

Alessandra Bussador

Localização Automática de Placas de Veículos em Fotos Digitais Utilizando Abordagem Granulométrica

Dissertação apresentada à Pontifícia Universidade Católica do Paraná para a obtenção do título de Mestre em Informática Aplicada.

Área de concentração:
Análise e Reconhecimento de Documentos

Orientador:
Prof. Dr. Jacques Facon

Curitiba
2009

Alessandra Bussador

Localização Automática de Placas de Veículos em Fotos Digitais Utilizando Abordagem Granulométrica

Dissertação apresentada à Pontifícia Universidade Católica do Paraná para a obtenção do título de Mestre em Informática Aplicada.

Curitiba
2009

Aos meus dois grandes amores, Miguel e Giovanna.

Agradecimentos

Ao meu marido Miguel, pelas grandes idéias, carinho e compreensão nas horas mais difíceis.

Aos meus pais, Jair e Elizabete, pelo constante apoio e carinho.

Ao meu orientador Jacques, pela amizade, ajuda e observações que foram determinantes para o término deste trabalho.

Aos meus colegas do Lardoc, pelo apoio dado e pelas suas grandes amizades.

A todos os amigos, que porventura esqueci de mencionar, que fazem ou fizeram parte deste trabalho.

A Pontifícia Universidade Católica do Paraná pelo imprescindível apoio financeiro.

Sumário

Agradecimentos	ii
Sumário	iii
Lista de Figuras	vi
Lista de Tabelas	viii
Lista de Símbolos	ix
Resumo	x
Abstract	xi
Capítulo 1	
Introdução	1
1.1 Motivação.....	2
1.2 Objetivos	3
1.3 Conteúdo da dissertação	3
Capítulo 2	
Estado da Arte	5
2.1 Definições básicas.....	5
2.1.1 Definição de imagem	6
2.1.2 Sistema de processamento e análise de imagens.....	6
2.2 Tipos de análises de layout de uma imagem	7
2.2.1 Análise estrutural.....	7
2.2.2 Análise funcional	8
2.3 Segmentação de imagens	9
2.4 Técnicas de segmentação de imagens binárias.....	10
2.4.1 Técnicas estatísticas.....	10

2.4.2 Segmentação por limiarização	11
2.5 Técnicas de segmentação de imagens em níveis de cinza	15
2.5.1 Textura e redes neurais	15
2.5.2 Textura e granulometria local	17
2.5.3 Filtros de Gibbs	17
2.5.4 Tophat diferencial.....	17
2.6 Técnicas de segmentação de placas de veículos.....	18

Capítulo 3

Morfologia Matemática	22
3.1 Operadores morfológicos binários	22
3.1.1 Dilatação	23
3.1.2 Erosão.....	23
3.1.3 Abertura.....	24
3.1.4 Fechamento.....	24
3.1.5 Reconstrução.....	25
3.1.6 Granulometria	25
3.2 Operadores morfológicos em níveis de cinza.....	26
3.2.1 Dilatação	27
3.2.2 Erosão.....	27
3.2.3 Abertura.....	28
3.2.4 Fechamento.....	29
3.2.5 Reconstrução dual	30
3.2.6 Granulometria linear em níveis de cinza	31
3.2.7 Resíduos morfológicos	33
3.2.8 Mapeamento a partir de resíduos morfológicos	33
3.2.9 Tophat por abertura	35
3.2.10 Tophat por fechamento.....	36

Capítulo 4

Metodologia e Desenvolvimento	38
4.1 Base de dados	38
4.1.1 Características da base de dados	39
4.2 Desenvolvimento	41
4.2.1 Aquisição da imagem.....	42
4.2.2 Conversão das imagens coloridas para níveis de cinza	43

4.2.3 Tophat por fechamento.....	44
4.2.4 Redução da complexidade da imagem	45
4.2.5 Granulometria em níveis de cinza	45
4.2.6 Reconstrução dual	46
4.2.7 Localização dos componentes.....	47
4.2.8 Filtragem de componentes	48
4.2.9 Localização da placa	49

Capítulo 5

Análise de Resultados	52
5.1 Análise da filtragem dos componentes.....	53
5.2 Análise da localização da placa	54
5.3 Comparação entre os resultados da filtragem dos componentes e da localização da placa.....	57
5.4 Comparação com o sistema SIAV.....	57
5.5 Casos particulares.....	59
5.5.1 Localização em veículos com pintura clara	59
5.5.2 Placas posicionadas incorretamente.....	60
5.5.3 Fator de luminosidade.....	60
5.5.4 Foto com mais de um veículo	61
5.5.5 Localização da placa em veículos com adesivos	62
5.5.6 Fotos tiradas em ângulos	63
Conclusão	64
6.1 Contribuição do trabalho	65
6.2 Perspectivas futuras	65
Referências Bibliográficas	66

Lista de Figuras

Figura 1 - Tipos de placas nacionais.....	2
Figura 2 - Análise estrutural e funcional [GOR95].....	8
Figura 3 - Estrutura física de uma página [SIM97]	11
Figura 4 - Seleção dos dígitos da placa	19
Figura 5 - Placa binarizada	19
Figura 6 - Placa segmentada - 7 entidades encontradas	19
Figura 7 – Exemplo de dilatação.....	27
Figura 8 - Exemplo de erosão.....	28
Figura 9 - Imagem aberta.....	29
Figura 10 - Imagem fechada.....	30
Figura 11 - Reconstrução Dual em níveis de cinza	31
Figura 12 – Relação do máximo de uma região.....	32
Figura 13 - Contribuições das zonas de máximo.....	32
Figura 14 - Linha de uma imagem em níveis de cinza	34
Figura 15 – Diferença entre a imagem original e a imagem aberta	36
Figura 16 – Diferença entre a imagem original e a imagem fechada	37
Figura 17 - Etapas do processamento.....	41
Figura 18 - Imagem exemplo de um carro estacionado	43
Figura 19 - Imagem convertida para níveis de cinza.....	44
Figura 20 - Processamento de tophat	45
Figura 21 - Redução de complexidade	45
Figura 22 - Imagem de marcadores	46
Figura 23 - Reconstrução dual	47
Figura 24 - Localização de componentes	48
Figura 25 - Filtragem de componentes	49
Figura 26 - Localização da placa	50
Figura 27 - Segmentação correta da placa	53
Figura 28 - Segmentação parcial da placa.....	54

Figura 29 - Localização correta da placa	55
Figura 30 - Localização parcial da placa	56
Figura 31 - Recorte para a retirada da cena	57
Figura 32 - Veículo com pintura clara	59
Figura 33 – Veículo com a placa inclinada	60
Figura 34 – Placa sombreada.....	61
Figura 35 – Imagem contendo duas placas	62
Figura 36 – Veículo com adesivos	62
Figura 37 – Foto com ângulo normal.....	63
Figura 38 – Foto com ângulo acentuado.....	63

Lista de Tabelas

Tabela 1 - Mapeamento dos pontos da Figura 12	35
Tabela 2 – Posicionamento dos veículos	40
Tabela 3 - Características de iluminação	40
Tabela 4 - Intensidade de iluminação das imagens homogêneas.....	40
Tabela 5 - Resultados da filtragem de componentes.....	53
Tabela 6 - Resultados da localização da placa.....	55

Lista de Símbolos

PDI	Processamento Digital de Imagens
RLSA	<i>Run-Length Smoothing Algorithm</i>
WWW	<i>World Wide Web</i>
ISO	<i>International Standards Organization</i>
dpi	<i>dots per inch</i>
PUC-PR	Pontifícia Universidade Católica do Paraná
CD-R	<i>Compact Disc Recordable</i>
GDOC	Gestão de Hiperdocumentos
Lardoc	Laboratório de Análise e Reconhecimento de Documentos
BROPI	Biblioteca de Rotinas de Processamento de Imagens
JPEG	Joint Photographic Experts Group
BMP	Bitmap – mapa de bits
OCR	Optical Character Recognition
SIAV	Sistema de Identificação Automática de Veículos
Contran	Conselho Nacional de Trânsito

Resumo

A análise de imagens é uma área importante na sociedade moderna, pois ajuda a solucionar problemas referentes à extração de informações a partir de imagens digitalizadas envolvendo diversas áreas, tais como: medicina, geologia, biologia, radar, satélite, automação industrial, dentre tantas outras. Devido à crescente necessidade de automação na resolução de problemas em diversas aplicações e com o crescente uso de processamento e análise de imagens, houve a necessidade de se estruturar uma metodologia de processamento. Esta metodologia é subdividida em aquisição, pré-processamento, segmentação, reconhecimento e interpretação.

Neste trabalho, consideremos especificamente as etapas de pré-processamento e segmentação visando a capacidade de localizar a placa de um veículo qualquer através da utilização de técnicas de morfologia matemática e reconstrução.

Para a realização deste trabalho foi necessária a criação de uma base de dados, contendo 180 imagens digitais de veículos estacionados e o desenvolvimento de técnicas algorítmicas para a utilização de ferramentas de morfologia matemática, tais como granulometria em níveis de cinza e tophat, para o processamento das imagens no intuito de segmentar e localizar as placas dos veículos. A aplicação do processo de segmentação resultou em uma localização 75% das placas dos veículos.

Abstract

Image analysis is an important area within modern society because helps to resolve problems such as information extraction from digital images encompassing several areas, such as: medicine, geology, biology, radar, satellite, industrial automation, and many other. Due to the increasing need to automation in the problem solving in several applications and with the images processing and analysis increasing use, there was the need to structure a processing methodology. This methodology is subdivided in acquisition, pre-processing, segmentation, recognition and interpretation.

In this work, will be considered namely the preprocessing and segmentation processes, looking forward to the ability to locate the license plate of a vehicles, via the application of techniques such as mathematical morphology and reconstruction.

For the accomplishment of this work it was required the creation of a data base, containing 180 digital images of parked cars and the development of algorithmic techniques to the application of mathematical morphology tools , such as gray level granulometry and tophat, in order to process the images to segment and find the license plates of the vehicles. The segmentation process achieved a 75% rate in correctly segmenting and finding the license plates.

Capítulo 1

Introdução

Devido à crescente necessidade de automação na resolução de problemas em diversas aplicações, e com o crescente uso de processamento e análise de imagens, houve a necessidade de se estruturar o desenvolvimento de algumas aplicações como:

- medição e planejamento do fluxo de tráfego;
- identificação de veículos para recuperação em caso de furto;
- controle automático de pedágios e estacionamentos pagos e aplicação das leis de trânsito (identificação automática de veículos infratores, estacionamento em áreas proibidas, etc.).

Neste trabalho, consideraremos especificamente os processos de pré-processamento e segmentação visando a capacidade de localizar a placa de um veículo particular através da utilização de técnicas de morfologia matemática, reconstrução, granulometria e mapeamento a partir de resíduos morfológicos. Estas imagens devem ser segmentadas para se identificar as regiões onde estão localizados os dígitos da placa e posteriormente tratar cada uma destas da devida forma. Por exemplo, as regiões encontradas nas fotos digitais podem ser submetidas a um OCR (*Optical Character Recognition*) para que o texto seja reconhecido.

1.1 Motivação

O Conselho Nacional de Trânsito (Contran) estabeleceu oito diferentes tipos de placas de carros, que tem a função de distinguir os diferentes usos dos veículos. A resolução n.º 46 é a mais abrangente, prevendo diferentes tipos de placas para veículos particulares (de passeio ou comerciais), aluguel, experiência (concessionárias), aprendizagem (auto-escolas), fabricantes (montadoras), missões diplomáticas, de coleção e oficiais [GAZ01]. Na Figura 1 é mostrado um exemplo de cada um dos diferentes de placas vigentes no País.



Figura 1 - Tipos de placas nacionais

Estas colocações motivaram a elaboração de uma dissertação de mestrado para a criação de um módulo para processar automaticamente fotos digitais de veículos em níveis de cinza, com o objetivo de segmentar os dígitos das placas de veículos particulares, que são a maioria dos veículos que circulam no país, utilizando para isto uma nova metodologia utilizando as técnicas de morfologia matemática, reconstrução, granulometria e mapeamento a partir de resíduos morfológicos. Este processo é realizado através de várias etapas, dentre as quais podemos citar:

aquisição ou digitalização, pré-processamento e segmentação, que serão descritas nos Capítulos 2 e 3.

1.2 Objetivos

O objetivo principal deste trabalho é a segmentação de placas de veículos particulares em imagens em níveis de cinza, sendo que esta segmentação é independente de ângulos.

Para desenvolver este trabalho, os estudos preliminares indicaram que a utilização das técnicas de granulometria e tophat por reconstrução deverão auxiliar na solução do problema acima descrito.

Assim, a contribuição desta dissertação é a apresentação de uma nova metodologia para a segmentação de imagens de veículos (fotos digitais) em níveis de cinza utilizando somente técnicas morfológicas.

1.3 Conteúdo da dissertação

Este trabalho contém os seguintes capítulos:

Capítulo 2: Estado da Arte

Apresentação de algumas técnicas de segmentação de imagens mais utilizadas, binários e em níveis de cinza e também técnicas de segmentação específicas para placas de veículos.

Capítulo 3: Morfologia Matemática

Estabelecemos a base teórica para o processo de morfologia matemática como: granulometria, tophat e reconstrução morfológica e descrevemos as técnicas utilizadas para a construção dos algoritmos de extração das características dos componentes da imagem.

Capítulo 4: Metodologia e Desenvolvimento

Descrevemos a base de dados utilizada para os testes e também o método de filtragem das imagens a partir de sua decomposição pela granulometria, definindo os atributos, as propriedades e o algoritmo para a filtragem.

Capítulo 5: Análise de Resultados

Mostramos os resultados obtidos por granulometria e as análises estatísticas do método utilizado para a segmentação e extração dos dígitos das placas.

Capítulo 6: Conclusão

Fazemos um breve resumo referente às conclusões do trabalho e alguns possíveis trabalhos futuros.

Capítulo 2

Estado da Arte

Neste capítulo serão apresentadas as técnicas utilizadas no processamento de imagens de documentos. Este estudo foi direcionado para as técnicas de segmentação mais utilizadas, caracterizando uns dos objetivos do trabalho que é somente a utilização de morfologia matemática para a segmentação e não uma composição destas técnicas. Também foi baseado no processamento de imagens de documentos pois este trabalho considera as imagens dos veículos como uma imagem de documento complexa, ou seja, uma imagem que não apresenta uma estrutura definida (*layout* do documento, conforme definido no item 2.2).

Fazendo uma análise superficial de uma imagem da placa de um veículo, tem-se a presença de caracteres. Assim foram pesquisadas técnicas que abordassem os problemas de localização e segmentação de caracteres em imagens, sendo estas específicas para texto ou não.

Este capítulo apresenta os tipos de análise de *layout*, técnicas para segmentação de imagens binárias e em tons de cinza e técnicas específicas sobre segmentação de placas de veículos.

2.1 Definições básicas

O processamento digital de imagens pode ser classificado segundo [FAC93] quanto ao grau de abstração em três níveis distintos: baixo, médio e alto, ocorrendo uma redução progressiva da quantidade de informações manipuladas à medida que

se passa por níveis crescentes de abstração. No processamento de baixo nível, os dados de entrada são pixels da imagem original e os dados de saída representam propriedades da imagem, na forma de valores numéricos associados a cada pixel. No processamento de nível médio este conjunto de valores produz como resultado uma lista de características. O processamento de alto nível produz, a partir destas características, uma interpretação do conteúdo da imagem. Estas especificações supõem uma diversificação dos processamentos.

2.1.1 Definição de imagem

Em geral, uma imagem pode ser descrita por uma função bidimensional $F(x,y)$, onde (x,y) é a coordenada espacial e $F(x,y)$ o valor da característica em (x,y) . Dependendo do tipo da imagem, o valor da característica pode ser intensidade de luz, profundidade, intensidade de onda de rádio ou temperatura. A imagem digital, por outro lado, é uma função discreta bidimensional $f(x,y)$ digitalizada pela coordenada espacial e pela magnitude do valor da característica. Uma imagem digital pode ser vista como uma matriz bidimensional, cujos índices linha e coluna identificam um pixel na imagem, e o valor do elemento da matriz correspondente identifica o nível de intensidade da característica [JES99].

2.1.2 Sistema de processamento e análise de imagens

Um sistema de processamento e análise de imagens [FAC93] pode ser constituído dos seguintes elementos:

Aquisição ou digitalização - A imagem capturada pelo sensor é transformada em uma imagem digital sobre a forma de uma matriz de valores discretos chamados pixels, abreviação de *picture element*, ou elemento de figura;

Pré-processamento - Esta etapa permite corrigir um certo número de imperfeições e defeitos que aparecem na imagem decorrentes de problemas na aquisição, cujas causas podem ser as características físicas do sistema, as condições

de iluminação, imperfeições no sistema de lentes da câmara, entre outras. O pré-processamento não é indispensável, mas na maioria dos casos é necessário;

Segmentação – O objetivo principal da segmentação é subdividir uma imagem em suas partes constitutivas. Em uma imagem natural a segmentação é realizada pela detecção de descontinuidades e/ou regiões na imagem. A maioria dos processos são baseados na pesquisa destas entidades, que são armazenadas sobre uma forma adequada, segmentos ou primitivas;

Representação – O objetivo da representação é elaborar uma estrutura adequada, agrupando os resultados das etapas anteriores e o armazenamento dos diversos padrões que contém o conhecimento a priori a respeito do documento. A representação é efetuada após a obtenção das primitivas, e as vezes, ao mesmo tempo. Algumas vezes a representação permite medir as propriedades das formas resultantes da segmentação;

Interpretação/Classificação – Este é o processo mais “inteligente” em um sistema de visão por computador, pois representa o nível mais alto na abstração e permite obter a descrição do fenômeno representado na imagem, fazendo uso do conhecimento a priori dos casos estudados, e também utilizando os conhecimentos adquiridos nas fases anteriores do processo.

2.2 Tipos de análises de layout de uma imagem

2.2.1 Análise estrutural

A análise estrutural do *layout* (pode ser chamado também de análise física ou geométrica do *layout*) obtém a segmentação física entre os grupos de componentes de uma imagem.

A análise estrutural pode ser realizada de modo *top-down* ou *bottom-up*. Para a análise *top-down*, inicia-se encontrando os limites dos grupos que existem na imagem e vai subdividindo sucessivamente em grupos menores até que todos os blocos sejam encontrados. Na análise *bottom-up*, os componentes conectados são

fundidos em objetos maiores e assim por diante, até que todos os blocos sejam encontrados. Outro método muito utilizado pelos autores é a combinação das duas técnicas citadas (*top-down* e *bottom-up*).

2.2.2 Análise funcional

A análise funcional do *layout* (também chamado análise sintática ou lógica do *layout*) utiliza informações de domínios dependentes consistindo de regras de *layout* para uma imagem particular para realizar a rotulação das estruturas de blocos dando algumas indicações das funções deste bloco [GOR95].

A Figura 2 apresenta um exemplo das análises funcional e estrutural em uma página de um artigo técnico.

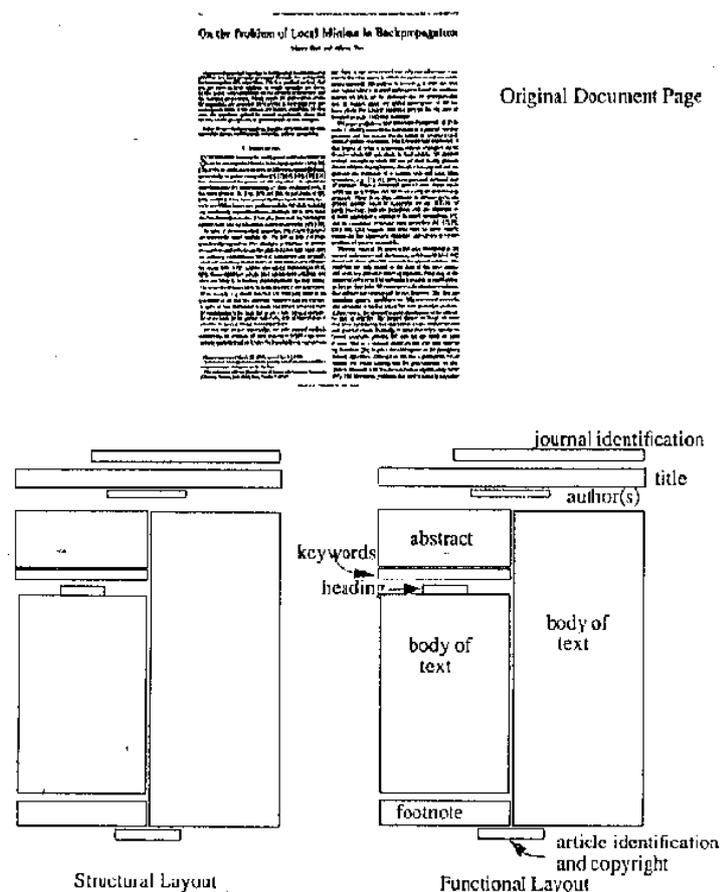


Figura 2 - Análise estrutural e funcional [GOR95]

2.3 Segmentação de imagens

A segmentação de imagens é um processo que tipicamente particiona o domínio espacial de uma imagem em subconjuntos mutuamente exclusivos, chamados regiões, onde cada região é uniforme e homogênea com respeito a algumas propriedades como tom ou textura e cujos valores diferem, em alguns aspectos e significados, das propriedades de cada região vizinha.

Haralick, em [HAR85], propôs as seguintes propriedades para um bom processo de segmentação:

- regiões de uma imagem segmentada devem ser uniformes e homogêneas com respeito a alguma característica. Por exemplo, nível de cinza ou textura;
- regiões adjacentes devem ter valores significativamente diferentes com respeito à característica segmentada;
- interior das regiões deve ser simples e sem falhas (buracos).

Entre as diferentes áreas de aplicação em que a segmentação atua, encontram-se variados tipos de imagem:

- Imagem de intensidade de luz: representa a variação da intensidade da luz na cena;
- Imagem de profundidade: é o mapeamento da informação de profundidade em diferentes pontos da cena;
- Imagem de ressonância magnética nuclear: representa a variação de intensidade de ondas de rádio geradas por sistemas biológicos, quando expostas aos pulsos de frequência de rádio;
- Imagem de temperatura: infravermelho;

- Imagem de ultra-sonografia: representa a emissão, reflexão e recepção de ondas.

Na literatura, encontram-se várias técnicas de segmentação, mas não existe um método simples que seja bom para todos os tipos de imagens, e nem todos os métodos são igualmente bons para um tipo particular de imagem.

2.4 Técnicas de segmentação de imagens binárias

2.4.1 Técnicas estatísticas

Simon em [SIM97] mostra uma abordagem *bottom-up* para processar documentos técnicos da área química. O primeiro passo do algoritmo é encontrar os componentes do documento. A árvore de custo mínimo é construída pela inserção da menor das distâncias entre componentes ainda não usada. Assim, em cada passo do algoritmo, o estado atual contém um certo número de componentes que possuem a menor distância interna no nível atual (inicialmente todos os vértices da árvore estão em componentes diferentes). Desta forma estes componentes possuem a maior coesão no nível corrente.

No ponto em que a próxima menor distância disponível dobrar ou triplicar a última distância utilizada ocorre uma mudança de nível. (O algoritmo de Kruskal é utilizado para montar a árvore de menor caminho no grafo de distâncias entre componentes). A classificação dos componentes é feita nos primeiros níveis (palavras, linhas, blocos), isto porque existem algumas heurísticas que são utilizadas de forma diferenciada para cada tipo de componente.

A Figura 3 apresenta uma representação da estrutura física de uma página de um documento.

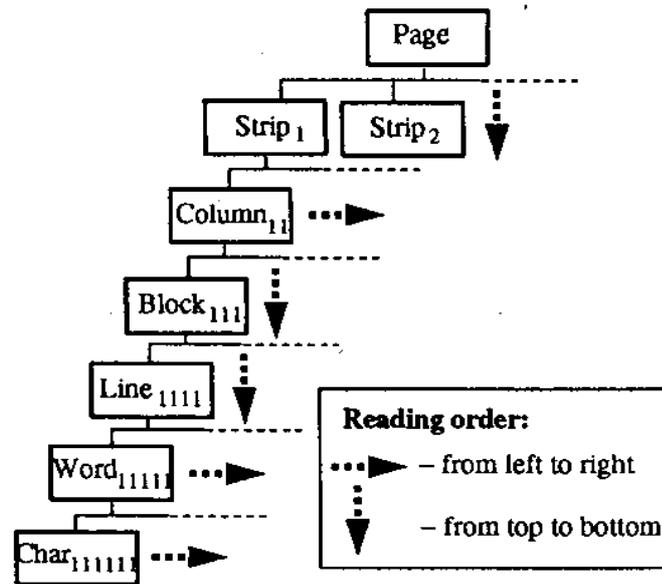


Figura 3 - Estrutura física de uma página [SIM97]

Jain em [JAI98] usou uma abordagem baseada na extração de componentes conectados para implementar a segmentação de imagens de documentos e identificação das suas regiões. Um novo modelo de documento o qual preserva informações geradas em top-down é proposto baseado no fato de que um documento é logicamente representado por edições iterativas, armazenamento, recuperação, transferência e análise lógica. Este método é aplicado para documentos de vários jornais e pode se ajustar a quantias moderadas de inclinação e ruído.

O algoritmo foi testado com 150 imagens (125 imagens com 300 dpi – banco de dados de documentos, 10 imagens com 100 dpi – texto japonês, e imagens de artigos com 100 e 300 dpi). Os principais erros na segmentação dos documentos foram: tabelas sem conexão, títulos/legenda em uma linha simples, colunas com aberturas pequenas, parágrafos começando com caracter grande, títulos/cabeçalhos escassos, espaços largos e desenhos escassos ou muito claros.

2.4.2 Segmentação por limiarização

A limiarização de uma imagem digital é um método que se baseia no histograma da imagem, buscando encontrar regiões bem definidas, afim de poder efetuar a divisão da imagem em objetos ou regiões. A continuidade dos níveis de

cinza é a primitiva de maior valor na segmentação por região. Assim, a limiarização efetua a subdivisão da imagem em função das regiões realmente significativas contidas no seu histograma [FAC96].

Seja uma imagem digital qualquer, onde $f(x,y)$ representa o nível de cinza de um ponto qualquer, supondo a existência de dois objetos bem definidos em função dos seus níveis de cinza: um fundo escuro sob um objeto claro. Se observamos o histograma, haverá a existência de um limiar L entre as suas regiões, representando os dois objetos da imagem.

A limiarização desta imagem pode ser representada por:

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) > L \\ 0 & \text{se } f(x, y) \leq L \end{cases}$$

onde $g(x,y)$ representa o ponto limiarizado.

O método de segmentação apresentado por [OTS79] baseia-se na análise discriminante. A operação de binarização é obtida pelo particionamento dos pixels de uma imagem em duas classes $C_0 = \{0, 1, 2, \dots, t\}$ e $C_1 = \{t+1, t+2, \dots, l-1\}$, isto é, objeto e fundo.

Otsu demonstrou que o valor ótimo do limiar t^* pode ser determinado pela maximização do critério de separabilidade da função $\eta = \frac{\sigma_B^2}{\sigma_T^2}$, onde η é a variância entre classes dividida pela variância total: $t^* = \text{ArgMax } \sigma_B^2$,

$$\begin{aligned} \sigma_T^2 &= \sum_{i=0}^{l-1} (i - \mu_T)^2 P_i & \mu_T &= \sum_{i=0}^{l-1} i p_i, \\ \sigma_B^2 &= \omega_0 \omega_1 (\mu_1 - \mu_0)^2 & \omega_1 &= 1 - \omega_0 & \omega_0 &= \sum_{i=0}^t P_i, \\ \mu_1 &= \frac{\mu_T - \mu_t}{1 - \omega_0} & \mu_0 &= \frac{\mu_T}{\omega_0} & \mu_t &= \sum_{i=0}^t i p_i, \end{aligned}$$

sendo que ω_0 e ω_1 correspondem às variâncias nas classes C_0 e C_1 ; μ_0 e μ_1 correspondem às médias das classes, respectivamente. σ_B^2 , σ_T^2 as variâncias inter-classes e total, respectivamente.

Em [SAH88] está previsto o histograma como uma estimativa da função densidade de probabilidade $p(g)$ da população, que compreende os níveis de cinza dos objetos e do fundo da seguinte maneira:

$$p(g) = \frac{1}{\sigma_1 \sqrt{2\pi}} * e^{-\left(\frac{(g-\mu_1)^2}{2\sigma_1^2}\right)} + \frac{1}{\sigma_2 \sqrt{2\pi}} * e^{-\left(\frac{(g-\mu_2)^2}{2\sigma_2^2}\right)}$$

Os parâmetros μ_i , σ_i^2 e as probabilidades a priori p_i com $i = 1,2,..$ da densidade de probabilidade $p(g)$, associados a uma imagem a ser binarizada, não são usualmente conhecidos. Kittler e Illingworth introduziram uma função critério $J(t)$:

$$J(t) = 1 + 2(p_1(t)\log(\sigma_1(t)) + p_2(t)\log(\sigma_2(t))) - 2(p_1(t)\log(p_1(t)) + p_2(t)\log(p_2(t)))$$

onde os parâmetros $\mu_i(t)$, $\sigma_i^2(t)$ e as probabilidades a priori $p_i(t)$ com $i = 1,2,..$ são definidos como:

$$p_1(t) = \sum_{g=0}^t h(g) \quad \text{e} \quad p_2(t) = \sum_{g=t+1}^{255} h(g)$$

$$\mu_1(t) = \frac{\sum_{g=0}^t gh(g)}{p_1(t)} \quad \text{e} \quad \mu_2(t) = \frac{\sum_{g=t+1}^{255} gh(g)}{p_2(t)}$$

$$\sigma_1^2 = \frac{\sum_{g=0}^t h(g)(g - \mu_1(t))^2}{p_1(t)} \quad \text{e} \quad \sigma_2^2 = \frac{\sum_{g=t+1}^{255} h(g)(g - \mu_2(t))^2}{p_2(t)}$$

onde g é o nível de cinza da imagem e $h(g)$ o valor do histograma para cada nível de cinza. Assim, o limiar T é obtido a partir da minimização da função $J(t)$.

Pun, em [PUN81] sugeriu um algoritmo que tenta achar o limiar t que maximiza a função $H = H_b + H_w$ onde: $H_b(t) = -\sum_{i=0}^t p_i \log(p_i)$ é a entropia associada, para uma imagem com 256 níveis de cinza, com os pixels pretos binarizados com um limiar t . De forma análoga, a entropia dos pixels brancos é $H_w(t) = -\sum_{i=t+1}^{255} p_i \log(p_i)$.

Pun demonstrou que basta maximizar a função $f(t)$:

$$f(t) = \frac{H_t}{H_T} \frac{\log(P_t)}{\log(\text{Max}\{p_0, p_1, \dots, p_t\})} + \left[1 - \frac{H_t}{H_T}\right] \frac{\log(1 - P_t)}{\log(\text{Max}\{p_{t+1}, p_{t+2}, \dots, p_{255}\})}$$

com

$$H_t = -\sum_{i=0}^t p_i \log(p_i), \quad H_T = -\sum_{i=0}^{255} p_i \log(p_i) \quad \text{e} \quad P_t = \sum_{i=0}^t p_i.$$

onde H_t , H_T e P_t representam, respectivamente, a entropia dos pixels pretos, a entropia total e a probabilidade cumulativa até o nível de cinza t .

O método de [KAP85] baseia-se no método de Pun. Kapur procura definir a probabilidade de distribuição A de um objeto contra a probabilidade de distribuição B do fundo, da seguinte forma:

$$\begin{aligned} \text{A:} & \quad \frac{p_0}{P_t}, \quad \frac{p_1}{P_t}, \quad \dots, \quad \frac{p_t}{P_t} \\ \text{B:} & \quad \frac{p_{t+1}}{(1-p_t)}, \quad \frac{p_{t+2}}{(1-p_t)}, \quad \dots, \quad \frac{p_{255}}{(1-p_t)} \end{aligned}$$

O valor de limiar ótimo é o valor de t que maximiza a expressão $H = H_b(T) + H_w(T)$, onde as entropias dos pixels pretos é $H_b(t) = -\sum_{i=0}^t \frac{p_i}{P_t} \log\left(\frac{p_i}{P_t}\right)$ e dos pixels brancos é $H_w(t) = -\sum_{i=t+1}^{255} \frac{p_i}{1-P_t} \log\left(\frac{p_i}{1-P_t}\right)$.

Todos esses métodos possuem uma desvantagem em comum: eles utilizam no cálculo somente a informação do histograma (ignorando os detalhes espaciais).

2.5 Técnicas de segmentação de imagens em níveis de cinza

2.5.1 Textura e redes neurais

Em [GOL96] é aplicada uma arquitetura onde a rede é particionada em várias sub-redes neurais. Cada sub-rede é uma estrutura de coluna na qual as características são representadas por meio de “float coding”. Os dados de entrada excita os “floats” de neurônios correspondentes nas sub-redes neurais. No processo de aprendizagem os pesos são trocados de tal forma que as montagens de Hebb são formadas nas estruturas de colunas. Todas as sub-redes são incorporadas em uma rede única por um sistema de controle de atividade neural. A simulação por computador da rede proposta foi realizada. Os resultados da simulação por computador mostra a possibilidade de sucesso para a montagem da rede neural para o problema de segmentação por textura.

São utilizados três processos: processo de aprendizagem, processo de diferenciação de montagem e processo de reconhecimento.

No processo de aprendizagem é extraído o conjunto de características para fazer as descrições das texturas. São 17 características de textura:

- 1: número total de pixels de borda dentro da janela de textura (11x11 pixels);
- 2 a 9: 8 tipos de detetores de contraste;
- 10 a 11: a diferença entre o brilho máximo e mínimo dentro da janela de textura;
- 12 a 16: histograma da distribuição de brilho dentro da janela de textura. A distância máxima possível de brilho é dividida em 5 intervalos iguais;
- 17: valor da coordenada do centro da janela de textura da imagem original.

No processo de diferenciação da montagem é feito o reconhecimento correto de cada textura encontrada na imagem original. O processo continua até que o retalho da imagem do conjunto de treinamento seja reconhecido corretamente.

No processo de reconhecimento, é extraído o conjunto de características de cada retalho da imagem ativando a rede neural. O processo acaba quando todos os retalhos da imagem do conjunto de teste são classificados.

Para os testes foi utilizada uma base de dados de imagens em 32 níveis de cinza. O total do conjunto de treinamento foi de 80 – 280 retalhos de imagens. Foram obtidos bons resultados onde a análise não precisava necessariamente conhecer o contexto do documento para reconhecer. A divisão da rede em sub-redes separadas aumentou a habilidade de reconhecimento, onde as pequenas diferenças de padrões de entrada foram suficientes para distinguir entre classes de textura diferentes com poucas iterações na rede. Esta estrutura também pode ser expandida facilmente para o reconhecimento de um número arbitrário de texturas simplesmente adicionando sub-redes idênticas na rede principal.

Em [JAI96] é citado um método de classificação de componentes de uma imagem por textura utilizando redes neurais. A abordagem é apresentada como uma generalização do método de filtragem multicanal. Ao invés de usar um banco de filtros gerais, a rede neural é treinada para encontrar o mínimo de filtros específicos, então ambas tarefas de extração de características e classificação são realizadas por uma mesma rede unificada. Comparado com o método de filtragem multicanal tradicional, a abordagem com as redes neurais permite realizar as tarefas de classificação e segmentação com mais eficiência.

A configuração da rede neural utilizada para a segmentação do *layout* do documento foi uma rede de 3 camadas com 20 neurônios em cada camada escondida. Esta rede foi treinada para classificar três classes – fundo, texto e gráficos.

Devido ao tamanho reduzido da máscara (11x11 pixels), a rede demonstrou exatidão em localizar as fronteiras de textura, encontrando até pequenas regiões, como números de páginas, mas teve problemas em discriminar entre texto e linhas de desenho.

O método baseado em textura para segmentação de *layout* de documentos não requer a limiarização dos documentos, e é invariante à rotação da imagem,

selecionando de forma adequada os padrões de textura. Não foi citada a base de testes utilizada.

2.5.2 Textura e granulometria local

Em [GRA94] é apresentado um método para segmentar imagens de biópsias de rins utilizando aberturas com elementos estruturantes sucessivos formando uma granulometria em toda a imagem. Os vetores definidos como o gradiente de volume serviram como entrada a uma rede neural. A rede foi treinada para distinguir entre vetores correspondentes a pixels do *interstitium* (região texturizada) ou pixels dos *tubules* (região não texturizada). O fator de correlação entre a área do *interstitium* e a função renal foi computada e comparada aos resultados obtidos através de processo manual e a outros dois processos automáticos.

2.5.3 Filtros de Gibbs

Em [SIV98] é apresentado uma nova classe de modelos espaciais randômicos, conhecidos como campos randômicos de Gibbs morfologicamente restringidos, que são capazes de modelar restrições geométricas em imagens por meio de morfologia matemática. Dois assuntos são considerados nestes modelos, simulação e inferência estatística. É apresentado uma variação do algoritmo Metrópolis, baseado em uma estratégia *multi-site updating*, para simulação dos campos randômicos de Gibbs morfologicamente restritos, que convertem substancialmente mais rápido que os tradicionais algoritmos *single-site updating*. Depois é considerado o problema de afinar os filtros de Gibbs para dados reais. É mostrado que, em condições naturais, a estimação de parâmetros de máxima semelhança podem ser implementados aproximadamente por meio de espectro de padrões.

2.5.4 Tophat diferencial

Em [GU98] é apresentado um sistema completo para extração de caracteres em imagens de cenas coloridas utilizando um novo algoritmo de segmentação morfológico – Tophat diferencial (DTT). Este novo método pode lidar com problemas de segmentação mais difíceis do que os outros algoritmos conhecidos, especialmente

com objetos existem em um fundo complicado, tamanhos irregulares e direções que são afetadas por condições externas variantes. A abordagem proposta para detecção de caracteres em imagens de cenas é robusta e adaptativa neste experimento. A segmentação é realizada a partir da imagem de entrada, onde é decomposta pelo DTT em uma série de sub-imagens, com diferentes tamanhos de objetos. O procedimento de segmentação começa com r_1B_{disk} e termina com r_8B_{disk} , porque em uma investigação estatística o maior caracter encontrado em uma imagem de cena é menor do que 17 (diâmetro do disco de r_8B_{disk}).

2.6 Técnicas de segmentação de placas de veículos

O algoritmo de localização da placa em [SOU00] pode ser dividido em dois procedimentos:

1) No primeiro procedimento temos a localização da placa através dos algoritmos abaixo:

- Procura por variação tonal padronizada através da análise do gradiente positivo na área da placa;
- Binarização local adaptativa sobre a região encontrada ou, se necessário, sobre a imagem inteira;

Nesta etapa, temos como resultado as coordenadas do provável local da placa.

2) No segundo procedimento há a confirmação do local através da análise do número de dígitos encontrados:

- Procura por dígitos de tamanhos pré-definidos;
- Análise dos dígitos selecionados verificando a formação de algum grupo válido;

Como resultado final temos a confirmação do local correto.

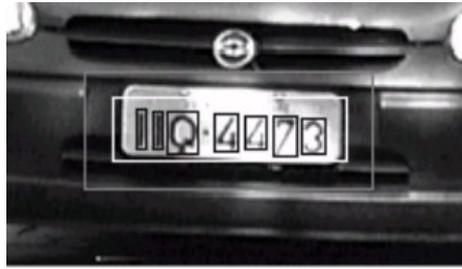


Figura 4 - Seleção dos dígitos da placa

Uma vez encontrada a placa é necessário segmentar os caracteres a fim de separá-los do resto da imagem e redimensioná-los para a rede neural. Foi utilizada uma técnica de crescimento controlado, dentro de cada entidade encontrada na área da placa, e um conjunto de heurísticas para descartar ruídos indesejáveis e selecionar corretamente os caracteres.

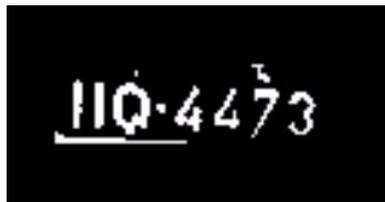


Figura 5 - Placa binarizada

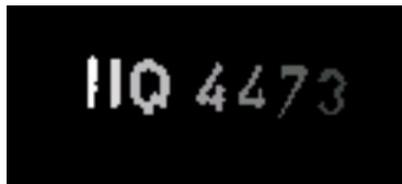


Figura 6 - Placa segmentada - 7 entidades encontradas

Em seguida é feito o reconhecimento dos caracteres através de uma rede neural com topologia *feedforward* utilizando o algoritmo *backpropagation* para treinamento foi escolhida. Foram desenvolvidas duas redes distintas, uma para os caracteres e outra para os algarismos, com 255x26 x26 e 255x10x10 neurônios nas camadas de entrada x escondida x saída respectivamente.

As taxas de acerto do sistema são:

- 98,7 % de sucesso na localização das placas;
- 87,1 % de sucesso na segmentação correta dos caracteres;
- 82,4 % de sucesso na localização das placas com a rotina preliminar (variação tonal);
- 87 % de sucesso no reconhecimento dos caracteres segmentados;
- 28 % de sucesso no reconhecimento correto das placas (7 dígitos).

Em [BRU99] é apresentado o sistema CLPR para identificação de veículos em vias públicas (VIPUR) onde baseia-se no regulamento de placas da Holanda. Este regulamento envolve uma pequena série de regras para a posição da placa no carro, o estilo, tamanho, cor da placa e dos caracteres, e fonte dos caracteres, etc.

O sistema CLPR consiste de quatro unidades principais: uma unidade de segmentação, uma de isolamento, uma de reconhecimento e uma unidade de análise sintática. A unidade de segmentação é completamente implementada pelo DT-CNNs e determina o local da placa baseado em características estruturais e algumas medidas de tamanho. A placa extraída é passada para o isolador de caracteres que isolará cada caracter da placa. Os caracteres isolados são processados pelo reconhecedor de caracteres (OCR). Esta parte do sistema utiliza informações padrão. O analisador sintático checa cada possível caracter retornado pelo reconhecedor um número satisfatório de regras sintáticas existentes nas placas holandesas. Se estas regras não forem satisfatórias, ou um dos caracteres for irreconhecível, a imagem é rejeitada.

As taxas de erro do sistema são:

- 5% com erro na aplicação das características das placas;
- 6% de rejeição de todas as imagens durante o estágio segmentação/isolamento.

Foi apresentado em [CUI97] uma abordagem utilizando campos randômicos de Markov e algoritmos genéticos para a binarização das letras das placas de veículos em movimento. Neste processo, é realizado a localização da placa baseado em informações de textura, essa região de interesse é caracterizada por um conjunto

de características extraídas da região, acompanhadas numa seqüência temporal de imagens. Em seguida, usando essa seqüência de objetos, são calculados parâmetros para corrigir a distorção de perspectiva das placas. Com este resultado, é aplicado o processo de binarização, que é baseado em informações espaciais e temporais simultaneamente. O modelo de Markov é otimizado através de um algoritmo genético.

Neste capítulo apresentamos o estado atual das técnicas de análise de imagens, bem como as principais abordagens dentro dessa área. Em seguida, apresentamos as principais técnicas de segmentação de imagens e técnicas específicas de segmentação e reconhecimento de placas de veículos. No capítulo seguinte será apresentado um estudo sobre processamento digital de imagens e morfologia matemática binária e em níveis de cinza, onde mostramos as operações básicas para as implementações dos algoritmos.

Capítulo 3

Morfologia Matemática

Morfologia vem das palavras gregas *morphê* (forma) e *logos* (ciência), portanto, a morfologia é a ciência que trata das formas que a matéria pode tomar, qualquer que seja a área em que se aplique (biologia, gramática, imagens, etc.) [FAC96].

O princípio da morfologia matemática consiste em extrair informações relativas à geometria e à topologia de uma imagem digital. Uma operação morfológica é determinada a partir de uma vizinhança ao redor do ponto central de um pixel. A grande potencialidade da morfologia matemática reside nos elementos estruturantes, que são conjuntos completamente definidos e conhecidos em forma e tamanho. Os elementos estruturantes são comparados, a partir de uma transformação, ao conjunto desconhecido da imagem. O resultado dessa transformação permite avaliar o conjunto desconhecido. O formato e o tamanho do elemento estruturante possibilitam testar e quantificar de que maneira, o elemento estruturante está ou não está contido na imagem.

3.1 Operadores morfológicos binários

Uma operação morfológica binária é completamente determinada através da vizinhança examinada ao redor do ponto central, da configuração dos pontos pretos (pixel ativo) e brancos (pixels inativo ou neutro) nessa vizinhança e do algoritmo.

3.1.1 Dilatação

A dilatação de um objeto aumenta sua área geométrica, marcando os pixels do fundo que são adjacentes ao contorno do objeto, com o valor dos pixels do objeto.

A dilatação simples é o processo onde todos os pontos que tocam o objeto pelo lado de fora são incorporados ao objeto, resultando em um aumento da área do objeto. A dilatação é muito usada no preenchimento de buracos e objetos segmentados.

A dilatação de um conjunto X pelo elemento estruturante B é definida por:

$$\mathcal{D}^B(X) = X \oplus \tilde{B} = \bigcup_{b \in \tilde{B}} X_b$$

onde \tilde{B} representa o transposto de B e \oplus representa a adição de Minkowski [FAC96].

Conforme a definição acima, o conjunto a ser dilatado, X , é deslocado em função das posições permitidas pelo elemento estruturante B . Os deslocamentos são realizados em relação ao ponto central de \tilde{B} .

3.1.2 Erosão

A erosão de um objeto reduz sua área geométrica, marcando os pixels do contorno do objeto com o valor dos pixels de fundo.

A erosão simples é o processo onde todos os pontos que estão no contorno do objeto são eliminados, resultando em uma diminuição da área do objeto. A erosão é muito utilizada para remover objetos que não são de interesse, de um segmento de imagem.

A erosão de um conjunto X pelo elemento estruturante B é definida por:

$$\mathcal{E}^B(X) = X \ominus \tilde{B} = \bigcap_{b \in \tilde{B}} X_b$$

onde \tilde{B} representa o transposto de B e \ominus representa a subtração de Minkowski [FAC96].

3.1.3 Abertura

A abertura de um objeto é definida como a erosão deste objeto, seguida de uma dilatação. A operação de abertura afeta a imagem de maneira a eliminar objetos pequenos e finos quebrando objetos nos pontos estreitos, geralmente deixando os contornos dos grandes objetos lisos e uniformes, sem mudanças bruscas em sua extensão [FAC96] .

A abertura de um conjunto X pelo elemento estruturante B é definida por:

$$\phi^B(X) = \delta^B(\varepsilon^{\tilde{B}}(X))$$

Nota: Sucessivas aberturas sobre uma imagem produzem o mesmo resultado. Se o efeito desejado é a eliminação de componentes maiores que o elemento estruturante, então algumas iterações do processo de erosão seguidas pelo mesmo número de iterações do processo de dilatação produzem o resultado esperado.

3.1.4 Fechamento

O fechamento de um objeto é definido como a dilatação deste objeto seguida de uma erosão. A operação de fechamento afeta a imagem de maneira a preencher buracos pequenos e finos presentes na extensão do objeto. É feita a conexão de objetos vizinhos em seu ponto mais próximo e geralmente os contornos dos objetos ficam lisos e uniformes sem mudanças bruscas em sua extensão [FAC96] .

O fechamento de um conjunto X pelo elemento estruturante B é definido por:

$$\gamma^B(X) = \varepsilon^B(\delta^{\tilde{B}}(X))$$

Nota: Sucessivos fechamentos sobre uma imagem produzem o mesmo resultado. O efeito desejado é conseguido variando-se o número de iterações. Às vezes, algumas interações do processo de dilatação seguida pelo mesmo número de interações do processo de erosão produz o efeito desejado.

3.1.5 Reconstrução

O processo de reconstrução consiste em recuperar um conjunto S a partir de um conjunto Z , onde S é um subconjunto de uma imagem binária X , e Z é uma amostra de S . Pode-se dizer que S é marcado por Z , portanto Z chama-se marcador.

A reconstrução binária $\rho_s(z)$ de uma máscara (conjunto finito) S a partir do marcador Z ($Z \subset S$), usando o elemento B , é:

$$\rho_s(z) = \lim_{n \rightarrow \infty} \underbrace{\delta_{cS}^B(\dots \delta_{cS}^B(Z))}_n$$

onde $\delta_{cS}^B(Z) = \delta^B(Z) \cap S$ representa a dilatação condicional relativamente a S [FAC96].

A reconstrução permite eliminar os padrões indesejáveis sem deformar os desejáveis.

3.1.6 Granulometria

A granulometria pode ser comparada com um processo de peneiramento. Peneirar consiste em separar substâncias reduzidas ao estado de fragmento, que apresentem vários tamanhos. Efetuar um bom peneiramento depende do tamanho da malha da peneira usada. Após peneirar, são obtidos dois conteúdos: o primeiro com os objetos menores que a malha da peneira, o segundo com o que sobrou do peneiramento, de tamanho maior que a malha.

Matheron em [MATH75] propôs um conjunto de regras que, se forem verificadas, permitem obter uma boa granulometria. Para isso, existem três axiomas:

Sejam X a imagem a ser analisada e $T^{(\lambda)}(X)$ a transformação que permite realizar uma análise granulométrica. $T^{(\lambda)}(X)$ representa o refugo da peneira de tamanho λ .

1. A transformação morfológica deve ser anti-extensiva; isso significa que o conjunto transformado deve ser menor que o de origem ou seja:
 $\forall \lambda > 0, T^{(\lambda)}(X) \subset X \quad \forall X$
2. A transformação morfológica deve ser crescente, ou seja:
 $\forall \lambda > 0, Y \subset X \Rightarrow T^{(\lambda)}(Y) \subset T^{(\lambda)}(X) \quad \forall X$
3. Consideremos a transformação de uma imagem X a partir de duas transformações morfológicas sucessivas de parâmetros respectivos λ_1 e λ_2 . O resultado final deve ser idêntico, qualquer que seja a seqüência de transformações empregadas. Além disso, o resultado deve ser idêntico ao obtido pela transformação de maior parâmetro λ :
 $\forall \lambda_1, \lambda_2 > 0, T^{(\lambda_1)}(T^{(\lambda_2)}(X)) = T^{(\lambda_2)}(T^{(\lambda_1)}(X)) = T^{\sup(\lambda_1, \lambda_2)}(X) \quad \forall X$

Suponha agora que as transformações consideradas são representadas por imagens binárias ou conjuntos discretos. Neste contexto, a granulometria é uma seqüência de aberturas ϕ_n , indexadas sobre um inteiro $n \geq 0$. Cada abertura (correspondente a um tamanho de malha) remove mais do que a anterior, até alcançar o conjunto vazio:

$$\forall X, \quad \forall n \geq m \geq 0, \quad \phi^{nB}(X) \subseteq \phi^{mB}(X)$$

A granulometria por abertura pode gerar deformações nos conjuntos, o que resulta em função e densidade de distribuição incorretas dificultando, portanto, a interpretação. Isto ocorre porque os conjuntos podem ter diversas formas, ou seja, a imagem contém vários padrões e/ou ruídos, dificultando o processo de peneiramento.

3.2 Operadores morfológicos em níveis de cinza

Nessa morfologia é necessário conhecer o valor do pixel mais escuro MIN e o valor do pixel mais claro MAX, na vizinhança de cada pixel ou numa parte da vizinhança da imagem original. Uma operação morfológica em níveis de cinza é completamente determinada através do tamanho e forma da vizinhança, das regiões de pesquisa de MIN e MAX e do algoritmo.

3.2.1 Dilatação

A dilatação de um sinal f por um elemento estruturante g é a seguinte:

$$\delta^g(f(x)) = \max\{f(y) + g(x - y) : y \in D[g]\}$$

onde a dilatação de f por g consiste em verificar se o elemento estruturante centrado em x está acima da função f . Max equivale a máximo, x é o ponto a ser processado na imagem original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante [FAC96].

Considerando a analogia da imagem em níveis de cinza com o relevo topográfico onde os padrões claros são picos e os escuros são vales, é possível determinar padrões de comportamento para a dilatação, como vemos na figura 7.

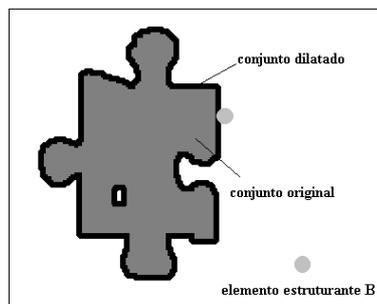


Figura 7 – Exemplo de dilatação

Os efeitos da dilatação em níveis de cinza são:

- clarear a imagem;
- alargar e engordar os picos (padrões claros);
- conectar picos próximos;
- reduzir e às vezes eliminar vales (padrões escuros);
- separar vales.

3.2.2 Erosão

A erosão de um sinal f por um elemento estruturante g é a seguinte:

$$\varepsilon^g(f(x)) = \min\{f(y) - g(x - y) : y \in D[g]\}$$

onde a erosão de f por g consiste em verificar se o elemento centrado em x está abaixo da função f . Min equivale a mínimo, x é o ponto a ser processado na imagem original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante [FAC96].

Uma imagem, de forma geral, apresenta um fundo que pode ser ou não uniforme e sobrepor padrões mais claros e/ou escuros. Portanto, uma imagem pode ser comparada a um relevo topográfico onde padrões claros são picos e escuros são vales. Através desta analogia, padrões de comportamento podem ser estabelecidos para a erosão, como exemplificado na Figura 8.

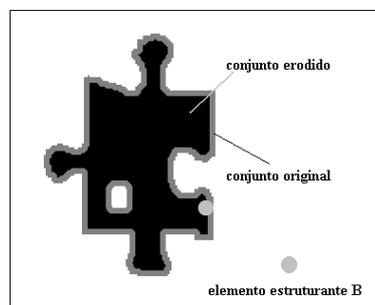


Figura 8 - Exemplo de erosão

Os efeitos da erosão em níveis de cinza são:

- escurecer a imagem;
- alargar e engordar os vales (padrões escuros);
- conectar vales próximos;
- reduzir e às vezes eliminar picos (padrões claros);
- separar picos próximos.

3.2.3 Abertura

A abertura consiste em erodir um conjunto f por g e depois dilatar esse conjunto erodido pelo mesmo elemento estruturante. Como em morfologia binária, a operação morfológica de abertura em níveis de cinza é definida, inicialmente como [FAC96]:

$$abe^g(f) = dil^g(ero^{\tilde{g}}(f))$$

Seguindo a analogia da imagem em níveis de cinza com o relevo topográfico, o comportamento da abertura é:

- separa picos próximos;
- elimina os picos inferiores em tamanho ao elemento estruturante;
- conserva vales afastados;
- emenda vales próximos;
- as entidades restantes após abertura ficam quase idênticas;
- a imagem aberta é mais regular que a imagem original;
- a imagem aberta é menos rica em detalhes que a imagem original.

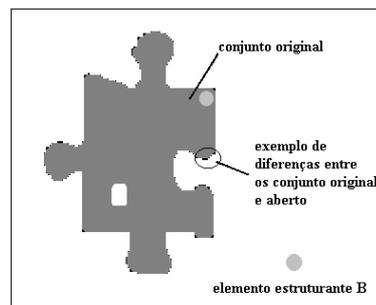


Figura 9 - Imagem aberta

3.2.4 Fechamento

O fechamento de um sinal f por um elemento estruturante g é [FAC96]:

$$fec^g(f) = ero^g(dil^{\tilde{g}}(f))$$

Explorando a propriedade de dualidade entre o fechamento e a abertura em níveis de cinza, podemos escrever:

$$fec^g(f) = -abe^g(-f)$$

Seguindo a analogia da imagem em níveis de cinza como o relevo topográfico, o comportamento do fechamento em níveis de cinza:

- separa vales próximos;
- elimina vales inferiores em tamanho ao elemento estruturante;
- conserva picos afastados;
- emenda picos próximos;
- as entidades restantes após o fechamento ficam quase idênticas;
- a imagem fechada é mais regular que a imagem original;
- a imagem fechada é menos rica em detalhes que a imagem original.

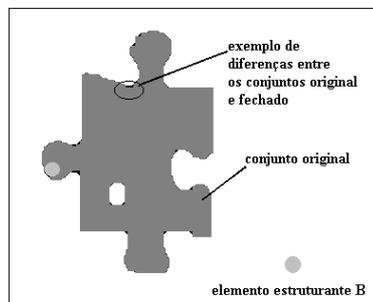


Figura 10 - Imagem fechada

3.2.5 Reconstrução dual

A reconstrução em níveis de cinza é a seqüência lógica da reconstrução binária. A meta é similar, tentar reconstituir uma imagem que sofreu modificações no decorrer do tratamento [FAC96].

Sejam duas imagens S e Z em níveis de cinza definidas no mesmo domínio, tomando os valores discretos $k \in [0, n]$ e respeitando a relação de ordem $Z \geq S$. Por dualidade pode-se definir a reconstrução dual em níveis de cinza (ou reconstrução por erosão):

$$\rho_s(Z) = \min_{n \geq 1} [\mathcal{E}_S^{nB_s}(Z)] = \min_{n \geq 1} \underbrace{\mathcal{E}_S^B(\dots \mathcal{E}_S^B(Z))}_n$$

onde B é o elemento estruturante e min é o mínimo.

A Figura 11 ilustra o processo de reconstrução dual em níveis de cinza da máscara S com o marcador Z e elemento estruturante quadrado.

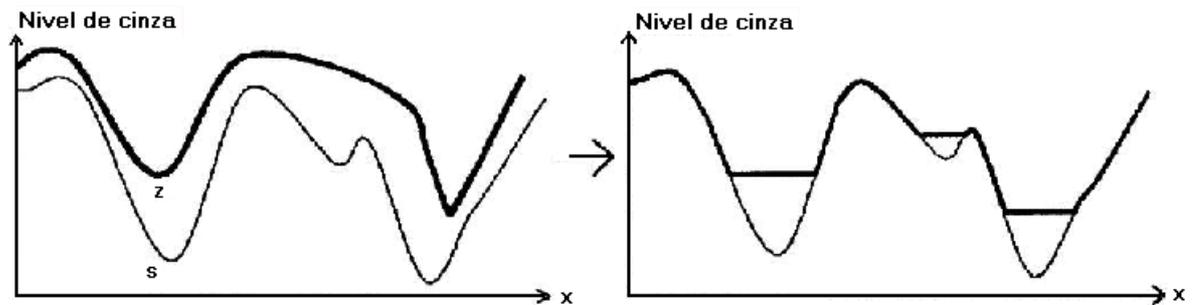


Figura 11 - Reconstrução Dual em níveis de cinza

3.2.6 Granulometria linear em níveis de cinza

Este método rápido de granulometria em níveis de cinza citado em [VIN94] analisa as linhas da imagem contabilizando os tamanhos dos segmentos de reta de acordo com os níveis de cinza conforme descrito a seguir. O resultado do algoritmo é um padrão de espectro que descreve a quantidade de segmentos de reta de cada tamanho para todos os níveis de cinza da imagem.

Considerando o caso da granulometria linear horizontal, com segmentos de linha L_n com $n \geq 0$, onde L_n é uma linha da imagem com $n+1$ pixels (as linhas podem ser processadas independentemente) onde as informações granulométricas estão contidas nas estruturas de picos (de máximo) de cada linha. Sendo $M = \{p_0, p_1, \dots, p_{n-1}\}$ um máximo da linha L , com altitude $L(M) = h$:

- qualquer abertura realizada em L por L_k com $k < n$ deixa M inalterado;
- realizando a abertura de L por L_n tráz todos os pixels de M para o valor de h' da vizinhança de M com o valor mais alto, conforme a Figura 12;
- a contribuição deste máximo para o n -ésimo elemento do padrão de espectro é igual a $n \times (h - h')$.

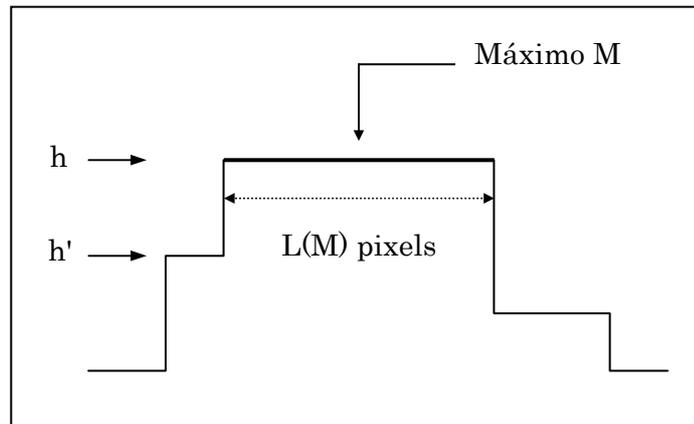


Figura 12 – Relação do máximo de uma região

Os princípios envolvidos na granulometria linear são:

1. procurar os máximos de cada linha, um após o outro;
2. computar recursivamente a contribuição do máximo M corrente ao espectro de padrões até que M não seja mais um máximo;
3. marcar as zonas de máximo já consideradas.

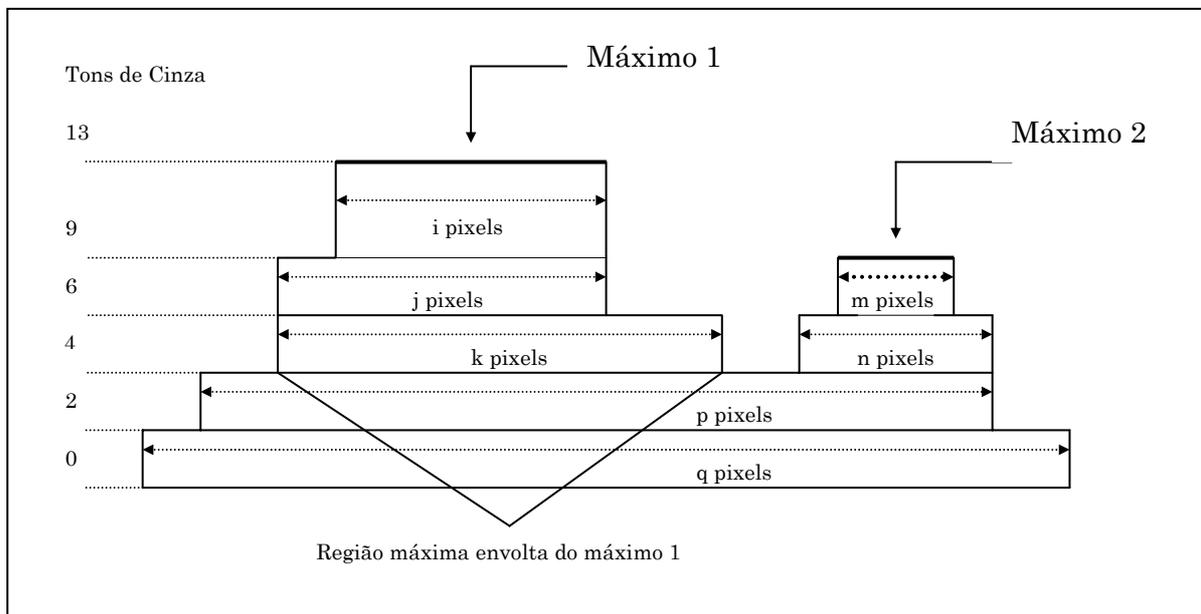


Figura 13 - Contribuições das zonas de máximo

O algoritmo de granulometria linear é aplicado em uma linha da imagem de cada vez, no caso horizontal, com um tempo linear de processamento de acordo com o número de pixels da linha. O espectro de padrão obtido descreve o efeito das aberturas de tamanho 1 até a largura total da imagem [VIN94].

O algoritmo é 3 ordens de magnitude mais rápido que os algoritmos clássicos (aberturas iterativas) e os requisitos de memória são pequenos.

Os tempos de execução dos algoritmos para uma imagem de 512x512, usando uma Sun Sparc Station 10 são:

- clássico - 204 segundos;
- linear – 0.206 segundos.

3.2.7 Resíduos morfológicos

O resíduo morfológico caracteriza a informação extraída de uma imagem a partir de uma série de transformações granulométricas. Este resíduo é dado pela diferença entre dois níveis granulométricos consecutivos.

Seja $T^{(\lambda)}(f)$ uma granulometria. O resíduo morfológico R_λ , de nível residual λ associado a parâmetros de tamanho λ , é definido como a diferença entre os resultados de dois níveis granulométricos consecutivos. Ou seja:

$$\forall \lambda \geq 1, f \in R^N, R_\lambda(f) = T^{(\lambda-1)}(f) - T^{(\lambda)}(f)$$

que define o resíduo morfológico para imagens em níveis de cinza, e representa os componentes preservados em um nível $(\lambda-1)$ que foram eliminados no nível granulométrico λ [GUI99].

3.2.8 Mapeamento a partir de resíduos morfológicos

Seja $(\phi R_\lambda)_{\lambda \geq 1}$ uma família de resíduos morfológicos binarizados. Para todo ponto $p \in S$, onde S é uma imagem, definimos um mapeamento contendo informações a respeito do parâmetro de tamanho $\lambda \in \mathcal{N}$, associado a cada nível residual λ .

$$M_\lambda(p) = \begin{cases} 1 & \text{se } \phi_{R_\lambda}(p) = 1 \\ 0 & \text{se } \phi_{R_\lambda}(p) = 0 \end{cases}$$

A partir desta informação, podemos considerar a informação de tamanho associada aos diferentes níveis em que os pontos da imagem são apagados (mudam de 1 para 0) [GUI99].

Definição de desaparecimento de ponto: O mapeamento de um ponto é representado por mudanças consecutivas do estado 1 para o estado 0 no mapeamento M , representando o instante em que um ponto deixa de pertencer ao resíduo morfológico.

Associado a esta informação de desaparecimento, podemos fazer duas possíveis considerações: uma a respeito da ordem de ocorrência, ρ , destes desaparecimentos, e a outra, sobre o número de ocorrência, η , destes desaparecimentos. A primeira está relacionada com os diferentes instantes em que um ponto desaparece nos diferentes níveis residuais, e a segunda com o número de transições. Normalmente o número de desaparecimentos pode ser associado à irregularidade (complexidade) das estruturas [GUI99].

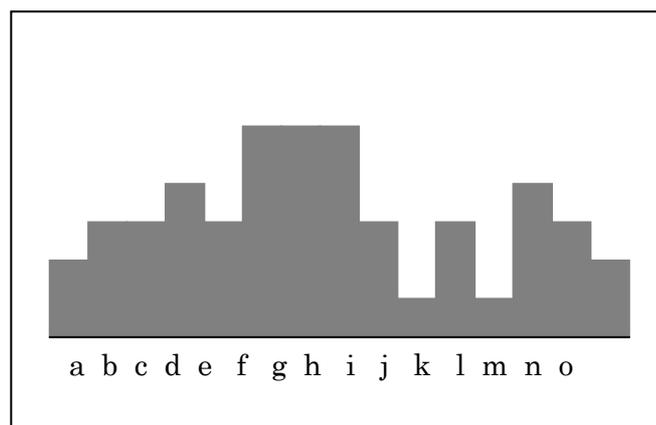


Figura 14 - Linha de uma imagem em níveis de cinza

Na Tabela 1 é apresentado um mapeamento M correspondente aos pontos da Figura 14, bem como o número de desaparecimentos η de cada ponto da imagem.

Tabela 1- Mapeamento dos pontos da Figura 14

λ	A	b	c	d	e	f	g	h	i	j	k	l	m	n	o
1	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0
2	0	0	0	0	0	1	1	1	0	0	0	0	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η	2	2	2	3	2	3	3	3	2	1	2	1	2	2	2

3.2.9 Tophat por abertura

A técnica de tophat por abertura é utilizada para a detecção de picos em uma imagem. O tophat por abertura $tophat_{abe}^g$ de um sinal f pelo elemento estruturante g é:

$$tophat_{abe}^g(f) = f - abe^g(f)$$

Como a abertura é um processo anti-extensivo, o seu resultado fica abaixo do sinal original, exemplificado na Figura 15.

Com o uso de um elemento estruturante adequado, o processo de abertura permite a eliminação dos picos. Fazer a diferença, entre os sinais original e resultado da abertura, permite tirar o ruído e eliminar a falta de homogeneidade, ou seja, ressaltar a informação dos picos da imagem [FAC96].

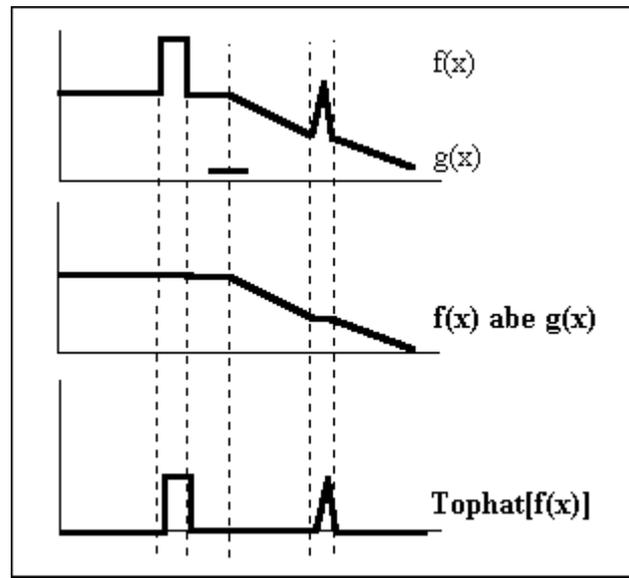


Figura 15 – Diferença entre a imagem original e a imagem aberta

3.2.10 Tophat por fechamento

De forma análoga, a técnica de tophat por fechamento é utilizada para a detecção dos vales em uma imagem. O tophat por fechamento $tophat_{fec}^g$ de um sinal f pelo elemento estruturante g é:

$$tophat_{fec}^g(f) = fec^g(f) - f$$

Como o fechamento é um processo extensivo, o seu resultado da transformação é sempre positivo, como vemos na Figura 16.

Com o uso de um elemento estruturante adequado, o processo de fechamento permite a eliminação dos vales. Fazer a diferença, entre os sinais original e resultado fechado, permite tirar o ruído e eliminar a falta de homogeneidade, ou seja, ressaltar a informação dos vales da imagem [FAC96].

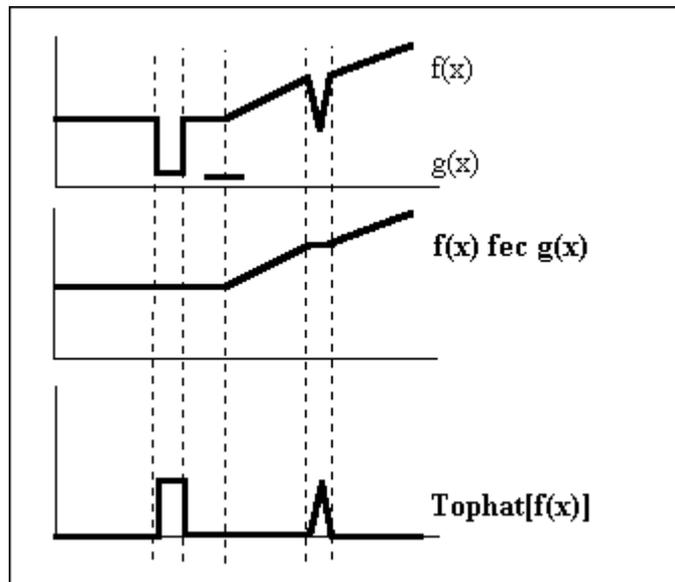


Figura 16 – Diferença entre a imagem original e a imagem fechada

Neste capítulo foram abordadas as técnicas sobre processamento digital de imagens e morfologia matemática binária e em níveis de cinza, onde mostramos as operações que nos darão sustentação para as implementações necessárias neste trabalho. No capítulo seguinte será apresentado a metodologia aplicada para desenvolver o trabalho proposto e a base de dados criada no Lardoc, e também o processamento necessário para se obter a segmentação dos dígitos da placa a partir de uma imagem em tons de cinza do veículo. Dentro deste processamento, destacamos o pré-processamento, para melhorar a qualidade da imagem obtida pela máquina fotográfica, a segmentação que tem o objetivo de fornecer os blocos para o reconhecimento e a codificação.

Capítulo 4

Metodologia e Desenvolvimento

Neste capítulo será apresentada a metodologia no desenvolvimento do trabalho e também a descrição de cada uma das etapas da qual este é constituído. O trabalho foi dividido em duas etapas: a primeira consiste em uma pesquisa qualitativa, ou seja, a busca de informações e métodos em trabalhos já realizados, o que foram vistos nos Capítulo 2 e Capítulo 3 e a segunda fase consiste de uma pesquisa quantitativa, isto é, confecção de uma base de dados, a experimentação do método mais relevante visto na pesquisa qualitativa, implementação e experimentação do método proposto.

Para a experimentação, foram utilizadas as rotinas de morfologia matemática disponíveis no software BROPI. A plataforma de desenvolvimento empregada na implementação deste trabalho foi o Microsoft Visual C++, utilizando-se a biblioteca de classes Microsoft Foundation Classes em ambiente Microsoft Windows 98.

O software utilizado como base do desenvolvimento deste trabalho, BROPI, foi implementado no LARDOC (Laboratório de Análise e Reconhecimento de Documentos), em parceria com o Polo de Software de Curitiba através do projeto GDOC (Gestão de Hiperdocumentos) [MAT98].

4.1 Base de dados

Para o desenvolvimento de um trabalho de análise de imagens se faz necessária a utilização de uma base de dados, pois a disponibilidade de uma

biblioteca contendo um número significativo de imagens é um fator crítico para qualquer projeto de pesquisa experimental nesta área.

Os testes e verificações das rotinas consideradas foram realizadas em uma base de dados contendo imagens de carros estacionados. Esta base conta com 180 imagens coloridas (as características das imagens serão apresentadas no item 4.1.1), adquiridas utilizando-se uma máquina fotográfica digital, escala 100% e ajustes de brilho e contraste padrões da câmera utilizada, marca Olympus, modelo D-360L.

A base foi dividida em duas bases menores, uma contendo 120 imagens (2/3) que foi utilizada para testes durante o desenvolvimento dos algoritmos e uma segunda base, contendo o restante das imagens (60 imagens – 1/3), foi utilizada para a validação final das rotinas e verificação do método proposto. Este processo de divisão da base de imagens em base de testes e base de validação foi repetido 5 vezes criando assim grupos distintos de bases de treinamento e validação. Em outras palavras, a base de imagens foi subdividida aleatoriamente cinco vezes, garantindo-se que em cada subdivisão os grupos de imagens gerados são diferentes, o que permitiu o treinamento, ou seja, a determinação do melhor conjunto de parâmetros para cada um dos cinco grupos de imagens de treinamento e a validação destes parâmetros com os cinco grupos de imagens de teste, ou validação.

4.1.1 Características da base de dados

As imagens utilizadas para a criação da base de dados possuem as seguintes características:

- colorida (RGB);
- 640x480 pixels;
- imagens da frente e costas dos veículos (em alguns casos com dois ou mais ângulos do mesmo veículo);
- distribuição de luz sobre a superfície da placa homogênea e heterogênea;
- variação de distância da câmera até o veículo: 2,0m – 3,5m.

Após a aquisição das imagens, estas foram convertidas para 256 níveis de cinza, através do software BROPI (Biblioteca de Rotinas para Processamento de Imagens).

No total de 180 imagens da base de dados, 33% são compostas de imagens tiradas da parte frontal do veículo e 67% são imagens da parte posterior (10 imagens foram tiradas com um ângulo superior a 30 graus com relação ao centro do veículo), pois o objetivo é que constem nesta base imagens que ilustrem todas as possibilidades, conforme as tabelas 2, 3 e 4. Na Tabela 2 são apresentadas as quantidades com relação a posição dos veículos nas imagens.

Tabela 2 – Posicionamento dos veículos

Imagens Frontais	Imagens Traseiras
58	122

As fotos foram tiradas em vários dias e situações de iluminação (chuva, sol e tempo nublado). Desta forma, temos na base diferentes situações de distribuição da iluminação incidente na placa do veículo; estas situações são: placa totalmente iluminada, parcialmente sombreada e totalmente sombreada. Esta distribuição de iluminação foi classificada na Tabela 3 da seguinte forma:

- distribuição homogênea;
- distribuição heterogênea.

Obs: Sendo a distribuição homogênea sub-classificada em imagens claras e imagens escuras, conforme Tabela 4.

Todo o processo de classificação das imagens foram feitas visualmente.

Tabela 3 - Características de iluminação

Imagens Homogêneas	Imagens Heterogêneas
143	37

Tabela 4 - Intensidade de iluminação das imagens homogêneas

Imagens Claras	Imagens Escuras
72	71

4.2 Desenvolvimento

Neste item será mostrado o processamento necessário para se obter a posição da placa do veículo a partir de uma imagem digital em níveis de cinza. Na Figura 17 descrevemos as etapas desde a aquisição da imagem do veículo passando pelo pré-processamento, processamento e segmentação, obtendo como resultado as coordenadas da placa na imagem.

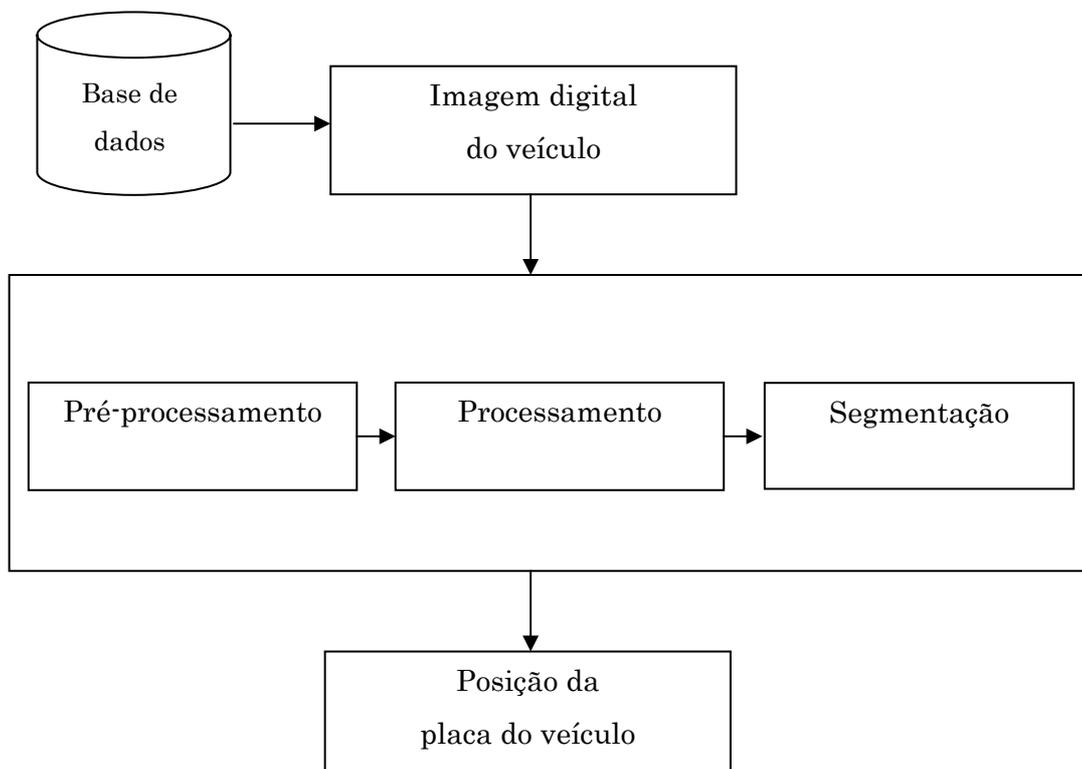


Figura 17 - Etapas do processamento

Cada uma das etapas do processo são compostas da seguinte forma:

Pré-processamento:

- Conversão das imagens coloridas para níveis de cinza;
- Tophat por fechamento;
- Redução da complexidade da imagem.

Processamento:

- Granulometria em níveis de cinza;
- Reconstrução dual.

Segmentação:

- Localização dos componentes;
- Filtragem de componentes.

Cada um destes itens serão abordados na seqüência sendo descritos o seu funcionamento e objetivo no processo como um todo.

4.2.1 Aquisição da imagem

O processo de aquisição das imagens foi descrito no item 4.1, que trata da criação da base de dados. Como já mencionado, a aquisição da base foi feita em RGB, isto para permitir que outros trabalhos sejam desenvolvidos utilizando a mesma base e abordando outros problemas. Outra explicação para a adoção do padrão RGB é que o padrão adotado pelo sistema operacional Windows é o sistema de cores RGB.

As imagens são adquiridas pela câmera digital em formato RGB, com compressão JPEG. Em seguida, foi utilizado o software da câmera para converter as imagens para o formato BMP (Bitmap, padrão do sistema operacional Windows 98). Estas imagens foram gravadas em CD-R (*Compact Disc Recordable*) nos formatos JPEG e BMP.

Na Figura 18 apresentamos um exemplo de imagem de documento que faz parte da base de dados.



Figura 18 - Imagem exemplo de um carro estacionado

4.2.2 Conversão das imagens coloridas para níveis de cinza

Na análise automática de imagens, o processamento da cor é de grande importância na identificação e extração de características e objetos. Este fato deve-se principalmente ao aumento da capacidade de discernimento dos objetos em imagens coloridas. Vários são os sistemas para a representação da cor encontrados na literatura; as duas principais abordagens são [JES99]:

- RGB: consiste na utilização de um espaço tridimensional onde cada cor é obtida a partir das cores primárias, vermelho, verde e azul, respectivamente, baixa, média e alta frequências do espectro visível pelo olho humano.
- HSI: H é a matiz (descreve a cor pura: vermelha, azul e amarela); S é a saturação da cor (grau de pureza); I é a componente que representa a intensidade de luz (brilho) da cor.

No processo de conversão das imagens RGB para níveis de cinza convencionou-se que o nível de cinza de um determinado pixel corresponde a média das três componentes de cor do pixel, calculado da seguinte maneira:

$$C = \frac{R + G + B}{3}$$

onde R, G e B são as componentes de cor.



Figura 19 - Imagem convertida para níveis de cinza

4.2.3 Tophat por fechamento

O objetivo da aplicação do processo de tophat na imagem é ressaltar os vales (regiões escuras) da imagem, facilitando assim a marcação das letras da placa do veículo no processo de granulometria que será aplicado na seqüência. Utilizou-se o elemento estruturante quadrado com 3 iterações (baseado na largura dos dígitos da placa). O resultado pode ser visto na Figura 20.



Figura 20 - Processamento de tophat

4.2.4 Redução da complexidade da imagem

A redução da complexidade da imagem se fez necessária para a redução do tempo de processamento da granulometria, isto porque quanto maior o número de elementos dispostos em cada uma das camadas de cor (níveis de cinza), maior será o tempo de processamento da granulometria. Portanto, optou-se por homogeneizar as regiões da imagem diminuindo-se o número de pequenos elementos e acelerando-se em muito o processo de granulometria.

O processo de redução da complexidade é realizado dividindo-se o nível de cinza de cada um dos pixels da imagem por um fator de redução da complexidade. Os seguintes valores foram testados: 64, 32, 16, 8, 4 e 2. O que apresentou melhores resultados foi o fator 64 e é apresentado na Figura 21.

**Figura 21 - Redução de complexidade**

4.2.5 Granulometria em níveis de cinza

A granulometria em níveis de cinza foi utilizada para a obtenção de marcadores os quais espera-se que sejam as letras da placa do veículo. O processo da granulometria utilizado é uma composição dos algoritmos apresentados por Vincent em [VIN94] e Guimarães em [GUI99].

O algoritmo de granulometria apresentado por Vincent é utilizado para obtermos um perfil dos tamanhos horizontais ou verticais dos objetos presentes na imagem. Este perfil de tamanhos é utilizado no lugar dos resíduos morfológicos no algoritmo de mapeamento descrito por Guimarães e assim obtemos a imagem com os marcadores das letras das placas dos veículos, mostrado na Figura 22.



Figura 22 - Imagem de marcadores

4.2.6 Reconstrução dual

Utilizando a imagem de marcadores, resultado do processo de granulometria descrito no item anterior, e como mascara a imagem resultante do processo de eliminação da complexidade da imagem, é aplicado o processo de reconstrução dual [FAC96].

Teoricamente, o resultado desta reconstrução mostrado na Figura 23 deve conter as letras que constituem as placas dos veículos.

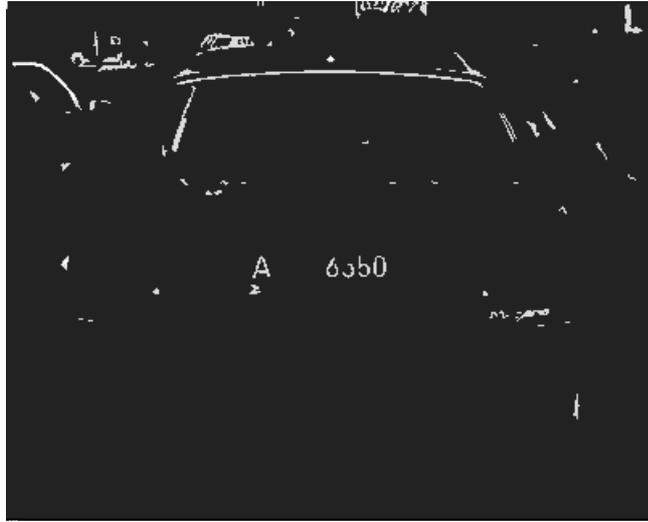


Figura 23 - Reconstrução dual

4.2.7 Localização dos componentes

Para a localização dos componentes da imagem, resultado da reconstrução, que indicarão o posicionamento da placa do veículo na imagem é adotado o processo descrito a seguir:

1. aproveitando-se o resultado da diminuição da complexidade da imagem (uma homogeneização das regiões da imagem), é calculado um histograma de distribuição de níveis de cinza;
2. este histograma é utilizado para determinar o nível de cinza correspondente à região de fundo da imagem, região com o maior número de pixels;
3. com base no nível de cinza encontrado no passo 2 (região com maior número de pixels) é realizada uma busca por pixels cujo nível de cinza seja de maior intensidade. Uma vez localizado um destes, todos os seus vizinhos, que possuem um nível de cinza maior que o nível de cinza correspondente a região de fundo da imagem, são considerados como parte do mesmo objeto;

4. a localização deste componente é determinada anotando-se a posição dos pixels mais a esquerda, a direita, acima e abaixo, que estão ligados pela 8-vizinhança ao primeiro pixel encontrado.



Figura 24 - Localização de componentes

4.2.8 Filtragem de componentes

O resultado obtido na localização de componentes normalmente apresenta um elevado número de elementos indesejáveis. Estes componentes estão presentes por serem semelhantes às letras das placas tanto em seus tamanhos horizontais, característica esta utilizada no processo de granulometria cinza, como em seus valores de níveis de cinza, característica utilizada no processo de mapeamento.

Portanto foram determinadas algumas regras para a filtragem dos componentes da imagem. Estas regras estão listadas a seguir:

1. componentes que estão localizados na borda da imagem são desconsiderados;
2. componentes que possuem uma das sua dimensões, altura ou largura, muito maior que a outra são desconsiderados (85% de diferença);
3. componentes cujas dimensões estão acima de 85% das médias de tamanho de todos os elementos da imagem são desconsiderados;

4. componentes cujas dimensões estão abaixo de 35% das médias de tamanho de todos os elementos da imagem são desconsiderados;
5. componentes que encontram-se alinhados com menos de três elementos na horizontal são desconsiderados, pois considera-se que pelo menos 3 caracteres da placa serão encontrados;
6. componentes cujo número de pixels não pertencente ao valor de referência para o fundo da imagem (calculado no processo de localização dos componentes) que estão acima de 85% ou abaixo de 35% são desconsiderados.

A aplicação destas regras no conjunto de componentes localizados na imagem resulta na manutenção dos componentes que provavelmente se encontram na região da placa do veículo, mostrado na Figura 25.



Figura 25 - Filtragem de componentes

4.2.9 Localização da placa

Após a filtragem dos componentes foi realizado o seguinte processo para a determinação da localização da posição da placa do veículo:

1. determinação do limiar de cada um dos componentes da imagem utilizando-se o algoritmo de OTSU [OTS79]. Este processo é realizado aplicando-se o algoritmo

- de OTSU na região da imagem que corresponde ao componente (possível letra ou dígito da placa);
2. determinação do nível de cinza de maior incidência no fundo da imagem para cada um dos componentes. A busca deste nível de cinza é realizada nos pixels cujos níveis de cinza são de menor intensidade que o limiar resultante da aplicação do algoritmo de OTSU;
 3. cálculo da média dos valores dos níveis de cinza de maior incidência em cada um dos componentes encontrados no passo 2;
 4. determinação da maior região definida por níveis de cinza com intensidade maior ou igual à média encontrada no passo 3.

A região determinada com este processo encerra a placa do veículo, como é mostrado na Figura 26.



Figura 26 - Localização da placa

A implementação das rotinas que compõem o método proposto não estão otimizadas, considerando-se utilização de memória e performance dos algoritmos, assim fica a sugestão para a realização de um estudo algorítmico para otimizar estas rotinas e desta forma melhorar o desempenho do processo de localização de placas.

Os resultados obtidos neste trabalho podem ser melhorados realizando-se uma etapa de pré-processamento na imagem que aumente o contraste da região da placa em relação ao veículo. Um segundo pré-processamento que pode melhorar significativamente os resultados é o aumento da qualidade e contraste dos caracteres da placa. Os estudos necessários para a determinação dos melhores métodos para o pré-processamento da imagem são apontados como continuação deste trabalho.

Neste capítulo foram apresentadas a metodologia aplicada para desenvolver o trabalho proposto e a base de dados criada para os testes, e também o processamento necessário para se obter a segmentação da placa de um veículo a partir de uma imagem em tons de cinza. Dentro deste processamento, destacamos o pré-processamento, para melhorar a qualidade da imagem obtida pela câmera fotográfica digital, o processamento da imagem e a segmentação que tem o objetivo de encontrar a localização das letras na placa de um veículo. No Capítulo 5 serão apresentados os resultados obtidos na aplicação do método aqui proposto sobre a base de dados criada para a execução do trabalho. Também serão apresentadas as análises estatísticas a respeito do desempenho do método para a discussão dos resultados alcançados.

Capítulo 5

Análise de Resultados

Neste capítulo apresentaremos os resultados obtidos pelo método proposto. Em seguida, serão feitos alguns comentários a respeito dos resultados encontrados e uma comparação com outros sistemas para processamento de imagens de placas.

De acordo com o método de processamento de imagens apresentado no Capítulo 4, foram considerados para efeito de análise duas de suas etapas, a saber: a filtragem de componentes, cujo o objetivo é a localização das letras da placa do veículo; a localização da placa, cujo o objetivo é a partir dos componentes já localizados demarcar a região correspondente à toda a placa do veículo.

O tempo de processamento para cada imagem foi de 28 segundos, considerando a execução da filtragem de componentes e também da localização da placa, utilizando um computador equipado com processador Pentium II de 350 MHz e 128 MB de memória RAM.

Para a validação do método, foram utilizados 5 grupos distintos de bases de validação contendo 60 imagens cada uma. Em outras palavras, a base de imagens foi subdividida aleatoriamente cinco vezes, garantindo-se que em cada subdivisão os grupos de imagens gerados são diferentes, o que permitiu validar o processo sem distorcer os resultados apresentados. As diferenças entre as taxas de localização correta para cada um dos grupos de imagens de validação permaneceram menores que 2%, indicando apenas a variação das imagens que formam os diferentes grupos e não diferenças no algoritmo ou nos parâmetros utilizados. Desta forma, para apresentar um resumo estatístico dos resultados obtidos a partir do processamento foi utilizada uma única base de validação.

Na seqüência, são apresentados alguns casos particulares nos quais são analisados os pontos fortes e também os pontos fracos do método proposto.

5.1 Análise da filtragem dos componentes

O resultado analisado como filtragem dos componentes é apresentado no item 4.2.8 como sendo a união dos componentes localizados na imagem.

Na etapa de filtragem dos componentes foi obtido o resultado de acordo com a Tabela 5 :

Tabela 5 - Resultados da filtragem de componentes

Segmentação Correta	Segmentação Parcial	Segmentação Incorreta
10 imagens – 16%	39 imagens – 65%	11 imagens – 19%

Foi considerado como uma segmentação correta quando todas as letras da placa foram localizadas corretamente, uma segmentação parcial quando uma ou mais letras da placa não foi localizada ou em casos que componentes não pertencentes a placa foram marcados. A segmentação incorreta caracteriza o caso no qual nenhuma das letras da placa foram encontradas.

A Figura 27 mostra o resultado correto da segmentação na filtragem dos componentes da placa.



Figura 27 - Segmentação correta da placa

A Figura 28 mostra a segmentação parcial na filtragem dos componentes da placa. Isto ocorre porque nem todas as letras são marcadas no processo de granulometria e mapeamento; este fato se dá em decorrência de pequenos defeitos na placa ou uma insuficiência de iluminação – estes fatos acarretam uma diferença de espessura nas várias letras que compõem a placa. Quando esta espessura torna-se muito menor que as demais letras, a letra é eliminada por não estar corretamente representada na imagem. Um outro fator é quando o nível de cinza pelo qual a letra está representada é muito próximo do nível de cinza que caracteriza o fundo da imagem, isto também acarreta na eliminação da letra no processo de redução da complexidade da imagem, fato bastante observado nas imagens escuras.



Figura 28 - Segmentação parcial da placa

5.2 Análise da localização da placa

O resultado analisado como localização da placa é apresentado no item 4.2.9 como sendo a busca da placa pelo nível de cinza que representa o fundo da imagem.

Na etapa de localização da placa foi obtido o resultado de acordo com a Tabela 6:

Tabela 6 - Resultados da localização da placa

Localização Correta	Localização Parcial	Localização Incorreta
36 imagens – 60%	9 imagens – 15%	15 imagens – 25%

Foi considerado como uma localização correta quando a região da placa foi identificada corretamente, uma localização parcial quando a região marcada corresponde a área da placa e uma pequena região do veículo a mais. A localização incorreta caracteriza o caso no qual nenhuma região da placa foi marcado.

A Figura 29 mostra o resultado correto da localização da placa.

**Figura 29 - Localização correta da placa**

A Figura 30 mostra a localização parcial da placa. A localização da placa baseia-se na busca da região cujo nível de cinza seja semelhante ao nível de cinza que corresponde ao fundo da placa. Esta busca apresenta um problema caracterizado por regiões, cuja a cor assemelha-se à cor de fundo da placa, situadas na vizinhança da placa. Este fato acarreta um efeito de expansão da região

marcada. Isto ocorre normalmente em carros de cor clara ou em situações de luminosidade muito intensa.



Figura 30 - Localização parcial da placa

Foi realizado um teste com uma base de 60 imagens produzidas a partir da base original. Estas imagens contém os recortes apenas do veículo, retirando-o da cena. Neste teste, a segmentação das placas apresentou uma melhora de 15% na localização correta da placa em relação ao processo aplicado na imagem completa.

Esta melhora de 15% ocorre porque não existem uma área de imagem cuja complexidade ou elementos possa interferir no resultado da granulometria, o que é mais comum no caso onde a imagem processada contém toda a cena (árvores, asfalto, placas, calçada, etc.).

É apresentado na Figura 31 a forma como foi feito o recorte e também o resultado do processamento.



Figura 31 - Recorte para a retirada da cena

5.3 Comparação entre os resultados da filtragem dos componentes e da localização da placa

Em situações nas quais a placa está delimitada por uma região escura, o processo de localização da placa melhora os resultados do processo de filtragem dos componentes significativamente, melhorando situações de segmentação parcial para localização correta e também mantendo a localização correta. Nas situações em que a placa não está delimitada por uma região escura, o processo de localização da placa apresenta um baixo desempenho, situações que podemos averiguar analisando as figuras 30 e 32.

5.4 Comparação com o sistema SIAV

Comparamos o método de segmentação proposto neste trabalho com os resultados apresentados pelo sistema SIAV descrito em [SOU00]. A seguir são colocadas as características apresentadas pelo sistema SIAV.

O sistema SIAV foi testado com um conjunto de 300 imagens. As imagens utilizadas para o teste possuem as seguintes características:

- 256 tons de cinza e 320x240 pixels;
- Dimensões da placa entre 70x20 e 120x40 pixels;

- Imagens da frente e costas dos veículos com distribuição de luz sobre a superfície da placa homogênea e heterogênea.

Comparação entre os resultados (SIAV x Método proposto):

- SIAV - 82,4% de sucesso na localização das placas com a rotina preliminar (variação tonal).
- Método proposto - 75% de sucesso na localização das placas;

O sistema SIAV apresenta um desempenho superior no processo de localização das placas de veículos, porém, é importante ressaltar que as características de posicionamento do veículo nas imagens processadas pelo sistema SIAV são muito mais restritas que as impostas neste trabalho.

- SIAV - 87,1% de sucesso na segmentação correta dos caracteres;
- Método proposto - 81% de sucesso na segmentação correta dos caracteres (filtragem de componentes).

O sistema SIAV realiza uma busca pela região provável da placa para em seguida realizar a segmentação dos caracteres da placa, inclusive utilizando técnicas de limiarização para conseguir um melhor índice de segmentação e também impondo restrições quanto ao tamanho dos dígitos da placa; o que não ocorre no método aqui proposto, que tenta primeiro localizar os caracteres e em seguida, a partir destes, localizar a região da placa.

O SIAV possui um bom desempenho em ambientes com diferentes tipos de iluminação. O tempo de processamento de cada imagem (320X240 pixels) é dependente da complexidade da imagem analisada e pode variar entre 1.4 e 130s (processador K6-2 400 MHz com 64 MB de RAM). O método aqui proposto apresentou tempo de processamento de 28 segundos por imagem (640x480 pixels) em um processador Pentium II 350 MHz com 128 MB de RAM.

5.5 Casos particulares

Nesta seção serão apresentadas imagens cujas características especiais apresentam situações que evidenciam a versatilidade do método proposto. Também serão apresentados casos nos quais o método falha.

5.5.1 Localização em veículos com pintura clara

É mostrado na Figura 32 o caso de carros que possuem uma pintura clara (ou branca). Nestes casos, no processo de localização da placa, que se baseia numa busca por nível de cinza, não há contraste suficiente para diferenciar a placa da pintura do veículo.

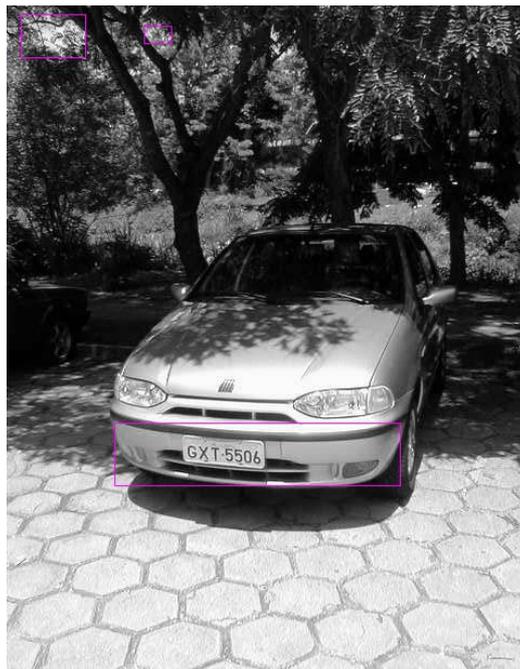


Figura 32 - Veículo com pintura clara

5.5.2 Placas posicionadas incorretamente

Nos casos em que a placa do veículo não está corretamente fixada (com inclinação) ou posicionada (não centralizada), o resultado da localização da placa não sofreram interferências, como é demonstrado na Figura 33.



Figura 33 – Veículo com a placa inclinada

5.5.3 Fator de luminosidade

Em situações que a placa do veículo se encontra na sombra projetada pelo próprio veículo, em dias ensolarados, a incidência de luminosidade na placa é muito pequena, fazendo com que a placa não fique destacada em relação ao veículo, como é demonstrado na Figura 34, o método não localiza a placa pela pequena diferença de tonalidade da placa em relação ao veículo.



Figura 34 – Placa sombreada

5.5.4 Foto com mais de um veículo

Em alguns casos em que aparece dois ou mais veículos na imagem, o método apresentou um bom desempenho, proporcional à distância da câmera até os veículos. Quando os veículos encontram-se em uma distância apropriada, as placas destes são localizadas corretamente e independentemente uma da outra. Este fato é demonstrado na Figura 35.



Figura 35 – Imagem contendo duas placas

5.5.5 Localização da placa em veículos com adesivos

Em casos nos quais os veículos possuem adesivos colados na pintura, o processo não é influenciado se o padrão destes adesivos não corresponder ao padrão das letras da placa do veículo, como pode ser visto na Figura 36.



Figura 36 – Veículo com adesivos

5.5.6 Fotos tiradas em ângulos

Para demonstrar que o método funciona com uma grande variabilidade de posicionamento da câmera em relação ao veículo, foram tiradas fotos do mesmo veículo em posições diferentes, lembrando que a distância da câmera até o veículo foi mantida, variando somente o ângulo de visão. Nas figuras 37 e 38 é mostrado um exemplo desta situação.



Figura 37 – Foto com ângulo normal



Figura 38 – Foto com ângulo acentuado

Vimos neste capítulo algumas discussões sobre o trabalho desenvolvido, e também algumas considerações sobre os resultados encontrados. No capítulo seguinte será apresentado a conclusão sobre este trabalho e as perspectivas para trabalhos futuros.

Conclusão

A abordagem apresentada neste estudo empregou a morfologia matemática para localizar e segmentar automaticamente placas de veículos de imagens em níveis de cinza. O objetivo principal do trabalho foi a segmentação de placas utilizando morfologia matemática, visando obter como resultado uma segmentação exata ou a mais exata possível para todas as situações e fugindo das abordagens como redes neurais e análise de textura.

Os resultados obtidos com a aplicação do método proposto neste trabalho em uma base de 60 imagens, 75% de acerto na localização das placas e 81% de acerto na segmentação dos caracteres corretamente localizados, podem ser considerados como satisfatórios quando são levadas em consideração as diferentes situações das imagens analisadas, conforme foi discutido no capítulo anterior e também levando-se em consideração a complexidade da imagem e a não utilização de algoritmos clássicos para a resolução de problemas que envolvem este nível de complexidade.

A versatilidade do método em relação às diversas condições de iluminação, posicionamento e distância da câmera é a melhor característica deste método. Além desta, ressalta-se a característica do método iniciar o processo de segmentação buscando os caracteres da placa na imagem da cena completa, e não em um recorte selecionado a partir de uma primeira segmentação da imagem para a localização da placa, método utilizado por outros sistemas similares.

Além destas características, devemos salientar também o bom desempenho apresentado em relação ao tempo de processamento deste método, 28 segundos por imagem.

6.1 Contribuição do trabalho

As principais contribuições do trabalho são:

- uma metodologia de segmentação utilizando granulometria em níveis de cinza;
- segmentação independente de ângulos;
- robustez em relação a iluminação dos veículos – dias ensolarados, nublados e chuvosos;
- bom desempenho em relação ao tempo de processamento.

6.2 Perspectivas futuras

Como a implementação das rotinas que compõem o método proposto não estão totalmente otimizadas, sugere-se a realização de um estudo algorítmico para otimizar estas rotinas e desta forma melhorar o desempenho do processo de localização de placas.

Os resultados obtidos neste trabalho também podem ser melhorados realizando-se uma etapa de pré-processamento na imagem que aumente o contraste da região da placa em relação ao veículo, o que aumentará a sucesso da localização das placas.

Um segundo processamento que pode melhorar significativamente os resultados é o aumento da qualidade e contraste dos caracteres da placa, isto antes de ser realizada a localização e filtragem dos elementos da imagem.

Os estudos necessários para a determinação dos melhores métodos para o pré-processamento da imagem são apontados como continuação deste trabalho.

Referências Bibliográficas

- ❖ [BRU99] Brugge, M. H. T., Nijhuis J. A. G., Spaanenburg L., Sdevens J. H., License Plate Recognition, Knowledge-Based Intelligent Techniques in Character Recognition, Edited by Lakhmi C. Jain and Beatrice Lazzerini, pp 263-295, 1999
- ❖ [CUI97] Cui, Y., Huang, Q.; Automatic License Extration from Moving Vehicles. Siemens Corporate Research, Inc, ICIP'97
- ❖ [FAC93] Facon, Jacques; Processamento e análise de imagens, CEFET-PR, julho, 1993
- ❖ [FAC96] Facon, Jacques; Morfologia Matemática: Teoria e Exemplos. Curitiba, Brasil, 1996
- ❖ [GAZ01] Jornal Gazeta do Povo, Curitiba, Paraná, 21/02/2001
- ❖ [GOL96] Goltsu, Alexander; An assembly neural network for texture segmentation. Neural Networks, IEEE, pp 643-653, 1996
- ❖ [GOR95] O’Gorman, Lawrence; Kasturi, Rangachar; Document image analysis, IEEE, 1995
- ❖ [GRA94] Gratin, C., Vitrià, J., Moreso, F., Serón, D.; Texture Classification using Neural Networks and Local Granulometries, Mathematical Morphology and Its Applications to Image and Signal Processing, Kluwer Academic Publishers, pp 309-316, 1994
- ❖ [GU98] Gu, L., Kaneko, T., Tanaka, N., Haralick, R. M.; Morphological Segmentation Applied to Character Extraction from Color Cover Images, Mathematical Morphology and Its Applications to Image and Signal Processing, Kluwer Academic Publishers, pp 367-375, 1998
- ❖ [GUI99] Guimarães, Silvio J. F.; Filtragem de imagens a partir da sua decomposição em resíduos morfológicos. Dissertação de Mestrado, Universidade Estadual de Campinas, 1999
- ❖ [HAR85] Haralick R. M. and Shapiro L. G., Image Segmentation Techniques, vol.29, pp. 100-132, 1985.

- ❖ [JAI96] Jain K., Anil; Karu, Kalle; Learning texture discrimination masks. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp 195-205, vol. 18, February, 1996
- ❖ [JAI98] Jain, Anil K.; Yu, Bin; Document representation and It's application to page decomposition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 20, March, 1998
- ❖ [JES99] Jesus, Andreia de; Uma abordagem morfológica para a segmentação de logotipos em cheques bancários brasileiros sem conhecimento a priori. Dissertação de mestrado, Pontifícia Universidade Católica do Paraná, 1999
- ❖ [KAP85] Kapur N. J., Sahoo P.K. and Wong A. K. C., A New Method for Gray-Level Picture Thresholding, Computer Vision, Graphics and Image Processing, vol.29, pp. 273-285, 1985.
- ❖ [MAT98] Matrakas, M. D.; Bortolozzi, F.; Projeto de uma Biblioteca para Processamento de Imagens de Documentos Orientada a Objetos, International Seminar on Document Management, pp 236-242, 1998
- ❖ [MATH75] Mathron G.; Dymanic Measurement of Computer Generated Image Segmentation, IEEE Trans. Pattern Analysis Mach. Intell. 7, pp 155-164, 1975
- ❖ [OTS79] Otsu N., A Threshold Selection Method from Gray-Level Histograms, IEEE Trans. Sys. Man and Cybernnetics, v. SMC9, nº1, pp. 62-66, 1979.
- ❖ [PUN81] Pun T., Entropic Thresholding – The New Approach, Computer Graphics and Image Processing, vol.16, pp.210-239, 1981.
- ❖ [SAH88] Sahoo P. K., Soltani S. and Wong A. K. C., A Survey of Thresholding Techniques, Computer Vision, Graphics and Image Processing, vol.41, pp 233-260, 1988
- ❖ [SIM97] Simon, Aniko; Pret, Jean-Christophe; Johnson A., Peter; A fast algorithm for bottom-up document layout analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp 273-277, vol. 19, March, 1997
- ❖ [SIV98] Sivakumar, K., Goutsias, J.; Monte Carlo Simulation and Statistical Inference of Morphologically Constrained GRFS, Mathematical Morphology and Its Applications to Image and Signal Processing, Kluwer Academic Publishers, pp 267-275, 1998
- ❖ [SOU00] Souza, F. P. C.; Susin, A. A.; Sistema de identificação automática de veículos. www.iee.ufrgs.br/iee/siav.htm, 2000

- ❖ [VIN94] Vincent, L.; Fast grayscale granulometrie algorithms. EURASIP Workshop ISMM'94. pp 265 – 272, Fountainebleau, France, 1994