
Figura 26 – Tratamento das imagens: (a) ângulo 90 graus completa e (b) ângulo 90 graus com redução de contexto.



Fonte: autoria própria

O passo seguinte ao da redução é a segmentação, processo pelo qual se extraí o plano de fundo, deixando apenas o objeto de interesse. A cor do fundo é substituída pela transparência. Na Figura 27 são apresentadas as imagens com redução e já segmentada.

Figura 27 – Tratamento das imagens: (a) com redução de contexto e (b) segmentada.



Fonte: autoria própria

Durante o processo segmentação, cada imagem gravada recebe um nome condizente com sua posição. Por exemplo, ao iniciar o processo de aquisição, a câmera é posicionada no ângulo 0° - Stream 1(um). O objeto estará no ângulo 0° em relação ao giro no próprio eixo – Frame

1(um).

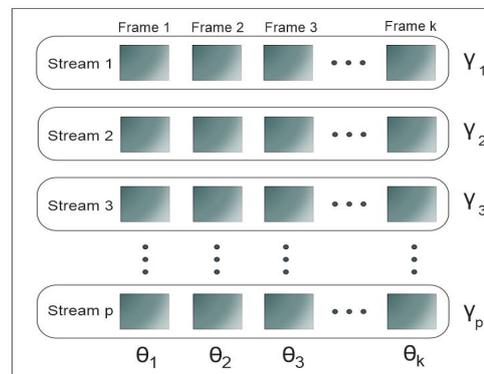
Essa imagem receberá o nome de “S1-F1001”, uma referência a “Stream 1” e “Frame 1”, contudo, por uma questão de organização em pasta, adicionou-se o algarismo “100”. Caso isso não fosse adotado, a organização em pasta seguiria o que o sistema operacional determinasse e poderia surgir algo como: “S1-F1” e logo em seguida “S1-F10”, “S1-F11”... “S1-F19”, “S1-F2”.

Com o sistema adotado, a sequência permanecerá correta, ou seja, será: “S1-F1001”, “S1-F1002”, “S1-F1003”... “S1-F1numFrames”. Onde, “numFrames” representa a quantidade total de *frames* estipulados no início da aquisição. Por exemplo, a configuração para a aquisição poderia ser 30 *streams* e 120 *frames*, assim “numFrames=120”.

Esses valores representam um ângulo de 3° para o deslocamento da câmera na semiesfera, tanto para os *frames* quanto para os *streams*. O total de imagens para essa configuração é de 3720 (três mil setecentos e vinte). Todos esses processos fazem parte do estudo do doutorando Diogo Roberto Olsen, por isso não foram detalhados neste trabalho.

O Stream2D tem uma constituição similar ao de uma matriz, onde as linhas representam os *streams* e as colunas representam os *frames*. Na Figura 28 é mostrado um esquema de como ele é montado.

Figura 28 – Composição do Stream2D.



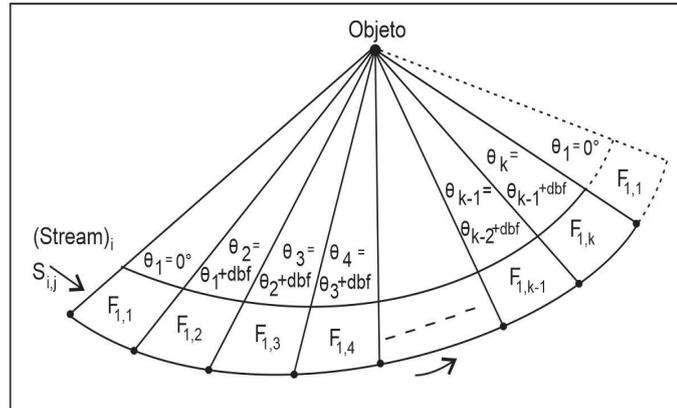
Fonte: adaptada de Diogo Roberto Olsen

Sendo γ o ângulo entre *streams* e θ o ângulo entre os *frames*, calculados a partir da definição no *software* dos valores de “*Degrees between frames*” e “*Degrees between streams*”. A utilização da variação no ângulo de deslocamento de 3° , entre *streams* e *frames*, deve-se em parte à capacidade do cartão de armazenamento (o laboratório tem um cartão de 128GB), pois não se consegue uma aquisição de 0.9° , 1° ou 2° . A quantidade de imagens seria superior a 1.1terabytes. Como as aquisições são em padrão RAW (será explicado no trabalho do doutorando Diogo Roberto Olsen), isso resulta em imagens com aproximadamente 30MB. Em uma aquisição de 1° , tem-se aproximadamente 36mil imagens, o que dá aproximadamente 1.1TB.

Para as formulações matemáticas foram adotados os valores de “*dbS*” e “*dbF*” como ângulos definidos na aquisição, respectivamente, “*Degrees between streams*” e “*Degrees between*

frames”, portanto eles podem assumir quaisquer ângulos permitidos pelo F2S2, por exemplo, os valores de “dbs” podem variar de 1° a 90° e “dbf” de 1° a 360° . O esquema de 1 *stream* esta descrito na Figura 29.

Figura 29 – Formalização do Stream2D



Fonte: autoria própria

Os valores de θ_1 e γ_1 serão 0(zero) no início de cada aquisição, a partir deles tanto θ quanto γ serão acrescidos dos valores determinados para “dbf” e “dbs”, respectivamente. Esse valor 0(zero) é referente ao primeiro ângulo de aquisição, assim sendo, θ_1 tem ângulo zero, mas contém uma imagem. O γ_1 tem ângulo 0(zero), mas contém os *frames* determinados por θ .

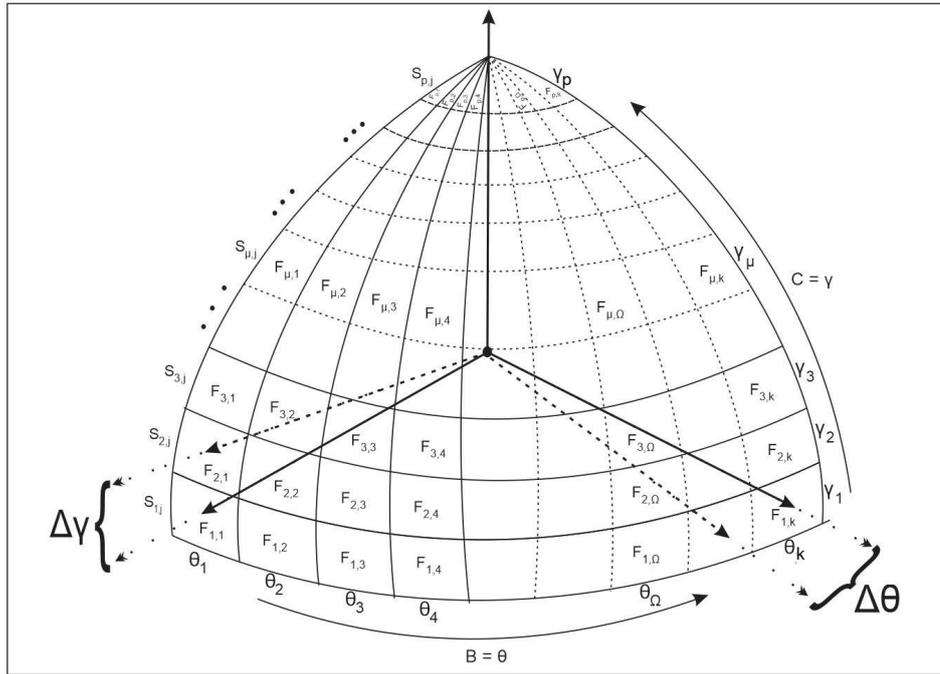
Os valores de “p” e “k”, total de *streams* e total de *frames*, respectivamente, são calculados a partir dessas duas variáveis (“dbs” e “dbf”). Assim, $k = \frac{360}{dbf}$ e $p = \frac{90}{dbs}$. Os valores de “i” e “j” variam de 1 a “p” e de 1 a “k”, respectivamente. Na Matriz 3.1 é apresentada a formação dos *streams*.

$$Stream_{(i,j)} = \begin{bmatrix} S_{1,j} = \{F_{1,1}, F_{1,2}, F_{1,3}, \dots, F_{1,k}\} \\ S_{2,j} = \{F_{2,1}, F_{2,2}, F_{2,3}, \dots, F_{2,k}\} \\ \vdots \\ S_{p,j} = \{F_{p,1}, F_{p,2}, F_{p,3}, \dots, F_{p,k}\} \end{bmatrix} \quad \begin{matrix} i = 1, 2, 3, \dots, p \\ j = 1, 2, 3, \dots, k \end{matrix} \quad (3.1)$$

Na Figura 30 é apresentado o posicionamento dos *frames* que constituem um *stream* e como o *stream* está posicionado no plano esférico para formar o Stream2D. Os símbolos μ e Ω foram escolhidos para representarem algum valor intermediário dos *streams* e *frames*.

Com esse esquema, pode-se determinar como o Stream2D é representado matricialmente. A Matriz 3.2 é a construção do Stream2D a partir do esquema apresentado na Figura 30 e da Matriz 3.1. Com ela pode-se identificar qualquer configuração de aquisição para o F2S2, inclusive se os ângulos de aquisição forem diferentes para os *streams* e *frames*.

Figura 30 – Esquema do Stream2D no plano esférico



Fonte: autoria própria

$$\begin{aligned}
 & \text{Frames} \\
 & \begin{matrix} & 1 & 2 & \dots & k \\ \text{Stream2D}_{(i,j)} = & \begin{bmatrix} F_{1,1} & F_{1,2} & \dots & F_{1,k} \\ F_{2,1} & F_{2,2} & \dots & F_{2,k} \\ F_{3,1} & F_{3,2} & \dots & F_{3,k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{p,1} & F_{p,2} & \dots & F_{p,k} \end{bmatrix} & \begin{matrix} i = 1, 2, 3, \dots, p \\ j = 1, 2, 3, \dots, k \end{matrix} \end{matrix} \quad (3.2)
 \end{aligned}$$

Assim, $i \sim \gamma_i = \{\theta_1, \theta_2, \theta_3, \dots, \theta_{k-1}, \theta_k\}$. Sendo cada uma das linhas um *stream* formado por *frames* (θ), então $\theta = \{F_1, F_2, F_3, \dots, F_k\}$, portanto quando θ assume uma posição dentro do Stream2D, tem-se $S_{(i,j)}$, por conseguinte, uma linha qualquer é dada por $S_{(i,j)} = \{F_{(i,1)}, F_{(i,2)}, F_{(i,3)}, \dots, F_{(p,k)}\}$.

3.3.3 Vantagens nas aquisições do F2S2 para estereoscopia

Como será apresentado no Capítulo 4, página 109, a visualização estereoscópica depende de duas imagens distintas entre si, o que, como já visto, está ligado à visão binocular. Existem outras formas de se construir uma visualização similar, no entanto é dependente de alterações nas imagens.

Um exemplo disso é pegar uma foto única, seja de uma paisagem seja de apenas um frame do objeto adquirido, replicá-la e colocá-la lado a lado, formando o par estéreo. Isso não garante a estereoscopia, mesmo que seja aplicada uma técnica, como o anáglifo. Não há sobreposição e diferença entre os pontos o suficiente para causar uma noção de profundidade e proporção. Para se conseguir isso, há a aplicação de alguma transformação geométrica em uma das imagens, contudo, isso pode não resultar em uma boa estereoscopia.

Em sistemas de aquisição como varredura a laser, luz estruturada ou câmeras paralelas e convergentes, a prática de transformação geométrica é aplicada para os ajustes necessários entre as duas imagens, para a geração de estereoscopia. No caso das aquisições por varredura e luz estruturada, na construção do objeto 3D ainda há a inserção de luz e câmera virtuais, para se ter um dos elementos que define a noção de proporção e profundidade, que é a sombra.

A inserção de luz, para geração de sombra, deverá ser calculada para cada ângulo do objeto, modificando-se a cada mudança de perspectiva, ainda assim, não garante a estereoscopia, apenas a noção de que o objeto é 3D. Voltando à teoria de estereoscopia, são necessárias duas imagens, portanto a geração de duas imagens desse objeto depende da inserção de outras câmeras e focos luminosos, em posições calculadas, para se obter o par estéreo.

A se considerar a visão binocular, essas posições são encontradas a partir da distância entre as pupilas, como já debatido. Então, presume-se que as duas imagens devam ser apresentadas com diferença de 6.5cm entre elas. Contudo, no Capítulo 4 será mostrado que apenas essa diferença não garante a estereoscopia quando se fala de câmeras, mesmo em uma aquisição convergente.

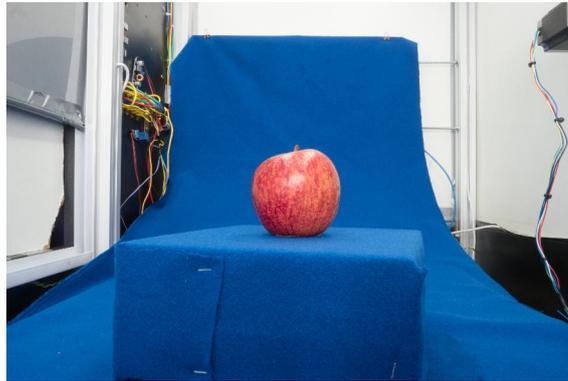
Os elementos que contribuem para uma visualização estereoscópica estão intrínsecos nas aquisições realizadas pelo F2S2, isso é uma vantagem sobre outros tipos de aquisição. Assim, deixa-se de realizar processos computacionais, que são onerosos, passando a ter uma concentração de esforços no carregamento das imagens em memória e acesso a disco. O modelo proposto neste trabalho poderia ser aplicado a outros tipos de aquisição, no entanto o F2S2 fornece todos os recursos para a visualização estereoscópica.

Ele provê a qualidade da imagem, a iluminação e sombra, capazes de proporcionar a visualização das características estereoscópicas, além dos pares de imagens para a característica tridimensional do objeto. Por conseguinte, consegue-se a interatividade por meio do Stream2D e sua manipulação, independentemente do ângulo do objeto que se quer observar. Contudo, para se chegar a essa simplicidade, houve o estudo de formação do par estéreo que melhor provê estereoscopia, sem causar desconforto visual.

Esses elementos são: perspectiva, iluminação e sombra, oclusão e gradiente de textura. A perspectiva proporciona a ilusão de espessura e profundidade de um objeto em relação a outro(s) (Figura 31).

Na Figura 31 é possível perceber a perspectiva entre a maçã e a base de elevação, entre

Figura 31 – Perspectiva na aquisição do F2S2



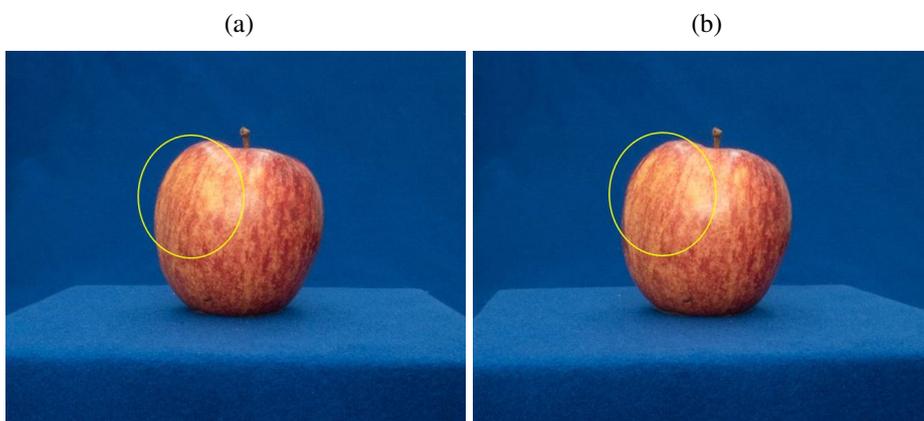
Fonte: autoria própria

outros. O gradiente de textura também está presente, ao se analisar a estrutura da base juntamente com as ondulações do pano do fundo infinito, identifica-se uma diferença que segue do início (parte inferior da foto) até o final (parte superior da foto). Essa textura dá a noção de que a parte inferior está mais próxima e a parte superior está mais distante do ponto de vista de quem está observando-a.

Na mesma foto, encontra-se a iluminação, a qual causa sombra da maçã na base e uma diferença entre a lateral visível da base e seu topo. Também, aparecem alguns pontos mais claros na maçã, isso proporciona a percepção de que ela é arredondada, ou seja, que tem profundidade e proporção.

Isso não seria possível se não houvesse essa iluminação ou se fosse utilizado um *flash*. Na Figura 32 os círculos em amarelo, na Figura 32.a e Figura 32.b, destacam uma saliência existente na maçã que só é percebida devido à iluminação incidente.

Figura 32 – Iluminação insidente apresenta diferença no objeto: (a) saliência pouco visível e (b) saliência mais aparente



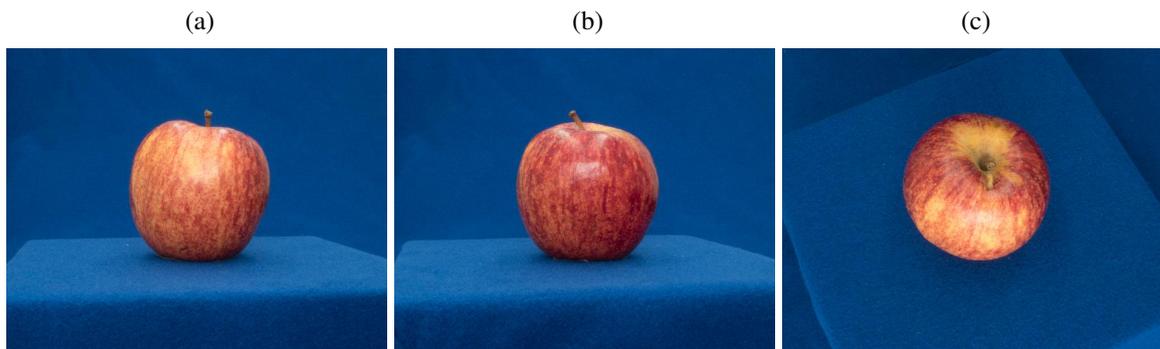
Fonte: autoria própria

As duas imagens são distintas, podendo ser percebido pela posição do talo, da saliência

destacada, mas principalmente pela posição da base.

Já a oclusão que é apresentada tem relação com duas vantagens alcançadas pelo F2S2, a primeira tem relação com a estereoscopia, pois a oclusão também é uma das condições para se perceber proporção e profundidade. No caso da maçã, isso é percebido pelos momentos em que o talo está oculto e momentos em que ele aparece mais. Um exemplo disso pode ser visto na Figura 33.

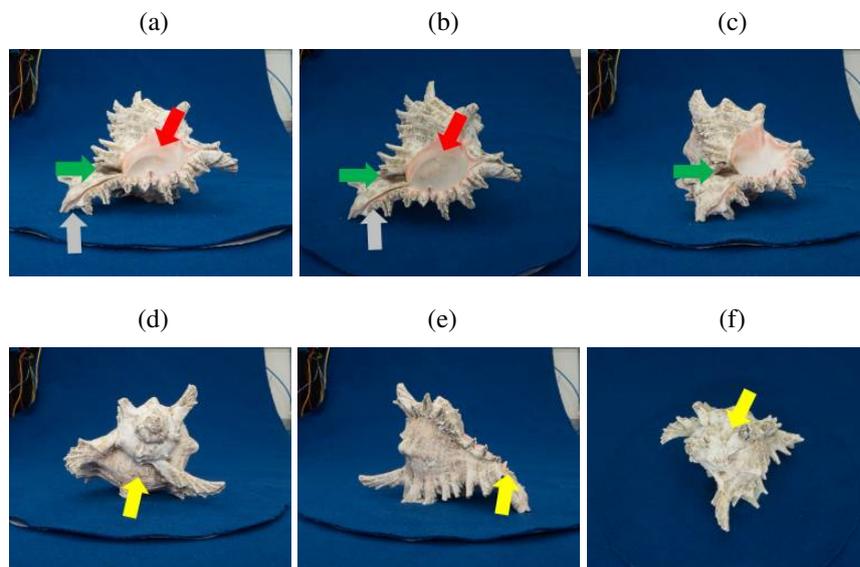
Figura 33 – Oclusão do talo da maçã



Fonte: autoria própria

O outro ponto a ser considerado, é que pela forma de aquisição realizada pelo F2S2, com exceção da parte de contato com a base, não há outra parte do objeto que sofra oclusão, ou seja, todas as partes do objeto tornam-se visíveis em algum momento. Na Figura 34 é apresentado um objeto que teria uma dificuldade adicional ao ser adquirida por outros métodos, mas que pelo F2S2 isso foi reduzido.

Figura 34 – Objeto sem oclusões de suas partes



Fonte: autoria própria

As setas indicam apenas alguns pontos que poderiam sofrer oclusões, mas que não sofrem. Inclusive, ao se visualizar com estereoscopia, percebe-se as reentrâncias, indicadas pelas setas verde e cinza (Figuras 34.a, 34.b e 34.c), a cavidade, indicada pela seta vermelha (Figuras 34.a e 34.b). As setas amarelas (Figuras 34.d, 34.e e 34.f) mostram partes que ficariam ocultas, nos lados, parte de trás e parte de cima da concha.

Em se tratando de um objeto todo irregular, a aquisição por varredura ou luz estruturada seria bastante custosa, não somente computacional, mas para obter todos os detalhes, assim como seria complexo adquirir essas reentrâncias e a cavidade da concha. Já pela aquisição de câmeras paralelas e convergentes, seria possível, no entanto ter-se-ia que analisar cada posição do objeto para se ter um conjunto de imagens capaz de produzir estereoscopia sem a necessidade de realizar as transformações geométricas de ajustes.

3.3.4 Limitações do F2S2

O F2S2 ainda apresenta alguns problemas, que podem ser solucionados com uma nova configuração, mas que nessa versão ainda permanecerá. Dentre elas está a iluminação, pois ela está acima do objeto, o que pode causar sombras, mesmo configurando algumas para estarem apagadas ou acesas. Essas sombras não estão relacionadas às características da estereoscopia, mas àquelas que podem ocultar detalhes do objeto, o que não é desejável.

Devido a este problema, um conjunto de quatro lâmpadas foi instalado na porta, contudo, há momento em que a bandeja e câmera causam sombra no objeto, pois ficam entre ele e o foco luminoso da porta. Uma configuração em que isso não ocorra, pode melhorar a questão de reflexos, sombras e mudanças entre *frames* e *streams*.

Os motores perdem precisão por não terem redução, ou seja, como os braços (H e V), a bandeja (C) estão ligados diretamente no eixo do motor, pode ocorrer imprecisão nos passos e ângulos de deslocamento. Isso pode acarretar uma perda acumulativa, que no final do curso não há garantias de que a configuração inicial tenha sido mantida. Por exemplo, quando se coloca o deslocamento de 3° em 3° entre os streams, pode acontecer de acumular um erro e nos últimos streams o grau seja diferente.

Em algumas aquisições realizadas, no momento de se criar a estereoscopia entre imagens de *streams* diferentes, isso é percebido, pois as imagens não geram uma boa visualização.

A base giratória (B), por também estar diretamente sobre o eixo do motor e este não ter um sistema de redução, acaba restringindo o tamanho e peso das peças. Tentou-se utilizar BE com 8 cm de altura e pouco mais de 20 cm de largura, não ultrapassando 2 kg, no entanto, juntamente com o peso do objeto, a aquisição apresentou problema. Esse problema foi devido à inércia no momento de parar ou iniciar o giro. Isso faz com que os frames, principalmente os iniciais, fiquem tremidos, sem qualidade na foto.

A altura do objeto também restringe a aquisição, pois, como já apresentado, se ele for

acima de 30cm, não há como adquiri-lo, uma vez que a bandeja se chocaria com o sistema de iluminação. Um objeto com altura superior a esse tamanho, faz com que a parábola de deslocamento ultrapasse os 65cm de altura, que é a máxima do F2S2.

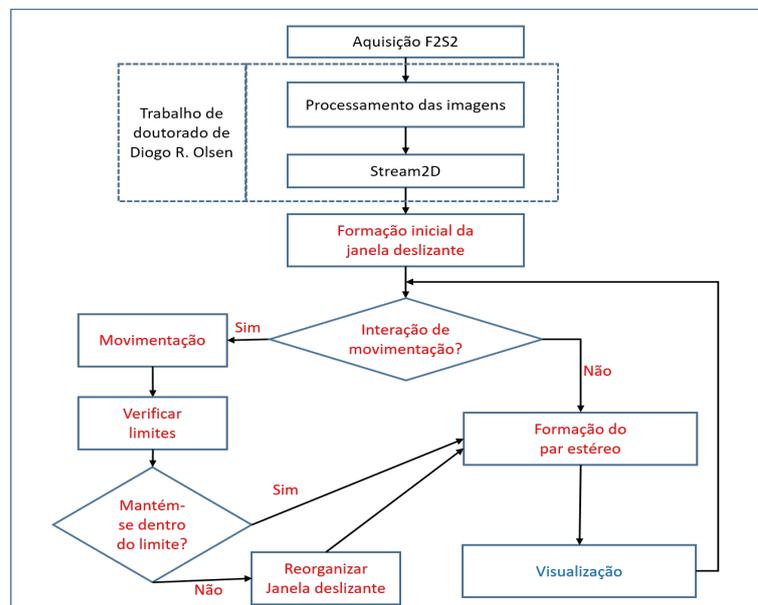
Os cartões de armazenamento, disponíveis no laboratório, para as aquisições, não permitem que se faça uma aquisição completa. Como o ideal é configurar a câmera para obter imagens no sistema RAW, que permite uma manipulação dos parâmetros após a aquisição, por exemplo, alterar o balanço de branco nas imagens, essas ficam com tamanhos de 28 a 30 megabytes. Para uma aquisição completa, que seria de 1° para os *streams* e *frames*, o que daria aproximadamente 32 mil imagens, seria necessário um cartão de 1 Terabyte ou superior.

A câmera utilizada tem apenas uma lente, o que não permite uma configuração para aquisição com boa profundidade de campo, importante para objetos que sejam largos, altos ou mesmo para macros (por ex.: aquisição de insetos). Essa configuração faz com que os objetos tenham foco em todo ele, o que não é possível com lentes sem profundidade de campo.

3.4 Modelo Proposto

Na Figura 35 apresenta-se o fluxograma desse modelo.

Figura 35 – Fluxograma do modelo proposto



Fonte: autoria própria

As etapas do modelo proposto iniciam na aquisição do objeto, passando pelo processamento das imagens e geração do Stream2D. A partir da montagem deste, um algoritmo faz o carregamento e descarregamento dos *frames* e *streams* em memória.

Os algoritmos para se ter a visualização passam pela montagem inicial de uma janela deslizante, verificação de movimentação, dependente da interação com o usuário, a verificação de limites da janela, reorganização da janela e formação do par estéreo, que representa o melhor par para a estereoscopia.

Após a visualização, retoma-se a execução, contudo cai em um laço de repetição, pois não havendo interação do usuário os mesmos *frames* não serão avaliados, o que faz com que sejam mantidos no ambiente de visualização.

3.5 Delineamento da pesquisa

A geração de projeções interativas estereoscópicas a partir de um conjunto de imagens, o Stream2D, é o foco deste trabalho. A interatividade vem da possibilidade de manipulação do objeto, para a visualização das suas perspectivas, cada uma apresentando a noção de proporção e profundidade, que é a estereoscopia.

3.5.1 Descrição das etapas

A primeira etapa é a verificação de problemas com aquisições de imagens com posicionamento da câmera em paralelo e convergente em relação ao objeto. Visando mostrar que nesse tipo de aquisição há a necessidade de alterar as imagens, ou seja, aplicar alguma correção da geometria epipolar para gerar estereoscopia. Ainda nessa etapa é produzida uma estereoscopia com uma imagem do F2S2, duplicando-a para verificar o resultado.

A segunda etapa é a verificação da distância entre a câmera e o objeto, pois não há garantias de que a pouca distância permita estereoscopia, baseando-se na visão binocular e na distância mínima, 25cm dos olhos, para um ser humano perceber proporção e profundidade nos objetos. Essas etapas darão base para a configuração das aquisições do F2S2 especificamente para estereoscopia.

Assim, antes de iniciar a aquisição de cada objeto, há a necessidade de posicioná-lo, centralizando-o em relação à câmera e giro da base. Essa, terceira, é a principal etapa para garantir que o objeto seja adquirido adequadamente, principalmente para a estereoscopia. A primeira, segunda e terceira etapas atendem ao objetivo específico “analisar a formação do par estéreo a partir do conjunto de imagens gerado pelo F2S2”.

Para a aquisição, quarta etapa, são definidos os parâmetros de altura e largura do objeto, se será ou não utilizada uma BE, o raio, quantidade de *streams* e *frames*, além dos tempos de intervalo entre os *frames* e os *streams*. Idem às etapas anteriores, quanto ao objetivo específico.

O processamento das imagens engloba processos definidos e planejados no trabalho do doutorando Diogo Roberto Olsen. Processos como correção de lentes, redução de contexto, segmentação, entre outros. Assim como, a montagem do Stream2D. Contudo, pelo trabalho do

doutorando citado estar caminhando em paralelo a este, realiza-se alguns dos processos desta etapa, para os experimentos planejados neste trabalho.

Após a montagem do Stream2D, copia-se ele para uma pasta específica para o objeto adquirido, dando-se um nome condizente com o objeto e dados para diferenciá-lo de qualquer outra aquisição do mesmo objeto. Esse passo é anterior a execução da visualização. Assim, dentro dele é possível configurar a leitura da pasta e apresentar a estereoscopia de acordo com os ângulos de aquisição. Idem às etapas anteriores, quanto ao objetivo específico.

Na quinta etapa, inicialmente é carregada uma janela deslizante, a qual contém um conjunto de *streams* e *frames*. Esse conjunto é dependente do arquivo de configuração armazenado dentro da mesma pasta que o objeto, ou seja, esse arquivo determina a quantidade de *streams* e *frames* da janela deslizante.

O tamanho da janela é fixo para a visualização de um determinado objeto. A sexta etapa é a movimentação, que significa interação do usuário com a visualização, que pode acarretar o descarregamento e carregamento de novos *frames* e *streams*. Essa etapa atende aos objetivos específicos “analisar o processo computacional de seleção das imagens que formam o par estéreo, considerando as quatro dimensões da interatividade”, “selecionar dentre as imagens do conjunto de perspectivas as quais melhor formam o par estéreo, considerando as quatro dimensões da interatividade” e “analisar o processo computacional de interação considerando as três transformações geométricas que caracterizam um objeto tridimensional - translação, escala e rotação”.

A sétima etapa é a verificação dos limites da janela. Isso ocorre quando o usuário executa uma movimentação e esta pode incorrer em erro na estereoscopia devido a falta de uma das imagens do par estéreo dentro da janela. Idem à etapa anterior, quanto aos objetivos específicos.

A verificação retorna um valor Verdadeiro ou Falso. O Falso, corresponde ao fato de que o movimento não formará o par estéreo, desencadeará a oitava etapa, que é a reorganização da janela deslizante. Essa reorganização faz o descarregamento e o carregamento de alguns *frames* e/ou *streams*. Idem às duas etapas anteriores, quanto aos objetivos específicos.

A nona e última etapa tem papel central neste trabalho, pois é a formação do par estéreo. Ela entrega ao visualizador as duas imagens, para que dentro dele se gere a visualização por qualquer uma das técnicas, anáglifo ou polarização da luz. Esta etapa atente ao último objetivo específico “desenvolver um processo computacional, que tenha baixo consumo de recursos como memória e acesso a disco, de geração de uma visualização interativa a partir de todas as perspectivas adquiridas de um objeto”

O Visualizador não faz parte deste trabalho, no entanto, ele foi desenvolvido como prova de conceito, por consequência foi mantido dentro do fluxograma. Idem à etapa anterior, quanto ao objetivo específico.

3.5.2 Definição dos testes

Os testes foram planejados a partir das etapas, portanto cada um será executado obedecendo a sequência delas.

Assim, o primeiro teste é realizar a captura de imagens em paralelo e convergente e gerar a estereoscopia sem alterar as imagens. Em seguida será aplicado um algoritmo de ajuste da geometria, para verificar se ocorre ou não alteração em uma das imagens do par estéreo.

Uma questão que ainda não se identificou e que pode interferir na estereoscopia é a distância entre o objeto e câmera, por isso o segundo teste é a verificação dessa distância e a estereoscopia formada por ela. Para todos os testes envolvendo a visualização até aqui, será utilizada a técnica de estereoscopia anáglifo e a polarização da luz (projektor DLP), sendo produzidas pelo *software* Gimp e também o StrereoPhoto Maker⁷. Este permite criar visualizações de duas imagens para as técnicas citadas, entre outras.

O terceiro teste é o de centralização do objeto. Uma aquisição em que o objeto não gire sobre seu próprio eixo de maneira adequada, pode prejudicar a visualização. Então, programamos a aquisição para apenas um conjunto de imagens, podendo ser um stream com 60 (6°) ou 120(3°) frames. Essa quantidade é suficiente para determinar se haverá problema.

As etapas que se seguem não fazem parte deste trabalho, mas foram executadas para obter alguns conjuntos de imagens, no entanto não foram realizados com os critérios necessários para obter um Stream2D com a qualidade almejada pelo trabalho do doutorando Diogo Roberto Olsen.

O teste a seguir diz respeito ao carregamento das imagens em memória. Ele terá início com a utilização de um algoritmo de força bruta, onde as imagens são carregadas todas em memória. Na sequência é utilizado um algoritmo de carregamento par-a-par, ou seja, encontra-se o melhor par estéreo e carrega-se as imagens para gerar a visualização. Por fim, o algoritmo de lista circular dinâmica, com o objetivo de gerenciar uma janela deslizante, é executado.

Para testar o funcionamento dos algoritmos, seus códigos foram desenvolvidos utilizando a biblioteca e ambiente citados na seção 3.2. Esses testes são de interação com o usuário, contudo a verificação da movimentação e reorganização da janela serão verificados via console em modo texto, inicialmente sem a apresentação de imagens em tela. Isso visa observar falhas na janela ou na chamada aos frames e streams.

Caso as chamadas aos streams e frames estejam corretas, a próxima verificação é se há reorganização da janela deslizante, permitindo que a movimentação seja sem interrupções ou falhas. Contudo, isso também será verificado via modo texto em console, apresentando os frames e/ou streams carregados.

Por fim, com os algoritmos de movimentação funcionando, o último teste é a criação

⁷ <<http://stereo.jpn.org/eng/stphmkr/index.html>>

da estereoscopia, onde há a determinação do melhor par estéreo, assim, após definição de quais ângulos privilegiam a estereoscopia sem desconforto, define-se como este par será encontrado dentro da janela e como será apresentado no visualizador. O qual, tem o papel apenas de prova de conceito, uma vez que será desenvolvido posteriormente com finalidade de ser colocado no projeto “*The Visible Human Table*”

3.5.3 Descrição dos testes

Nesta seção serão descritos alguns detalhes dos testes a serem realizados.

A captura das imagens paralelas e convergentes levam em consideração a forma como a comunidade acadêmica e de entretenimento fazem para adquirir imagens para a geração da estereoscopia. Contudo, ao contrário do uso de duas câmeras distintas ou mesmo uma câmera com duas lentes separadas pela distância média interpupilar humana, optamos por realizar o teste com a mesma câmera. Ela é posicionada segundo a medição de réguas e lasers, para garantir a exatidão nas distâncias e no foco, na distância média interpupilar. São utilizados três lasers e duas réguas, uma para a distância entre objeto e sensor da câmera, a outra para a medição do posicionamento da câmera para simbolizar as imagens dos olhos esquerdo e direito.

Este teste, como informado, será verificado pelo uso de um software de ajuste da geometria, programa para encontrar a matriz fundamental entre imagens. Além desse, também será utilizado um software de edição de imagens, para a geração de estereoscopia anáglifo ou polarizada. As imagens aos pares serão abertas pelos softwares e alteradas, no caso do anáglifo, para os canais vermelho e ciano, aplicando-se a elas o modo de adição entre pixels. Isso gerará uma imagem que poderá ser observada com os óculos anáglifos (vermelho e ciano). Para a imagem polarizada, as imagens serão colocadas lado a lado e projetadas com o projetor DLP.

Os mesmos softwares serão utilizados para verificar a questão de distância entre objeto e câmera. Para a realização deste teste, o mesmo objeto será adquirido em diferentes distâncias, mas utilizando o F2S2 e o ângulo que melhor gera a estereoscopia. O propósito dessa aquisição é verificar se a distância mínima para estereoscopia também influencia nas aquisições de imagens. Esse teste pode determinar qual distância inviabiliza a estereoscopia ou não. Portanto, a aquisição será apenas do mesmo objeto, distanciando a câmera a cada 10cm, iniciando em 25cm e indo até o máximo permitido no F2S2.

Para a criação do Stream2D, apenas as aquisições com 3° entre os frames e streams serão considerados, pois é um processo em lote, ou seja, nessa configuração de aquisição são 3720 imagens, divididas em 31 streams, pois inicia em 0° e termina em 30°, contendo 120 frames cada. Inicialmente é realizada redução de contexto, após é realizada a segmentação com determinação do nome de cada frame.

O algoritmo de força bruta terá como parâmetros os dados do arquivo de configuração do Stream2D, onde estão os dados da aquisição. A partir da leitura desse arquivo, faz-se o

carregamento de todo o conjunto de imagens. Somente após o carregamento é que se permite a movimentação. No algoritmo par-a-par, faz-se a busca pelo par estéreo e as duas imagens são carregadas em memória, em caso de movimentação, encontra-se o novo par, descarrega-se o antigo e carrega-se o novo.

Já o algoritmo de lista circular dinâmica, juntamente com a janela deslizante foram idealizadas para diminuir o custo de carregar todas as imagens em memória ou carregar e descarregar a todo instante que se faz uma movimentação. A ideia é minimizar os custos, uma vez que não terá acesso constante a disco e ao carregamento de imagens em memória. Seu processo não é complexo, pois há uma definição de uma estrutura contendo os nós para seus vizinhos, total de oito, além do número do stream e do frame, separadamente.

Como a aquisição é feita apenas com o objeto em uma posição, que foi chamada de “em pé”, optamos por colocar um valor booleano chamado “oposto”. Quando a visualização chega em uma posição em que a próxima seria a aquisição do objeto ao contrário, que se chamou de “ponta cabeça”, faz-se uma inversão via software simulando essa aquisição. Não há garantia de manter esse parâmetro, pois foi apenas para estudo. O carregamento inicial, que também é dependente do tamanho do Stream2D, coloca em memória uma quantidade de streams e frames, o suficiente para algumas movimentações não causarem sua reorganização, assim também diminui-se o custo computacional.

Entretanto, a movimentação desencadeia uma verificação de limite da janela. Essa verificação parte do princípio de que o par estéreo não pode ser desfeito com uma movimentação futura, ou seja, se um passo além do qual o usuário requisitou resultar em uma proximidade do limite da janela, ela terá que ser reconfigurada naquela direção. Pode ocorrer de voltar o passo e o carregamento ter sido em vão, contudo a volta não provocará os processos de descarregar e carregar novos streams e/ou frames, uma vez que a avaliação de limite na nova direção não atingirá, mesmo que em um passo futuro, a proximidade do limite.

Essa análise é diferente para cada tamanho de janela. Pode ocorrer de se ter uma quantidade pequena de streams e frames, por exemplo, havendo apenas 3 streams e 120 frames cada, o total de imagens será 360, portanto, não há necessidade de se criar uma janela. Com essa quantidade do Stream2D é possível carregá-las todas em memória. A janela se tornaria única e totalmente conectada. Já com um tamanho 10 vezes maior, a janela pode ser algo em torno de, por exemplo, 11 streams por 21 frames, o que dá 231 nós em memória. A quantidade ímpar é pelo motivo de manter a imagem base fora da formação do par, sendo um parâmetro utilizado apenas para encontrar os frames que o formarão.

A estereoscopia pode ser formada pela análise dos ângulos de aquisição, por isso a imagem base, em casos onde a aquisição é com ângulos menores do que 4, 5 e 6 graus, não faz parte do par. Por meio dela é possível calcular a distância angular dos frames, esquerdo e direito, além de permitir uma verificação prévia dos limites da janela deslizante. Uma requisição de movimentação será realizada pelo usuário, independentemente para qual direção foi o movimento,

a imagem base será o frame adjacente ao atual, portanto, antes do visualizador receber o par estéreo, o programa verifica os limites, não se atingindo o limite com o passo futuro, encontra-se o par e o retorna para formar a estereoscopia.

Em aquisições com 4, 5 e 6 graus, a imagem base é parte do par, mas a análise de limite da janela é mais crítica, pois a imagem base pode chegar mais próxima do limite da janela e ainda formar o par estéreo. Uma vez reorganizada, pode ocorrer de um próximo passo necessitar de outra reorganização, contudo a quantidade de frames e/ou streams a serem carregados é menor do que as aquisições com ângulos menores do que estes apresentados neste parágrafo.

3.5.4 Processo de realização dos testes

O primeiro teste será realizado escolhendo uma peça para sua aquisição. Essa peça terá um contraste com o fundo escolhido, para que seja possível sua extração sem comprometer o objeto. Com um laser posicionado na lateral do F2S2 e passando pelo centro da base (Eixo B), esse objeto será posicionado, porém sem o critério do alinhamento pelo laser, uma vez que não sofrerá rotação. No entanto, ele será colocado exatamente sobre o eixo do motor, que é a parte central do eixo B.

Com o objeto e o laser posicionados, outro laser será colocado transversalmente ao primeiro, assim a primeira régua poderá ser corretamente posicionada, ela determinará a distância entre a câmera e o objeto. O terceiro laser será posto transversalmente ao segundo, passando pela distância determinada para a aquisição paralela. Sobre esse laser será colocada a segunda régua. Ela marcará a distância de 6.5cm entre as aquisições.

Para garantir a exatidão nessa distância, o primeiro laser será retirado da posição que estava e será utilizado juntamente com o segundo para marcar a distância entre as capturas da esquerda e direita, imagens que representam a visão dos olhos esquerdo e direito, respectivamente. As distâncias variarão de 25cm até o máximo do F2S2 (55cm).

A diferença entre a aquisição paralela e a convergente é o posicionamento dos lasers que marcam o posicionamento da câmera. Nas lentes existe um ponto que determina a parte central da lente, então os lasers esquerdo e direito devem passar por esse ponto vindo da parte traseira da câmera. Com isso, há garantias de que a câmera e, conseqüentemente, sua lente estará direcionada para o objeto. O objeto não será tirado da posição em que foi colocado para a aquisição paralela.

Na aquisição convergente, a câmera ficará em um ângulo determinado pelo ângulo de 90° , que forma a linha reta entre o objeto e a metade da distância do posicionamento das aquisições. Portanto, tem-se a distância de altura (h), a distância entre os dois pontos de aquisição e o ângulo reto, o que permite calcular as outras medidas e os ângulos. Essas medidas e ângulos poderão dar base de comparação com a aquisição do F2S2 e a estereoscopia que ele provê. Com um transferidor de graus e os lasers, posicionar-se-á a câmera nos ângulos encontrados para que a

lente fique direcionada para o encontro dos lasers esquerdo e direito, que se cruzavam na base do objeto.

Para cada distância estipulada para a aquisição, o mesmo procedimento de encontrar o ângulo em que a câmera deveria ser posicionada será realizado, esses valores serão calculados previamente no programa Excel da Microsoft e com o posicionamento verificado por meio dos lasers, réguas e transferidor de graus. O conjunto de imagens será criado seguindo o critério de distância entre o objeto e câmera, tanto para a paralela, como para a convergente e para a aquisição no F2S2.

As imagens receberam um nome contendo a palavra que representa o lado (esquerda ou direita), a distância focal (16mm, 35mm ou outra) e a distância entre posição do sensor da câmera e o objeto (25cm, 35cm ou outro). As aquisições foram fixadas para não haver diferença entre as três, portanto as imagens foram adquiridas com a distância focal em 16mm e outra com 50mm. Contudo, a análise foi utilizando a paralela, convergente e F2S2 em distância focal igual, ou seja, 16mm nas três ou 50mm nas três, para garantir que as imagens tivessem o mínimo de diferença possível.

Com as aquisições realizadas, utilizou-se os softwares de edição de imagens para gerar a estereoscopia, sem alteração de qualquer uma delas. As imagens foram abertas aos pares (imagem esquerda e direita) para a aquisição com distância de 25cm, inicialmente com as aquisições em paralelo. Foram geradas as imagens estereoscópicas anáglifo e projetadas também em par polarizado para observação. Além disso, em todos os pares foi aplicado um algoritmo de cálculo de diferença. Esse cálculo utiliza cada pixel da imagem esquerda e da direita, na mesma posição, para calcular a diferença entre eles e identificar quão diferente uma imagem é da outra. Isso indica que a uma delas deve ser aplicado o ajuste de geometria.

As distâncias em visão humana para estereoscopia são conhecidas, mas em fotografias isso não é garantido, por isso encontrá-las para guiar as aquisições para as visualizações estereoscópicas é um dos experimentos deste trabalho. Por esse motivo é que foram adquiridas as imagens paralelas, convergentes e do F2S2 em diferentes distâncias. Contudo, existe um parâmetro das câmeras que impede a afirmação com absoluta certeza, o que se fez foi uma aproximação. O parâmetro é a distância focal, que é a distância física real entre o ponto de convergência da lente e a superfície do sensor da câmera.

Para a estereoscopia com as imagens adquiridas pelo F2S2, a criação do Stream2D será feita com aquisições de 3° entre os frames e streams, essa precisão permite identificar todos os detalhes do objeto. Contudo, esses testes foram realizados para se ter um conjunto completo de imagens para a criação da visualização interativa estereoscópica, uma vez que os processos compõem a criação do Stream2D fazem parte do trabalho do doutorando Diogo Roberto Olsen. Foram feitos neste trabalho apenas com o intuito de ter o Stream2D em que se pudesse navegar e testar as teorias aqui debatidas.

Os algoritmos de redução de contexto e segmentação, utilizados para os testes, foram desenvolvidos no Matlab, sendo o segmentador baseado em SVM e morfologia matemática (BROBOUSKI, 2020). O processamento dessas imagens é em lote, ou seja, nessa configuração definida para as aquisições serão 3720 imagens, onde abre-se uma e delimita-se o tamanho de uma janela de corte, a qual deve conter todo o objeto e ainda manter a integridade até a última.

Isso pode requerer um teste inicial, pois um objeto pode ocupar um espaço maior na imagem quando adquirido em 90° . Por exemplo, um objeto que tenha altura de 15cm e largura de 35cm, obviamente ocupará uma porção maior da imagem quando visto de cima (ângulo de 90° entre câmera e objeto).

Para a segmentação, abre-se uma ou mais imagens e colhe-se uma quantidade de pixel do fundo e do objeto. Esses pixels são separados em duas classes, para treinamento do algoritmo SVM. Após o treinamento, executa-o para todas as imagens do Stream2D, por isso recebeu o nome de segmentador em lote.

Ao término desses processos, tem-se o Stream2D montado. Isso permite a execução dos próximos testes. Sendo o primeiro deles o algoritmo de força bruta, que recebeu esse nome pelo fato de carregar todo o Stream2D em memória. Nesse teste pretende-se verificar se haverá sobrecarga de memória, mas também se o tempo decorrido para o carregamento de todas as imagens não será um problema para a visualização.

Para o carregamento será criada uma lista dinâmica totalmente conectada com o número de streams e o número de frames. Cada imagem será carregada para dentro de um nó e sua posição na lista indica sua posição no seu stream. Não foi optado pela criação de uma matriz fixa devido a complexidade na movimentação, principalmente quando a imagem base estiver na primeira ou última coluna da matriz. Isso desencadearia um cálculo para se encontrar a próxima imagem base e o par estéreo.

O algoritmo par-a-par manterá em memória apenas o par estéreo e a imagem base fará parte do par. Encontra-se a imagem base e a partir dela a imagem que forma o par, carrega-se em memória, gera-se a estereoscopia, seja anáglifo, seja polarizada. Quando um movimento é requisitado, primeiro encontra-se a próxima imagem base e seu par, descarrega-se o antigo par e carrega-se o novo, retomando o processo de geração estereoscópica.

O movimento pode ser rápido, por exemplo, estar olhando a parte frontal do objeto e querer passar a olhar a parte de trás. O usuário fará quantas requisições de movimento forem necessárias para olhar a perspectiva de interesse, isso fará com que, a cada requisição, todo o processo desde o encontro da imagem base até a formação da estereoscopia, inclusive carregar e descarregar imagens, seja executado. Pode ocorrer um acúmulo de requisições, provocando uma interrupção na visualização.

O terceiro algoritmo a ser testado é a janela deslizante implementada por meio de uma lista circular dinâmica. Ser circular não significa que todos os nós estarão em memória, mas ao se

atingir os streams ou frames das extremidades, os próximos carregados serão os da extremidade contrária. Por exemplo, se a lista atingir o frame da extremidade final de um stream, o próximo frame a ser carregado será o primeiro desse mesmo stream. Isso é aplicado a todos os streams que fazem parte da janela. O mesmo ocorre se o movimento estiver no sentido de mesmo frame, mas em streams diferentes, o que levará ao frame, por exemplo, do último stream, assim o próximo stream, que é o primeiro, será carregado, contudo apenas os frames que fazem parte da janela.

Em resumo, sendo o stream 1 o primeiro ($Stream2D_{(1,j)}$) e 31 o último ($Stream2D_{(31,j)}$), o frame 1001 o primeiro ($Stream2D_{(i,1001)}$) e 1120 o último ($Stream2D_{(i,1120)}$), a janela deverá ser composta por uma submatriz da Matriz $Stream2D_{(i,j)}$, com um subconjunto de streams, $\varsigma \subseteq i$, por exemplo $\varsigma = \{1, 2, 3, \dots, 7\}$, com um subconjunto de frames, $\tau \subseteq j$, por exemplo $\tau = \{1109, 1110, 1111, \dots, 1120\}$. Um movimento na horizontal para a direita, seguindo a ideia do parágrafo anterior, fará com que os frames 1001, que estão fora de τ , sejam carregados e os frames 1109 de τ de ς sejam descarregados, $Stream2D_{(\varsigma,1109)}$. Com isso, a nova composição de τ da submatriz com streams ς , será com os frames $\{1110, 1111, \dots, 1120, 1001\}$.

Em um movimento na vertical para cima, utilizando a mesma submatriz ($Stream2D_{(\varsigma,\tau)}$), com os mesmos subconjuntos ($\varsigma = \{1, 2, 3, \dots, 7\}$ e $\tau = \{1109, 1110, 1111, \dots, 1120\}$), fará com que os frames de mesma posição, pertencentes a τ do stream 31, que estão fora de ς , sejam carregados e os frames τ do stream 7 ($Stream2D_{(\tau,\tau)}$) sejam descarregados. A partir desse movimento a configuração da submatriz terá os subconjuntos: $\varsigma = \{31, 1, 2, \dots, 6\}$ e $\tau = \{1109, 1110, 1111, \dots, 1120\}$.

Esse exemplo apresenta duas possibilidades de reorganização da janela, as outras possibilidades estão ligadas ao movimento em diagonal, o qual faz com que a janela seja reorganizada seguindo a lógica do que foi apresentado no parágrafo anterior. Então, quando se faz o movimento diagonal e se chega a um ponto que necessita a reorganização carregando e descarregando frames e streams, o comportamento do algoritmo é executar os dois processos em sequência, descritos no parágrafo anterior. Isso fará com seja mantida a integridade da janela.

A janela será criada com um número definido pela quantidade de streams e frames do Stream2D. A primeira janela será carregada no início da execução do visualizador e mantida com o mesmo tamanho até o final da execução. Em caso de mudança de objeto, mesmo em tempo de execução, o pedido para selecionar a pasta do objeto que será visualizado acarretará na leitura do novo arquivo de configuração e conseqüentemente a alteração do tamanho da janela, se o Stream2D for de tamanho diferente do anterior. Senão for, a janela manterá o mesmo tamanho, mas com nova execução para o carregamento do novo objeto.

Para que essa integridade seja preservada, um algoritmo entrará em execução, a cada movimento, para verificar os limites da janela. Os processos para manter a integridade são iguais para qualquer tamanho da janela, o que diferencia é a imagem base fazer ou não parte do par estéreo. No caso, ela fará parte apenas em aquisições em que, utilizando ao menos três frames sequenciais, a distância angular entre os dois da extremidade não ultrapasse o valor de conforto

visual na estereoscopia.

A partir da imagem base, verifica-se a próxima requisição de movimentação, seja para qual direção, o algoritmo analisa uma posição futura, ou seja, mais uma requisição na mesma direção. Nessa análise, se a imagem base assumir uma posição de risco, que é não existir uma das imagens do par estéreo, haverá a reorganização da janela deslizante.

Quando o usuário interage com o visualizador, ele requisita um movimento em qualquer direção, isso faz com a nova imagem base passe a ser a vizinha imediata. Em movimentos na horizontal, tanto esquerda quanto direita, o algoritmo analisa uma posição após a segunda vizinha. Então, imaginando o subconjunto $\tau = \{1001, 1002, 1003, \dots, 1010\}$, se a imagem base atual fosse o frame 1003, com uma movimentação para a direita, a nova passaria a ser o frame 1004. A verificação de limite analisaria se o frame 1007 faria parte da janela e se isso fosse falso, haveria reorganização dela.

A mesma análise é aplicada ao movimento para a esquerda. Os movimentos para cima e para a baixo também analisam com esses critérios. Os movimentos nas diagonais são um pouco diferentes, pois devem analisar nos dois sentidos, horizontal e vertical, para que se um deles for falso, ocorra a reorganização. Essa ocorre carregando e descarregando os frames e streams nos dois sentidos também.

Por fim, a formação estereoscópica depende do par estéreo correto, o qual é passado ao algoritmo de geração da estereoscopia. Esse algoritmo recebe também um comando de apresentação, que pode ser gerar uma imagem anáglifo ou posicionar o par lado a lado na tela, para que seja uma projeção polarizada. Esta depende do monitor prover imagens polarizadas e do uso de óculos passivos ou ativos. Para este trabalho, os óculos são ativos, os quais são sincronizados com o projetor.

3.5.5 Análise de Resultados dos testes

A primeira análise será sobre a estereoscopia com imagens adquiridas em paralelo e convergente, pois este trabalho se pauta na teoria de que esse tipo de aquisição, comum em trabalhos e entretenimento, não produz estereoscopia sem que haja algum ajuste da geometria da imagem. Entretanto, a forma como será analisado é por experimento, gerando uma imagem anáglifo e também projetando em 3D pela técnica de polarização. Em seguida, haverá a aplicação de dois algoritmos, sendo o SURF, para determinar a matriz fundamental e ocorrência de ajustes. O segundo será para determinar a diferença entre as imagens, isso indicará quão diferentes são e onde influencia isso na formação estereoscópica.

O posicionamento do objeto na formação da estereoscopia também é analisado, pois um objeto que gire elipticamente ou que se desloque lateralmente durante a aquisição pode causar desconforto ou mesmo não representar fielmente o objeto na visualização. Então, essa análise também é visual, com a projeção de um objeto adquirido sem o rigor de posicionamento

centralizado e outros com esse rigor.

A limitação da visão binocular para observação de proporção e profundidade é analisada por meio das aquisições em diferentes distâncias, para as aquisições convergentes, paralelas e do F2S2. Procura-se determinar se há uma distância mínima de posicionamento da câmera, isso advém dos estudos que apresentam as distâncias de 25cm e 700m, mínima e máxima respectivamente, em que o ângulo de percepção de estereoscopia é de aproximadamente 120° dos 180° que um humano tem de campo visual. Apesar de existirem cálculos bem definidos para a estereoscopia, o F2S2 se diferencia das aquisições onde esses cálculos são aplicados, porém calcular também para as imagens do F2S2 pode corroborar com as teorias levantadas neste trabalho.

Sobre a redução de contexto e segmentação não haverá aprofundamento nas análises, apenas apontamentos do que pode ser melhorado, para também melhorar a visualização estereoscópica.

O programa utilizado para desenvolvimento dos algoritmos tem a apresentação de consumo de memória, tempo de execução, entre outros. Portanto, pode-se extrair essas informações para análise do comportamento dos algoritmos de força bruta, par-a-par e também da lista dinâmica com janela deslizante. Isso proverá uma base de comparação, para se determinar qual terá melhor desempenho. Isso pode ser em forma de gráfico ou análise em tabela. Porém, em outro momento será analisada a questão de movimentação e reorganização da janela, inclusive com seu desempenho.

Uma análise exclusiva do algoritmo de lista e da janela será realizada, pois em um segundo momento, após o carregamento inicial, ocorre a movimentação e possível reorganização da janela. Por isso, será feita a verificação do comportamento em situações de reorganização.

Como todo o processo de carregamento de streams e frames, além da reorganização da janela são realizados por um algoritmo específico para esse fim, encontrar o par estéreo é uma tarefa que não existe o mesmo esforço computacional do que montar ou reorganizar a janela. Assim como, a formação da estereoscopia também não exige esforço computacional. Portanto, serão avaliados quanto a formação de um par que não causa desconforto e se a visualização estereoscópica é percebida pelo usuário.

Neste capítulo foram apresentados a metodologia e infraestrutura para desenvolvimento da pesquisa. No próximo capítulo será abordado o detalhamento do modelo proposto.

Capítulo 4

Detalhamento do modelo proposto

4.1 Introdução

O modelo desenvolvido neste trabalho não tem relação com reconstrução 3D, pois não cria a malha de pontos, por isso adotou-se a designação de projeção interativa estereoscópica. Seu significado tem relação com a possibilidade de visualizar o objeto adquirido em todas as suas perspectivas (ângulos), além de possibilitar a manipulação do mesmo. Adicionado ao modelo tem-se a visão estereoscópica, a qual permite ter a noção de proporção e profundidade. Assim, cria-se uma visualização estereoscópica para todas as perspectivas, o que dá também a noção de estar visualizando o objeto em 3D.

O diferencial deste trabalho está em uma junção da aquisição e da geração da estereoscopia, uma vez que a forma de adquirir, por meio do F2S2, provê um conjunto de imagens que permite realizar a visualização tridimensional e estereoscópica sem causar desconforto ao observador. Além de permitir que este possa manipulá-la e direcionando-se isso para a área educacional, tem-se um OA que auxilia no processo de ensino/aprendizagem.

Para isso, neste capítulo são abordados a aquisição paralela e convergente na formação da estereoscopia, a centralização do objeto e preparação das imagens, o Stream2D na formação da estereoscopia, as janelas deslizantes no Stream2D e a projeção interativa estereoscópica .

4.2 Aquisição paralela e convergente para formação da estereoscopia

Um diferencial encontrado neste trabalho é a formação do par estéreo para um conjunto de imagens, não sendo apenas de um par de imagens. Isso leva a um planejamento de como montar uma visualização sem desconforto, o que é dependente da aquisição e definição dos melhores ângulos.

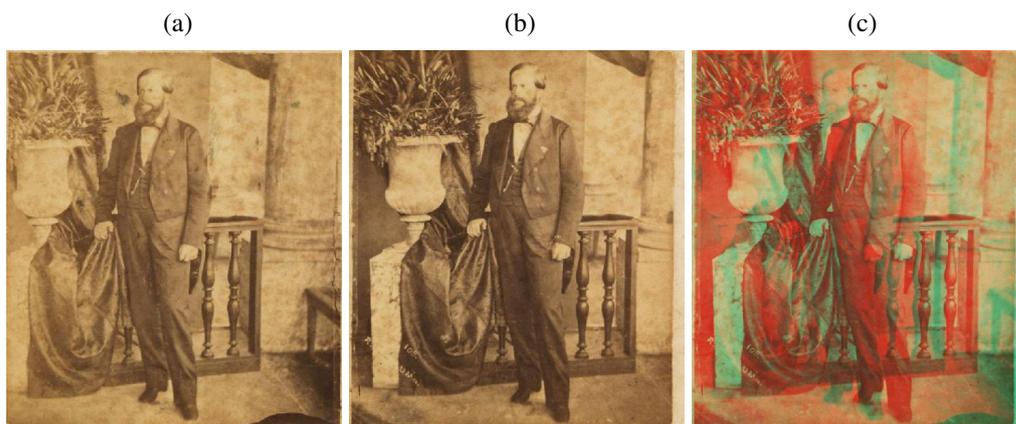
Para tanto, buscamos determinar a diferença entre a formação da estereoscopia utilizando imagens do Stream2D e adquiridas com duas técnicas de aquisição conhecidas, que são a paralela e convergente. Quando se faz uma aquisição paralela, utilizando uma ou duas câmeras para capturar duas imagens do mesmo objeto, tomando como parâmetro a distância interpupilar, para se aproximar da visão humana, existem algumas transformações geométricas, ajustes da geometria epipolar, que devem ser consideradas.

Quando se gera uma imagem estereoscópica utilizando duas imagens paralelas o resultado é uma visualização com paralaxe positiva, ou seja, a noção de profundidade adentra o plano de visualização. Um exemplo disso são as imagens desenvolvidas para visualização no Estereoscópio como o desenvolvido por Brewster (Figura 10, página 62).

Duas imagens paralelas podem formar a estereoscopia, contudo a sensação é diferente do que se tem nos cinemas ou televisões, onde a projeção “salta” do plano e se aproxima do espectador. No caso dessas imagens, a profundidade inicia no plano e estende-se para o sentido contrário ao do espectador.

Imagens dessa natureza foram distribuídas juntamente com os estereoscópios. A invenção chegou inclusive no Brasil, onde Dom Pedro II foi fotografado por Revert Henrique Klumb. A sensação de profundidade e proporção podem ser percebidas com a criação do anáglifo ou na projeção polarizada. Isso foi testado, nas duas visualizações, contudo aqui é mostrada apenas em anáglifo com as imagens de Dom Pedro II (Figura 36).

Figura 36 – Imagens de Dom Pedro II: (a) imagem esquerda; (b) imagem direita; e (c) Estereoscopia anáglifo



Fonte: Revert Henrique Klumb - Brasileira Fotográfica Digital

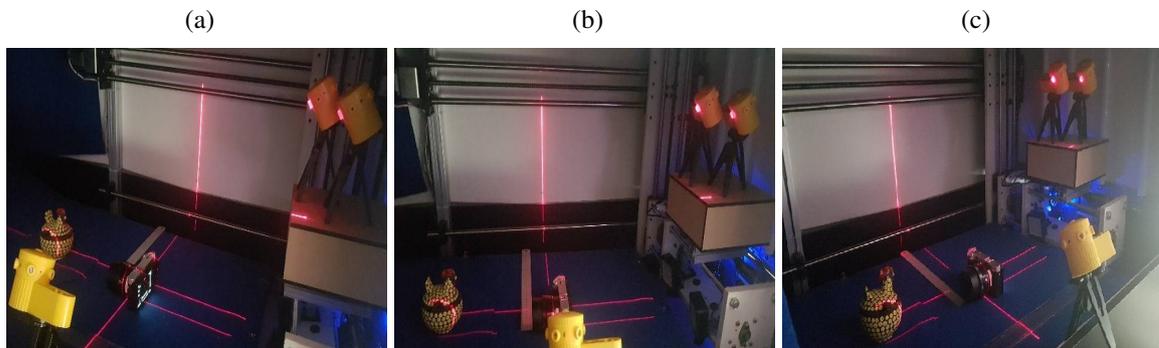
As imagens (Figura 36.a e 36.b) foram obtidas a partir da original encontrada no acervo da Brasileira Fotográfica Digital¹. A partir da original, fizemos a separação em imagem direita e esquerda para a criação da imagem estereoscópica anáglifo (Figura 36.c).

Para a comparação com as aquisições do F2S2, utilizamos a mesma câmera, o mesmo plano de fundo, as distâncias do objeto, as configurações da câmera, como abertura, ISO e tempo de exposição, parâmetros utilizados nas aquisições para geração do Stream2D. Na Figura 37 é apresentada a preparação para a aquisição.

Dois lasers foram colocados a 6.5cm de distância, para garantir a exatidão no posicionamento da câmera. Outro laser foi posicionado transversalmente aos dois anteriores, passando pelo ponto que marca a posição do sensor da câmera (Figura 38).

¹ <<http://brasilianafotografica.bn.br/?tag=historia&paged=3>>

Figura 37 – Preparação para aquisição em paralelo: (a) posições dos lasers, câmera e objeto; (b) régua de distância entre a esquerda e direita; e (c) régua de distância entre objeto e câmera



Fonte: autoria própria

Figura 38 – Câmera com os pontos: seta vermelha é o centro da lente e seta amarela é a posição do sensor



Fonte: autoria própria

Na Figura 37 está a aquisição paralela com distância de 25cm entre o laser transversal e o início do objeto, a medição foi realizada com laser transversal tocando o objeto e na régua na marca 0cm. Posteriormente o mesmo laser foi deslocado para tocar a régua na marca de 25cm (régua transparente logo abaixo do laser, visto na Figura 37.c).

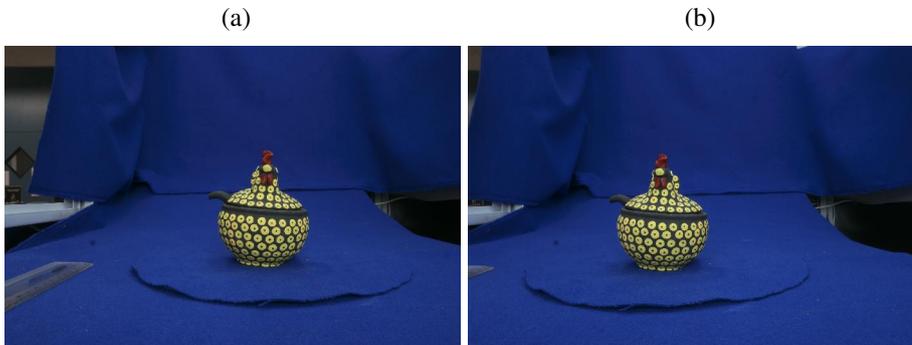
Outra régua, uma de metal, também vista na Figura 37, foi colocada para o posicionamento dos dois lasers na distância média interpupilar. Essa régua foi posicionada inicialmente ao lado do objeto, onde o laser transversal a tocava em toda sua extensão, para garantir que estava alinhada.

Em seguida foram posicionados os lasers de alinhamento da câmera. Essa mesma régua foi deslocada por toda a distância calculada na régua transparente, até o posicionamento final da câmera, para garantir que a distância entre os dois lasers era exatamente 6.5cm em toda essa distância.

Com essas especificações e buscando garantir uma aquisição com todo o rigor necessário

para as fotos terem a distância de 25cm e 6.5cm entre cada foto, obteve-se as imagens como mostrado na Figura 39.

Figura 39 – Aquisição paralela: (a) imagem Esquerda e (b) imagem direita

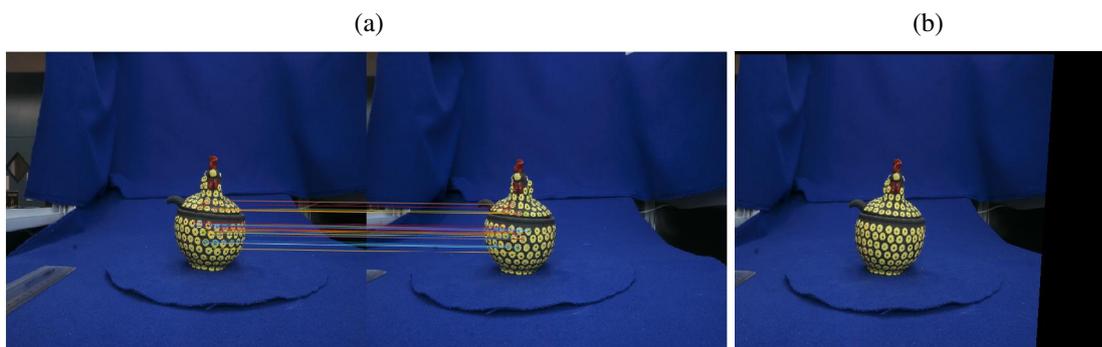


Fonte: autoria própria

As fotos foram adquiridas com distância focal em 16mm e abertura do diafragma em $f/22$, ISO 100 e tempo do obturador em 1 segundo. Essas são as configurações utilizadas nas aquisições do Stream2D para as visualizações testadas neste trabalho.

Para determinar o quanto esse par estéreo forma a estereoscopia, aplicamos o algoritmo de matching denominado SURF, encontrado na literatura e disponível como biblioteca do Matlab. O objetivo foi encontrar uma matriz fundamental entre as imagens e a partir dela realizar os ajustes da geometria epipolar, caso necessário. O resultado encontrado após aplicação do algoritmo é apresentado na Figura 40.

Figura 40 – Algoritmo de Matching SURF: (a) matriz fundamental e (b) imagem da direita após ajustes da geometria - Aquisição Paralela



Fonte: autoria própria

Sobre as mesmas imagens também foi aplicado um método de identificação de diferença entre elas. O cálculo consiste em percorrer as imagens pixel a pixel e realizar a diferença entre eles. Quando um pixel é exatamente igual ao outro, ou seja, quando o pixel(x_0, y_0) da imagem esquerda é igual ao pixel(x_0, y_0) da imagem direita, o resultado da diferença é 0(zero), portanto a imagem resultante nesse ponto receberá um pixel preto($R=0, G=0, B=0$).

Não havendo igualdade entre os pixels, a cor na imagem resultante será igual a diferença. Com isso, pode-se determinar o quanto uma imagem é diferente da outra e também a necessidade de ajustes na geometria epipolar para gerar a estereoscopia. Na Figura 41 é apresentado o resultado da diferença entre as imagens paralelas da Figura 40.

Figura 41 – Imagem resultante da diferença entre as imagens paralelas



Fonte: autoria própria

O cálculo da diferença segue a Equação 4.1, onde duas imagens foram carregadas paralelamente em memória e percorridas pixel a pixel, verificando as diferenças entre os canais vermelho, verde e azul.

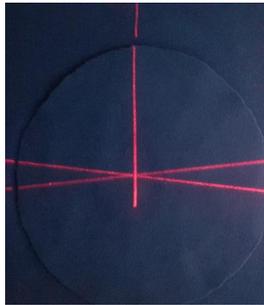
$$ImgDiff_{resultante_\eta} = \begin{cases} R_{resultante_\eta} \\ G_{resultante_\eta} \\ B_{resultante_\eta} \end{cases} = \begin{cases} \| RimgE_\eta - RimgD_\eta \| \\ \| GimgE_\eta - GimgD_\eta \| \\ \| BimgE_\eta - BimgD_\eta \| \end{cases} \quad (4.1)$$

Sendo $\eta = \{1, 2, 3, \dots, (altura * largura)\}$, altura e largura são iguais para as duas imagens, pois é uma condição da aquisição e da redução de contexto para o Stream2D. Portanto, se a imagem direita tiver 2500x2500 pixels, a imagem esquerda também terá esse tamanho. Assim como, a imagem resultante terá esse tamanho. Os canais de cores vermelho, verde e azul estão representados pelas letras da sigla RGB, as siglas imgE e imgD representam, respectivamente imagem esquerda e imagem direita. A palavra “resultante” representa o valor encontrado no cálculo entre os canais para um determinado pixel.

Na Figura 41 pode-se verificar que uma pequena parte do objeto se sobrepõe, isso é percebido no encontro entre as imagens. Este estudo mostra que uma sobreposição entre as imagens diminui o esforço na geração da estereoscopia, contudo somente a sobreposição não garante, pois utilizando a mesma imagem e duplicando-a, gerando o par estéreo e utilizando um software para colocá-las lado a lado ou gerando um anáglifo, não apresenta estereoscopia. Portanto, uma angulação entre elas deve ocorrer para que haja visualização de proporção e profundidade.

O mesmo processo foi realizado em imagens com captura convergente. Para isso, foram utilizados os lasers também posicionados a 6.5cm, direcionados para o centro do objeto, mas cruzando exatamente no início, determinado pelo laser transversal. Na Figura 42 é apresentada a preparação para a aquisição.

Figura 42 – preparação para aquisição convergente.

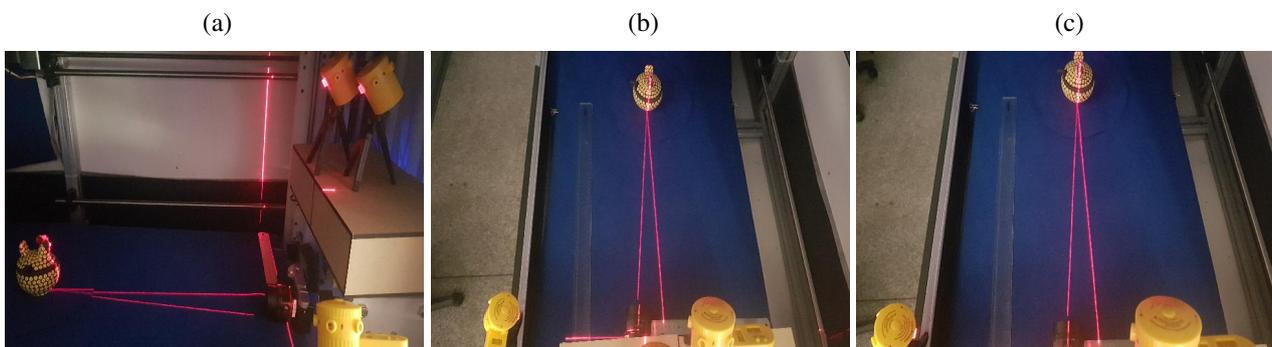


Fonte: autoria própria

Na Figura 42 é apresentado o posicionamento dos lasers fazendo a convergência e o terceiro laser (transversal) marcando onde o objeto será posicionado para se determinar a distância entre ele e a câmera. Após esse posicionamento, ela foi deslocada para marcar a distância entre as aquisições.

As três situações são mostradas na Figura 43, na primeira pode ser visto o laser transversal passando pela câmera e pela régua, a segunda é o laser passando pelo centro da câmera e também o laser transversal sobre a marca de 55cm na régua transparente, além do laser da direita na régua metálica na marca de 6.5cm. A terceira imagem é uma imagem mais aproximada, para mostrar mais detalhes que as outras duas.

Figura 43 – Posicionamento do objeto, lasers e réguas para a aquisição convergente

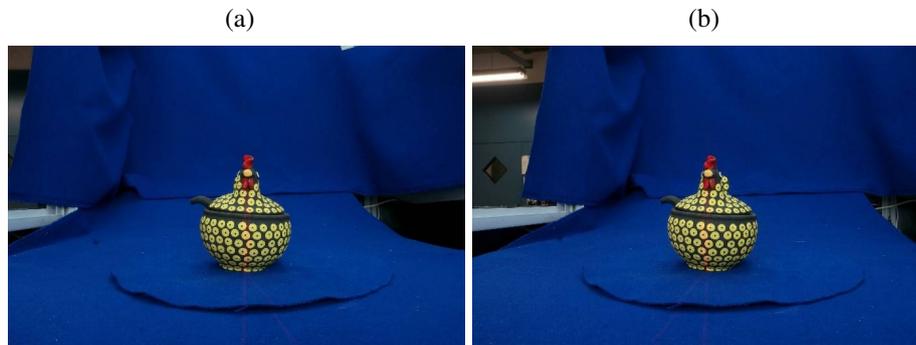


Fonte: autoria própria

As duas réguas marcam a distância entre o objeto e o sensor da câmera, régua transparente, e a distância de 6.5cm entre os lasers direito e esquerdo, régua metálica. O laser transversal marca a distância entre o objeto e o sensor da câmera, podendo ser visto na imagem intersectando

o laser esquerdo e sobre o ponto vermelho na lente da câmera. Isso garante que a lente está direcionada para o ponto de convergência dos lasers no objeto. Na Figura 44 é mostrado o par convergente, respectivamente, imagem direita e imagem esquerda.

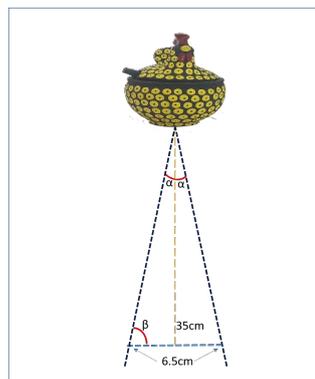
Figura 44 – Imagens convergentes: (a) esquerda e (b) direita.



Fonte: autoria própria

A aquisição convergente é a que mais se aproxima daquela realizada pelo F2S2. Essa proximidade é percebida pelo fato de que as imagens apresentam uma angulação entre elas. Os ângulos obtidos nesse experimento são apresentados na Figura 45, para efeito de comparação, principalmente nos ajustes que foram realizados nesse tipo de aquisição para gerar a estereoscopia.

Figura 45 – Ângulos na aquisição convergente



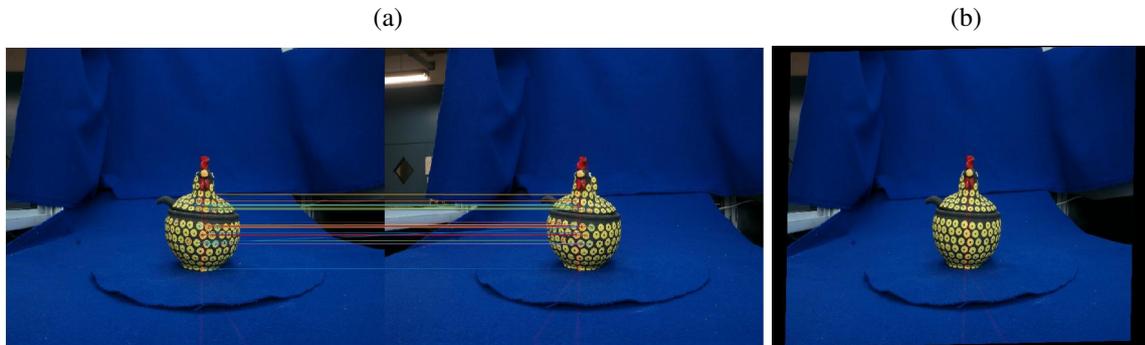
Fonte: autoria própria

O exemplo utiliza uma distância de 35cm entre a câmera e o objeto e com diferença de 6.5cm entre o posicionamento esquerdo e direito da câmera, por esse motivo os ângulos encontrados foram $\alpha \approx 5^\circ$ e $\beta \approx 85^\circ$. O ângulo β representa a inclinação ideal para a aquisição convergente para a distância câmera-objeto. Para a distância entre o objeto e câmera de 55cm os valores dos ângulos são $\alpha \approx 4^\circ$ e $\beta \approx 86^\circ$

A partir dessas imagens foi calculada a matriz fundamental por meio do algoritmo citado, utilizado na captura das imagens paralelas. O resultado com a matriz e a geometria epipolar

aplicada à imagem da direita pode ser vista na Figura 46.

Figura 46 – Algoritmo de Matching SURF: (a) matriz fundamental e (b) imagem da esquerda após ajustes da geometria - Aquisição Convergente



Fonte: autoria própria

Assim como ocorreu na aquisição em paralelo, o algoritmo apresenta uma situação de ajuste que indica haver pouca identificação entre as duas imagens para formar a estereoscopia, portanto, necessita de ajustes da geometria epipolar. O mesmo processo de diferença, mostrado na Equação 4.1, foi aplicado sobre essas imagens e obteve o resultado apresentado na Figura 47.

Figura 47



Fonte: autoria própria

Em relação ao resultado da diferença das imagens paralelas, na convergência há pouca diferença, no entanto ao se aplicar a estereoscopia, tanto anáglifo quanto na polarização, percebeu-se um pouco de desconforto.

Pelo lado da visão humana, quando se observa um objeto de forma paralela, o foco não está diretamente nele. Uma forma de avaliar isso é elegendo um objeto qualquer próximo a você, porém direcionar o olhar para um foco além. Isso simulará a aquisição paralela, pois o objeto estará no seu campo de visão e próximo a você, assim será possível identificar o objeto, inclusive sua tridimensionalidade, contudo pode ocorrer de não haver nitidez.

A mesma ideia é utilizada para a visão convergente, porém o foco é direcionado ao objeto, então fazendo o mesmo experimento descrito para a visão paralela, deve-se olhar diretamente para algum ponto no objeto e será possível perceber a profundidade e proporção, ou seja, terá a visão estereoscópica.

Em visualizações por meio de imagens, a percepção da proporção e profundidade pode não ocorrer apenas pela junção de duas imagens, como visto nas explicações das aquisições convergentes e paralelas. Essa percepção pode vir a ser percebida com ajustes da geometria epipolar em uma das imagens, contudo, a visualização será de paralaxe positiva, ou seja, dentro do plano de projeção.

4.3 Centralização do objeto e preparação das imagens

Como apresentado na subseção 3.3.2, inicialmente faz-se um planejamento para qualquer aquisição. O objetivo desse planejamento é obter imagens com qualidade e acuidade visual, onde seja possível aplicar os processos de escala, principalmente de aumento, a ponto de apresentar todos os detalhes que a olho nu não seja percebido, mas antes de chegar ao nível dos pixels. Outras transformações geométricas também podem ser aplicadas, como translação e rotação, isso garante em parte as características tridimensionais.

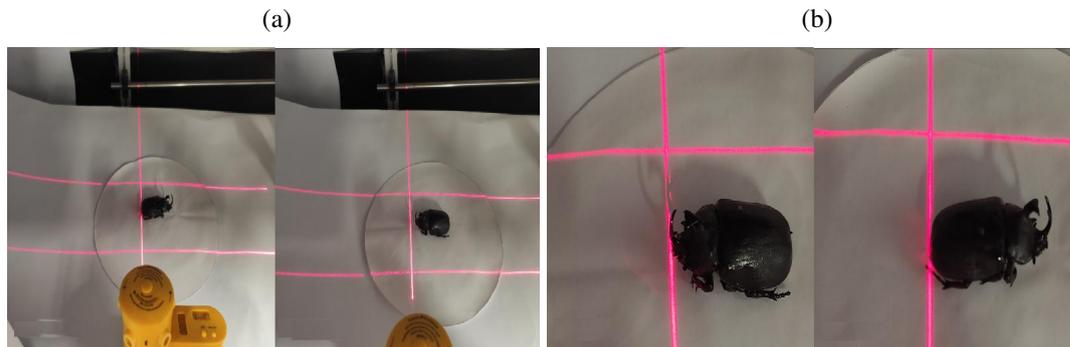
Um dos processos que permite alcançar a qualidade desejada é a centralização do objeto. Como informado, isso não está ligado à centralização baseada nas dimensões e nem no centro de gravidade, pois peças totalmente irregulares, como no caso da concha do mar (Figura 3.3.3, página 92), as condições para se determinar com exatidão esses valores são difíceis e toma tempo. Buscando contornar esse problema, adotamos a medição por laser.

A medição é realizada pelos mesmos três lasers utilizados nas aquisições paralela e convergente. Eles devem tocar o objeto em pontos que contribuem para o ajuste. Um dos lasers foi posicionado para a medição da parte frontal e traseira do objeto, isso garante que ele esteja centrado em relação a esse sentido. Após esse ajuste, posicionamos os outros dois lasers nas laterais do objeto, verificando visualmente a intensidade que estão tocando-o.

O processo todo foi realizado por meio de rotações do objeto em 180° e observando o quanto os lasers estão tocando o objeto, o que foi supracitado como intensidade. A centralização ocorreu fazendo ajustes do laser e do objeto, obedecendo ao critério de, caso não haja centralização, mover o objeto na metade da distância e o(s) laser(s) na outra metade, para o lado que for observado uma divergência maior. Na Figura 48 é mostrado o processo de centralização do objeto na parte dianteira e traseira.

O besouro foi posicionado em cima do eixo do motor, que é a parte central da base, mas como já dito, não garante que o objeto esteja centralizado. Isso serve apenas para também posicionar o laser da parte frontal/traseira.

Figura 48 – Processo de centralização por laser da parte frontal e traseira



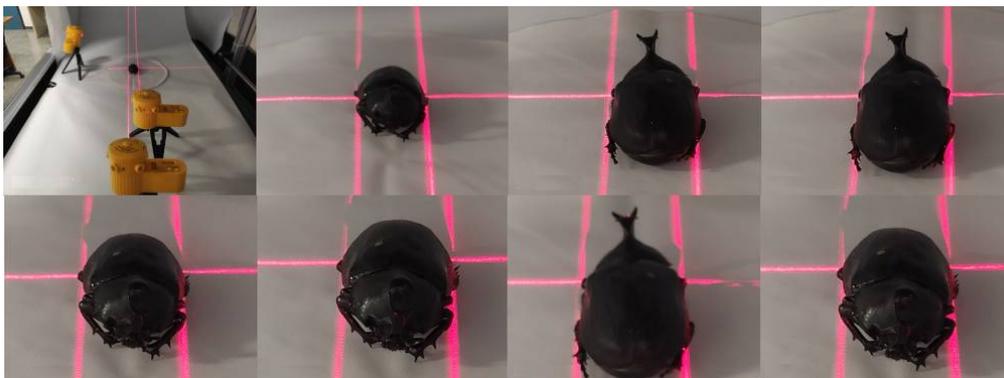
Fonte: autoria própria

O laser foi posicionado para tocar no objeto, que no exemplo é a parte traseira do besouro. Percebe-se pela Figura 48.a que o laser toca o besouro, então executa-se um giro de 180° e ao terminar verifica-se a situação, que no exemplo fica perceptível que não há toque na parte frontal.

Para ajustar isso, deslocamos o besouro na metade da distância entre seu limite frontal (por exemplo, o chifre) e o laser. Após essa movimentação, deslocamos o laser na outra metade. Isso faz com que o laser volte a tocar o besouro. Então, giramos novamente em 180° para determinar se o laser continuará tocando na mesma intensidade a parte de trás. Isso pode ser visto na Figura 48.b.

O processo de regular a parte frontal/traseira terminou assim que a mesma intensidade do laser foi percebida nas duas partes. Tendo-se essa percepção, regulamos os lasers da lateral. A sequência apresentada na Figura 49 diz respeito a essa regulagem.

Figura 49 – Ajustes dos lasers da lateral



Fonte: autoria própria

A primeira imagem mostra o posicionamento inicial dos dois lasers; a segunda imagem é o zoom aplicado para mostrar onde os lasers tocam o besouro; na terceira imagem aplicamos um giro de 180° e percebe-se que o laser da direita tocou com mais intensidade do que o da esquerda.

Portanto, deve-se executar o mesmo processo que foi realizado no ajuste da frontal/traseira, que é deslocar cada um dos lasers e besouro na metade da distância do ponto divergente, ou seja, o besouro é deslocado para a esquerda, o laser da esquerda vai para a direita e o laser da direita vai para a esquerda. Isso pode ser visto na quarta imagem.

Ao se realizar os ajustes, gira-se em 180° , para observar novamente se não houve mudança, caso ainda persista a diferença nos dois lados, regula-se novamente. Pela primeira imagem na parte inferior, pode-se ver que ainda não estava ajustado. Na segunda vê-se que houve um novo ajuste. O processo foi repetido até se ter a centralização do objeto. O que pode ser visto nas duas últimas imagens.

Por ter havido uma modificação de posicionamento nas laterais, pode ser que tivesse ocorrido um deslocamento na parte frontal e traseira, então fez-se uma verificação. Contudo, o processo de ajuste lateral do objeto foi realizado observando o laser da parte frontal/traseira, para que não se alterasse esse ajuste. Portanto, há um cuidado para deslocar lateralmente o objeto, sem tirá-lo do alinhamento na parte frontal e traseira.

Além da centralização, os parâmetros de tamanho (altura e largura) do objeto devem ser anotados e também da base, quando houver, para lançá-los no software de controle do F2S2. Esses parâmetros possibilitam a determinação do raio mínimo para a aquisição, como foi apresentado na subseção 3.3.2. Somente após todos os parâmetros definidos é que se inicia a aquisição.

A formação do par estéreo parte do princípio de não causar desconforto no momento da visualização estereoscópica e isso está ligado diretamente ao grau entre as imagens, portanto este estudo mostra que as aquisições, para este fim, devem se manter entre 1° e 6° . A partir de 7° começa a necessitar de ajustes da geometria epipolar para gerar uma visualização sem causar o desconforto.

Considerando-se os ângulos para de B e C entre 1° e 6° , a aquisição segue como descrito pela Matriz 4.2, a qual foi definida na subseção 3.3.2, página 92. Utilizando para exemplificação o ângulo de 6° para “dbf” e “dbs”, os valores de p e k são encontrados pelo cálculo $p = \frac{90}{6}$ e $k = \frac{360}{6}$. Portanto, esse Stream2D tem p = 15 streams com k = 60 frames.

$$Stream2D_{(i,j)} = \begin{bmatrix} F_{1,1} & F_{1,2} & \dots & F_{1,k} \\ F_{2,1} & F_{2,2} & \dots & F_{2,k} \\ F_{3,1} & F_{3,2} & \dots & F_{3,k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{p,1} & F_{p,2} & \dots & F_{p,k} \end{bmatrix} \quad \begin{matrix} i = 1, 2, 3, \dots, p \\ j = 1, 2, 3, \dots, k \end{matrix} \quad (4.2)$$

Os movimentos dos eixos B (θ) e C (γ), como descrito na subseção 3.3.2, Figura 22, página 86, possibilitam o posicionamento da câmera e conseqüentemente a angulação para obter cada uma das linhas (streams). Considerando a aquisição do primeiro stream, ele tem os valores

de $i = 1$ e $j = 1, 2, 3, \dots, 60$, $Stream2D_{(1,j)} = \{F_{1,1}, F_{1,2}, F_{1,3}, \dots, F_{1,60}\}$ e o último tem os valores de $i = 15$ e $j = 1, 2, 3, \dots, 60$, $Stream2D_{(15,j)} = \{F_{15,1}, F_{15,2}, F_{15,3}, \dots, F_{15,60}\}$.

O comportamento do F2S2 para as aquisições segue uma sequência, então após as definições dos parâmetros de aquisição e o aceite para iniciar, a base B é movida para a posição inicial, determinada por um sensor colocado abaixo dela, a câmera é posicionada na altura e distância determinadas nos parâmetros, regula-se o foco, fecha-se as portas e novamente no software clica-se no botão para continuar.

A primeira foto é com o objeto no ângulo $\theta_1 = 0$, $i = 1$ e $j = 1$. A segunda é o deslocamento angular de B (dbf), o qual foi determinado no exemplo como 6° , assim sendo, o novo ângulo de $\theta_2 = (\theta_1 + 6)$, $i = 1$ e $j = 2$ e assim por diante até uma posição anterior a se completar 360° , ou seja, $\theta_{60} = 354^\circ$ e $j = 60$, ficando inalterado o valor de $i = 1$. O motivo de não haver a aquisição do ângulo 360° é que essa perspectiva do objeto foi adquirida em θ_1 . Ao se terminar essa aquisição, o ângulo de deslocamento da câmera também é acrescido do valor determinado para γ , portanto, $\gamma_1 = 0$ e $\gamma_2 = \gamma_1 + 6$, tendo o processo de aquisição recomeçado com $\theta_1 = 0$, $i = 2$ e $j = 1$.

Todas as imagens capturadas são no formato “RAW”, contudo cada câmera que adquire nesse formato tem seu padrão, por exemplo, a câmera utilizada da Samsung tem o formato e extensão proprietária “SRW”. No entanto, existem bibliotecas da OpenCV que permite a abertura e tratamento a partir desse padrão. Os tratamentos, correções e conversões para “PNG” são parte do trabalho do doutorando Diogo Roberto Olsen.

Para este trabalho cabe como citação os processos que são realizados, como correção de lente, correções de balanço de branco, redução de contexto, mas principalmente a segmentação. A segmentação proporciona ter apenas o objeto de interesse na imagem e com isso possibilita a criação da visualização com ou sem fundo. Por exemplo, a imagem ser apresentada com um fundo preto ou colocar um ambiente, que seria uma foto.

Isso serviria para uma aplicação onde deseja-se colocar o objeto adquirido dentro do seu “habitat”, como uma abelha em colmeia ou coletando pólen em plantas. Nesse panorama, adquiria-se a abelha, coletaria fotos do seu habitat e durante a visualização mesclaria a visualização estereoscópica da abelha com o fundo estático do habitat. Ao término dos processos citados, tem-se a formação do Stream2D, como apresentado na Figura 51, página 122. Tendo isso, parte-se para a criação da visualização interativa estereoscópica.

4.4 Stream2D na formação da estereoscopia

Em aquisições por câmeras fotográficas existe a propriedade de não haver garantias da distância do objeto na foto, podendo ter uma imagem que represente uma distância menor do que 25cm que não permita a visão estereoscópica, assim como ter aquela que permita. Uma câmera

pode estar próxima ao objeto, contudo sua configuração para aquisição traga a sensação de que o objeto esteja mais afastado.

Essa distância não está ligada ao processo de escala, pois uma imagem que produza estereoscopia, mesmo tendo sua escala aumentada, ela mantém a estereoscopia até um certo ponto, relacionado ao plano de projeção, ou seja, chegará um momento em que a imagem extrapolará o plano e dificultará a visualização estereoscópica.

As aquisições do F2S2 com a montagem do Stream2D para o propósito de estereoscopia mostram-se diferenciadas de outras, uma vez que não há necessidade de ajustes da geometria epipolar nas imagens. Em relação ao posicionamento da câmera a menos de 25cm do objeto, pode-se configurar parâmetros, como profundidade de campo e distância focal, na tentativa de se gerar a estereoscopia.

Essa característica das visualizações por imagem permitem estereoscopia com uma distância menor do que 25cm, entre a câmera e o objeto, foi avaliada com a aquisição de alguns insetos como no caso do escaravelho. Ele foi adquirido com 15cm de raio e 3° entre frames. Na Figura 50 é apresentado um par que ditam 6° entre os frames.

Figura 50 – Imagem com 15cm na aquisição: (a) frame esquerdo e (b) frame direito

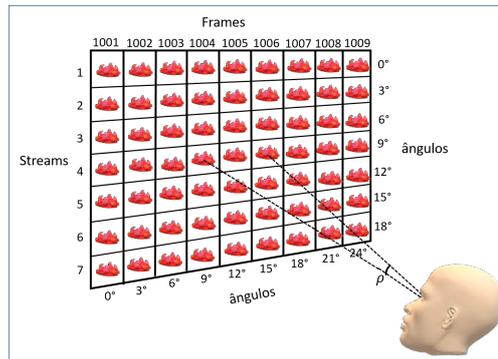


Fonte: autoria própria

Segundo a teoria da acomodação visual, essa estereoscopia não pode ocorrer, pois entra no espaço de conforto dos olhos. Isso será visto na Seção 5.3, página 150, com a análise do escaravelho e de outros insetos. Aqui, discutir-se-á o Stream2D e se ele pode ser utilizado para gerar estereoscopia independentemente da distância entre a câmera e o objeto.

O Stream2D representado pela matriz $Stream2D_{(i,j)}$ da subseção 3.3.2, onde os streams são as linhas e os frames as colunas, pode ser disposto em um plano bidimensional, onde pode ser percorrido para encontrar o par que provê estereoscopia. Esta é dependente do ângulo de aquisição, pois é por meio desse valor que se calcula os componentes do par estéreo. Na Figura 51 é apresentada uma situação de aquisição onde o ângulo estipulado para os streams e frames foi de 3°.

O valor de ρ será debatido no decorrer desta seção, pois ele diz respeito ao valor de

Figura 51 – Submatriz da $Stream2D_{(i,j)}$ com 3° entre streams e frames

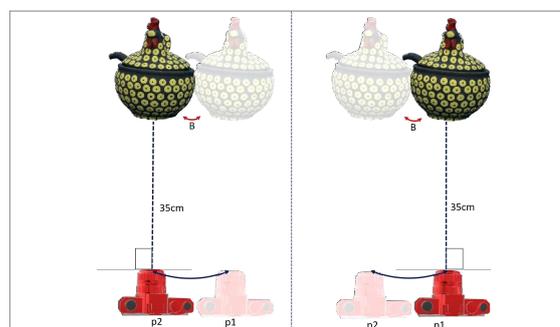
Fonte: autoria própria

formação de estereoscopia para as aquisições do F2S2. Fazendo uma comparação com as aquisições debatidas, esse ângulo não está ligado ao deslocamento da câmera em 6.5cm, mas à rotação angular do objeto. Assim, tem-se que a câmera está direcionada para uma face ou perspectiva do objeto para um determinado ângulo, o qual, em um segundo momento, é alterado, girando-se θ graus em relação à aquisição anterior.

Diferentemente das aquisições paralela e convergente, não se faz mudança de posicionamento da câmera ou utilização de duas. Levando em consideração um comportamento fictício da câmera, ela faz um movimento circular com raio fixo, determinado nas configurações de início de cada aquisição. No entanto, a cada perspectiva, o ângulo entre a câmera e o objeto será sempre 90° , mantendo-se o fundo e laterais sempre inalterados.

Na Figura 52 é apresentada essa situação fictícia de deslocamento da câmera, pois em p_1 o ângulo do objeto era qualquer entre 0° e 120° , considerando uma aquisição em 3° .

Figura 52 – Posição fictícia da câmera durante a aquisição



Fonte: autoria própria

Sendo p_1 um ângulo qualquer no stream e p_2 determinado por $p_2 = p_1 + dbf$, então p_2 é o p_1 que foi acrescido de 3° , por exemplo, se p_1 fosse o ângulo $\theta = 15^\circ$, p_2 seria $\theta = 18^\circ$. Quem se moveu foi o objeto e não a câmera, por isso é uma situação fictícia. O que se quis apresentar

com isso é que, se fosse a câmera que se movesse ao redor do objeto, teria-se o mesmo efeito, contudo não seria igual a uma aquisição convergente.

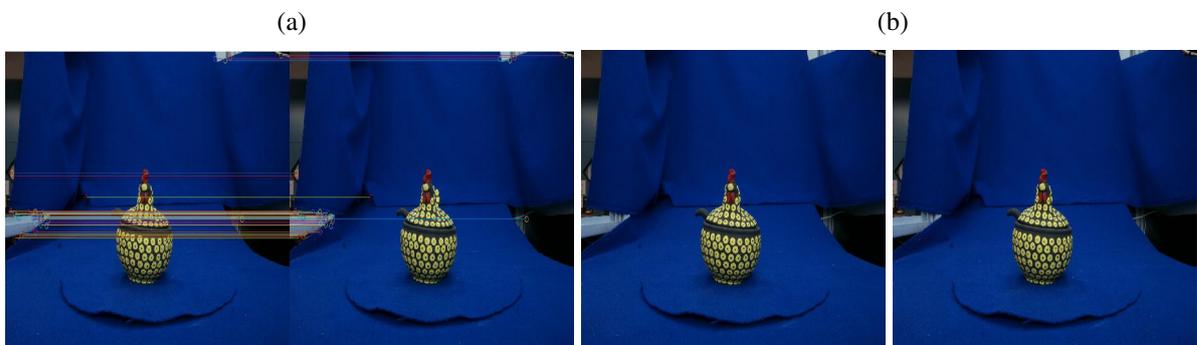
Na aquisição convergente a câmera assume uma angulação em relação ao objeto, a qual varia de acordo com a distância entre a câmera e ele. Já nas aquisições do F2S2, como citado anteriormente, o ângulo sempre será de 90° .

Em termos de análise das capturas, a matriz fundamental e geração dos ajustes pelo algoritmo SURF não apresentou modificação na imagem. O teste foi realizado nos mesmos moldes utilizados para a aquisição paralela e convergente.

A câmera foi posicionada em quatro diferentes distâncias (25, 35, 45 e 55cm) do objeto e este foi rotacionado em 3° , produzindo quatro imagens (0° , 3° , 6° e 9°) para cada uma das distâncias.

No entanto, para efeito de comparação entre o resultado das aquisições paralela e convergente, na Figura 53 é apresentada a aquisição na mesma distância utilizada nessas outras duas apresentadas, ou seja, em 25cm e com imagens em 0° e 6° .

Figura 53 – Algoritmo SURF em Frames do F2S2: (a) matriz fundamental e (b) imagem original e imagem ajustada



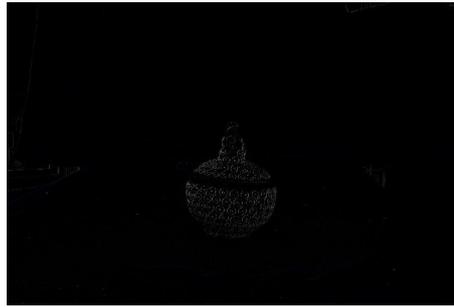
Fonte: autoria própria

A imagem resultante sofreu um ajuste da epipolar quase imperceptível, permanecendo praticamente similar à original. Isso também foi identificado utilizando imagens com distância angular menor do que 6° . Esse ajuste mínimo pode ser visto na Figura 54, onde é apresentado o cálculo de diferença entre as duas imagens, original e a que sofreu o ajuste da epipolar.

Em comparação aos ajustes necessários nas imagens paralelas e convergentes, as imagens produzidas pelo F2S2 sofreram menos distorções. Isso também pode ser percebido quando se faz a diferença entre os frames utilizados para encontrar a matriz fundamental. O ângulo entre esses dois frames é igual a 6° e a diferença entre eles é apresentada na Figura 55.

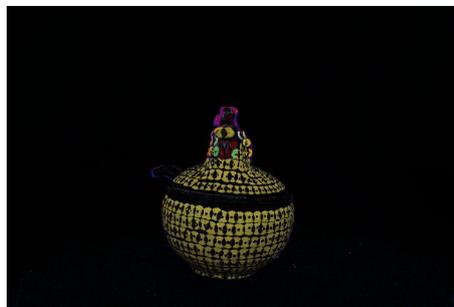
Em imagens com distância angular próxima a 6° , por exemplo 9° , a diferença entre elas se aproxima do resultado encontrado nas imagens convergentes. Em imagens com distância superior

Figura 54 – Cálculo da diferença entre imagens da Figura 53.b



Fonte: autoria própria

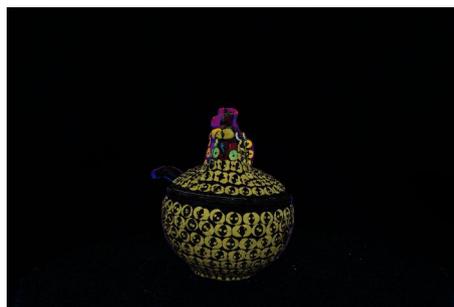
Figura 55 – Diferença entre os frames 0° e 6°



Fonte: autoria própria

a 9° a diferença se aproxima do resultado encontrado nas imagens paralelas. Um exemplo disso foi o resultado entre as imagens 0° e 9°, o qual é apresentado na Figura 56 .

Figura 56 – Diferença entre os frames 0° e 9°



Fonte: autoria própria

Como citado na seção 4.2, quando se realiza o cálculo de diferença entre uma mesma imagem, o que tem de resultado é uma imagem totalmente preta, pois um pixel qualquer que seja $RGB=[145,135,100]$, diminuído dele mesmo resulta $RGB=[0,0,0]$. Então, onde há uma pequena diferença, o pixel é inserido na imagem com a cor referente a essa divergência e quanto mais pixels diferentes de preto, menos igualdade há entre as imagens. Essa diferença também indica o provável ajuste da geometria epipolar em uma das imagens, para se ter estereoscopia com o par

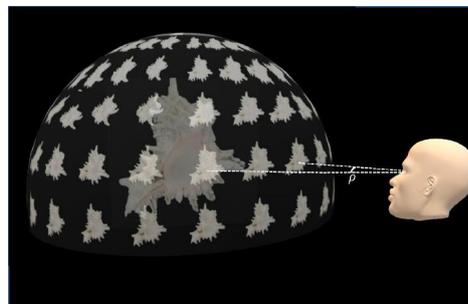
estéreo escolhido.

4.4.1 Par estéreo no Stream2D

A aquisição semiesférica do F2S2 proporciona a criação de um plano esférico onde as imagens são dispostas, isso significa transformar a matriz do plano cartesiano em uma estrutura esférica. Com isso, cada perspectiva projetada para dentro do plano e sobrepondo pontos em comum representará o objeto real tridimensionalmente. Além disso, cada perspectiva estará exatamente na distância angular de suas adjacentes (esquerda, direita, superior e inferior) determinada no início da aquisição.

Utilizando o mesmo esquema apresentado na Figura 51, mas agora no plano esférico, o ângulo ρ permanece o mesmo, que significa o usuário estar olhando para duas perspectivas dentro do Stream2D que distam valores angulares, os quais são determinados na aquisição, contudo pode ocorrer de uma visualização utilizar um par em que sua distância seja maior do que a determinada na aquisição. Para o caso da estereoscopia, essa distância angular é igual ou menor do que 6° , como visto na seção 4.3. Na Figura 57 é apresentada como se dá a visualização do par estéreo no Stream2D no plano esférico.

Figura 57 – Formação de par estéreo no plano esférico do Stream2D



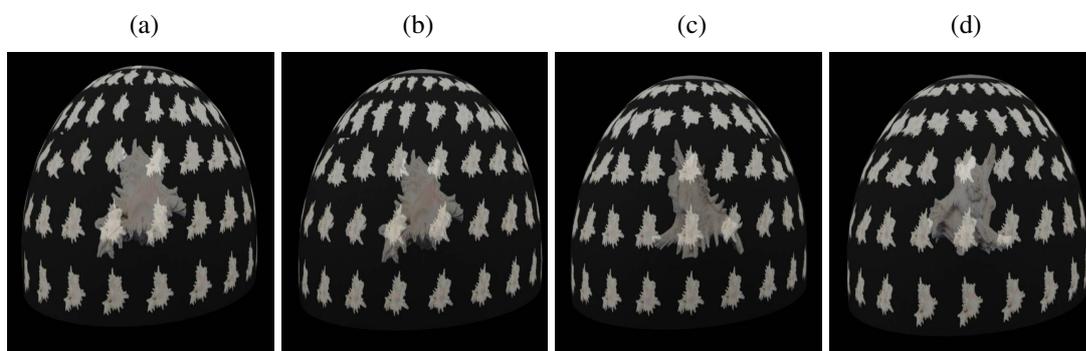
Fonte: autoria própria

O valor de 6° foi determinado inicialmente levando em consideração as aquisições comuns em filmagens cinematográficas e nos textos dos autores citados, onde acreditava-se que ele tinha relação com a diferença interpupilar. O que se percebeu é que a relação é tênue, uma vez que o ângulo entre as imagens na aquisição convergente vai variar de acordo com a distância do objeto às câmeras de aquisição. Contudo, aquisição apresenta algumas semelhanças, principalmente para as distâncias testadas neste trabalho.

Nos testes realizados foi identificado que, para as mesmas distâncias utilizadas na aquisição do F2S2, o ângulo para o método convergente é de: $\approx 7^\circ$ para 25cm, $\approx 5^\circ$ para 35cm, $\approx 4^\circ$ para 45cm e $\approx 3^\circ$ para 55cm. No entanto, esses valores são apenas a metade do ângulo, ou seja, eles correspondem ao α (Figura 45) e que para as duas posições (esquerda e direita) o valor é o dobro ($2 \times \alpha$).

Já nas aquisições do F2S2, independentemente da distância que se estipule para a captura, o ângulo será o mesmo, ou seja, se a posição da câmera for estipulada para 30cm, 40cm ou qualquer outra dentro dos limites do F2S2, pode-se determinar um ângulo que permita criar um par estéreo, que ao ser visualizado em estereoscopia não cause desconforto. Assim, mesmo o F2S2 podendo realizar uma aquisição entre valores angulares dentro do intervalo de $< 1^\circ$ até 360° para os frames, esse valor permanecerá o mesmo do início ao fim da aquisição. Então, se o ângulo escolhido for 2° para θ , este não será alterado durante o processo. Na Figura 58 é apresentada a forma semiesférica da aquisição.

Figura 58 – Stream2D no plano esférico, visão: (a) frontal; (b) direita; (c) esquerda; e (d) traseira



Fonte: Autoria de Diogo Roberto Olsen e Flávio de Almeida e Silva (outro deste trabalho) com design de Felipe Teixeira de Almeida e Silva

O ângulo ρ , apresentado na Figura 51 e na Figura 57, não é igual ao ângulo de aquisição θ , a não ser quando este for superior a 3° e igual ou inferior a 6° . Para os casos em que θ assumia valores inferiores ou igual a 3° , calcula-se a distância entre o frame chamado de “imagem base” e os frames que formarão o par estéreo. A imagem base nem sempre fará parte do par e isso foi adotado para a navegação na janela deslizante, como será apresentado no decorrer deste trabalho.

Para determinar o par estéreo por meio do ângulo de aquisição e da formação do Stream2D, inicialmente foi utilizada a técnica de anáglifo e o software Gimp, para verificar qual paralaxe era possível ser gerada. Essa hipótese surgiu pelo fato de que a aquisição não segue o mesmo formato do encontrado na literatura.

Nos trabalhos citados na subseção 2.3.1, página 55, as aquisições são com câmeras diferentes, em formação de cúpula, dispostas em arco ou mesmo por ângulos, mas isso gera, em muitos casos, paralaxe positiva. Para comprovar a eficácia do F2S2 na produção de conteúdos estereoscópicos, com paralaxe negativa, foram realizados experimentos utilizando duas imagens, iniciando pela construção do par com a menor distância, ou seja, a utilização da mesma imagem.

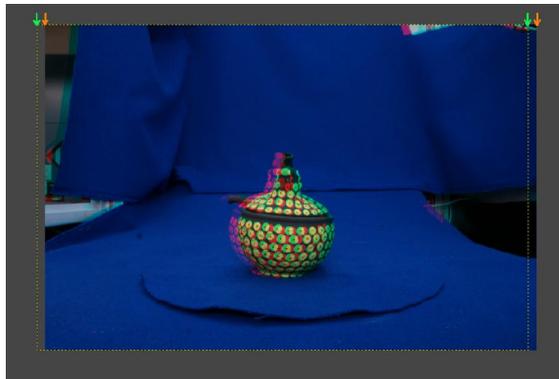
Esse experimento produziu a estereoscopia com paralaxe positiva. No entanto, não produz se apenas houver uma sobreposição das imagens, necessita-se de um deslocamento horizontal de uma delas. Assim, utilizando o Gimp, abriu-se a mesma imagem nomeadas como *ImgE* e *ImgD*,

deslocou-se a primeira para a esquerda e aplicou-se as configurações de canais e o processo de adição para geração do anáglifo.

O processo de adição é o contrário do que foi aplicado para encontrar a diferença entre as imagens, pois onde os pixels se sobrepõe a cor volta a da imagem original, onde não há sobreposição, ficam nas cores vermelho e ciano.

Esse deslocamento é uma das formas mais simples de ajuste, pois não faz alteração de posicionamento dos pixels, como ocorre nos ajustes em que a imagem é rotacionada e/ou inclinada. O resultado do processo descrito é visto na Figura 59. Pode-se observar a estereoscopia utilizando os óculos anáglifo no padrão vermelho e ciano.

Figura 59 – Estereoscopia Anáglifo da mesma imagem



Fonte: autoria própria

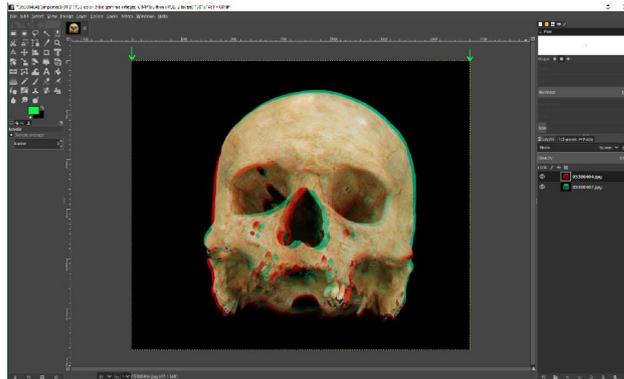
Essas imagens, *ImgE* e *ImgD*, que formam o par estéreo, são, respectivamente, imagem para olho esquerdo e direito. As setas em verde claro e laranja, indicam o quanto de deslocamento foi aplicado entre as imagens para que se produzisse corretamente a estereoscopia com paralaxe positiva.

Esse deslocamento, sendo as setas verdes os limites da imagem esquerda e as setas laranjas os limites da imagem direita, deve ser desconsiderado no momento de gravar a imagem em disco, ou seja, corta-se a imagem a partir da linha pontilhada, marcada pela seta verde da direita até a o limite da imagem marcada pela seta laranja da esquerda.

O corte ocorre devido à região em que não há sobreposição das imagens e não gera estereoscopia. Com imagens que não são as mesmas isso não ocorre, pois a diferença angular entre elas fará o papel do deslocamento ou mesmo dos ajustes da geometria, como visto no início desta seção (4.4).

Para a sequência dos experimentos, determinou-se uma diferença de 2° entre cada imagem para verificação da necessidade de ajustes horizontais ou da geometria. A imagem formada é apresentada na Figura 60. As setas em verde mostram os limites das imagens esquerda e direita.

Figura 60 – formação da estereoscopia com imagens que distam 2°



Fonte: autoria própria

Como houve sobreposição das duas imagens em toda a extensão, descartou-se os ajustes. Também não foi necessário o corte, uma vez que não ocorreram partes excedentes ou sem sobreposição. A estereoscopia gerada e que pode ser observada com os óculos anáglifo, foi com paralaxe negativa. Isso significa que a noção de proporção e profundidade está fora do plano de projeção.

Este mesmo processo foi realizado para frames com ângulos de diferença entre 3° e 6°. Nas aquisições entre 4°, 5° e 6° a formação do par estéreo, que não causa desconforto, deve ser entre um frame qualquer e seu adjacente. A explicação para isso é que a diferença angular maior do que 6° obtida entre o frame esquerdo e o frame direito provoca um imediato desconforto com a estereoscopia anáglifo, já na polarizada ela não aparece sem ajustes das imagens. Por exemplo, ao selecionar duas imagens adquiridas em 4° e sendo estas não adjacentes, a diferença angular será de no mínimo 8°, extrapolando os limites da estereoscopia sem desconforto.

O par estéreo, independentemente do ângulo de aquisição, dentro dos parâmetros para a estereoscopia, pôde ser formado dentro do Stream2D. Não somente entre frames do mesmo stream, mas também entre frames de streams diferentes, contudo a visualização é mais interessante em dispositivo que permita a identificação do direcionamento da visualização, ou seja, em um celular onde tem acelerômetro dá para identificar a inclinação da cabeça para mostrar a perspectiva. Isso permite uma estereoscopia na diagonal e vertical, além da navegação com essa inclinação. Essas visualizações são possíveis em projetor e monitor, contudo para ver na inclinação desejada tem que inclinar a imagem na tela.

4.5 Janelas deslizantes no Stream2D

A navegação no Stream2D com formação do par estéreo tem o intuito de não permitir interrupções indesejadas, como tempo de espera no carregamento dos streams e frames. A visualização deve ocorrer de forma que ao mudar um frame, seja imediatamente visto com

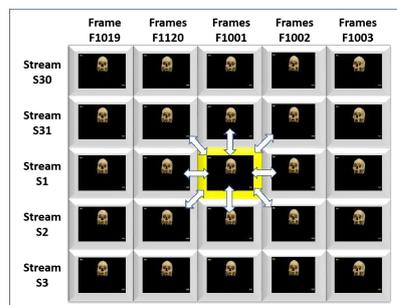
estereoscopia.

Para isso, os streams e frames podem ser colocados em memória, permitindo um acesso rápido. Contudo, uma aquisição completa, com pouco mais de 32 mil imagens, tornaria o carregamento em memória custoso computacionalmente. Mesmo que não seja completo, como nas aquisições de 3° em 3° , o que gera um Stream2D de 3720 (três mil e setecentos e vinte) imagens, pode ocorrer de um computador não ter capacidade para armazenar todo o Stream2D em memória.

Com esta visão, optamos por desenvolver uma forma de carregar um conjunto de imagens, como uma submatriz ($Stream2D_{(s,\tau)}$) da Matriz $Stream2D_{(i,j)}$. Para tanto, criamos uma lista circular totalmente conectada, onde a navegação pode ocorrer sem a necessidade de carregar e descarregar frames, mas ocorrendo esses dois processos, que eles causem interrupções na visualização.

A denominação “totalmente conectada” vem do fato de que um frame está conectado com todos os seus adjacentes superior, inferior, esquerda, direita e também nas diagonais. O termo circular é devido à mudança na visualização do último frame para o primeiro do mesmo stream e vice-versa, assim como de um frame, em uma posição qualquer, do último stream para o frame, na mesma posição, no primeiro stream e vice-versa. Na Figura 61 é apresentada como são as ligações entre um frame e seus adjacentes.

Figura 61 – Ligação entre um frame e seus adjacentes



Fonte: autoria própria

Os processos, descarregar e carregar, que envolvem a mudança na janela são dependentes da posição da imagem base. Avaliar essa posição determina se haverá ou não uma alteração na composição da janela.

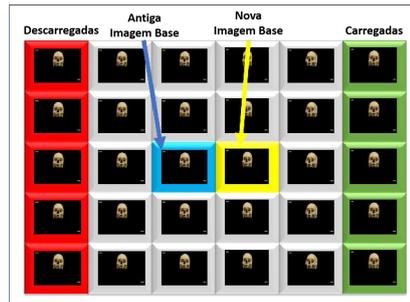
Essa composição pode sofrer alteração de três formas: a) descarrega-se e carrega-se um conjunto de frames de um stream, carregamento horizontal; b) descarrega-se e carrega-se um conjunto de frames de alguns streams, carregamento vertical; ou c) descarrega-se e carrega-se um subconjunto de frames de ambas as formas (“a” e “b”).

Essas formas são independentes da direção da movimentação, ou seja, se ela for da esquerda para a direita ou da direita para a esquerda, o processo “b” será ativado. Sendo da

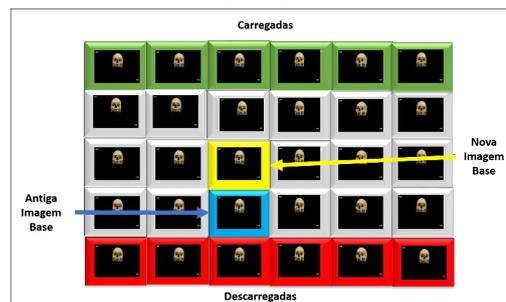
parte superior para a inferior ou o contrário, o processo “a” será ativado. O processo “c” só será ativado se a movimentação for para alguma das diagonais. Na Figura 62 são apresentas essas formas.

Figura 62 – Processos de carregar e descarregar um subconjunto da $Stream2D_{(i,j)}$

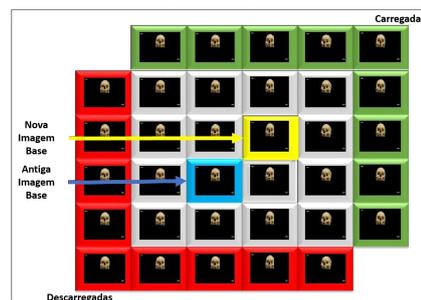
(a) Imagem base na linha i na $Stream2D_{(i,j)}$ para $Stream2D_{(i,j+1)}$



(b) Imagem base na coluna j na $Stream2D_{(i,j)}$ para $Stream2D_{(i-1,j)}$



(c) Imagem base na diagonal (i,j) na $Stream2D_{(i,j)}$ para $Stream2D_{(i-1,j+1)}$



Fonte: autoria própria

O algoritmo de deslocamento, carregamento e descarregamento de frames e streams recebe um comando de navegação do usuário, analisa a direção do movimento e se a imagem base atingirá um limite dentro da janela.

O intuito é verificar se o par estéreo poderá ou não ser formado com o próximo movimento na mesma direção. Caso a formação do par não ocorra com a movimentação na mesma direção,

faz-se os processos descritos (a, b ou c). O pseudocódigo do algoritmo da janela deslizante é apresentado no Algoritmo 1.

Algoritmo 1: Pseudocódigo da janela deslizante

```

1 Função Janela (Movimento)
2   Recebe comando de movimentação;
3   Verifica limites da janela;
4   Analisa próximo movimento;
5   se limite atingido no próximo movimento então
6     Descarrega frames e/ou streams;
7     Carrega frames e/ou streams;
8   fim
9   Retorna Janela;

```

O comando de movimentação vem do usuário, por manipulação de teclado, mouse, tela touch ou até mesmo acelerômetro de celular em uma visualização com óculos de realidade virtual. A função recebe esse comando e verifica os limites da janela, isso se dá por meio do nome do frame, já que ao carregar a janela em memória, também se guarda as informações acerca de cada um dos streams e frames. Assim, se a imagem base for a da Matriz $Stream2D_{(1,1001)}$, a janela é composta de um número de streams e frames a partir dela.

Se a configuração da janela for de 11 streams por 33 frames, um total de 363 frames são armazenados em memória. Então, a montagem inicial da janela, para uma aquisição de 3° em 3° , é a submatriz $Stream2D_{(\varsigma,\tau)}$ (Matriz 4.3), sendo $\varsigma = \{27, 28, \dots, 1, \dots, 5, 6\}$ e $\tau = \{1105, 1106, \dots, 1001, \dots, 1016, 1017\}$.

$$Stream2D_{(\varsigma,\tau)} = \begin{bmatrix} S27 - F1105 & \dots & S27 - F1001 & \dots & S27 - F1017 \\ S28 - F1105 & \dots & S28 - F1001 & \dots & S28 - F1017 \\ \vdots & \dots & \vdots & \dots & \vdots \\ S1 - F1105 & \dots & S1 - F1001 & \dots & S1 - F1017 \\ \vdots & \dots & \vdots & \dots & \vdots \\ S5 - F1105 & \dots & S5 - F1001 & \dots & S5 - F1017 \\ S6 - F1105 & \dots & S6 - F1001 & \dots & S6 - F1017 \end{bmatrix} \quad (4.3)$$

A quantidade descrita está relacionada ao total de streams e frames da submatriz da matriz $Stream2D_{(i,j)}$, portanto, havendo mudança na quantidade, refletirá no tamanho da janela. Por exemplo, em uma aquisição de 1° para streams e frames, o tamanho da janela deverá ser aumentado, pois as imagens do par estéreo, para uma boa formação da estereoscopia, estarão a 6 frames de distância. Caso seja mantido o tamanho de 11×33 , ela será reconfigurada com maior frequência.

Considerando a descrição dada da janela (3° entre streams e frames) e assumindo um movimento horizontal para a direita e a imagem base “S1-F1001” (destacada em vermelho na

Matriz 4.3), esse movimento tornará o frame “S1-F1002” a nova base, ela estará na mesma linha ς , mas em outra coluna τ da submatriz $Stream2D_{(\varsigma,\tau)}$, com isso esse movimento alterará a coluna τ em uma posição, mantendo a mesma linha ($Stream2D_{(\varsigma,\tau+1)}$).

A formação do par estéreo será entre a sua adjacente da esquerda e da direita, ou seja, frames “S1-F1001” e “S1-F1003”, que na matriz estarão, respectivamente, nas posições $Stream2D_{(\varsigma,\tau-1)}$ e $Stream2D_{(\varsigma,\tau+1)}$. Com esse movimento a janela não se reconfigura, pois fazendo um movimento ainda para a direita, levando a imagem base mais uma posição ($Stream2D_{(\varsigma,\tau+1)} = S1 - F1003$), o par estéreo estará dentro da janela e o frame que formará o par ainda estará longe do limite direito ($Stream2D_{(\varsigma,33)} = S1 - F1017$).

Ao tornar o frame “S1-F1003” a nova imagem base ($Stream2D_{(\varsigma,\tau)}$), o par estéreo é montado com “S1-F1002” e “S1-F1004”, respectivamente $Stream2D_{(\varsigma,\tau-1)}$ e $Stream2D_{(\varsigma,\tau+1)}$. Como descrito, o frame “S1-F1017” ($Stream2D_{(\varsigma,33)}$) é o limite direito e uma reconfiguração da janela ocorreria quando a imagem base se tornasse “S1-F1014” ($Stream2D_{(\varsigma,\tau)}$) e os frames “S1-F1013” e “S1-F1015” o par estéreo, que seriam $Stream2D_{(\varsigma,\tau-1)}$ e $Stream2D_{(\varsigma,\tau+1)}$. Nessa situação, $Stream2D_{(\varsigma,\tau+1)} = S1 - F1015$ ficaria apenas a 2 posições do limite.

Nesse panorama, um próximo movimento tornaria o frame “S1-F1015” a imagem base ($Stream2D_{(\varsigma,\tau)}$) e o seu adjacente da direita, “S1-F1016” ($Stream2D_{(\varsigma,\tau+1)}$), estaria muito próximo do limite da janela. Com isso, duas movimentações, na mesma direção, provocaria um erro, já que a imagem base seria “S1-F1017” ($Stream2D_{(\varsigma,\tau)}$) o novo frame da direita que formaria o par ($Stream2D_{(\varsigma,\tau+1)} = S1 - F1018$) não pertenceria à janela. Antes que isso ocorra, uma reorganização da janela se faz necessária.

Os movimentos para a esquerda, para cima e para baixo obedecem aos mesmos critérios. Já os movimentos diagonais verificam os dois limites, superior e esquerda ou direita, inferior e esquerda ou direita. Em caso de proximidade com um desses limites, a janela é reorganizada, um exemplo de uma aproximação com os limites superior e direito foi apresentado na Figura 62.c.

O algoritmo/método proposto, com janelas deslizante por meio de uma lista dinâmica, possibilitou a estereoscopia com uso de recursos computacionais modestos. Pois, o carregamento de algumas imagens, como determinado pela janela deslizante, pôde ser executado em computadores pessoais com 16MB de memória RAM e disco SSD. Isso é demonstrado em resultados experimentais.

4.6 Projeção interativa estereoscópica para visualização objetos reais em ambiente virtual

Nas seções 4.3, 4.4 e 4.5 foram descritos os processos que envolvem a aquisição, formação do Stream2D e como ocorre o carregamento de streams e frames em memória. Todos esses processos são componentes do modelo de projeção interativa estereoscópica.

Optamos pelo termo “interativa” e não tridimensional, uma vez que não se faz a reconstrução do objeto por meio de malha ou nuvem de pontos. Simula-se a tridimensionalidade quando se faz a navegação pelo Stream2D, apresentando a cada movimento uma das perspectivas do objeto adquirido.

A interatividade também é um diferencial neste trabalho, por contribuir para o desenvolvimento de objetos de aprendizagem que se buscam uma proximidade com a realidade. Os objetos não perdem qualidades de textura, cor e detalhes, pelo fato da aquisição ser em 8k. Assim, ao se aplicar alguma escala ou mesmo cortes na imagem original, não se insere artefatos inexistentes ou com redução de escala, atinge-se uma qualidade que ainda supera a capacidade da visão humana.

A aproximação com a realidade ocorre devido à visualização estereoscópica, inerente à visão normal binocular humana. Ela proporciona a noção de proporção e profundidade dos objetos, dando a sensação de se estar olhando o objeto real ao alcance das mãos, independentemente da distância que se está da tela de projeção.

Utilizando uma visualização com controle via teclado, o usuário pode manipular o objeto em qualquer direção pressionando as teclas definidas para isso. Esse processo é similar a pegar o objeto real em mãos e girá-lo para observá-lo por completo. Entretanto, em objetos reais, caso não se tenha alguma lente de aumento, como uma lupa, para observar mínimos detalhes, isso se torna difícil, pois se a aproximação do objeto for menor do que 15 ou 20 centímetros, perde-se o foco e há dificuldade em identificar com clareza o que está observando.

Além disso, o objeto real não contém informações acerca dele. Com a construção de objetos de aprendizagem é possível inserir informações que permitam ao usuário adquirir um conhecimento em detalhes. Por exemplo, se o objeto fosse um vaso, seria possível ter informações como período histórico da manufatura, material, forma como foi moldado, entre outras. Se fossem ossos, teriam informações de posição no corpo, função etc.

Para tornar a visualização interativa, utilizando o teclado, configura-se teclas que podem navegar em todas as direções já citadas. Por exemplo, em teclados que contém separadamente as teclas numéricas (estilo calculadora), pode-se configurar a navegação horizontal, vertical e diagonal nos números, a alteração de escala (zoom in e out) nas teclas “+” e “-” respectivamente. Além dessas, pode-se configurar a translação e rotação, mudanças de visualizações estereoscópicas, como anáglifo e polarização da luz, em outras teclas.

Em visualização com tela sensível ao toque, pode-se criar botões para todas as interações já definidas para o teclado. O mesmo pode ser feito para a interação com óculos de realidade virtual, trocando apenas as teclas ou botões para a identificação do acelerômetro e simulação do mouse, como ocorre em determinadas marcas de óculos ou aplicativos para celular. Na Figura 63 é apresentada uma possibilidade de configuração para a tela sensível ao toque.

As informações sobre o objeto ou partes dele podem ser colocadas de forma dinâmica,

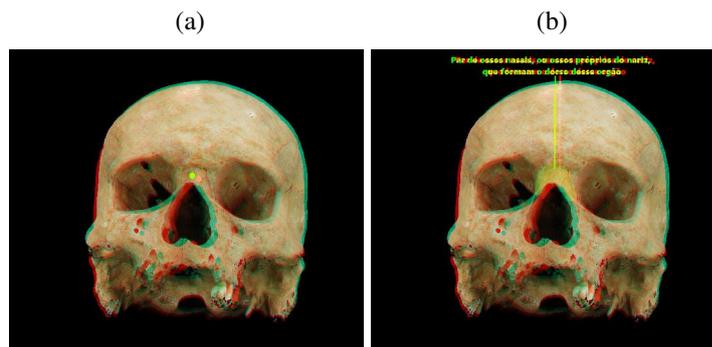
Figura 63 – Janelas do possível visualizador: (a) tela principal com botões e (b) secundária com projeção do par estéreo para polarização da luz.



Fonte: autoria própria

deixando a cargo de quem for produzir o material instrucional. Uma das principais características dos objetos de aprendizagem é a reusabilidade, portanto deixar o objeto adquirido sem informações é desejável, uma vez que o mesmo poderia ser utilizado por diversas áreas do conhecimento e até mesmo em outros ambientes instrucionais. Na Figura 64 é apresentado um exemplo de inserção de informações pertinentes ao objeto. Na Figura 64.a tem um estilo de etiqueta (tag), em amarelo, que pode aparecer para ser acionada. No caso, o objeto é um crânio e a parte selecionada foi um osso, ou par de ossos, denominado Osso Nasal. Na Figura 64.b tem as informações sobre o osso selecionado.

Figura 64 – Crânio estereoscópico com: (a) etiqueta para informações e (b) conteúdo sobre o osso selecionado



Fonte: autoria própria

Este exemplo é simplista, mas dá uma dimensão do que se pode fazer com qualquer objeto adquirido pelo F2S2. As figuras estão estereoscópicas no padrão anáglifo, podendo ser observadas com óculos em vermelho e ciano. O objetivo foi apresentar uma possibilidade de inserção de dados acerca do objeto ou parte dele e com estereoscopia também nas informações.

A interatividade não é encontrada dessa forma descrita em outras apresentações, como em filmes cinematográficos estereoscópicos. O espectador não tem contato com a execução do

filme. Neste trabalho desenvolveu-se uma forma de apresentar e permitir que usuários interajam, podendo ser utilizado por diferentes pessoas ao mesmo tempo, bastando cada computador executar o visualizador.

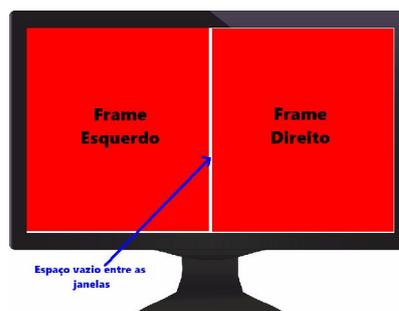
Como a visualização não depende de equipamento, o projeto citado no início da seção 3.2, página 77, a “The visible human table” também se beneficiará da visualização interativa estereoscópica. Na mesa a interatividade é por tela sensível ao toque e pode-se acoplar à visualização do corpo os objetos como ossos, adquiridos no F2S2.

Com isso, pode-se criar uma visualização a partir da seleção de partes do corpo que está na mesa, com a utilização de etiquetas para seleção de peças digitalizadas no F2S2, as quais serão projetadas com a técnica de polarização da luz, para apresentar estereoscopia.

A estereoscopia por polarização da luz utilizando um projetor DLP pode ser de diferentes configurações, ou seja, duas imagens podem ser enviadas em sequência, lado a lado ou cima e baixo. Para este trabalho, testou-se o envio lado a lado, uma vez que não se identificou como realizar o envio em sequência utilizando a linguagem C++ e SDL.

Nessa configuração, divide-se a tela do computador em duas, calculando-se o tamanho das janelas de cada imagem e deixando uma distância de aproximadamente 1 centímetro entre elas. O par estéreo será colocado nessas janelas obedecendo ao critério descrito na subseção 4.4, página 120. Na Figura 65 é apresentado como a janela do computador será dividida.

Figura 65 – Divisão das janelas para estereoscopia polarizada lado a lado



Fonte: autoria própria

Os cálculos para que isso seja genérico leva em consideração a resolução máxima do monitor onde o visualizador está executando. Na equação 4.4 é apresentado o cálculo para isso. Considera-se um espaço vazio entre as janelas de aproximadamente meio centímetro.

Essa relação de centímetros com a janela pode ser encontrada pegando a largura e altura máximas e estipulando uma resolução de pixels por centímetro. Por exemplo, em um monitor que suporta 1920x1080 pixels e tem 15 polegadas, pode ter 141,212 pixels por polegada. Isso dará aproximadamente 27.8 pixels para meio centímetro. Então, o cálculo está levando em

consideração que serão tirados 28 pixels, sendo 14 de cada janela.

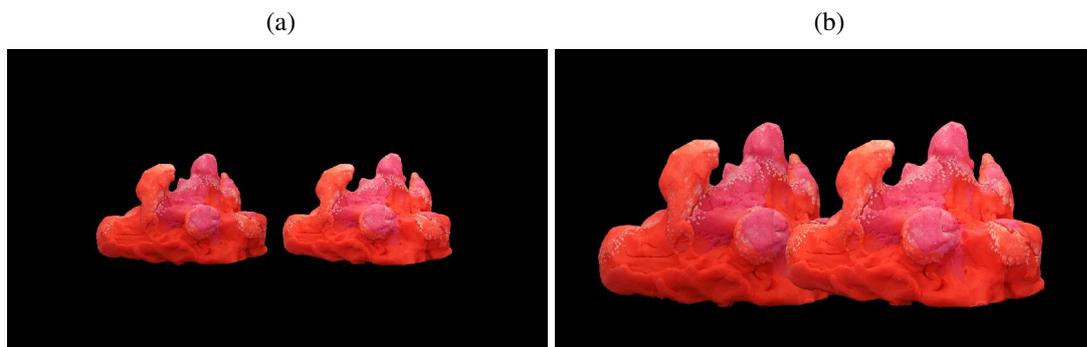
$$\begin{aligned} \text{posxJan1} &= \left(\frac{\text{LarguraMaxima} - (2 * \text{LarguraImagem})}{2} \right) - 14 \\ \text{posxJan2} &= \text{posxJan1} + \text{LarguraImagem} + 14 \\ \text{posyJan2} &= \text{posyJan1} = \left(\frac{\text{AlturaMaxima}}{2} \right) - \left(\frac{\text{AlturaImagem}}{2} \right) \end{aligned} \quad (4.4)$$

Sendo posxJan1 e posyJan1 as posições iniciais da janela da esquerda, posxJan2 e posyJan2 as posições iniciais da janela da direita. LarguraMaxima é a resolução máxima configurada para o monitor, no caso 1920 pixels. AlturaMaxima é a configuração do monitor, aqui é 1080 pixels. Os parâmetros LarguraImagem e AlturaImagem são iguais para as duas janelas, uma vez que as imagens adquiridas e após a redução de contexto, manterão o mesmo tamanho para todo o Stream2D.

Como as imagens são gravadas com fundo transparente e o fundo da visualização é preferencialmente preto, o espaço vazio não aparece, dando a impressão que as janelas estão conectadas. Não importa a resolução do monitor, o cálculo e dimensionamento das imagens ocorrerão e se ajustarão às resoluções maiores ou menores do que apresentado no exemplo.

A configuração da apresentação da estereoscopia com inserção de um espaço vazio entre as janelas permite que se faça a alteração de escala sem perda da estereoscopia, seja diminuindo ou aumentando a escala dos frames. A janela dupla corrige um problema que estava ocorrendo com a apresentação do par estéreo em escala com sobreposição. Isso atrapalhava a visualização conforme o objeto aumentava. Na Figura 66 é apresentado o problema da sobreposição quando a escala era aumentada.

Figura 66 – Sobreposição da estereoscopia polarizada com janela dupla sem divisão: (a) escala normal e (b) escala aumentada



Fonte: autoria própria

Na visualização com o projetor DLP a imagem da Figura 66.b se sobrepõe não apenas no que está sendo visto nessa figura, mas também na visualização com os óculos polarizados, onde as partes da esquerda e direita acabam extrapolando o ambiente e invadindo cada um o espaço do outro. Isso só é percebido com a visualização no projetor. Assim, a divisão da janela

não permitirá esse problema, sendo que cada objeto será “cortado” conforme atinge os limites da janela, seja na esquerda ou direita, seja na parte superior ou inferior.

Esses cortes têm o objetivo de não prejudicarem a visualização estereoscópica. O aumento da escala representa a aproximação do observador com o objeto, por exemplo pegar um algo na mão e trazê-lo para mais próximo dos olhos. No entanto, como já apresentado, quanto mais próximo dos olhos, menos foco se tem e também há a perda de estereoscopia, problema que não ocorre na visualização proposta neste trabalho e que será discutida no capítulo 5.

O modelo de projeção interativa estereoscópica é a junção das técnicas debatidas até aqui, onde a aquisição tem o papel de captura do objeto com qualidade e acuidade visual, seguida pela pós-aquisição em que se faz o tratamento das imagens.

Após esse tratamento, cria-se o Stream2D, o qual possibilita a identificação de posicionamento de cada perspectiva por meio do nome dado a cada uma das imagens. O Stream2D deve ser colocado em uma pasta específica para cada um dos objetos adquiridos, preferencialmente tendo o nome, data e hora da aquisição, pois pode ocorrer de terem sido realizadas mais de uma aquisição do mesmo objeto.

Dentro da pasta do Stream2D também é colocado um arquivo texto contendo dados como número de streams e frames, assim como o nome do objeto e o ângulo de aquisição. Esses dados vêm do arquivo XML resultante da configuração da aquisição.

O arquivo gerado com esses dados recebe uma extensão denominada “.cfg”, a qual é lida pelo software de visualização para estipular como será montada a janela deslizante e também determina como encontrar o par estéreo. No Quadro 6 é apresentado um exemplo de dados encontrados no arquivo de configuração.

Quadro 6 – Arquivo de configuração para o visualizador

Nome dos campos	Valores
Nome do objeto	Massa de Modelar
Streams	31
Frames	120
Ângulo Stream	3
Ângulo Frame	3

Como apresentado, o cálculo do encontro do par estéreo não é complexo e ao contrário das técnicas estudadas para a geração de estereoscopia, não há qualquer ajuste da geometria epipolar, assim parte-se da determinação do que foi chamada de imagem base, que foi estipulada para sempre, no início da execução, ser o frame “S1-F1001”, que de acordo com o tamanho da janela de 11x33, estaria na posição ($Stream2D_{(6,17)}$). A forma de seleção do par estéreo é apresentada no Algoritmo 2.

Algoritmo 2: Pseudocódigo da definição do par estéreo

```

1 escolha (graus)
2   caso <1 ou 1:
3     FrameEsq = (FrameAtual)->PosEsquerda->PosEsquerda->PosEsquerda;
4     FrameDir = (FrameAtual)->PosDireita->PosDireita->PosDireita;
5   fim
6   caso 2:
7     FrameEsq = (FrameAtual)->PosEsquerda;
8     FrameDir = (FrameAtual)->PosDireita->PosDireita;
9   fim
10  caso 3:
11    FrameEsq = (FrameAtual)->PosEsquerda;
12    FrameDir = (FrameAtual)->PosDireita;
13  fim
14  caso 4 ou 5 ou 6:
15    FrameEsq = (FrameAtual);
16    FrameDir = (FrameAtual)->PosDireita;
17  fim
18 fim

```

Sendo *graus* o valor lido do arquivo de configuração referente ao valor de *Frame*. A mesma proposta serve para *streams*, pois pode ocorrer de se determinar *graus* diferentes para *streams* e *frames*. No entanto, para a visualização interativa realizou-se as aquisições com ângulos iguais. *FrameEsq* é o frame que fará parte do par estéreo e está a esquerda da imagem base, assim como *FrameDir* é o frame que também faz parte do par e está a direita.

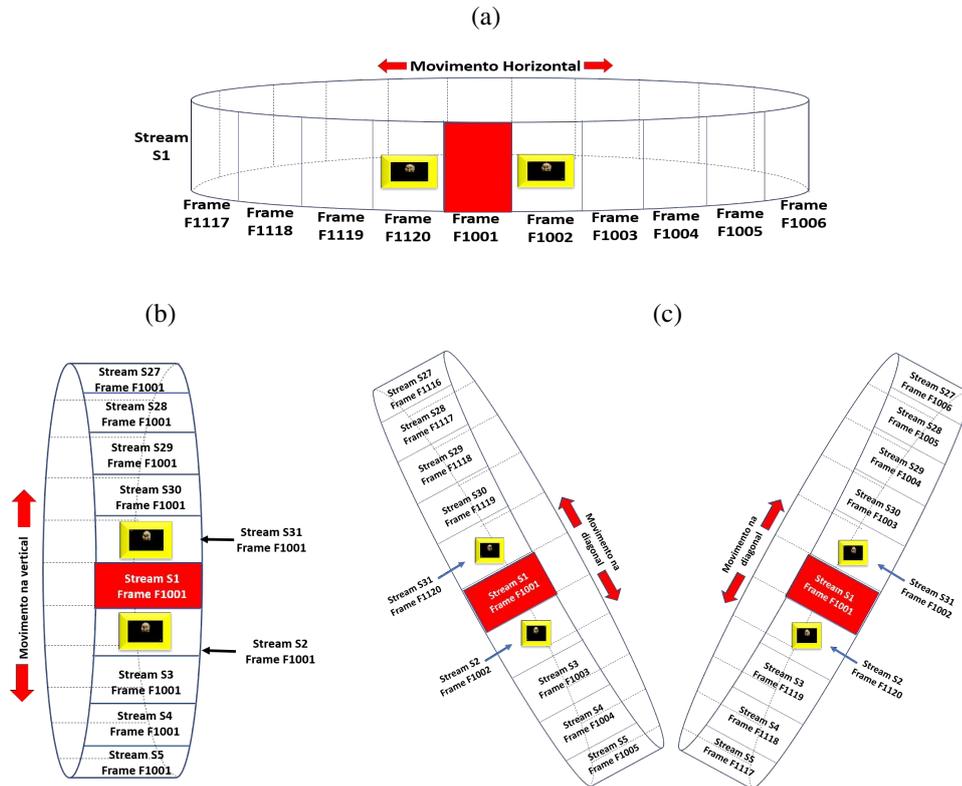
FrameAtual é a imagem base (definida como $Stream2D_{(6,17)}$), a posição anterior é definida como *PosEsquerda* ($Stream2D_{(s,\tau-1)}$) e a posição posterior é definida como *PosDireita* ($Stream2D_{(s,\tau+1)}$) são posições onde se encontram os frames que estão dentro da janela deslizante e que formam o par estéreo, portanto “(FrameAtual)→PosEsquerda” significa, por exemplo, “S1-F1001”→“S1-F1120”($Stream2D_{(s,\tau-1)}$) e “(FrameAtual)→PosDireita” significa “S1-F1001”→“S1-F1002”($Stream2D_{(s,\tau+1)}$). Com isso, tem-se que os frames da extremidade estão a 6° de distância um do outro, para a aquisição de 3°, os quais formam o melhor par estéreo para a visualização sem desconforto.

Os casos especiais são as aquisições em 4°, 5° e 6°, pois utilizam a própria imagem base para formar o par estéreo, pois se não fosse assim, a distância angular ultrapassaria o ângulo de conforto visual. Tomando o mesmo exemplo apresentado no parágrafo anterior, o par seria montado por “S1-F1001”→“S1-F1120” e “S1-F1001”→“S1-F1002”, sendo que a distância entre os frames da extremidade seria 8°, 10° e 12° respectivamente aos graus 4°, 5° e 6° para a aquisição.

Ao passo que utilizando a imagem base para formação do par, a distância angular sempre estará dentro do limite da visualização com conforto. Por exemplo, “(FrameAtual)” e

“(FrameAtual)→PosDireita” seriam os frames “S1-F1001” e “S1-F1001”→“S1-F1002” que obedecem ao critério, pois estão a 4°, 5° e 6° respectivamente aos graus 4, 5 e 6 da aquisição, além de estarem dentro da zona de conforto visual. Exemplos de movimentos possíveis são apresentados nas Figuras 67.a, 67.b e 67.c.

Figura 67 – Movimento na janela deslizante: (a) horizontal; (b) vertical e (c) diagonal



Fonte: autoria própria

Sendo (a) o deslocamento sequencial na linha i dos elementos da matriz $Stream2D_{(i,j)}$ para $Stream2D_{(i,j+1)}$ ou $Stream2D_{(i,j-1)}$. (b) É o deslocamento do sequencial dos elementos na coluna j da matriz $Stream2D_{(i,j)}$ para $Stream2D_{(i+1,j)}$ ou $Stream2D_{(i-1,j)}$. (c) É o deslocamento sequencial dos elementos na diagonal (linha e coluna: $(i-1,j+1)$, $(i+1,j+1)$, $(i+1,j-1)$ ou $(i-1,j-1)$) da $Stream2D_{(i,j)}$ para $Stream2D_{(i-1,j+1)}$, $Stream2D_{(i+1,j+1)}$, $Stream2D_{(i+1,j-1)}$ ou $Stream2D_{(i-1,j-1)}$.

Com o deslocamento e a reorganização da janela deslizante definida por meio de uma lista dinâmica totalmente conectada, além da possibilidade de encontrar o par estéreo por meio da imagem base, a criação do visualizador torna-se também genérica, pois passa a ser uma chamada de função e passagem de parâmetros, os quais são os dados do arquivo de configuração.

Neste capítulo foi apresentado o detalhamento do modelo proposto para o desenvolvimento da pesquisa. No próximo capítulo serão abordados os resultados experimentais e as discussões.

Capítulo 5

Resultados Experimentais e Discussão

5.1 Introdução

Neste capítulo serão apresentados os resultados alcançados para a geração dos objetos interativos estereoscópicos. Alguns experimentos foram empíricos, para se determinar a viabilidade do caminho adotado. A partir do Stream2D de alguns objetos adquiridos foram desenvolvidas visualizações de apenas um par estéreo, utilizando tanto a técnica de anáglifo quanto a polarização.

Para se determinar a validade das aquisições do F2S2 para a estereoscopia, foram realizadas comparações com aquisições padrões, encontradas na literatura, como as aquisições em paralelo e convergente. O intuito foi verificar se essas aquisições produziriam estereoscopia sem que houvesse necessidade de alterar o par estéreo, mudar a geometria de uma das imagens.

A partir desses testes, focou-se nas aquisições do F2S2 e em como as aquisições poderiam influenciar na visão estereoscópica, como distância dos objetos à câmera, centralização e montagem do Stream2D. Após a montagem do Stream2D, partiu-se para a parte de formação da estereoscopia com pares variando a distância angular, para se encontrar os ângulos mínimos e máximos a produzirem uma visualização sem causar desconforto.

Ao encontrar esses mínimos e máximos, passou-se a focar no programa responsável em carregar e descarregar imagens da memória, verificando se não provocam interrupções na visualização ou demora em apresentação da estereoscopia. Por fim, a produção de um visualizador teste para comprovar se a visualização seria interativa e estereoscópica. Para isso, as seções que se seguem abordam cada um desses temas isoladamente.

5.2 Aquisições em paralelo e convergente

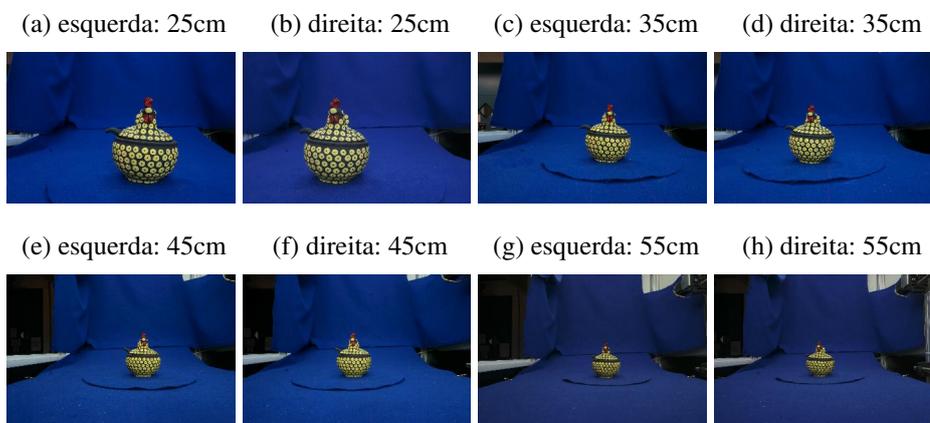
As aquisições em paralelo e convergente foram realizadas para se verificar a estereoscopia por meio das técnicas de anáglifo e polarização. Como são formas de aquisições utilizadas, principalmente na cartografia e interpretação aerofotométrica, em que a estereoscopia ou reconstrução tridimensional é utilizada, viu-se que poderia ser uma forma de aquisição que colocaria a viabilidade do F2S2, para algumas aquisições, em cheque.

Nos últimos anos a fotogrametria aérea, por meio de satélites em órbita, alterou subs-

tancialmente técnicas como a cartografia e a interpretação aerofotométrica. Porém, os satélites não são únicos nessa captura, hoje também são utilizados drones, além de aviões. Contudo, a aquisição realizada por esses mecanismos é em paralelo, havendo um tratamento das imagens para ajustes, sobreposição de partes que estão presentes nas fotos adjacentes, além de descartes de partes que não produzem estereoscopia.

Para os testes foram adquiridas algumas imagens com distâncias variando de 25cm a 55cm, como apresentado na Figura 68. As aquisições seguiram os critérios definidos na seção 4.2.

Figura 68 – Aquisição do par estéreo paralelo



Fonte: autoria própria

Com isso, foi possível avaliar um dos intuitos deste trabalho, que é a distância entre os objetos e a câmera e sua influência na geração da estereoscopia. Essa avaliação também, apesar de, em alguns casos, ser com imagens não adquiridas pelo F2S2, é parte do objetivo específico 1. Pois, analisou-se as imagens paralelas e convergentes comparando-as com as aquisições do F2S2.

Essa análise não diz respeito à formação do melhor par estéreo e nem à questão das diferenças entre as imagens direita e esquerda não produzem qualquer tipo de estereoscopia, o que na realidade não é falso. A estereoscopia é formada em aquisições paralelas e convergentes, o que se percebeu foi a necessidade de alguns ajustes para que não causasse desconforto na visualização.

Além disso, existem duas diferenças ligadas ao conteúdo dentro da imagem, onde a primeira é a imagem completa, com o fundo azul e a segunda é a imagem segmentada, onde o fundo é transparente ou preto.

No caso em que a imagem não foi segmentada, mantendo-se o conteúdo intacto, a estereoscopia criada foi de paralaxe positiva. Contudo, ajustes foram necessários para que as imagens produzissem a visão estereoscópica.

Não houve um cálculo implementado para avaliar esses ajustes, utilizou-se o software, citado na seção 3.2, StereoPhoto Maker, para a geração da estereoscopia e avaliação dos ajustes. A explicação do seu funcionamento pode ser encontrada no site do programa na parte *easy adjustment*¹.

Assim como ocorreram ajustes com a imagem completa, ocorreu também com a imagem segmentada, no entanto a estereoscopia apresentada foi de paralaxe negativa, com um pouco de projeção do objeto para fora do plano de visualização (percepção subjetiva dependente do observador). Na Figura 69 é apresentada a abertura do par estéreo e a transformação da imagem esquerda para o canal vermelho e a direita para o ciano e aplicado o processo de geração do anáglifo.

Figura 69 – Preparação do par estéreo com distância de 25cm entre câmera e objeto - StereoPhoto Maker



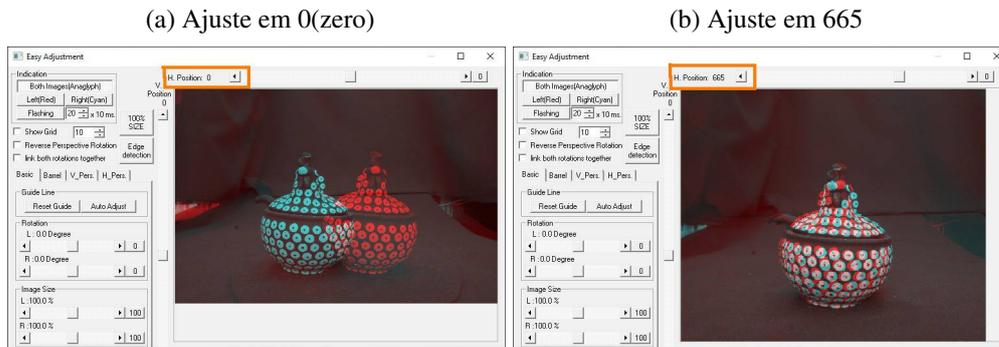
Fonte: autoria própria

Pela situação apresentada na Figura 69.b, onde a imagem esquerda e direita estão invertidas devido ao processo de geração do anáglifo, demonstra-se a necessidade do ajuste entre as duas. Além disso, elas também mostram o mesmo resultado conseguido com o cálculo de diferença, apresentado na Figura 41, página 113.

¹ <https://stereo.jp/eng/stphmkr/help/adjust_01.htm>

O ajuste para produzir a estereoscopia foi horizontal, aplicado à uma das imagens. Não houve necessidade de um ajuste de rotação ou outro, apenas o deslocamento de uma das imagens por uma quantidade de pixels. O processo e a quantidade de deslocamento aplicado à imagem pode ser visto na Figura 70.

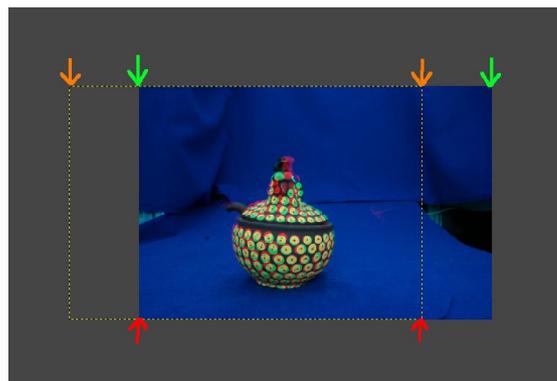
Figura 70 – Ajuste horizontal aplicado ao par estéreo



Fonte: autoria própria

Um resultado similar foi alcançado utilizando o software Gimp, contudo, nele foi necessário ajustar as imagens e realizar o corte do excesso que não provê estereoscopia. O excesso no StereoPhoto Maker foi automático, ou seja, assim que a imagem foi sendo colocada em ajuste, o programa ia cortando a parte que não interessava para a visualização. Na Figura 71 apresenta-se a imagem deslocada e os pontos onde deve ocorrer o corte.

Figura 71 – Ajuste horizontal aplicado ao par estéreo no Gimp



Fonte: autoria própria

As setas em laranja e verde indicam o deslocamento das imagens. As setas em vermelho indicam onde ocorreu o corte para poder manter somente a imagem que apresenta estereoscopia. O deslocamento em relação às imagens convergentes é maior, isso pode ser visto na apresentação da diferença entre imagens convergentes, na seção 4.2.

O resultado do anáglifo para esse par estéreo, com aquisição paralela, foi similar para os

dois softwares utilizados na geração da imagem estereoscópica. As duas imagens são apresentadas na Figura 72

Figura 72 – Imagens estereoscópicas

(a) StereoPhoto Maker

(b) Gimp



Fonte: autoria própria

As imagens das aquisições convergentes também necessitaram de ajustes. Contudo, existem algumas considerações a serem levantadas. Uma delas é a questão de trabalhar com as imagens contendo fundo e sem fundo.

Na Figura 73 mostra-se a aquisição para as imagens convergentes e assim como ocorreu com a aquisição em paralelo, foram realizadas 8 aquisições com variação de distância entre a câmera e o objeto.

Figura 73 – Aquisição do par estéreo convergente

(a) esquerda: 25cm

(b) direita: 25cm

(c) esquerda: 35cm

(d) direita: 35cm

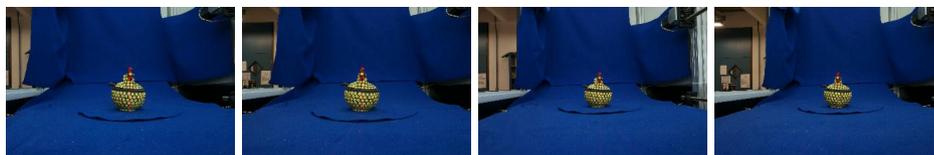


(e) esquerda: 45cm

(f) direita: 45cm

(g) esquerda: 55cm

(h) direita: 55cm



Fonte: autoria própria

As aquisição em paralelo e convergente foram segmentadas para garantir uma comparação coerente, pois o Stream2D é segmentado, portanto todas as imagens contêm apenas o objeto de interesse e o fundo transparente, o que alterou alguns resultados experimentais.

Então, a partir desse ponto são apresentados os resultados das análises de diferença com essas imagens segmentadas. Na Figura 74 é apresentada a diferença sobre o par estéreo da Figura 68.a e 68.b .

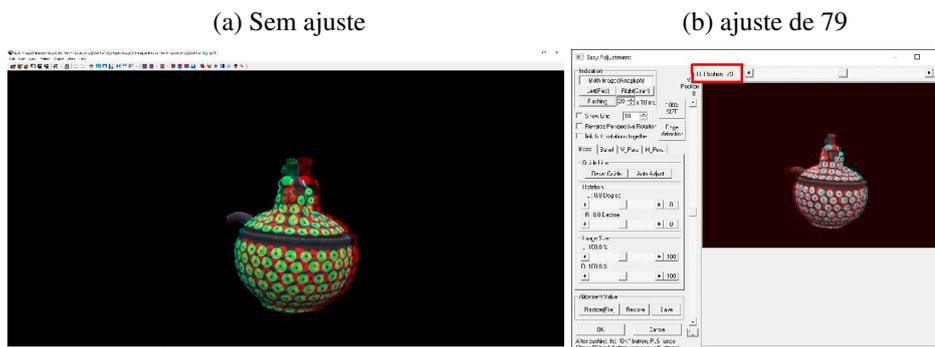
Figura 74 – Par estéreo paralelo segmentado - diferença



Fonte: autoria própria

As imagens convergentes também foram manipuladas nos softwares de edição e geração da estereoscopia, onde o par de distância similar ao paralelo (25cm entre objeto e câmera) foi aberto e verificado. O ajuste realizado foi menor do que aquele aplicado ao par paralelo. Isso pode ser visto na Figura 75.

Figura 75 – Par estéreo convergente - StereoPhoto Maker



Fonte: autoria própria

Após esse ajuste, a imagem anáglifo resultante apresentou uma estereoscopia com paralaxe negativa, no entanto parte da imagem ficou com uma borda vermelha e o que se projetou para fora do plano de visualização foi o bico da galinha. A parte de baixo, pote/barriga, estava com uma percepção de paralaxe positiva. Essa visualização pode variar de observador, mas no geral será vista dessa forma.

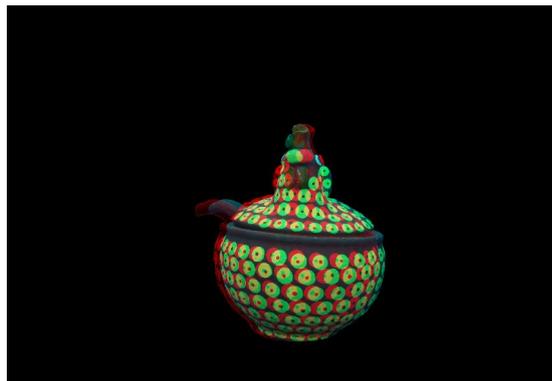
Mesmo com o ajuste aplicado, as estereoscopias, com a técnica anáglifo, deixam uma sensação desconfortável, pois pontos avermelhados são perceptíveis e que provocam a perda de estereoscopia, pois tenta-se focar a visão onde há noção de proporção e profundidade, mas as partes que não estão bem formadas acabam atrapalhando.

Ajustes com mais critérios do que apenas deslocar horizontalmente deveriam ser aplicados, contudo não é o objetivo deste trabalho, aqui cabe a verificação de geração de estereoscopia com paralaxe negativa sem a necessidade de alteração do par estéreo. O que até o presente

momento no estudo, não se verificou.

A visualização pela técnica de polarização minimiza o problema do desconforto pelas cores avermelhadas, entretanto, houve a necessidade de ajustar as imagens dentro do ambiente de visualização, para que elas fossem sobrepostas e apresentassem a estereoscopia. Na Figura 76 apresenta-se a imagem resultante do ajuste do par estéreo convergente e a geração de estereoscopia.

Figura 76 – Imagem anáglifo resultante do ajuste do par convergente



Fonte: autoria própria

Outra análise foi realizada utilizando o cálculo da diferença entre as imagens e também a comparação com as aquisições do F2S2 nas mesmas distâncias, para alguns ângulos. Assim, posicionou-se a câmera na mesma altura das aquisições paralela e convergente, programou-se o giro da base (B) para 3°, 6° e 9°. A aquisição iniciou com a primeira imagem em 0°, depois nos três outros ângulos citados.

O ângulo de 9° apresentou desconforto, necessitando de ajuste, assim como as outras duas técnicas utilizadas para aquisição. Por conseguinte, a análise foi realizada com os ângulos que não causaram problemas durante a visualização. No Quadro 7 apresenta-se o resultado da diferença entre os canais de cores para as imagens paralelas, convergentes, F2S2 em 3°, 6° e 9° para as imagens segmentadas. Cada uma das imagens foi adquirida com largura e altura igual, 3264x2176 pixels.

Quadro 7 – Diferença entre os canais

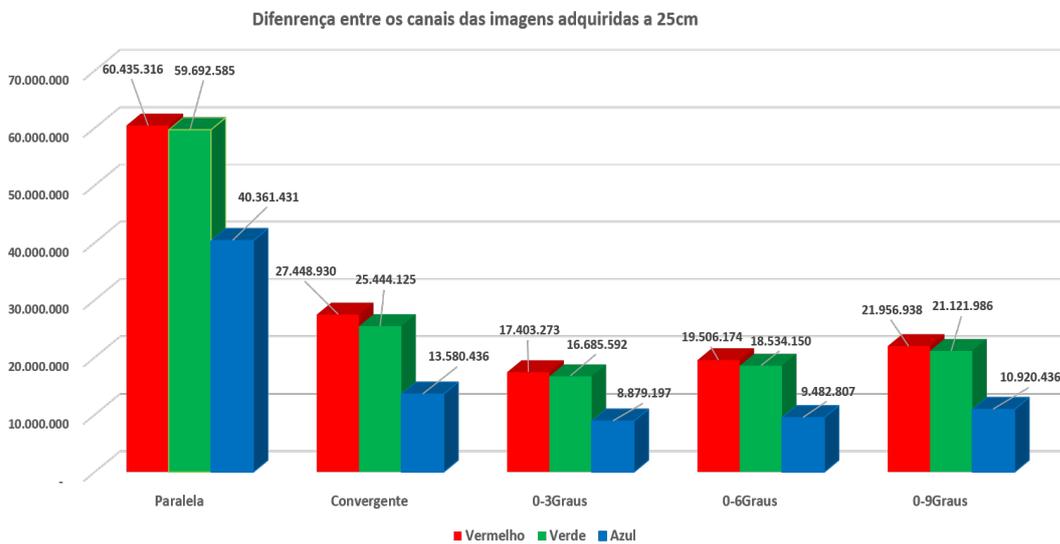
Total de Pixels			7.102.464		
Máxima diferença entre canais			7.102.464*256	1.818.230.784	1.818.230.784
Canais	Paralela	Convergente	0-3Graus	0-6Graus	0-9Graus
Vermelho	60.435.316	27.448.930	17.403.273	19.506.174	21.956.938
Verde	59.692.585	25.444.125	16.685.592	18.534.150	21.121.986
Azul	40.361.431	13.580.436	8.879.197	9.482.807	10.920.436

Percebe-se uma desigualdade entre as aquisições paralelas, convergentes e F2S2, mesmo em imagens segmentadas. Enfatiza-se que são segmentadas, pois em imagens com o fundo, a

diferença é consideravelmente maior, devido ao que tem em uma imagem e não tem em outra. Por exemplo, na aquisição paralela, como mostrado na Figura 71, o excedente de cada lado, que pode ser aproximadamente 35% somando os dois cortes, são pixels diferentes entre o par estéreo, o que aumentaria o resultado da diferença.

Na convergente, esse excedente é menor, chegando a aproximadamente 10% do total da imagem. Essa porcentagem também influencia no resultado final. Por esses motivos, optou-se por realizar os cálculos com as imagens segmentadas, pois o fundo se torna igual para todas as aquisições, deixando apenas o objeto de interesse. Na Figura 77 mostra-se o gráfico relacionado ao Quadro 7.

Figura 77 – Diferença entre os canais em valores de quantidade de pixels



Fonte: autoria própria

A partir desses dados foram calculadas as probabilidades de cada canal em relação à maior diferença encontrada para duas imagens, ou seja, quanto o pixel(x,y) da imagem esquerda é diferente do pixel(x,y) da imagem direita. Esse valor é acumulado para toda a imagem, obtendo um total da diferença de cada canal. Isso pode ser visto na Equação 5.1

$$\begin{aligned}
 totalDifR &= totalDifR + \|PIimageE(Lin, Col, 1) - PIimageD(Lin, Col, 1)\| \\
 totalDifG &= totalDifG + \|PIimageE(Lin, Col, 2) - PIimageD(Lin, Col, 2)\| \\
 totalDifB &= totalDifB + \|PIimageE(Lin, Col, 3) - PIimageD(Lin, Col, 3)\| \\
 probTotal &= LinImg * ColImg * 256 \\
 probDifR &= totalDifR/probTotal \\
 probDifG &= totalDifG/probTotal \\
 probDifB &= totalDifB/probTotal
 \end{aligned} \tag{5.1}$$

Sendo, totalDifR, totalDifG e totalDifB o acúmulo da diferença entre os pixels vermelhos, verdes e azuis da imagem esquerda (PImageE) e da direita (PImageD) nas posições Linha (Lin) e Coluna (Col) correspondentes entre elas. Os parâmetros 1, 2 e 3 são referentes à leitura na matriz, onde linha e coluna são as posições e eles são, respectivamente, os canais vermelho, verde e azul.

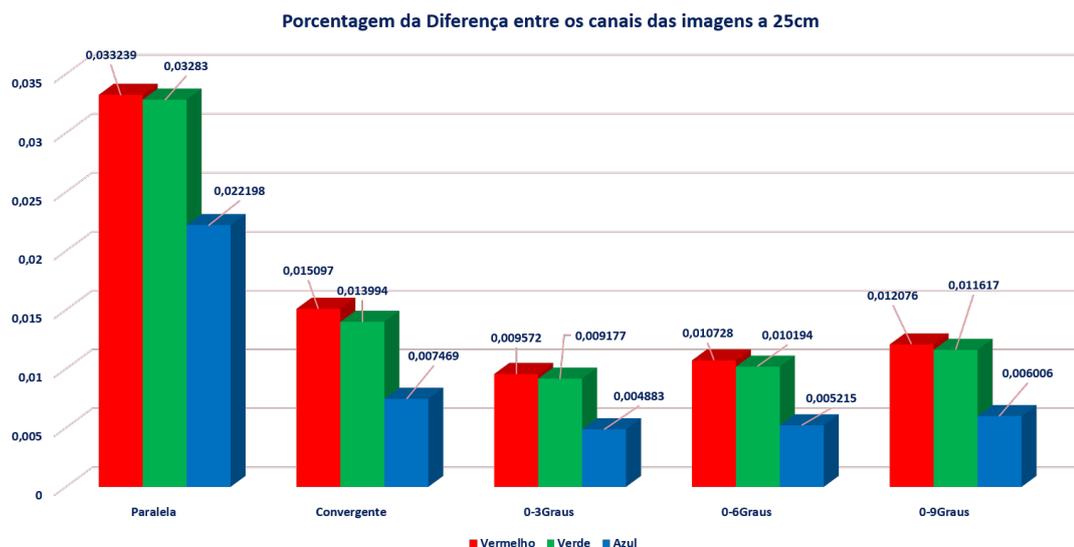
O valor de probTotal (probabilidade de todos os pixels serem diferentes) é encontrado pela multiplicação das linhas e colunas da imagem, sendo que as duas têm tamanhos iguais, multiplicado por 256 (valores possíveis para cada canal). Assim, encontra-se as probabilidades de cada canal (probDifR, probDifG, probDifB) dividindo o valor da diferença total pela probabilidade total. No Quadro 8 e na Figura 78 apresentam-se as probabilidades encontradas.

Quadro 8 – Probabilidades diferença entre os canais

Canais	Paralela	Convergente	0-3Graus	0-6Graus	0-9Graus
Vermelho	0,033239	0,015097	0,009572	0,010728	0,012076
Verde	0,03283	0,013994	0,009177	0,010194	0,011617
Azul	0,022198	0,007469	0,004883	0,005215	0,006006

Ao analisar os dados do Quadro 8 percebe-se que, entre as imagens adquiridas pelo F2S2, há uma diferença menor nas imagens que distam 3° entre elas. A explicação para isso está no fato de que uma imagem que girou 3° tem mais pontos em comum do que uma imagem que girou 6°, contudo essa diferença conta a favor desta imagem e não daquela. Como será apresentado neste capítulo, a estereoscopia tem níveis de projeção em relação ao plano e isso está ligado à distância angular.

Figura 78 – Probabilidade da diferença entre os canais



Fonte: autoria própria

Como o capítulo se destina a apresentar os resultados alcançados com o F2S2, nesta seção serão colocadas duas imagens que foram utilizadas para a comparação das aquisições

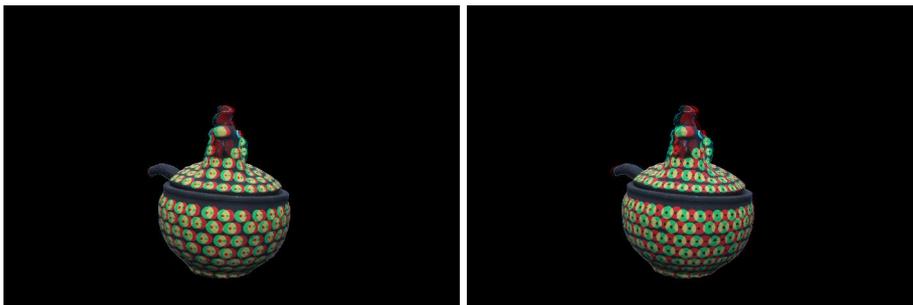
paralela, convergente e do F2S2 e em outras seções serão utilizados outros objetos adquiridos. Essa decisão foi baseada na ideia e que aqui deve-se utilizar o mesmo objeto que foi adquirido nas outras técnicas.

Aqui é apresentada, para comparação com as outras técnicas, Figura 79, as duas imagens anáglifo geradas a partir dos ângulos presentes no cálculo de diferença.

Figura 79 – Anáglifo do par estéreo gerado pelo F2S2

(a) distância angular: 3°

(b) distância angular: 6°



Fonte: autoria própria

5.3 Distâncias nas aquisições de objetos e formação da estereoscopia

Como visto no Seção 2.3.1, a estereoscopia era produzida com imagens duplicadas ou com as aquisições paralelas, com a câmera sendo posicionada a poucos centímetros de diferença nas aquisições. Em alguns casos, as câmeras tinham duas lentes posicionadas na distância interpupilar, para que a aquisição fosse similar à visão binocular. O primeiro filme a utilizar duas câmeras e dois projetores foi “*The Power of Love*”. As duas câmeras foram colocadas no mesmo equipamento, como se fosse uma filmadora de lente dupla.

Atualmente, pode-se ter apenas um filme, que é projetado por dois projetores ou um com sistema 3D, no caso do primeiro, há uma emissão de muitos quadros por segundo, intercalados nos projetores em uma tela. Com o uso de óculos polarizado passivo, o cérebro é enganado para ter a noção de profundidade. Em outros casos, o projetor faz a sincronização com os óculos e também produz a sensação de profundidade.

Em nenhum dos casos se discute a distância de captura das imagens, pois normalmente não se tem captura a curta distância, mas uma aproximação de escala (zoom). Como também visto na Seção 2.3.1, o problema da distância tem haver com a acomodação visual ou “descanso nos olhos”. Como o objetivo do F2S2 é digitalizar objetos de toda natureza, existe o problema nas aquisições de insetos ou animais pequenos. Por exemplo, um grilo, onde, para apresentar todos os detalhes, pode ser adquirido com uma lente macro, para uma aproximação.

Uma lente macro, dependendo do tipo, pode realizar uma aquisição a uma distância aproximada de 1cm entre a lente e o objeto, portanto identificar se há estereoscopia, para esses objetos, é uma questão de se ter ou não conteúdos para o desenvolvimento de OA, para a entomologia/biologia. Na Figura 80 mostra-se duas imagens adquiridas no F2S2 de um grilo, uma com lente macro e outra com a lente de 16-50mm a 25cm do objeto.

Figura 80 – Imagens do Grilo

(a) 25cm sem lente macro

(b) com lente macro



Fonte: (a) autoria própria e (b) Diogo Roberto Olsen

O resultado visual não é diferente, pois os detalhes podem ser percebidos nas duas imagens, por exemplo, com a imagem em tamanho original consegue-se visualizar os pelos nas patas, os olhos, entre outros. Um problema para aquisições tão próximas é a falta de profundidade de campo, pois ela permite que se tenha foco em todo o objeto.

Uma lente macro não tem tanta profundidade, com isso um objeto longo pode ter desfoque na parte que fique mais distante da lente. Um exemplo é o próprio grilo, que quando está com a cabeça de frente para a câmera, suas patas traseiras ficam desfocadas. Já com uma lente que tenha profundidade de campo, isso pode não ocorrer. Na Figura 81 são apresentadas duas imagens do grilo, onde pode-se perceber como a pata traseira na aquisição com lente macro ficou um pouco mais desfocada do que a aquisição sem macro.

Figura 81 – Profundidade de campo em aquisições

(a) 25cm sem lente macro

(b) com lente macro



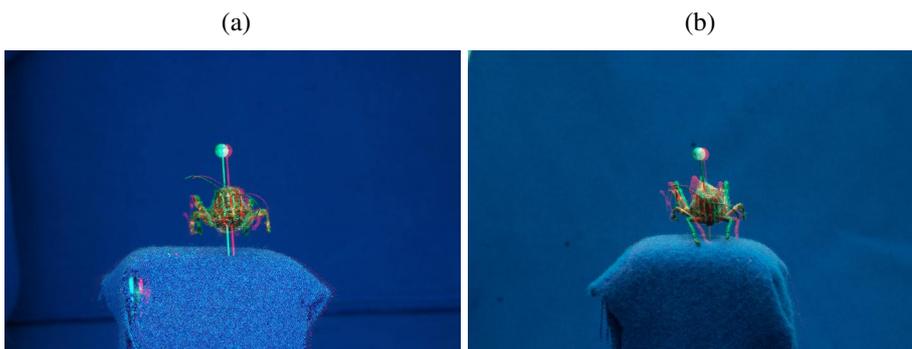
Fonte: (a) autoria própria e (b) Diogo Roberto Olsen

A lente macro traz uma maior nitidez na parte focada do objeto, mas perde o foco na parte contrária. Já com a lente sem macro e com o foco calibrado para o objeto todo, a nitidez está mais na parte central. Essa questão de como realizar a melhor aquisição para todo o tipo de objeto é parte do trabalho do doutorando Diogo Roberto Olsen. Para este trabalho a análise foi a formação da estereoscopia, por isso não houve tanto critério na aquisição.

Para a estereoscopia, a distância não influenciou e as aquisições geraram a noção de profundidade. Contudo, na imagem com macro, a parte da cabeça, onde está o foco, apresenta estereoscopia, já a parte de trás que está desfocada não está bem definida. Na imagem sem macro, a estereoscopia é percebida das patas dianteiras até o final das asas.

Essas estereoscopias em anáglifo têm um resultado diferente da polarizada, pois a alteração das cores para vermelho e ciano acabam prejudicando a visualização. Como na polarização não há modificação das cores, fica mais nítida a imagem. Na Figura 82 apresenta-se as duas estereoscopia.

Figura 82 – Estereoscopia anáglifo: (a) imagem macro e (b) sem macro



Fonte: autoria própria

O mesmo teste de estereoscopia foi realizado com anáglifo e na polarização para os objetos adquiridos em 25cm, 35cm, 45cm, e 55cm. Em todos os casos a estereoscopia foi formada, descartando-se a hipótese de que a aquisição a menos de 25cm poderia não formar a visualização estereoscópica, por entrar na faixa de descanso do olhos. Nas outras distâncias a visualização foi percebida, isso leva à conclusão de que o F2S2 pode produzir materiais didáticos em forma de OA para diversas áreas do conhecimento, inclusive a entomologia, mesmo com a aquisição a menos de 25cm entre a câmera e o objeto. Isso atende ao objetivo específico 1, pois o intuito foi verificar se as distâncias influenciavam na estereoscopia.

5.4 Centralização do objeto para estereoscopia

As aquisições em que o objeto foi centralizado tendo por base seu possível centro ou centro de massa, acabam provocando uma diferença, principalmente entre frames de streams diferentes. Isso também altera a visualização estereoscópica, mesmo que a distância angular

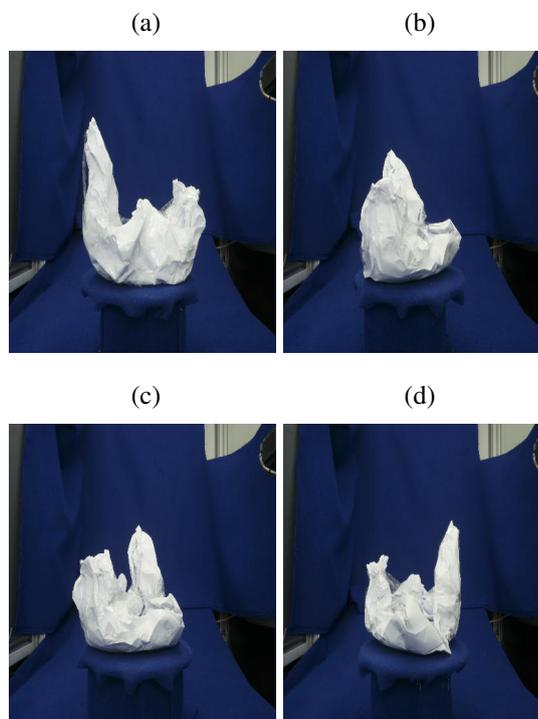
seja obedecida, não há um encaixe dos frames e a estereoscopia acaba provocando desconforto visual.

O procedimento era rudimentar, consistindo em posicionar o objeto sobre o eixo da base giratória, contudo alguns objetos são difíceis de se encontrar o centro, como no caso da concha do mar. Ela tem protuberâncias em praticamente todas as partes, o que dificulta saber exatamente qual é seu centro. Mesmo sendo totalmente irregular, poder-se-ia localizar o centro de massa para centralizá-la. Contudo, isso dependeria de tentativas e erros.

Outros objetos não seriam tão simples assim, por exemplo, realizar a aquisição de um martelo com cabo de madeira, deitado sobre a base, onde o maior peso provavelmente seja a parte de ferro. Seu centro de massa estaria na chamada “cabeça do martelo”, então, em uma aquisição ele giraria sobre o eixo dela. Essa aquisição não seria boa, pois o foco também estaria na cabeça do martelo e quando o cabo estivesse virado para a câmera, ele ficaria desfocado.

Uma das formas de fazer a aquisição do martelo seria tirando as medidas do início do cabo até a cabeça, pegando a média. Sendo o martelo não tão irregular, a aquisição poderia ser realizada por essa medida. No entanto, existem objetos que esse procedimento não resolve e um deles foi uma folha de papel, a qual foi amassada, dobrada e enrolada com uma fita adesiva para se manter na forma que foi moldada. Na Figura 83 apresenta-se o papel.

Figura 83 – Centralização do objeto papel

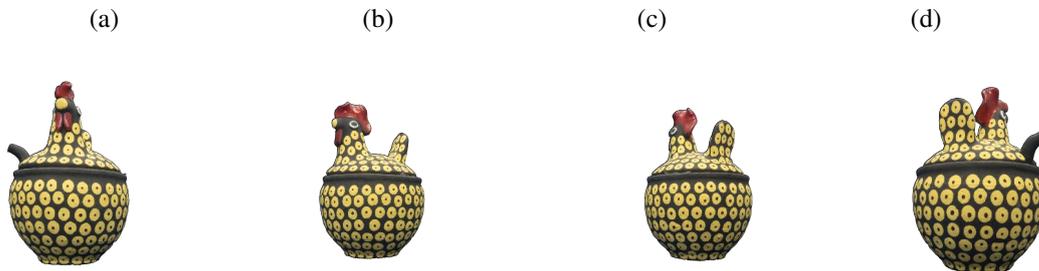


Fonte: autoria própria

Não houve como colocá-lo centralizado pelo centro de massa e nem pelo seu centro devido sua irregularidade. Assim, utilizou-se os lasers para os ajustes de posicionamento. Na

Figura 84 é apresentada uma aquisição da galinha de barro sem centralização.

Figura 84 – Aquisição sem centralização do objeto

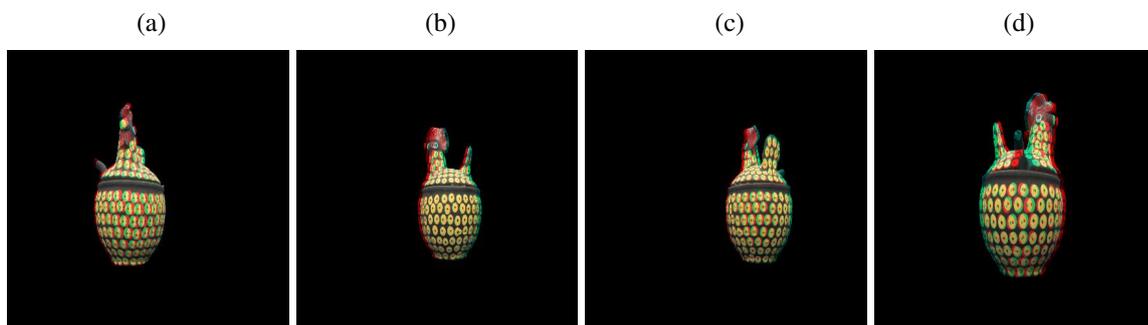


Fonte: autoria própria

O posicionamento do objeto nas Figuras 84.a, 84.b, 84.c e 84.d deixa claro como o movimento circular interfere, pois nas figuras (a) e (d) a imagem está maior, representando uma proximidade com o observador, já nas outras duas, (b) e (c), a impressão é que estão mais distantes. Isso é verdadeiro, pois o movimento fez com que o objeto ficasse mais próximo ou mais distante da câmera.

Em relação à estereoscopia, não se alterou, pois como a distância angular entre um frame e outro permaneceu a mesma, o par estéreo foi selecionado conforme o ângulo estipulado na aquisição. Contudo, a mesma sensação de distanciamento do objeto ocorre durante a visualização, com uma constatação de que a paralaxe se alterna entre a positiva e a negativa. Um efeito interessante se o objetivo for criar um movimento no estilo carrossel para os objetos. Esse efeito pode ser visto na Figura 85, onde estão quatro imagens em anáglifo.

Figura 85 – Estereoscopia sem centralização do objeto



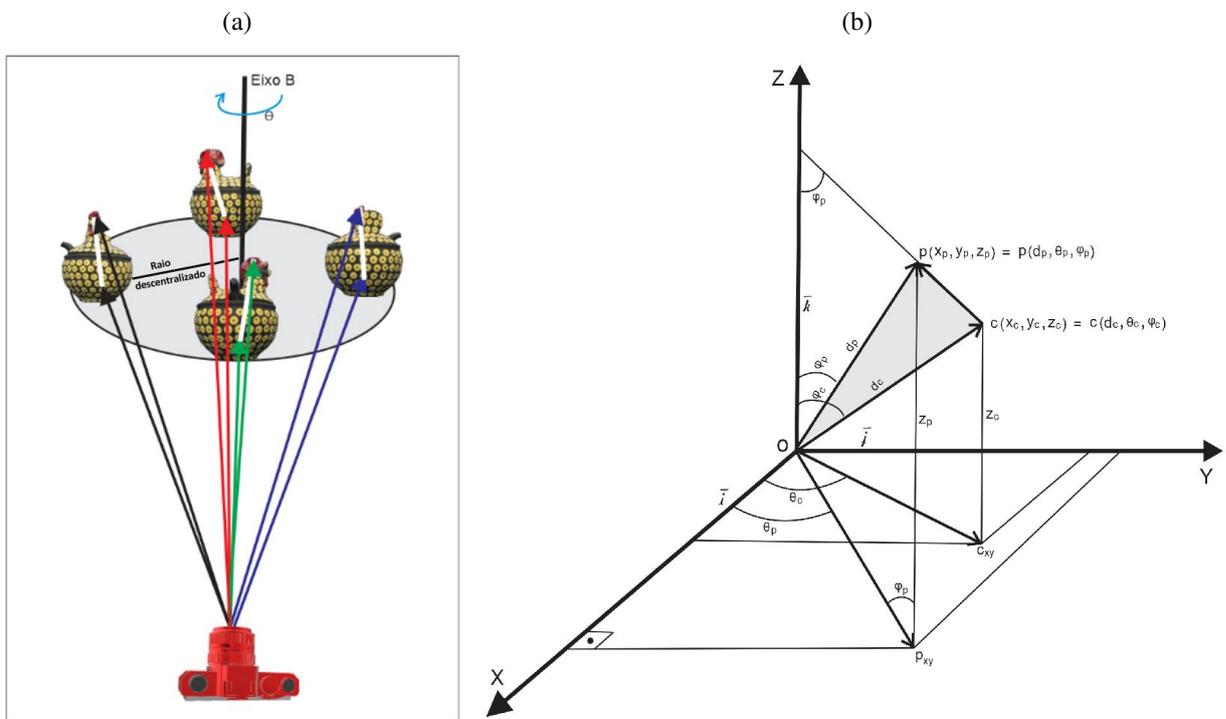
Fonte: autoria própria

A centralização visa proporcionar uma observação como se o usuário estivesse manipulando em suas mãos o objeto real. Isso não ocorre em casos onde o objeto está descentralizado, pois o movimento é circular ou elíptico, como se o observador pegasse algo e o aproximasse e afastasse com giros do braço e ao mesmo tempo mudasse a face do objeto.

Esse movimento está ligado diretamente ao raio definido na aquisição, portanto, ao não centralizar o objeto, altera-se o valor de “r” para cada ângulo de aquisição. Esse movimento pode ser analisado como planetário, porém não girando em torno da câmera, mas com ela posicionada a uma distância do centro do eixo “B”.

O objeto, em relação ao centro do eixo “B” não altera sua distância, ela é fixa. Com isso, pode-se calcular o giro circular para uma aquisição e obter a estereoscopia deste objeto, no entanto a distância entre frames e streams deve ficar dentro da distância angular que não causa desconforto. Na Figura 86 apresenta-se os dados para se encontrar a fórmula que define essa configuração de movimentação circular.

Figura 86 – Definição dos eixos e variáveis para o movimento circular/elíptico



Fonte: autoria própria

Na Figura 86.a está representado o movimento com quatro posições. A câmera tem sua posição definida na aquisição, assim cada uma das posições tem duas setas que indicam o foco da câmera, que é na metade da altura do objeto e outra que está em uma ponto que pode ser encontrado por medição. Esse ponto fica acima do ponto de foco e será utilizado para determinar a distância do objeto à câmera para cada ângulo θ .

Os segmentos em branco, na Figura 86.a, indicam a distância entre o ponto de foco e o ponto que será utilizado para encontrar a distância entre a câmera e qualquer ponto do objeto, mesmo com o giro circular. Na Equação 5.2 apresenta-se a fórmula para posicionamento da câmera.

$$\begin{aligned}
\text{C\~{a}mera} &= C(dc, \theta_c, \varphi_c) \\
x_c &= dc * \text{sen}\varphi_c * \cos\theta_c \\
y_c &= dc * \text{sen}\varphi_c * \text{sen}\theta_c \\
z_c &= dc * \cos\varphi_c
\end{aligned} \tag{5.2}$$

A Equação 5.3 diz respeito à posição de um ponto arbitrário do objeto, representado na Figura 86.c.

$$\begin{aligned}
\text{Ponto} &= P(dp, \theta_p, \varphi_p) \\
x_p &= dp * \text{sen}\varphi_p * \cos\theta_p \\
y_p &= dp * \text{sen}\varphi_p * \text{sen}\theta_p \\
z_p &= dp * \cos\varphi_p
\end{aligned} \tag{5.3}$$

Assim, encontra-se a distância de P à câmera pela definição na Equação 5.4.

$$d^2(p, c) = (x_p - x_c)^2 + (y_p - y_c)^2 + (z_p - z_c)^2 \tag{5.4}$$

Para a primeira Equação (Eq. 5.2), tem-se “C” como a posição da câmera no deslocamento angular φ , “dc” é a distância da câmera ao ponto determinado no objeto. O φ , assim como os eixos (x, y e z) foram definidos na Seção 3.3.2. Na segunda equação (Eq. 5.3), “P” o ponto arbitrário no plano esférico, “dp” é a distância até um ponto P, “zp” (sendo a distância entre o eixo Pxy e o ponto, como catetos. Na terceira equação (Eq. 5.4), “ $d^2(p, c)$ ” representa a distância calculada entre a câmera e o ponto “P”.

Ainda sobre a Figura 86.a, para se descobrir o deslocamento linear do raio entre a câmera e o objeto, que representa o giro circular, utiliza-se a Equação 5.5. Ela apresenta a fórmula para o deslocamento no Eixo “X”, para qualquer raio de descentralização do objeto.

$$\begin{aligned}
&\text{C\~{a}lculo para o Eixo X} \\
\text{RaioVerdadeiro} &= \text{RaioAquisicao} - (\cos\theta * \text{RaioDescentralizado})
\end{aligned} \tag{5.5}$$

Sendo “RaioVerdadeiro” o raio entre a câmera e o objeto durante o deslocamento descentralizado. O “RaioAquisicao” é determinado durante a configuração da aquisição do F2S2, ele é igual ao “r” e o ângulo θ foram definidos na Seção 3.3.1. O “RaioDescentralizado” é o deslocamento proposital do objeto em relação ao centro do Eixo “B”.

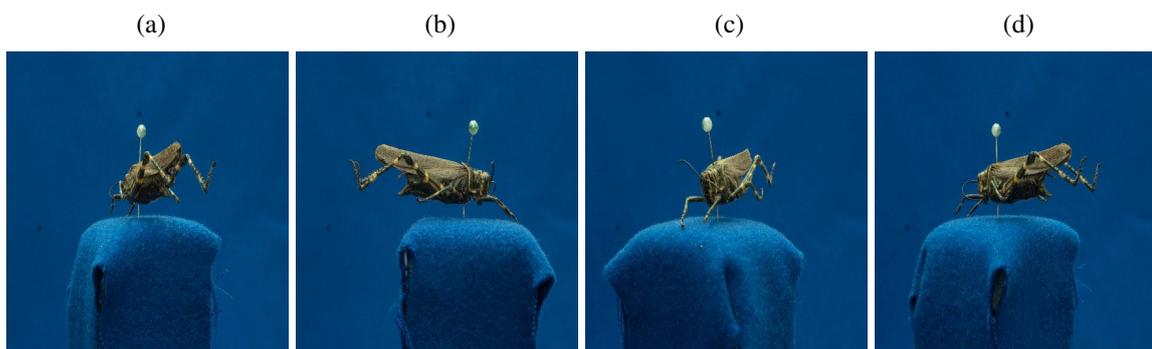
Quando o objeto está centralizado, pelo método utilizado neste trabalho, a distância câmera para objeto não se altera, portanto o objeto girará sobre seu próprio eixo, independentemente se ele estiver exatamente sobre o eixo “B”. No geral, quando se centraliza pelos lasers

e o objeto é irregular, pode ocorrer um deslocamento em relação ao eixo “B”, contudo é algo imperceptível, pois serão milímetros, mas o objeto girará sobre seu eixo.

Um exemplo de aquisição onde o objeto está centralizado, mas a base onde foi colocado e que está sobre o eixo “B”, foi a do grilo. A centralização foi realizada a partir dos limites das suas patas dianteiras e traseiras. O laser para posicionamento da parte de trás, para centralizar na distância do raio configurado para a câmera, foi colocado para tocar o limite da pata traseira e ao girar o grilo em 180° e tocar a dianteira também no limite.

Os lasers laterais foram posicionados para tocarem as patas no sentido da frente para trás e vice-versa quando girado em 180°. Assim, ao visualizar o grilo, percebe-se que ele gira em torno do seu eixo, o que aparentemente é o centro do seu corpo. Já a base onde está posicionado, assim como a agulha que o prende na base, gira elípticamente. Na Figura 87 apresenta-se quatro posições do grilo, onde é possível perceber sua centralização e a descentralização de sua base.

Figura 87 – Grilo com centralização



Fonte: autoria própria

Com isso, conclui-se que a centralização do objeto está ligada ao objetivo que se tem, no entanto colocá-lo sobre seu centro trás uma visualização mais próxima do “natural”. Assume-se aqui que um giro “natural” é aquele em que o objeto é girado sobre o seu eixo a uma distância qualquer, podendo ser afastado ou aproximado, em uma linha reta frontal aos olhos.

Portanto, um giro “não natural” é aquele onde a pessoa gira o objeto em sua mão e ao mesmo tempo distancia-o e aproxima-o de seus olhos, fazendo um movimento circular com os braços, causando um giro circular ou elíptico. A centralização trás o benefício de não causar oclusões nas aquisições, isso também atende ao objetivo específico 1.

5.5 Algoritmo de força bruta

O algoritmo de força bruta foi desenvolvido pensando em uma utilização na *The Visible Human Table*, pois sua capacidade de processamento e armazenamento comporta essa carga de carregamento em memória e processamento. O computador que controla a mesa tem processador

i9 7920X, com 128GB de RAM DDR4, placas de vídeo Nvidia GTX1080ti/RTX2080ti, placa mãe x299 da gigabyte, um HD de 1TB, um SSD M.2 de 1TB e outro SSD M.2 de 250GB.

Em um computador pessoal que não tenha essa capacidade de processamento e armazenamento, o algoritmo não é indicado. O motivo é a quantidade de imagens presentes em apenas um Stream2D, que pode chegar a 32 mil. Um teste foi realizado no computador descrito na Seção 3.2, o qual tem 16GB de memória e dois espaço de armazenamento, sendo o principal um SSD M.2 de 500GB.

Nesse teste foi carregado em memória todas as imagens do objeto crânio, contendo 1 stream com 400 frames. O que se observou foi a crescente ocupação de memória e acesso a disco. No início da execução do algoritmo, com apenas 7% do carregamento, já se atingiu 2GB de memória de processamento.

O algoritmo abre a pasta onde o Stream2D está. Após isso, uma matriz do tamanho relativo à quantidade de streams e frames dentro da pasta, por isso são colocados apenas as imagens e o arquivo de configuração (ex: “concha.cfg”), é alocada em memória. Porém, o programa desconsidera o arquivo de configuração, considerando apenas as extensões de imagens, podendo ser “.jpg” ou “.png”, no caso elas eram “.jpg”, pois foram segmentadas colocando um fundo preto.

Essa foi uma das primeiras segmentações realizadas, não se preocupando com manter o fundo transparente, para uso com outras cores no visualizador. Uma imagem com fundo preto tem praticamente o mesmo tamanho em bytes do que uma imagem com fundo transparente, por conseguinte não acarreta mais ocupação de memória.

Conforme o carregamento ocorria, a memória era ocupada e com metade das imagens carregadas já se tinha 10GB de ocupação. O tempo de carregamento, com essa metade carregada, estava em 47 segundos. O tempo total de carregamento para as 400 imagens foi de 1 minuto e 35 segundos, o que não é algo a ter como problemático, contudo a quantidade de imagens não condiz com uma aquisição onde o objeto possa ser observado por completo.

Uma aquisição mínima para se gerar estereoscopia do objeto em semiesfera, todas as perspectivas, exceto aquela que está em contato com a base, é de 6°. Com isso, o Stream2D conterà $\frac{90}{6}$ streams e $\frac{360}{6}$ frames, tendo um total de 900 imagens, mais do que o dobro da quantidade usada no teste.

Ao final do carregamento, a memória ocupada se estabilizou em 20GB e mesmo com a interação do usuário, isso não se alterou, contudo houve lentidão e interrupções. Levando em consideração que são apenas 400 imagens, um Stream2D com 32mil seria praticamente inviável no computador utilizado para o teste.

Até mesmo um Stream2D com aquisição em 3° para streams e frames causou problemas, pois esse Stream2D contém 30 streams e 120 frames por stream e o carregamento falhou ao atingir 60GB de memória de processamento em 33% das imagens carregadas.

No Algoritmo 3 apresenta-se o pseudocódigo implementado para o teste.

Algoritmo 3: Pseudocódigo do Algoritmo Força Bruta

```

1 Função ForcaBruta (PastaStream2D)
2   | Abre Arquivo de Configuração (.cfg);
3   | Ler Número de Stream do arquivo;
4   | Ler Número de Frames do arquivo;
5   | repita
6   |   | repita
7   |   |   | Carrega imagem na matriz;
8   |   |   | até Número de Frames;
9   |   | até Número de Streams;
10  | Retorna ForcaBruta;

```

O carregamento dos streams e frames são em uma estrutura bidimensional homogênea, que comumente é denominada de matriz, do tamanho total da aquisição. Neste caso, como foram 30 streams de 120 frames cada, perfazendo um total de 3720 imagens, a matriz foi determinada com 30 linhas por 120 colunas. Em cada uma das células foi armazenada uma imagem. A Equação 5.6 representa o pseudocódigo.

$$MatForcaBruta(i, j) = \sum_{i=0}^{NumStream-1} \sum_{j=0}^{NumFrames-1} \quad (5.6)$$

Sendo NumStream a quantidade de streams e NumFrames a quantidade de frames. Como as matrizes, na linguagem utilizada, iniciam em 0(zero) para linha e 0(zero) para as colunas, o total de elementos será o NumStream-1, assim como o NumFrames também. Isso se dá pelo fato de que, como apresentado na Seção 3.3.2, os streams iniciam em 1 e frames em 1001.

Como a ideia é a disponibilização para execução em outros computadores, inclusive com menor capacidade que este utilizado, o algoritmo poderia provocar travamentos ou até mesmo não carregar todas as imagens e acabar tendo um problema de interrupção da execução. A avaliação da execução deste algoritmo está relacionada aos objetivos específicos 2, 3 e 4.

5.6 Algoritmo par-a-par

Diferentemente do algoritmo de força bruta, o algoritmo Par-a-Par não carregou antecipadamente nenhuma imagem em memória. Ele acessou a pasta onde o Stream2D estava armazenado e abriu o arquivo de configuração “.cfg”, de onde tirou as informações para se configurar.

Para este algoritmo não existe a chamada “imagem base”, uma vez que ele navega diretamente dentro da pasta e seleciona o arquivo da sequência, baseado no movimento requisitado pelo usuário. Cada imagem do par estéreo foi carregada em memória, após isso ocorrer, o

algoritmo esperava a interação para determinar se apresentaria o par em forma de anáglifo ou polarização. Isso também ocorreu no algoritmo de força bruta.

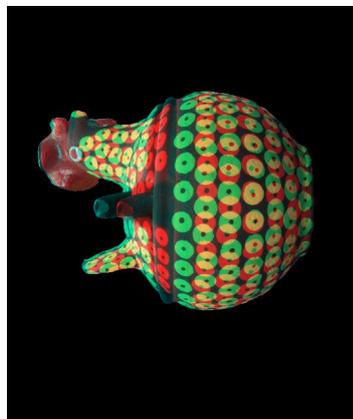
Após a apresentação do par, caso ocorresse outra requisição de movimentação, uma das imagens era descartada e outra colocada no lugar. A movimentação poderia ser para qualquer direção, contudo se ela ocorresse dentro do mesmo stream, por exemplo, sendo o par estéreo formado pelos frames $Stream2D_{(i,j)}$, a imagem esquerda, e $Stream2D_{(i,j+1)}$, a imagem direita, se uma movimentação fosse para a direita, acarretava uma mudança de posicionamento. Então, a imagem esquerda passaria a ser o frame $Stream2D_{(i,j+1)}$ e a direita $Stream2D_{(i,j+2)}$, retirando-se da memória o frame da posição $Stream2D_{(i,j)}$.

Para efeito de exemplificação, não se está considerando as distâncias angulares da aquisição, apenas indicando o que ocorre em termos de mudança de posicionamento dos frames. O esforço de carregamento das imagens diminui, pois não se carregam dois frames, mas apenas um. O movimento para a esquerda provoca o mesmo comportamento, sendo que o frame da posição $Stream2D_{(i,j)}$ passaria a ser a imagem da direita e o frame da posição $Stream2D_{(i,j-1)}$ a imagem da esquerda.

Uma movimentação na vertical ou diagonal dependerá do dispositivo de visualização. Por exemplo, se ela for em óculos RV, onde a inclinação da cabeça poderá indicar a movimentação, o par poderá ser montado com um dos frames que já está em memória.

Como o óculos identifica a inclinação, o movimento seguirá isso, apresentando um par e sua estereoscopia com a mesma inclinação. Imaginando então uma inclinação em que a cabeça seja girada 90° para a direita, o resultado será colocar o frame $Stream2D_{(i,j)}$ como imagem esquerda e o frame $Stream2D_{(i+1,j)}$ como imagem direita. O resultado disso pode ser visto na Figura 88, contudo a imagem foi girada, para que a observação não necessite da inclinação da cabeça.

Figura 88 – Inclinação da cabeça em 90°



Fonte: autoria própria

Na imagem da Figura 88, o frame $Stream2D_{(1,1001)}$ é observado pelo olho esquerdo e

o frame $Stream2D_{(3,1001)}$ pelo olho direito, isso também ocorre em caso contrário, ou seja, se a inclinação da cabeça for para a esquerda, haverá uma inversão entre esses dois frames para serem apresentados para cada olho. A discussão sobre a formação do melhor par estéreo em qualquer direção será na seção 5.8.

Na execução da visualização Par-a-Par, o consumo de memória, considerando uma imagem de 3859x2940 pixels (2.41MB), que é superior a 2K (3840x2160), foi de aproximadamente 585MB, mesmo com a constante interação do usuário.

No entanto, foi percebido uma leve interrupção a cada interação. Não visualmente, mas ao se pressionar uma tecla de movimentação, essa ocorria milissegundos após, perceptível apenas por quem estava manipulando.

Ao se pressionar repetidamente alguma tecla de movimentação, houve um acúmulo, resultando em movimentos do objeto mesmo após serem cessadas as interações do usuário. Isso não foi um comportamento desejado, pois ocorria de passar o ponto desejado ou ter uma espera para que o movimento parasse. Esse problema ocorreu mesmo com um consumo considerado pequeno, quando comparado com o algoritmo de força bruta, pois o consumo dele para 400 imagens atingiu 20GB.

No Algoritmo 4 apresenta-se o pseudocódigo para a execução Par-A-Par, no entanto o processo de movimentação, que está explicado nos parágrafos iniciais desta seção, está apenas representado por uma função de chamada ao que seria o código.

Algoritmo 4: Pseudocódigo do Algoritmo Força Bruta

```

1 Função ParAPar (StreamFrameAtual e PastaStream2D)
2   repita
3     Encontrar ImagemPar para StreamFrameAtual de acordo com
4       ComandoMovimento;
5     Carregar em memória StreamFrameAtual como imagem da esquerda;
6     Carregar em memória ImagemPar como imagem de Direita;
7     ApresentarEstereoscopia(StreamFrameAtual e ImagemPar);
8     Movimentação(ComandoMovimento);
9     se ComandoMovimento diferente de ESC então
10      Avaliar direção da movimentação(novoStreamFrame);
11      Atualizar StreamFrameAtual com valor do novoStreamFrame;
12   fim
13 até ComandoMovimento igual a ESC;
14 Retorna 0;

```

O algoritmo, apesar do consumo baixo, em relação ao força bruta, também apresentou o problema de acúmulo de interação. Isso pode atrapalhar em momentos de execução de um OA dentro de um ambiente instrucional.

Não é desejável, ao visualizar um objeto interativo estereoscópico, ter que esperar as interações que já ocorreram, mas que ainda estão em memória, cessarem para executar outras.

O movimento tem que ser contínuo e também aceitar outras interações como rotação, escala e translação, sem que haja alguma interrupção na visualização. Assim como o algoritmo Força Bruta, este atende aos objetivos específicos 2, 3 e 4.

5.7 Algoritmo de janelas deslizantes na formação da estereoscopia

A utilização de uma janela deslizante para controlar a visualização interativa estereoscópica foi pensada para minimizar os problemas de carregamento e descarregamento de imagens na memória, uma vez que o Stream2D pode ter até 32mil imagens.

Não é todo computador que pode carregar uma quantidade dessa em memória, execução em força bruta, em pouco tempo e ainda manter o desempenho de execução sem interrupções.

Como citado, o computador da “*The Visible Human Table*” tem capacidade e se mostrou viável para carregar um conjunto, pouco mais do que três mil e quinhentas imagens, levando aproximadamente 40 segundos.

Em outro computador, um AVELL i7-4910MQ 2.90GHz, Nvidia Quadro K3100M 4GB, 32GB de RAM e HD 2TB, ocorreu o mesmo problema que o descrito na Seção 5.5 com o computador usado para testes. No entanto, a falha foi com pouco mais de 50% das imagens carregadas.

O algoritmo inicia com a construção de uma lista dinâmica totalmente conectada com uma quantidade de nós ímpares tanto para os frames quanto para os streams. Definiu-se assim para que tenha um elemento central e a partir dele navegar na lista.

Esse elemento é chamado de “pivô”, o qual também representa inicialmente a imagem base. Aquela que em determinadas aquisições não fará parte do par estéreo e em outras será uma das imagens dele. O nó da lista foi definido como uma estrutura na linguagem C++ (struct), Algoritmo 5.

O noLista contém os ponteiros para cada nó adjacente, mesmo os da diagonal, assim como duas variáveis para armazenarem o número do stream e do frame, um ponteiro para a textura, onde a imagem é armazenada, por fim, foi definida uma variável booleana para indicar se a visualização é uma imagem oposta. A utilização dessa imagem oposta ainda não está definida, ela será debatida na Seção 5.8.

A construção da lista parte do princípio que, para a estereoscopia, o Stream2D será montado com um ângulo máximo entre os frames, principalmente, mas também entre streams. Esse ângulo, como já discutido, é 6°, portanto serão 60 frames e 15 streams, assim ficou definido que a janela terá 21x11 nós. Sendo o primeiro número referente aos frames e o segundo aos streams.

Algoritmo 5: Estrutura na lista dinâmica da janela deslizante

```

1 estrutura noLista
2   inteiro numStrm;
3   inteiro numFrm;
4   noLista *superior;
5   noLista *inferior;
6   noLista *esquerda;
7   noLista *direita;
8   noLista *diagSupDir;
9   noLista *diagInfDir;
10  noLista *diagSupEsq;
11  noLista *diagInfEsq;
12  SDL_Texture* f2s2Textura;
13  booleano oposto;
14 fim

```

Viu-se que, com esse tamanho de janela, o deslocamento dentro dela permitiu aproximadamente sete passos para a direita ou para a esquerda na horizontal, antes de se reorganizar. Na vertical foram três passos para cima ou para baixo. Nas diagonais também foram três passos para qualquer uma delas.

Com a aquisição no ângulo máximo, a estereoscopia foi gerada utilizando dois frames adjacentes. Esses frames foram $Stream2D_{(i,j)}$ e $Stream2D_{(i,j+1)}$, com “i” chegando a no máximo “i+3” e “j” chegando a “i+7”.

Assim sendo, no início da execução, o pivô (imagem base) é determinado como $Stream2D_{(i,j)}$, com “i = 1” e “j=1001”, portanto um deslocamento na horizontal resultou na mudança do pivô para $Stream2D_{(i,j)}$, com “i=1” e “j=1008”, para a direita ou “i=1” e “j=1114” no movimento para a esquerda.

A reorganização ocorreu quando o pivô se tornou um desses valores de “j”, pois em mais dois passos o par estéreo não se formaria. Na direção vertical o valor de “j” foi sempre o mesmo, ou seja, assumindo o mesmo pivô inicial, o valor é “j=1001” e quem se alterou foi “i”, o qual chegou a “i=4” ou “i=27”.

A diagonal é um caso particular, uma vez que ela é o resultado dos dois deslocamentos (horizontal e vertical), portanto os valores para “i” e “j” foram os quatro máximos (“i=4” ou “i=27”, “j=1008” ou “j=1114”).

Para centralizar o pivô na matriz, adotou-se a estratégia de encontrar o frame e stream que estivessem na distância de dez posições antes do $Stream2D_{(1,1001)}$. Contudo, como a contagem para trás do primeiro não há, o que existe é um movimento circular, esse frame e stream torna-se o último na aquisição, por conseguinte ele é o da posição $Stream2D_{(1,1120)}$.

A partir desse frame, fez-se um cálculo de subtração até o limite esquerdo da janela. O

mesmo procedimento foi utilizado para o stream, ou seja, não há um anterior ao primeiro, mas como é circular, pega-se o último stream, portanto acima do pivô é $Stream2D_{(31,1001)}$, a partir dele fez-se uma subtração até o limite superior da janela.

Os limites foram calculados por meio do tamanho definido para a janela. O tamanho total foi de 21×11 , portanto os limites esquerdo e direito foram determinados por esses números. Então, pega-se as metades dos dois valores, sendo para frames $\frac{21}{2}$ e streams $\frac{11}{2}$. O resultado é arredondado para baixo, independentemente do decimal, o que nestes casos deram “10.5” e “5.5”, arredonda-se para “10” e “5”. No Algoritmo 6 apresenta-se o pseudocódigo do carregamento da janela deslizante inicial.

Algoritmo 6: Pseudocódigo do Algoritmo Janela Deslizante Inicial

```

1 Função ListaDinamicaInicial (ListaJanDesl)
2   repita
3     Encontrar Stream;
4     repita
5       Encontrar Frame;
6       Aloca Memória para a Nova Estrutura noLista;
7       Armazena Nova Estrutura noLista com a imagem “Stream-Frame”;
8       Atualiza ponteiros da Nova Estrutura noLista armazenada;
9       Atualiza ponteiros da ListaJanDesl;
10      Atualiza valorframe;
11    até valorframe igual a  $\tau$ ;
12    Atualiza valorstream;
13    Coloca valorframe com valor inicial;
14  até valorstream igual a  $\varsigma$ ;
15  Retorna ListaJanDesl;

```

As Equações 5.7 e 5.8 representam os cálculos para encontrar os componentes da janela deslizante inicial, independentemente do grau de aquisição. Por exemplo, em aquisição de 2° para streams e frames, tem-se que o último stream será 45 e o último frame será 180. Se a aquisição for em 6° para streams e frames, tem-se que o último stream será 15 e o último frame será 60.

$$Frame = \begin{cases} (1000 + k) - \text{int}(\frac{\tau}{2}) + m & \text{caso } 1 \leq m \leq \text{int}(\frac{\tau}{2}) \\ 1000 + (m - \text{int}(\frac{\tau}{2})) & \text{caso } \text{int}(\frac{\tau}{2}) < m \leq \tau \end{cases} \quad (5.7)$$

$$Stream = \begin{cases} p - \text{int}(\frac{\varsigma}{2}) + n + 1 & \text{caso } 1 \leq n \leq \text{int}(\frac{\varsigma}{2}) \\ n - \text{int}(\frac{\varsigma}{2}) & \text{caso } \text{int}(\frac{\varsigma}{2}) < n \leq \varsigma \end{cases} \quad (5.8)$$

Sendo “m” o contador que varia de 1 até τ , “k” é o valor encontrado pelo cálculo de $\frac{360}{dbf}$ e τ é a quantidade total das colunas da janela. “n” é o contador que varia de 1 até ς , “p” é o valor

encontrado pelo cálculo de $\frac{90}{dbs}$ e ς é a quantidade total de linhas da janela. Com isso, tem-se a Equação 5.9 que representa o pseudocódigo de carregamento da janela deslizante inicial.

$$Stream2D(\varsigma, \tau) = \sum_{n=1}^{\varsigma} \sum_{m=1}^{\tau} F_{Stream, Frame} \quad (5.9)$$

$Stream2D(\varsigma, \tau)$ representa um nó que é armazenado em memória e $F_{Stream, Frame}$ representa um nome dentro do Stream2D, por exemplo “S1-F1001”. Pelo tamanho da janela em 11 streams e 21 frames, define-se que o pivô terá “10” elementos na esquerda e outros “10” na direita, assim como ele terá “5” acima e outros “5” abaixo. Na Matriz 5.10 apresenta-se a configuração inicial da janela deslizante. O pivô está destacado em cor vermelha.

$$Stream2D_{\varsigma, \tau} = \begin{bmatrix} F_{27,1111} & \dots & F_{27,1120} & F_{27,1001} & \dots & F_{27,1011} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ F_{31,1111} & \dots & F_{31,1120} & F_{31,1001} & \dots & F_{31,1011} \\ F_{1,1111} & \dots & F_{1,1120} & F_{1,1001} & \dots & F_{1,1011} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ F_{6,1111} & \dots & F_{6,1120} & F_{6,1001} & \dots & F_{6,1011} \end{bmatrix} \quad (5.10)$$

O desempenho do carregamento inicial se mostrou bom, quando comparado com os outros dois algoritmos apresentados, uma vez que levou aproximadamente 50 segundos para completar a tarefa.

As imagens foram as mesmas utilizadas nos outros testes, com 3589x2940 pixels e 2.41MB. As reorganizações não consumiram mais do que milésimos de segundo, o que não foi percebido durante a execução em tempo real, apenas visualizado no passo a passo no visual studio.

Não ocorreram interrupções na visualização, já que as imagens estavam em memória, deixando de realizar constantes acessos a disco para buscá-las e carregá-las, como ocorreu no algoritmo par-a-par. O consumo de memória de processamento também foi bom, em comparação aos outros dois algoritmos.

A memória ficou constante durante o carregamento da janela, fixada em 276MB. Com a visualização do par estéreo e interação passou a 322MB, não se alterando durante a reorganização, pois a janela também é de tamanho constante, ou seja, carrega-se na mesma quantidade do que é descarregado.

Os valores de consumo foram em média 55.14% menores do que o algoritmo Par-a-Par, que consumiu 585MB de memória de processamento. Em relação ao algoritmo de força bruta, mesmo tendo um comportamento semelhante, pois tanto nele quanto na janela há um carregamento prévio em memória, antes de se visualizar o objeto, os valores são discrepantes.

No força bruta, para uma quantidade de 400 frames foram 20GB de memória, já na janela deslizante para um total de 231 frames e com interação do usuário foram 322MB.

Na Tabela 1 estão os dados das execuções dos algoritmos. A última linha de dados diz respeito a interação do observador, levando a janela a uma reconfiguração, no caso, a maior reorganização possível, ou seja, quando a movimentação é para qualquer uma das diagonais, com isso, a reorganização conta com 15 imagens, sendo 11 em coluna e mais 4 nas linhas.

Tabela 1 – Consumo de memória de processamento

Dados da execução dos algoritmos e resultado de consumo de memória de processamento						
Algoritmo	quantidade de imagens	porcentagem carregada	tempo de carregamento	memória de processamento	Executou	
Força Bruta	400	100%	≈ 1m:35s	20 GB	sim, mas com lentidão e interrupções	
Força Bruta	3720	33%	≈ 4m	60 GB	não, deu estouro de memória	
Par a Par	2	100%	≈ 0.5s sem interação	585 MB	sim	
Par a Par	2	100%	≈ 2.5s com interação	585 MB	sim, mas com acúmulo de interação e lentidão	
Janelas deslizantes	231	100%	≈ 1m	276 MB	sim	
Janelas deslizantes	231	houve reorganização	100%	≈ 10 ⁻³ s com interação	322 MB	sim, sem interrupções ou lentidão

Fonte: Autoria Própria

Na Figura 62.c, página 130, mostra-se um exemplo, contudo, lá foi apenas uma representação e não conta com a quantidade que foi utilizada neste teste. Há um ponto de intersecção na diagonal, pois a imagem de encontro com a linha e coluna na extremidade é carregada apenas uma vez, ou no carregamento das linhas ou no carregamento das colunas.

Esse algoritmo atende aos objetivos específicos 2, 3, 4 e 5. Uma vez que foi possível criar um processo computacional que permitiu a visualização interativa estereoscópica a partir de todas as perspectivas adquiridas de um objeto.

5.8 Movimentação e formação estereoscópica

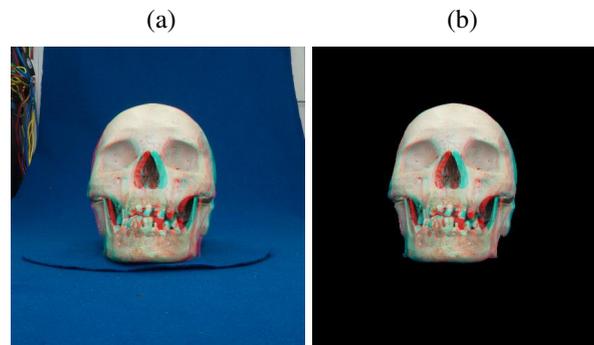
No início do estudo da estereoscopia, guiado pelas formas de se produzir a noção de profundidade, acreditou-se que a diferença interpupilar era o que regia e que uma aquisição que não seguisse esse padrão não produziria a visualização estereoscópica.

O F2S2 mostrou que uma aquisição em ângulos e com apenas uma câmera que captura todas as perspectivas, mantendo um ângulo de 90° em relação ao objeto, produz a sensação com

paralaxe positiva e negativa, que depende apenas de como se monta o par estéreo.

Isso pode ser percebido na Figura 89, onde o mesmo par é montado com as imagens originais, com o fundo do F2S2 e o outro com o par segmentado.

Figura 89 – Estereoscopia com paralaxe positiva e negativa



Fonte: autoria própria

Aqui foram utilizadas as imagens com fundo, mantendo a mesma imagem para o olho esquerdo e direito, no entanto se houver a inversão do par, consegue-se a paralaxe positiva com a imagem segmentada.

Porém, isso é mais facilmente identificado com a projeção em polarização, em anáglifo há um leve desconforto, pois a alteração das cores para vermelho e ciano provoca essa sensação. No projetor e com os óculos com duas opções de visualização, esse desconforto não ocorre.

O padrão adotado nas visualizações produzidas neste trabalho foi colocar o stream $Stream2D_{(i,j)}$ como imagem da esquerda e o $Stream2D_{(i,j+w)}$ como imagem da direita. Imaginando um objeto parado, isso representaria um movimento do observador da esquerda para a direita.

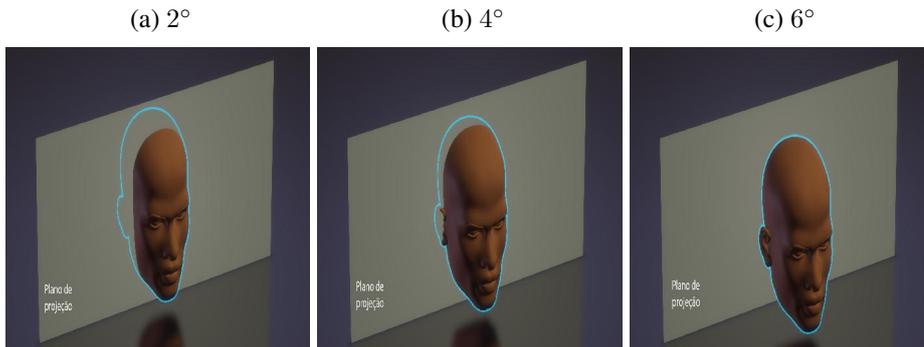
Imaginando o objeto na mão do observador, ele estaria girando-o da direita para esquerda. Lembrando que “w” está ligado ao ângulo de aquisição θ , ou seja, se θ for 2° , “w” será igual a 3, pois a diferença entre as imagens do par é de até no máximo 6° .

Não importa se o observador girar o objeto da esquerda para a direita, o par é o mesmo. A sensação de profundidade está ligada aos graus na aquisição, o que também não foi debatido em trabalhos estudados.

O nível de paralaxe, entende-se nível pela distância que se percebe a estereoscopia em relação ao plano de projeção, é diferente para um par que tenha 2° , 3° até 6° . Uma representação disso é mostrada na Figura 90.

Ao se visualizar a estereoscopia, quanto maior o grau de distância entre o par, até 6° , maior é a sensação de que o objeto está fora do plano. Essa sensação também está ligada a como o objeto é observado, pois nesses exemplos o par estéreo foi formado por uma distância angular

Figura 90 – Estereoscopia em relação ao plano de projeção



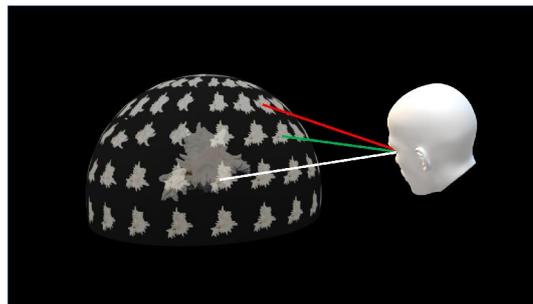
Fonte: autoria própria

de até 6° na horizontal. Essa percepção é subjetiva, pois depende do observador, por isso pode-se avaliar por um grupo de pessoas, como será indicado nos trabalhos futuros.

A visualização na vertical, quando se inclina a cabeça, obedece a mesma regra, isso foi apresentando na Figura 88, portanto a estereoscopia é formada com duas imagens na mesma distância angular que as na horizontal. Um dos desafios deste trabalho foi encontrar um par estéreo que proporcionasse a noção de profundidade em imagens diagonais.

Essa formação de par representa uma inclinação da cabeça ou do objeto, como se estivesse com algo na mão e sem girá-lo quisesse olhar com a cabeça um pouco inclinada. Na Figura 91 apresenta-se essa hipotética situação.

Figura 91 – Inclinação da cabeça na diagonal



Fonte: autoria própria

Nela foram colocadas três linhas, uma branca, uma verde e uma vermelha, pois até o início dos estudos, em uma visualização desse tipo, supunha-se que a formação do par seria com a distância angular entre os frames da diagonal. Portanto, seguindo essa ideia, o par deveria ser montado com $Stream2D_{(i,j)}$ e $Stream2D_{(i-w,j+w)}$, por exemplo, $Stream2D_{(3,F1001)}$ e $Stream2D_{(1,1003)}$ em uma aquisição onde $\theta = 3^\circ$. A distância entre esses dois frames é 6° .

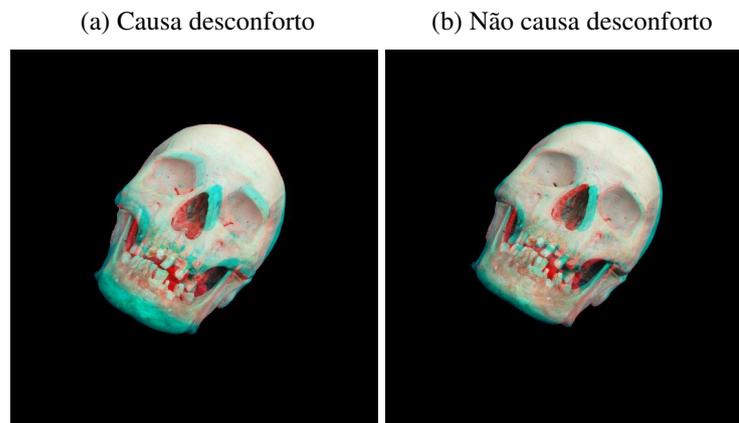
Após a formação de pares com essa distância para frames e streams, linha branca com vermelha, percebeu-se que a estereoscopia ocorria, mas causando desconforto, no anáglifo. O

que levou a uma nova suposição que a distância na diagonal se comporta diferentemente da horizontal e vertical.

Dessa suposição surgiu a investigação de qual seria o melhor par, quando se trata de uma movimentação na diagonal. Isso influencia a visualização em óculos RV, porém em outros visualizadores pode ser requisitado isso. Assim, outras distâncias foram testadas, como 9° na diagonal, contudo o desconforto foi mais intenso do que o outro. Testou-se com apenas 2° entre os frames, que resultou na mesma estereoscopia apresentada na Figura 90.a.

O melhor resultado foi com os mesmos 6° , no entanto entre os frames $Stream2D_{(i,j)}$ e $Stream2D_{(i-(w-1),j+w)}$ (linha branca com verde), para a aquisição $\theta = 3^\circ$. Confirmado pelo teste em anáglifo usando o GIMP, mostrado na Figura 92, onde a imagem da esquerda (92.a) é a estereoscopia desconfortável e a da direita (92.b) apresenta uma estereoscopia correta.

Figura 92 – Estereoscopia na diagonal



Fonte: autoria própria

Os valores de “i” e “j” para esse exemplo são encontrados pelo movimento do observador da esquerda para a direita indo de baixo para cima, ou giro do objeto da direita para a esquerda. Os valores desses índices se alteram em movimentos para as outras diagonais, por exemplo, se o movimento do giro do objeto for na diagonal da esquerda para a direita de cima para baixo, ou seja, diagonal inferior esquerda, os frames seriam $Stream2D_{(i+(w-1),j-w)}$ e $Stream2D_{(i,j+w)}$, por exemplo, $Stream2D_{(2,1020)}$ e $Stream2D_{(1,1022)}$. Esses frames e suas posições foram explicados na montagem da janela deslizante, na Seção 5.7, na Matriz 5.10, página 165.

Até o momento não se faz aquisição de um objeto invertido, ou seja, posiciona-se o objeto com uma das partes em contato com a base giratória e somente assim. A aquisição invertida seria colocar o lado contrário a esta posição inicial, portanto seu topo é que passaria a estar em contato com a base giratória. Em linhas gerais, pode-se dizer que a aquisição é com o objeto “em pé” e não “de ponta cabeça”.

Para a visualização desses objetos adquiridos somente “em pé”, deveria-se travar quando é acessado o primeiro stream e o último stream. Não seria uma janela deslizante com uma lista

dinâmica circular. Foi realizado um teste criando uma opção chamada “oposta”, na qual se apresenta um objeto de forma inversa sem que este tenha sido adquirido.

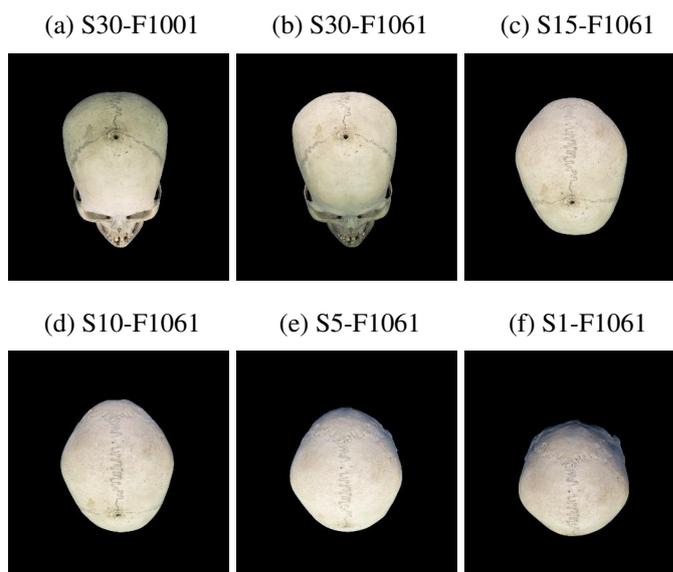
Em termos de graus, a ideia é pegar um frame que esteja na posição inversa àquela que está sendo apresentada. Então, todo frame dentro de um stream tem o seu oposto e este fica a 180° do outro. Por exemplo, o frame $Stream2D_{(1,1001)}$ tem como oposto o $Stream2D_{(1,1061)}$, para uma aquisição de $\theta = 3^\circ$. Em uma aquisição de $\theta = 1^\circ$, os opostos seriam frame $Stream2D_{(1,1001)}$ e $Stream2D_{(1,1181)}$.

Sabendo quem são seus opostos, baseados em θ , pode-se invertê-los conforme a movimentação. Assim, imaginando um observador manipulando um objeto em suas mãos, onde ele faz um movimento de baixo para cima, até atingir o topo. Para o F2S2, em uma aquisição normal, o topo é o último stream, por conseguinte, a visualização deveria parar nele e não permitir movimento, a não ser que permita movimentar dentro desse stream.

Com a ideia do oposto, o próximo stream e frame seria o oposto ao visível neste instante. Voltando ao exemplo de $\theta = 3^\circ$ em navegação do primeiro para o último, a sequência de movimentos partiria de $Stream2D_{(1,1001)}$ e chegaria em $Stream2D_{(31,1001)}$. Chegando neste e mantendo o movimento na vertical, deve-se acessar o oposto ao $Stream2D_{(31,1001)}$, que é o frame $Stream2D_{(31,1061)}$.

Ainda mantendo o movimento, esse objeto deverá sofrer uma rotação também de 180° para representar um movimento dito como “natural” em objetos reais. Na Figura 93 apresenta-se o resultado da aplicação da ideia dos Frames opostos.

Figura 93 – Imagens em sentido oposto



Fonte: autoria própria

A primeira imagem (Figura 93.a) é a única que não sofreu rotação de 180° . Um problema com essa visualização oposta é que o F2S2 não adquire um objeto a partir de uma posição abaixo

do objeto, com a câmera inclinada para cima, sempre inicia do meio do objeto e vai até o topo. Por isso, a última imagem, Figura 93.f, corresponde ao ângulo mais baixo do objeto. Se houvesse uma aquisição mais abaixo disso, o que seria visto é a base do crânio e um pouco do maxilar.

Essa mesma ideia é aplicada no caso em que a visualização está no primeiro stream (p.ex.: $Stream2D_{(1,1001)}$) e a movimentação faz um acesso tentando visualizar a parte que está em contato com a base giratória, ou seja, não foi adquirida. O movimento faz um salto, que neste caso, mantém-se no primeiro stream e vai ao frame oposto $Stream2D_{(1,1061)}$.

Os frames que se seguem são invertidos e a lista é percorrida no frame $Stream2D_{(i,1061)}$, como se olhasse o objeto da base ao topo. Faz o movimento contrário ao apresentado na Figura 93, indo da imagem (f) até (a), depois descendo os streams do 31 ao 1 no mesmo frame (1001, oposto ao 1061).

Como citado, isso ainda é um objeto de estudo, mas já implementado no software, contudo um debate com o grupo de pesquisa e com o prof. orientador Edson J. R. Justino é que se vai disponibilizar para o uso na “*The Visible Human Table*”.

A formação da estereoscopia por meio das aquisições do F2S2 se mostrou promissora, pelo fato de não necessitar ajustes das imagens que formam o par estéreo. Os objetos adquiridos foram submetidos à visualização pelas técnicas de anáglifo e polarização, sendo que na primeira há o problema de alteração das cores e em alguns casos causa um desconforto. Esse não é pela má qualidade da estereoscopia, mas pelo fato de algumas cores não serem indicadas para a técnica.

Um objeto que seja vermelho, preto, branco ou azul perde qualidade, pois o anáglifo utiliza algumas delas para criar a noção de profundidade. Por exemplo, se a cor do objeto é vermelha e o anáglifo for nas cores vermelho e ciano, o objeto perderá qualidade. Já na estereoscopia polarizada, não há problema, pois a cor do objeto é aquela que será projetada. Além de manter a cor e textura intactas, os óculos agridem menos do que o de duas cores. Contudo o anáglifo é menos dispendioso do que o polarizado, pois não exige equipamento próprio, a não ser o óculos, que hoje pode ser encontrado em diversos lugares a preços acessíveis.

O custo computacional para a formação do par estéreo, em comparação ao uso dos softwares de ajustes da epopilar, como o SURF, foi desconsiderado, pois não há ajustes. A visualização interativa estereoscópica ocorre com a seleção do par estéreo e com a seleção da técnica de apresentação. Para a polarização, faz-se a seleção, cria-se uma duplicação de tela e apresenta o par. O projeto faz o entrelaçamento, assim como ocorre em TVs 3D. Então, qualquer objeto adquirido nas especificações feitas neste trabalho, poderá apresentar estereoscopia, desde que selecione o par correto.

O visualizador implementado como teste foi capaz de produzir a estereoscopia e permitir a interação do usuário, assim como a janela deslizante permitiu o carregamento dos frames sem que houvesse interrupções na visualização e na formação estéreo. A movimentação pela

janela manteve estável o consumo de memória de processamento, não onerando o processador ou fazendo constantes acessos ao disco rígido, o que permitiu uma visualização suave. Com isso, acredita-se que a criação de OA seja possível, sem que complicações, inclusive com estereoscopia do conteúdo ou qualquer adição de informação que ocorra.

Neste capítulo foram apresentados os resultados experimentais e discussões. O próximo capítulo destina-se às considerações finais.

Capítulo 6

Considerações finais

Esta pesquisa objetivou o desenvolvimento de uma visualização interativa estereoscópica, a qual permita a criação de objetos de aprendizagem ou conteúdos para ambientes instrucionais. Foram debatidos temas como a aquisição de objetos por diferentes meios, como laser e fotos. Qual a diferença entre cada técnica de aquisição e quais poderiam contribuir para a criação da estereoscopia.

A formação das visualizações estereoscópicas também foi discutida, pois é uma técnica que remonta o século XIX e que de lá para cá pouca coisa foi mudada na teoria de como gerar a noção de profundidade utilizando duas imagens. No geral, são realizados ajustes em uma das imagens para que, ao ser visualizada, cause o menor desconforto possível.

Outro ponto abordado foi como se carregar em memória uma quantidade, considerada grande, de imagens, mantendo o desempenho do computador e não causando interrupções durante a visualização. Assim, alguns algoritmos foram testados, mas que não necessariamente sejam únicos que podem ser utilizados.

Abordou-se como a aquisição do *Full Frames Semi-Spherical Scanner* gera um conjunto de perspectivas de um objeto. Além de como criar uma estrutura, denominada Stream2D, que representa o posicionamento de cada perspectiva em um plano esférico. Este plano simboliza o objeto completo, ou seja, há uma representação tridimensional.

Obter uma visualização ou representação tridimensional de um objeto não garante que o mesmo possa ser visualizado com estereoscopia. Por isso, foi estudada uma técnica que pudesse ser empregada na geração dessa visualização para todas as perspectivas. Descobriu-se qual a influência dos ângulos na aquisição, quais são aqueles que permitem a criação do objeto estereoscópico e os que não causam desconforto durante a visualização.

Assim, acredita-se que a originalidade foi conseguida por meio da simplicidade da criação das visualizações, uma vez que se descarta a utilização de algoritmos de correção da epipolar, prática comum em aquisições de imagens paralelas ou convergentes, como ocorre em capturas para cartografia ou aerofotogrametria. Além de permitir uma percepção de profundidade com paralaxe negativa, que significa observar o objeto fora do plano de projeção, como ocorre em projeções cinematográficas.

O algoritmo de janelas deslizantes, por meio de uma lista dinâmica totalmente conectada, mostrou-se eficaz no carregamento de um conjunto de imagens, permitindo a interação do usuário, com reorganização da janela sem causar interrupções ou onerar o processamento.

Com isso, os objetivos específicos foram alcançados, pois na visualização se navega e observa o objeto por completo, além de permitir que o usuário faça a translação, rotação e escala do objeto. Garantiu-se a interatividade do usuário, já que ele pode mudar as perspectivas e cada uma se apresenta com noção de profundidade, não importa se em movimento horizontal, vertical ou diagonal.

Atingindo-se os objetivos específicos, também se atingiu o objetivo geral, que foi o desenvolvimento de uma projeção interativa estereoscópica. Por conseguinte, a questão de pesquisa pode ser respondida, uma vez que, as visualizações foram apresentadas em duas técnicas de estereoscopia, a anáglifo e polarização da luz, sendo que na segunda não se perde qualidade na textura e cor, tendo uma representação fiel do objeto real no ambiente virtual. Contudo, alguns pontos que foram identificados e não abordados neste trabalho, podem ser objetos de estudos. Eles estão elencados como trabalhos futuros.

6.1 Trabalhos Futuros

Na teoria da estereoscopia há uma afirmação de que não se tem estereoscopia em distâncias menores do que 25cm. Isso é percebido quando se olha um objeto real posicionado abaixo dessa distância na frente dos olhos. No entanto, em imagens fotográficas isso pode não ser correto, então, sugere-se:

- uma investigação com a determinação da distância correta do objeto ao sensor das câmeras, pode ser um objeto de estudo. Isso exige alguns levantamentos acerca da câmera e dos sensores.

Abordou-se aqui que não foi necessário aplicar correções nas epipolares das imagens, pois os ângulos de aquisição proporcionam as correções. No entanto, pode-se explorar as aquisições com ângulos superiores a 6° . Por isso, sugere-se:

- aplicar os softwares de geometria epipolar, encontrados na literatura, nas imagens com graus que já proveem a estereoscopia, para descobrir seus valores. Posteriormente aplicar sobre imagens com os graus superiores, aqueles que causam desconforto, buscando um padrão. Havendo padrão, desenvolver uma técnica específica para o F2S2 de estimação da epipolar, com a aplicação de correções para estas perspectivas e geração da estereoscopia.

Mesmo tendo a possibilidade de visualizar o objeto com a representação tridimensional, esse não tem nuvem ou malha de pontos, algo que pode ser explorado para a geração de impressões 3D. Assim, a sugestão é de:

- criar a malha ou nuvem de pontos do objeto adquirido, colocando sobre ele a textura e cor a partir das imagens, além da exportação para padrões aceitáveis por impressoras 3D.

Para uma parte mais prática, visando o usuário final, deve-se focar em programas específicos para o F2S2, pois possibilitar a visualização foi um dos propósitos deste trabalho, no entanto o visualizador e suas funcionalidades não foi. Assim, sugere-se:

- criar um visualizador que permita a inserção de informações acerca do objeto ou de parte dele, pois isso é almejado em um ambiente virtual de aprendizagem.

Além do ambiente virtual de aprendizagem, também há a possibilidade de um programa específico para a criação dos objetos de aprendizagem, portanto a sugestão é de:

- criar programa em que se possa abrir a pasta do Stream2D, selecionar cada uma das perspectivas do objeto em um “canvas” e clicar nos pontos onde inserirá as informações. Em outro “canvas”, ao lado, inserir as informações desejadas. Isso trará funcionalidades ao visualizador, para a criação de objetos de aprendizagem.

Outro ponto inexplorado foi a geração de estereoscopia com fundo dinâmico. Por exemplo, adquirir um inseto e colocá-lo tendo ao fundo seu habitat natural. Portanto, sugere-se:

- criar no visualizador a possibilidade de inserção do “habitat”, em caso de objetos como órgãos, isso seria o corpo, em caso de insetos, a natureza. Contudo, a estereoscopia pode se tornar de paralaxe positiva, assim, caberá um estudo de como transformar o fundo em algo dinâmico e estereoscópico, para que interaja com o objeto. Com isso, haverá uma junção das paralaxes, positiva e negativa.

Como a visualização esteresocópica é subjetiva, depende do observador, indicasse uma estudo mais aprofundado utilizando grupos de pessoas. Portanto, sugere-se:

- um trabalho de pós-graduação dentro da área de educação para avaliação da projeção interativa estereoscópica, onde se avalie o objeto de aprendizagem e níveis de estereoscopia por grupos de estudantes de diferentes áreas e faixa etária.

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 2 vezes nas páginas 38 e 77.

ALAM, A.; ULLAH, S.; ALI, N. The effect of learning-based adaptivity on students' performance in 3d-virtual learning environments. *IEEE Access*, IEEE, v. 6, p. 3400–3407, 2017. ISSN 21693536. Citado na página 37.

ALVARENGA, C. E. A.; GINESTIÉ, J.; BRANDT-POMARES, P. How and why Brazilian and French teachers use learning objects. *Education and Information Technologies*, Education and Information Technologies, v. 22, n. 5, p. 1973–2000, 2017. ISSN 15737608. Citado na página 37.

ALYAMAN, M.; ABD-RAHEEM, A.; ALDEIRI, F. Design of an automated extraoral photogrammetry 3d scanner. *International Arab Journal of Information Technology*, IEEE, v. 16, n. 3ASpecial Issue, p. 533–539, 2019. ISSN 23094524. Citado 2 vezes nas páginas 48 e 53.

ARDILA, C. C. et al. Study for development of a patient-specific 3D printed craniofacial medical device : Design based on 3D virtual biomodels/CAD/RP. *Procedia CIRP*, Elsevier B.V., v. 70, p. 235–240, 2018. ISSN 2212-8271. Disponível em: <<https://doi.org/10.1016/j.procir.2018.03.335>>. Citado 2 vezes nas páginas 52 e 53.

BAŞGÖZE, Z.; MACKEY, A. P.; COOPER, E. A. Plasticity and Adaptation in Adult Binocular Vision. *Current Biology*, v. 28, n. 24, p. R1406–R1413, 2018. ISSN 09609822. Citado na página 39.

BARONE, S. et al. Structured light stereo catadioptric scanner based on a spherical mirror. *Optics and Lasers in Engineering*, Elsevier Ltd, v. 107, n. March, p. 1–12, 2018. ISSN 01438166. Disponível em: <<https://doi.org/10.1016/j.optlaseng.2018.03.004>>. Citado na página 38.

BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. In: LEONARDIS, A.; BISCHOF, H.; PINZ, A. (Ed.). *Computer Vision – ECCV 2006*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 404–417. ISBN 978-3-540-33833-8. Citado 2 vezes nas páginas 67 e 73.

BERNARDO, A. Virtual Reality and Simulation in Neurosurgical Training. *World Neurosurgery*, Elsevier Inc, v. 106, p. 1015–1029, 2017. ISSN 18788769. Disponível em: <<http://dx.doi.org/10.1016/j.wneu.2017.06.140>>. Citado 2 vezes nas páginas 52 e 53.

BERRIER, S. et al. Improved appearance rendering for photogrammetrically acquired 3D models. *2015 Digital Heritage*, p. 255–262, 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7413879/>>. Citado 2 vezes nas páginas 66 e 73.

BESL, P. J.; JAIN, R. C. Three-dimensional object recognition. *ACM Comput. Surv.*, Association for Computing Machinery, New York, NY, USA, v. 17, n. 1, p. 75–145, mar. 1985. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/4078.4081>>. Citado 2 vezes nas páginas 46 e 53.

- BESSA, S.; COSTA, V. G. d. Apropriação do Conceito de Divisão por meio de Intervenção Pedagógica com Metodologias Ativas. *Bolema: Boletim de Educação Matemática*, v. 33, n. 63, p. 155–176, 2019. ISSN 0103-636X. Citado na página 37.
- BROBOUSKI, W. J. P. *Segmentação de Imagens Coloridas de Corpos Humanos (Base VHP) em Ultra-Alta-Definição usando Classificadores SVM*. 117 p. Tese (Doutorado) — Pontifícia Universidade Católica do Paraná, 2020. Citado na página 105.
- CABRAL, B.; CAM, N.; FORAN, J. Accelerated volume rendering and tomographic reconstruction using texture mapping hardware. *Proceedings of the 1994 Symposium on Volume Visualization*, p. 91–98, 1994. Disponível em: <<http://portal.acm.org/citation.cfm?doid=197938.197972>>. Citado 2 vezes nas páginas 51 e 53.
- CANCLINI, A. et al. Evaluation of low-complexity visual feature detectors and descriptors. *2013 18th International Conference on Digital Signal Processing, DSP 2013*, p. 0–6, 2013. Citado 2 vezes nas páginas 69 e 73.
- CATELLI, F. et al. Um problema didático: como determinar ângulos de paralaxe trigonométrica. *Revista Brasileira de Ensino de Física*, v. 40, n. 1, p. 1–9, 2017. ISSN 1806-1117. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172018000100406&lng=p>. Citado 2 vezes nas páginas 57 e 72.
- COHEN, A.; REISMAN, S.; SPERLING, B. B. Personal spaces in public repositories as a facilitator for open educational resource usage. *International Review of Research in Open and Distance Learning*, v. 16, n. 4, p. 157–176, 2015. ISSN 14923831. Citado na página 37.
- CURTI, S.; SIRTORI, D.; VELLA, F. 3D effect generation from monocular view. *Proceedings - 1st International Symposium on 3D Data Processing Visualization and Transmission, 3DPVT 2002*, n. Cid, p. 550–553, 2002. Citado 2 vezes nas páginas 60 e 72.
- DARDON, U. et al. Modelagem 3D e suas aplicações na pesquisa paleontológica. *Gaea - Journal of Geoscience*, v. 6, n. 2, p. 76–89, 2010. ISSN 1983-3628. Citado na página 37.
- DIESEL, A.; BALDEZ, A.; MARTINS, S. Os princípios das metodologias ativas de ensino: uma abordagem teórica. *Revista Thema*, v. 14, n. 1, p. 268–288, 2017. ISSN 2177-2894. Citado na página 37.
- DO, P. N. B.; NGUYEN, Q. C. A Review of Stereo-Photogrammetry Method for 3-D Reconstruction in Computer Vision. *Proceedings - 2019 19th International Symposium on Communications and Information Technologies, ISCIT 2019*, IEEE, p. 138–143, 2019. Citado 2 vezes nas páginas 47 e 53.
- Dong Hui; Han Dian Yuan. Research of image matching algorithm based on surf features. In: *2012 International Conference on Computer Science and Information Processing (CSIP)*. [S.l.: s.n.], 2012. p. 1140–1143. ISSN null. Citado 2 vezes nas páginas 68 e 73.
- DREBIN, R. a.; CARPENTER, L.; HANRAHAN, P. Volume rendering. *SIGGRAPH '88: Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, v. 22, n. 4, p. 65–74, 1988. ISSN 00978930. Citado 2 vezes nas páginas 50 e 53.
- DREWS, P.; De Bem, R.; De Melo, A. Analyzing and exploring feature detectors in images. *IEEE International Conference on Industrial Informatics (INDIN)*, n. May, p. 305–310, 2011. ISSN 19354576. Citado 2 vezes nas páginas 69 e 73.

- EYRE, M. et al. Integration of Laser Scanning and Three-dimensional Models in the Legal Process Following an Industrial Accident. *Safety and Health at Work*, Elsevier Ltd, v. 8, n. 3, p. 306–314, 2017. ISSN 20937997. Disponível em: <<http://dx.doi.org/10.1016/j.shaw.2016.11.005>>. Citado 2 vezes nas páginas 49 e 53.
- FAN, K.; BAEK, S. J. A robust proposal generation method for text lines in natural scene images. *Neurocomputing*, Elsevier B.V., v. 304, p. 47–63, 2018. ISSN 18728286. Disponível em: <<https://doi.org/10.1016/j.neucom.2018.03.041>>. Citado na página 70.
- FARIA, J. W. V. d. et al. Virtual and stereoscopic anatomy: when virtual reality meets medical education. *Journal of Neurosurgery*, v. 125, n. 5, p. 1105–1111, 2016. ISSN 0022-3085. Disponível em: <<http://thejns.org/doi/10.3171/2015.8.JNS141563>>. Citado 2 vezes nas páginas 66 e 73.
- FASANO, A. et al. Exploiting mirrors for laser stripe 3D scanning. *Proceedings of International Conference on 3-D Digital Imaging and Modeling, 3DIM*, v. 2003-Janua, p. 243–250, 2003. ISSN 15506185. Citado na página 49.
- FATTAL, R.; LISCHINSKI, D. Variational classification for visualization of 3D ultrasound data. *Proceedings Visualization 2001*, p. 403–410, 2001. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=964539>>. Citado 2 vezes nas páginas 51 e 53.
- FAU, M.; CORNETTE, R.; HOUSSAYE, A. Apport de la photogrammétrie à la numérisation 3D d'os de spécimens montés : potentiel et limites. *Comptes Rendus - Palevol*, Académie des sciences, v. 15, n. 8, p. 968–977, 2016. ISSN 16310683. Disponível em: <<http://dx.doi.org/10.1016/j.crpv.2016.08.003>>. Citado na página 37.
- FENG, J.; SONG, B.; ZHOU, B. Bottom and concave surface rendering in image-based visual hull. *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry - VRCAI '08*, p. 1, 2008. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1477862.1477866>>. Citado 2 vezes nas páginas 65 e 73.
- FISCHER, A.; MANOR, A. Utilizing image processing techniques for 3D reconstruction of laser-scanned data. *CIRP Annals - Manufacturing Technology*, v. 48, n. 1, p. 99–102, 1999. ISSN 00078506. Citado 2 vezes nas páginas 48 e 53.
- GOODARZI, A. et al. Effect of Stereoscopic Anaglyphic 3-Dimensional Video Didactics on Learning Neuroanatomy. *World Neurosurgery*, Elsevier Inc, v. 107, p. 35–39, 2017. ISSN 18788769. Disponível em: <<https://doi.org/10.1016/j.wneu.2017.07.119>>. Citado na página 38.
- GORTLER, S. J. et al. The Lumigraph. *Siggraph*, p. 43–54, 1996. ISSN 00978930. Citado 2 vezes nas páginas 66 e 73.
- GOSSOW, D.; DECKER, P.; PAULUS, D. An evaluation of open source SURF implementations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 6556 LNAI, p. 169–179, 2011. ISSN 03029743. Citado 2 vezes nas páginas 68 e 73.
- GROVE, P. M.; HARROLD, A. The range of fusible horizontal disparities around the empirical horopters Philip M. Grove and Ashleigh Harrold School of Psychology, The University of Queensland, Australia. 2013. Citado 2 vezes nas páginas 59 e 72.

- GUARNIERI, M. An Historical Survey on Light Technologies. *IEEE Access*, IEEE, v. 6, p. 25881–25897, 2018. ISSN 21693536. Citado na página 39.
- HAN, B. Fast Calculating Orientation in Surf. *Test*, p. 165–168, 2010. Citado 2 vezes nas páginas 68 e 73.
- HEINZLE, S. et al. Computational stereo camera system with programmable control loop. *ACM SIGGRAPH 2011 papers on - SIGGRAPH '11*, v. 1, n. 212, p. 1, 2011. ISSN 07300301. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1964921.1964989>>. Citado 2 vezes nas páginas 63 e 72.
- HERNANDEZ, C.; VOGIATZIS, G.; CIPOLLA, R. Multiview photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 3, p. 548–554, 2008. ISSN 01628828. Citado 2 vezes nas páginas 66 e 73.
- HINCKLEY, K. et al. Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction*, v. 5, n. 3, p. 260–302, 1998. ISSN 10730516. Disponível em: <<http://portal.acm.org/citation.cfm?doid=292834.292849>>. Citado 2 vezes nas páginas 52 e 53.
- HUARD, S. Polarization of Light. n. 1, p. 1–9, 1996. Citado 3 vezes nas páginas 61, 62 e 72.
- HUI, J. Approach to the Interior Design Using Augmented Reality Technology. *Proceedings - 2015 6th International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2015*, p. 163–166, 2016. Citado na página 37.
- ILLINGWORTH, J.; HILTON, A. Looking to build a model world: automatic construction of static object models using computer vision. *Electronics and Communication Engineering Journal*, n. June, p. 103–113, 1998. ISSN 09540695. Citado 2 vezes nas páginas 64 e 72.
- IZQUIERDO, M. A. G. et al. Sub-pixel measurement of 3D surfaces by laser scanning. *Sensors and Actuators, A: Physical*, v. 76, n. 1-3, p. 1–8, 1999. ISSN 09244247. Citado 2 vezes nas páginas 48 e 53.
- JANKOVIC, M. P.; JOVICIC, N. S. Design and implementation of the setup for laser 3D scanning. *2017 25th Telecommunications Forum, TELFOR 2017 - Proceedings*, v. 2017-January, p. 1–4, 2018. Citado na página 37.
- JAVIDI, B.; HONG, S. H.; MATOBA, O. Multidimensional optical sensor and imaging system. *Applied Optics*, v. 45, n. 13, p. 2986–2994, 2006. ISSN 15394522. Citado na página 38.
- JUNG, Y. J.; SOHN, H.; RO, Y. M. Visual discomfort visualizer using stereo vision and time-of-flight depth cameras. *IEEE Transactions on Consumer Electronics*, v. 58, n. 2, p. 246–254, 2012. ISSN 00983063. Citado 2 vezes nas páginas 63 e 72.
- KASPER, M.; HECKMAN, C. Multiple Point Light Estimation from Low-Quality 3D Reconstructions. *Proceedings - 2019 International Conference on 3D Vision, 3DV 2019*, IEEE, p. 738–746, 2019. Citado 2 vezes nas páginas 38 e 39.
- KAZOKA, D.; PILMANE, M. 3D dissection tools in Anatomage supported interactive human anatomy teaching and learning. *SHS Web of Conferences*, v. 68, p. 02015, 2019. Citado na página 37.

KHADKA, A. R.; REMAGNINO, P.; ARGYRIOU, V. Object 3D Reconstruction Based on Photometric Stereo and Inverted Rendering. *Proceedings - 14th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2018*, IEEE, p. 208–215, 2018. Citado na página 38.

KIGKA, V. I. et al. 3D reconstruction of coronary arteries and atherosclerotic plaques based on computed tomography angiography images. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 40, p. 286–294, 2018. ISSN 17468108. Disponível em: <<https://doi.org/10.1016/j.bspc.2017.09.009>>. Citado na página 38.

KIM, C. et al. Multi-perspective stereoscopy from light fields. *ACM Transactions on Graphics*, v. 30, n. 212, p. 1, 2011. ISSN 07300301. Citado 2 vezes nas páginas 55 e 72.

KING, R. D. A brief history of stereoscopy. *Wiley Interdisciplinary Reviews: Computational Statistics*, v. 5, n. 4, p. 334–340, 2013. ISSN 19395108. Citado na página 39.

KLECKA, J.; BOSTIK, O. Effects of environment model parametrization on photogrammetry reconstruction. *Mendel*, v. 24, n. 1, p. 151–158, 2018. ISSN 25713701. Citado na página 38.

KOH, N. A. et al. Kohan N et al. Barriers of self-directed learning in virtual environment Self-directed learning barriers in a virtual environment: a qualitative study. *J Adv Med Educ Prof*, v. 5, n. 3, p. 116, 2017. Citado na página 37.

KOUTSOMITROPOULOS, D. A.; SOLOMOU, G. D. A learning object ontology repository to support annotation and discovery of educational resources using semantic thesauri. *IFLA Journal*, v. 44, n. 1, p. 4–22, 2018. ISSN 17452651. Citado na página 37.

KUMAR, J. MALUS ' S LAW OF POLARISATION. n. 2, p. 125–134, 2010. Citado 2 vezes nas páginas 61 e 72.

KUTULAKOS, K. N.; VALLINO, J. R. Calibration-Free Augmented Reality. v. 4, n. 1, p. 1–20, 1998. Citado 2 vezes nas páginas 64 e 72.

LACOTTE, B. Elimination of keystone and crosstalk effects in stereoscopic video. n. 95-31, 1995. Citado 4 vezes nas páginas 57, 58, 59 e 72.

LAGES, J.; GIUST, R.; VIGOUREUX, J. M. Composition law for polarizers. *Physical Review A - Atomic, Molecular, and Optical Physics*, v. 78, n. 3, p. 1–14, 2008. ISSN 10502947. Citado 2 vezes nas páginas 61 e 72.

LAMBOOIJ, M. et al. Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. *Journal of Imaging Science and Technology*, v. 53, n. 3, p. 030201, 2009. ISSN 10623701. Disponível em: <<http://www.ingentaconnect.com/content/ist/jist/2009/00000053/00000003/art00001>>. Citado 2 vezes nas páginas 63 e 72.

LERTWIRAM, S. et al. A design and implementation for ultrasound practice system for fetal echocardiography. *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry - VRCAI '11*, p. 367, 2011. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2087756.2087817>>. Citado 2 vezes nas páginas 51 e 53.

LEVOY, M.; HANRAHAN, P. Light Field Rendering. 1996. Citado 2 vezes nas páginas 64 e 72.

LEVOY, M. et al. The Digital Michelangelo Project: 3D Scanning of Large Statues. *Conference on Computer Graphics and Interactive Techniques*, p. 131–144, 2000. ISSN 1581132085. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.5681>>. Citado 4 vezes nas páginas 48, 49, 53 e 54.

LIVATINO, S. et al. Stereoscopic visualization and 3-D technologies in medical endoscopic teleoperation. *IEEE Transactions on Industrial Electronics*, v. 62, n. 1, p. 525–535, 2015. ISSN 02780046. Citado 2 vezes nas páginas 61 e 72.

LORENSEN, W. E.; CLINE, H. E. Marching cubes: A high resolution 3D surface construction algorithm. *Proceedings of the 14th annual conference on Computer graphics and interactive techniques - SIGGRAPH '87*, v. 21, n. 4, p. 163–169, 1987. ISSN 00978930. Disponível em: <<http://portal.acm.org/citation.cfm?doid=37401.37422>>. Citado 2 vezes nas páginas 50 e 53.

LUCHT, J. V. D. et al. STRUCTURED-LIGHT BASED 3D LASER SCANNING OF SEMI-SUBMERGED STRUCTURES. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 4, n. 2, p. 287–294, 2018. ISSN 21949050. Citado na página 38.

MARKIEWICZ, J.; ZAWIESKA, D.; PODLASIAK, P. Comparing multi-source photogrammetric data during the examination of verticality in a Monumental tower. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, v. 42, n. 2W3, p. 475–480, 2017. ISSN 16821750. Citado 2 vezes nas páginas 47 e 53.

MARTÍN, J. G. Possibilities for the use of Anatomage (the Anatomical Real Body-Size Table) for Teaching and Learning Anatomy with the Students. *Biomedical Journal of Scientific & Technical Research*, v. 4, n. 4, 2018. Citado na página 37.

MATUSIK, W. et al. Image-based 3D photography using opacity hulls. *Proceedings of the 29th annual conference on Computer graphics and interactive techniques - SIGGRAPH '02*, p. 427, 2002. ISSN 07300301. Disponível em: <<http://portal.acm.org/citation.cfm?doid=566570.566599>>. Citado 2 vezes nas páginas 65 e 73.

MEESTERS, L. M.; IJSELSTEIJN, W. A.; SEUNTIËNS, P. J. A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 14, n. 3, p. 381–391, 2004. ISSN 10518215. Citado 2 vezes nas páginas 62 e 72.

MENDES, C. C. T.; OSÓRIO, F. S.; WOLF, D. F. Real-time obstacle detection using range images: Processing dynamically-sized sliding windows on a GPU. *Robotica*, v. 35, n. 1, p. 85–100, 2017. ISSN 14698668. Citado 2 vezes nas páginas 71 e 73.

MOON, I.; JAVIDI, B. Three-dimensional visualization of objects in scattering medium by use of computational integral imaging. *Optics Express*, v. 16, n. 17, p. 13080, 2008. ISSN 1094-4087. Citado na página 38.

MOTA, A. R.; ROSA, C. T. W. d. Ensaio sobre metodologias ativas: reflexões e propostas. *Revista Espaço Pedagógico*, v. 25, n. 2, p. 261–276, 2018. ISSN 0104-7469. Citado na página 37.

MOYA, E. C. Using Active Methodologies: The Students' View. *Procedia - Social and Behavioral Sciences*, v. 237, n. June 2016, p. 672–677, 2017. ISSN 18770428. Citado na página 37.

- MURRAY, J. Some perspectives on visual depth perception. *ACM SIGGRAPH Computer Graphics*, v. 28, n. 2, p. 155–157, 1994. ISSN 00978930. Disponível em: <<http://portal.acm.org/citation.cfm?doid=178951.178985>>. Citado 2 vezes nas páginas 55 e 72.
- NABIZADEH, A. H. et al. Adaptive learning path recommender approach using auxiliary learning objects. *Computers and Education*, Elsevier Ltd, v. 147, n. June 2019, p. 103777, 2020. ISSN 03601315. Disponível em: <<https://doi.org/10.1016/j.compedu.2019.103777>>. Citado na página 37.
- NAZARENKO, A. L. Blended Learning vs Traditional Learning: What Works? (A Case Study Research). *Procedia - Social and Behavioral Sciences*, Elsevier B.V., v. 200, n. October, p. 77–82, 2015. ISSN 18770428. Disponível em: <<http://dx.doi.org/10.1016/j.sbspro.2015.08.018>>. Citado na página 37.
- NICOLAE, C. et al. Photogrammetry applied to problematic artefacts. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, v. 40, n. 5, p. 451–456, 2014. ISSN 16821750. Citado 2 vezes nas páginas 66 e 73.
- NIOLA, V.; ROSSI, C.; SAVINO, S. A new real-time shape acquisition with a laser scanner: First test results. *Robotics and Computer-Integrated Manufacturing*, Elsevier, v. 26, n. 6, p. 543–550, 2010. ISSN 07365845. Disponível em: <<http://dx.doi.org/10.1016/j.rcim.2010.06.026>>. Citado 2 vezes nas páginas 49 e 53.
- NUR, G. et al. Sensitivity of the HVS for Binocular disparity Cue in 3D displays under different ambient illumination conditions. *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, n. 1, p. 459–460, 2012. ISSN 0747668X. Citado 2 vezes nas páginas 55 e 72.
- 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 1: CSEDU*, [S.l.]: SciTePress, 2018. p. 259–266. ISBN 978-989-758-291-2. Citado na página 79.
- OSORIO-ZULUAGA, G. A.; DUQUE-MENDEZ, N. D. Search and selection of learning objects in repositories: A review. *Proceedings - 13th Latin American Conference on Learning Technologies, LACLO 2018*, p. 513–520, 2018. Citado na página 37.
- OWDA, A.; BALSABARREIRO, J.; FRITSCH, D. Methodology for digital preservation of the cultural and patrimonial heritage: generation of a 3D model of the Church St. Peter and Paul (Calw, Germany) by using laser scanning and digital photogrammetry. *Sensor Review*, v. 38, n. 3, p. 282–288, 2018. ISSN 02602288. Citado 2 vezes nas páginas 47 e 53.
- PONZONI, F. J.; DISPERATI, A. A. Estereoscopia e paralaxe na determinação de distâncias verticais mediante o uso de fotografias aéreas. *Portal do Monitoramento de Queimadas e Incêndios*, n. 1, p. 1–21, 1992. Citado 2 vezes nas páginas 59 e 72.
- RAMOS, A. P.; PRIETO, G. R. 3D virtualization by close range photogrammetry indoor gothic church apses. the case study of church of San Francisco in Betanzos (La Coruña, Spain). *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, v. 40, n. 5W4, p. 201–206, 2015. ISSN 16821750. Citado na página 38.

- REDMOND, C. et al. Using reusable learning objects (RLOs) in wound care education: Undergraduate student nurse's evaluation of their learning gain. *Nurse Education Today*, Elsevier, v. 60, n. August 2017, p. 3–10, 2018. ISSN 15322793. Disponível em: <<https://doi.org/10.1016/j.nedt.2017.09.014>>. Citado na página 37.
- REMONDINO, F. From point cloud to surface: the modeling and visualization problem. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV, p. 24–28, 2003. Disponível em: <https://static.aminer.org/pdf/PDF/000/291/298/surface_reconstruction_from_large_point_clouds_using_virtual_shared_memory.pdf>. Citado 2 vezes nas páginas 47 e 53.
- ROMANONI, A. et al. Automatic 3D reconstruction of manifold meshes via delaunay triangulation and mesh sweeping. *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016*, 2016. Citado na página 37.
- RUANO, R.; JOUBIN, L. Fetal Lung Volume Estimated by 3-Dimensional Ultrasonography and Magnetic Resonance Imaging in Cases With Isolated Congenital Diaphragmatic Hernia Rodrigo. ... of *Ultrasound in ...*, p. 353–358, 2004. Disponível em: <<http://www.jultrasoundmed.org/content/23/3/353.short>>. Citado 2 vezes nas páginas 51 e 53.
- SCHWARTZ, C.; KLEIN, R. Acquisition and Presentation of Virtual Surrogates for Cultural Heritage Artefacts. *Electronic Visualisation and the Arts*, v. 7, p. 50–57, 2012. Citado 2 vezes nas páginas 66 e 73.
- SCHWARTZ, C.; RUITERS, R. WebGL-based streaming and presentation of objects with bidirectional texture functions. *Journal on Computing and ...*, v. 6, n. 3, p. 1–21, 2013. ISSN 15564673. Disponível em: <<http://dl.acm.org/citation.cfm?id=2499932>>. Citado 2 vezes nas páginas 66 e 73.
- SCOTT, E.; SORIA, A.; CAMPO, M. Adaptive 3D Virtual Learning Environments - A Review of the Literature. *IEEE Transactions on Learning Technologies*, IEEE, v. 10, n. 3, p. 262–276, 2017. ISSN 19391382. Citado na página 37.
- SHEN, Y.; MIAO, Z. Oriented Gradients for Human Action Recognition. *Proceedings of the Second International Conference on Internet Multimedia Computing and Service*, p. 175–178, 2010. Disponível em: <<http://doi.acm.org/10.1145/1937728.1937770>>. Citado 2 vezes nas páginas 70 e 73.
- SHIBATA, T. et al. The zone of comfort: Predicting visual discomfort with stereo displays. v. 11, n. 8, 2012. Citado 2 vezes nas páginas 63 e 72.
- SHUM, H. Y.; SZELISKI, R. Stereo Reconstruction from Multiperspective Panoramas. *Proc. ICCV*, v. 26, n. 1, p. 45–62, 1999. Citado 2 vezes nas páginas 64 e 72.
- SILVA., F. de A 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 na página 80.
- ŚLUZEK, A.; PARADOWSKI, M. Real-time retrieval of near-duplicate fragments in images and video-clips. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 6474 LNCS, n. PART 1, p. 18–29, 2010. ISSN 03029743. Citado 2 vezes nas páginas 68 e 73.

SON, J. Y.; JAVIDI, B.; KWACK, K. D. Methods for displaying three-dimensional images. *Proceedings of the IEEE*, v. 94, n. 3, p. 502–522, 2006. ISSN 00189219. Citado 2 vezes nas páginas 62 e 72.

SOUZA, H. A. G. d. Processos técnicos e artísticos para realização de filme documentário 3d estereoscópico autor. *Congresso SET - Broadcast&Cable*, v. 03, p. 312–330, 2009. Citado 2 vezes nas páginas 67 e 73.

SÁRA, R. How to Teach Stereoscopic Matching. n. September, p. 239, 2010. ISSN 13342630. Disponível em: <<http://books.google.com/books?id=gICNQAACA AJ{&}pgi>>. Citado 2 vezes nas páginas 55 e 72.

SRIHARI, S. Representation of Three-Dimensional Digital Images. *Surveys*, v. 13, n. 4, p. 399–424, 1981. ISSN 15577341. Disponível em: <<http://portal.acm.org/citation.cfm?id=356862>>. Citado 2 vezes nas páginas 50 e 53.

STYTZ, M. R.; FRIEDER, G.; FRIEDER, O. Three-dimensional medical imaging: algorithms and computer systems. *ACM Computing Surveys*, v. 23, n. 4, p. 421–499, 1991. ISSN 03600300. Citado 2 vezes nas páginas 50 e 53.

TAKIMOTO, R. Y. et al. 3D reconstruction and multiple point cloud registration using a low precision RGB-D sensor. *Mechatronics*, v. 35, p. 11–22, 2016. ISSN 09574158. Citado 2 vezes nas páginas 49 e 53.

TANG, J. et al. Structured dynamic time warping for continuous hand trajectory gesture recognition. *Pattern Recognition*, Elsevier Ltd, v. 80, p. 21–31, 2018. ISSN 00313203. Disponível em: <<https://doi.org/10.1016/j.patcog.2018.02.011>>. Citado 2 vezes nas páginas 70 e 73.

TEKES, A. et al. 2D and 3D MRI features of classic bladder exstrophy. *Clinical Radiology*, The Royal College of Radiologists, v. 69, n. 5, p. e223–e229, 2014. ISSN 1365229X. Disponível em: <<http://dx.doi.org/10.1016/j.crad.2013.12.019>>. Citado 2 vezes nas páginas 52 e 53.

TERZIĆ, K.; HANSARD, M. Methods for reducing visual discomfort in stereoscopic 3D: A review. *Signal Processing: Image Communication*, v. 47, p. 402–416, 2016. ISSN 09235965. Citado 2 vezes nas páginas 63 e 72.

THOMAS, D. B. et al. 3D scanning and printing skeletal tissues for anatomy education. *Journal of Anatomy*, v. 229, n. 3, p. 473–481, 2016. ISSN 14697580. Citado na página 37.

TORI, R.; KIRNER, C.; SISCOOTTO, R. *Fundamentos e tecnologia de Realidade Virtual e Aumentada*. [S.l.: s.n.], 2006. 422 p. ISSN 0100-6991. ISBN 8576690683. Citado 4 vezes nas páginas 57, 59, 60 e 72.

TU, K. Y.; LIN, H. Y.; LEE, T. T. Rapidly constructing a simple 3-D graphic model using a consumer-user digital still camera. *IEEE Transactions on Consumer Electronics*, v. 52, n. 3, p. 819–824, 2006. ISSN 00983063. Citado 2 vezes nas páginas 65 e 73.

VASCO, M. A. A. et al. Utilização de tomografias computadorizadas de baixa resolução para construção de modelos geométricos detalhados de mandíbulas com e sem dentes. *Revista Internacional de Metodos Numericos para Calculo y Diseno en Ingenieria*, v. 32, n. 1, p. 1–6, 2016. ISSN 1886158X. Citado 2 vezes nas páginas 51 e 53.

- Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. [S.l.: s.n.], 2001. v. 1, p. I–I. ISSN 1063-6919. Citado 2 vezes nas páginas 67 e 73.
- VOGIATZIS, G. et al. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 12, p. 2241–2246, 2007. ISSN 01628828. Citado 2 vezes nas páginas 65 e 73.
- WANG, H. et al. 3-D Point Cloud Object Detection Based on Supervoxel Neighborhood With Hough Forest Framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 8, n. 4, p. 1570–1581, 2015. ISSN 21511535. Citado na página 37.
- WANG, L. et al. A segmentation and classification scheme for single tooth in MicroCT images based on 3D level set and k-means++. *Computerized Medical Imaging and Graphics*, Elsevier Ltd, v. 57, p. 19–28, 2017. ISSN 18790771. Disponível em: <<http://dx.doi.org/10.1016/j.compmedimag.2016.05.005>>. Citado na página 51.
- WANG, L. et al. Intermediate shadow maps for interactive many-light rendering. *Visual Computer*, Springer Berlin Heidelberg, v. 34, n. 10, p. 1415–1426, 2018. ISSN 01782789. Citado na página 38.
- WANG, S.; DU, Y.; HUANG, Z. Ear detection using fully convolutional networks. *Proceedings of the 2nd International Conference on Robotics, Control and Automation - ICRCA '17*, p. 50–55, 2017. Disponível em: <<http://dl.acm.org/citation.cfm?doid=3141166.3141168>>. Citado 2 vezes nas páginas 71 e 73.
- WANG, Z. et al. Wearables Can Afford: Light-weight Indoor Positioning with Visible Light. *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '15*, p. 465–465, 2015. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2742647.2745924>>. Citado 2 vezes nas páginas 61 e 72.
- WHEATSTONE, C. Xviii. contributions to the physiology of vision. —part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, v. 128, p. 371–394, 1838. Disponível em: <<http://rstl.royalsocietypublishing.org/content/128/371.short>>. Citado na página 39.
- WOOD, D. N. et al. Surface light fields for 3D photography. *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, p. 287–296, 2000. ISSN 0097-8930. Disponível em: <<http://doi.acm.org/10.1145/344779.344925>>. Citado 2 vezes nas páginas 64 e 73.
- WU, Z. et al. Direct information retrieval after 3D reconstruction in grating-based X-ray phase-contrast computed tomography. *Journal of Synchrotron Radiation*, v. 25, n. 4, p. 1222–1228, 2018. ISSN 16005775. Citado na página 38.
- YALNIZ, I. Z. et al. Ottoman archives explorer. *Journal on Computing and Cultural Heritage*, v. 2, n. 3, p. 1–20, 2009. ISSN 15564673. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1658346.1658348>>. Citado 2 vezes nas páginas 70 e 73.
- YANG, L. et al. Sliding window-based region of interest extraction for finger vein images. *Sensors (Switzerland)*, v. 13, n. 3, p. 3799–3815, 2013. ISSN 14248220. Citado 2 vezes nas páginas 71 e 73.

YEMEZ, Y.; SCHMITT, F. 3D Reconstruction of Real Objects With High Resolution Shape and Texture. *Image and Vision Computing*, v. 22, n. 13, p. 1137–1153, 2004. ISSN 02628856. Citado na página 65.

YEO, C. T. et al. Utility of 3D Reconstruction of 2D Liver Computed Tomography/Magnetic Resonance Images as a Surgical Planning Tool for Residents in Liver Resection Surgery. *Journal of Surgical Education*, Elsevier, v. 75, n. 3, p. 792–797, 2018. ISSN 18787452. Disponível em: <<http://dx.doi.org/10.1016/j.jsurg.2017.07.031>>. Citado na página 38.

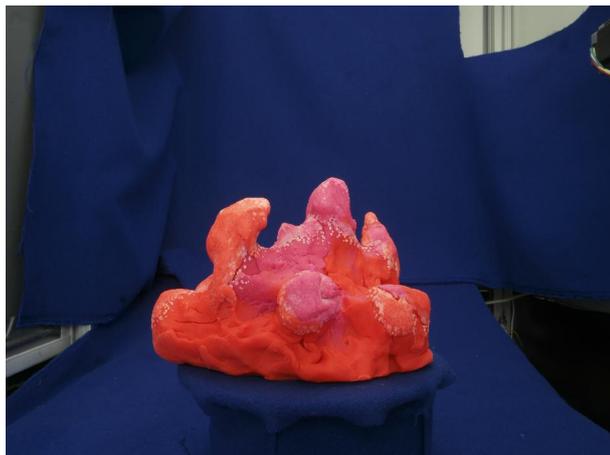
ZHANG, D. Study on interior decoration system design based on 3D scene modeling technology. *Proceedings - 2017 International Conference on Smart Grid and Electrical Automation, ICSGEA 2017*, v. 2017-January, p. 380–383, 2017. Citado na página 37.

ZONE, R. *Stereoscopic Cinema and the Origins of 3-D Film, 1838-1952*. University Press of Kentucky, 2007. ISBN 9780813124612. Disponível em: <<http://www.jstor.org/stable/j.ctt5hk0fn>>. Citado 2 vezes nas páginas 62 e 72.

Capítulo A – Objetos Adquiridos

Objetos adquiridos na terceira versão do F2S2

1) Massa de modelar



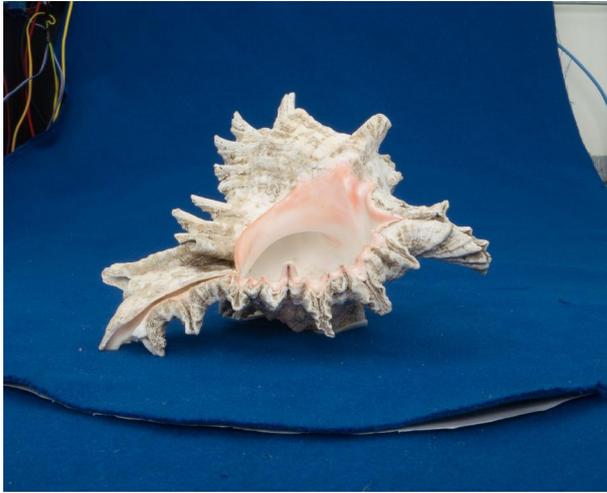
2) Papel



3) Boneco - Darth Maul - Star WarsTM



4) Concha do Mar - disponibilizada pelo doutorando Diogo Roberto Olsen - origem Moçambique



5) Boneco de Madeira com uma bola vermelha de plástico nos pés



6) Maçã



7) Parte interna de um motor de arranque - Peça de automóvel cedida pela Renault Curitiba-PR



8) Vaso de cerâmica



9) Metade da parte superior da perna de um cavalo, corte medial, realizado pela acadêmica da veterinária da PUCPR, Daniela Góes Turchenski, como parte do seu trabalho de conclusão de curso.



10) Metade da parte inferior da perna de um cavalo, corte medial, realizado pela acadêmica da veterinária da PUCPR, Daniela Góes Turchenski, como parte do seu trabalho de conclusão de curso.



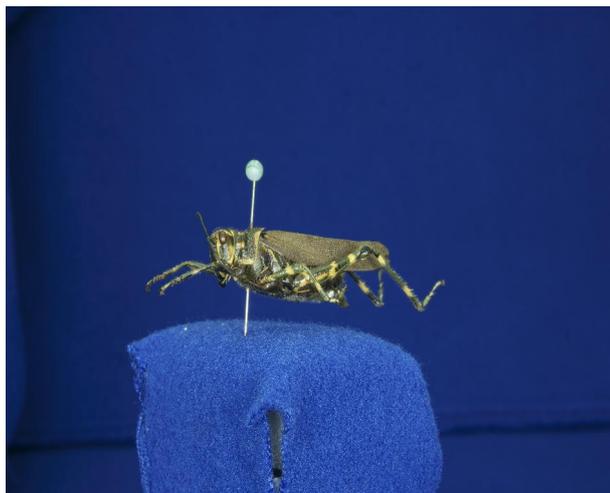
11) Carro metálico de brinquedo, FerrariTM



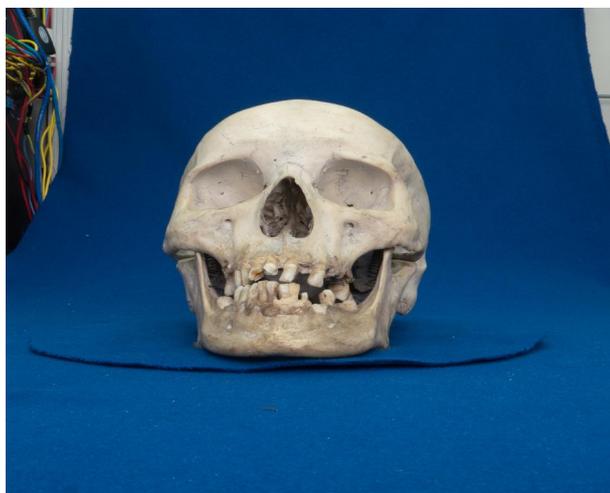
12) Açucareiro de barro em forma de galinha



13) Grilo



14) Crânio completo



Objetos adquiridos na segunda versão do F2S2

15) Crânio humano com a parte superior extraída (Crânio Aberto)



16) Crânio humano sem a mandíbula (Crânio Fechado)

