

Marcelo Cherem Ramalho

*Sistema de identificação da ordem de
produção estampada em tarugos de aço*

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientador:

Prof. Dr. Antônio de Pádua Braga

Co-orientador:

Prof. Dr. Walmir Matos Caminhas

UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
ÁREA DE CONCENTRAÇÃO: ENGENHARIA DE COMPUTAÇÃO E TELECOMUNICAÇÕES
LINHA DE PESQUISA: INTELIGÊNCIA COMPUTACIONAL

Belo Horizonte – MG

2005

Agradecimentos

Expresso a minha gratidão às pessoas e às instituições que contribuíram para a realização desta dissertação:

Agradeço ao Prof. Antônio de Pádua Braga e ao Prof. Walmir Matos Caminhas pelos ensinamentos e orientações;

Ao Marcos Vinícius, ao Prof. Guilherme, ao Leandro, ao Leonardo Claudino, ao Douglas, ao Turíbio e ao Célio pelas discussões relevantes e soluções desenvolvidas;

Aos meus pais, José Márcio e Telma, aos meus irmãos, Luciana e André, à minha família e aos meus amigos pelo apoio e compreensão em todos os momentos.

À minha esposa, Fernanda, pelo amor e carinho;

À Fernanda e ao Vinícius pela ajuda na impressão dos trabalhos;

Ao Frederico pelas dicas estratégicas na elaboração do texto e da apresentação;

À Açominas e à FCO (Fundação Cristiano Otoni) pelo apoio financeiro.

Conteúdo

Lista de Figuras

Lista de Tabelas

Resumo

Abstract

1	Introdução	p. 14
1.1	Motivação do projeto	p. 19
1.2	Especificações do Projeto	p. 19
1.3	Contribuição da Dissertação	p. 21
1.4	Organização do Trabalho	p. 22
2	Revisão da Literatura	p. 23
2.1	Introdução	p. 23
2.2	Definição das Etapas do Problema	p. 23
2.3	Pré-processamento	p. 26
2.3.1	Limiarização	p. 26
2.3.1.1	Método de Niblack	p. 28
2.3.1.2	Método de Otsu	p. 29
2.3.2	Detecção de Bordas	p. 31
	Sobel	p. 31
	Prewitt	p. 32

	Frei-Chen	p. 32
	Roberts	p. 32
2.3.3	Detecção de Linhas	p. 32
	Transformada de Hough	p. 32
	RANSAC	p. 33
2.3.4	Esqueletização	p. 33
2.4	Segmentação	p. 33
2.4.1	Análise de Textura	p. 34
	Estatística	p. 34
	Aplicação de técnicas de filtragem	p. 34
2.5	Extração de Características	p. 35
	Casamento de Padrões	p. 35
	Momentos de Zernike	p. 36
2.5.1	Casamento de Padrões	p. 36
2.6	Classificação	p. 40
2.6.1	Classificador de Bayes	p. 41
2.7	Conclusões	p. 42
3	Metodologia	p. 43
3.1	Introdução	p. 43
3.2	Definição das Etapas	p. 43
3.3	Identificação da Rotação do Tarugo	p. 45
3.3.1	Análise das Linhas Extremas	p. 46
3.3.2	Análise de Regiões	p. 49
3.3.3	Decisor Baseado em Blocos	p. 52
3.4	Reconhecimento do Número da OP	p. 54
3.4.1	Limiarização	p. 57

3.4.2	Casamento de Padrões	p. 65
3.5	Conclusões	p. 68
4	Teste em Campo	p. 70
4.1	Introdução	p. 70
4.2	Condicionamento	p. 70
4.3	Identificação da Rotação do Tarugo	p. 73
4.4	Isolamento do Tarugo	p. 75
4.5	Conclusão	p. 78
5	Conclusão	p. 79
	Trabalhos Futuros	p. 81
	Referências	p. 83

Lista de Figuras

1	Figuras reais de tarugos. Em (A): Detalhe da seção reta do tarugo. Em (B): Detalhe da extensão do tarugo. Figuras adaptadas do site da Açominas [1].	p. 15
2	Processo de produção dos tarugos.	p. 16
3	Exemplos de tarugos estampados com OP (retângulo azul) e localização (retângulo vermelho). A tinta está presente em um tarugo, realçando a informação da localização para o cliente.	p. 16
4	Arquitetura de um processo de automação industrial.	p. 17
5	Sistema de inspeção dos tarugos. Os defeitos detectados são reparados. Os tarugos são então empacotados.	p. 18
6	Integração do sistema de visão com os sistemas da automação.	p. 18
7	Descrição e definição de possíveis posicionamentos da câmera na LIT.	p. 21
8	Etapas do problema de reconhecimento de caracteres, com base nas divisões propostas por Gonzales [18] e Taxt [57].	p. 24
9	Tipos de figuras utilizadas em algoritmos de reconhecimento de caracteres. Em (A): Imagem em escala de cinza. Em (B): Imagem limiarizada. Em (C): Esqueleto da imagem. Em (D): Contorno da imagem.	p. 25
10	Exemplo de imagem com histograma bimodal. Em (A): Imagem de caracteres coreanos. Imagem adaptada da base de dados disponível em [27]. Em (B): Histograma da Imagem em (A).	p. 27
11	Exemplo de uma imagem com iluminação não uniforme. Em (A): Resultado da aplicação do método global de Otsu [41]. Em (B): Resultado da aplicação do método local de Niblack [38].	p. 28
12	Exemplo de correlação cruzada normalizada entre três imagens. Os resultados estão na Tabela 2.	p. 39

13	Divisão das etapas do problema, definindo, também, os possíveis planos de trabalho.	p. 44
14	Plano de trabalho para reconhecimento da rugosidade e identificação da rotação.	p. 44
15	Plano de trabalho para extração de características e classificação. . . .	p. 45
16	Demonstração das duas regiões distintas no tarugo.	p. 45
17	Definição da posição dos caracteres.	p. 46
18	Disposição das linhas do tarugo.	p. 47
19	Regiões analisadas no método e os parâmetros considerados.	p. 50
20	Separação conseguida pela análise das regiões. Azul e Verde = Rugoso. Vermelho e Preto = Liso.	p. 51
21	Comparação da variância de regiões em margens opostas do mesmo tarugo. <i>Azul = abaixo, vermelho = acima, verde = esquerda e preto = direita</i> . Símbolos iguais pertencem ao mesmo tarugo. As regiões rugosas (verde e azul) apresentam maior variância do que as regiões lisas (vermelho e preto).	p. 51
22	Configuração do reconhecimento da rotação pela variância de blocos. . .	p. 52
23	Parâmetros do decisor baseado em blocos.	p. 52
24	Otimização da configuração do reconhecimento da rotação pela variância de blocos.	p. 53
25	Montagem final do classificador. Dois decisores determinam a localização da rugosidade. Decisor 1: acima ou abaixo. Decisor 2: esquerda ou direita. . .	p. 53
26	Imagens utilizadas na implementação de rotinas para identificação. . . .	p. 55
27	Grupo 1 - Imagens com melhor contraste e pouca deformação.	p. 55
28	Grupo 2 - Imagens com baixo contraste.	p. 56
29	Grupo 3 - Imagens com deformações.	p. 56
30	Grupo 4 - Imagens com reflexos.	p. 56
31	Imagem com baixo contraste. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.	p. 57

32	Imagem com deformação. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.	p. 57
33	Imagem com bom contraste. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.	p. 58
34	Imagem com reflexos. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.	p. 58
35	Comparação dos histogramas da imagem original e do caractere.	p. 59
36	Comparação dos histogramas da imagem equalizada e do caractere. Em (A): comparação dos histogramas. Em (B): imagem limiarizada com o limiar $\mu(i, j) - \sigma(i, j)$. Em (C): imagem limiarizada com o limiar de Otsu.	p. 59
37	Comparação dos histogramas da imagem equalizada após a aplicação de um filtro de médias. Em (A): comparação dos histogramas. Em (B): imagem limiarizada com o limiar $\mu(i, j) - \sigma(i, j)$. Em (C): imagem limiarizada com o limiar de Otsu.	p. 60
38	Imagens do grupo 1 limiarizadas pelo método de filtragem seguida de equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - \sigma(i, j)$	p. 61
39	Imagens do grupo 1 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$	p. 61
40	Imagens do grupo 2 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$	p. 62
41	Imagens do grupo 3 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$	p. 62
42	Imagens do grupo 4 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$	p. 63
43	Imagens da Figura 39 após aplicação do filtro de médias de 3×3 <i>pixels</i>	p. 63
44	Imagens da Figura 40 após aplicação do filtro de médias de 3×3 <i>pixels</i>	p. 64
45	Imagens da Figura 41 após aplicação do filtro de médias de 3×3 <i>pixels</i>	p. 64
46	Imagens da Figura 42 após aplicação do filtro de médias de 3×3 <i>pixels</i>	p. 64

47	O método utilizado permite que o número esteja deslocado do centro da figura. Em (A): Padrão. Em (B): Imagem que contém o número a ser identificado.	p. 65
48	Esquema do algoritmo de identificação por correlação.	p. 66
49	Banco de padrões utilizados na correlação, com a presença de fundo de imagem.	p. 66
50	Resultado da correlação das imagens da Figura 43 com os padrões da Figura 49.	p. 67
51	Iluminação para realçar detalhes no plano do tarugo. Em (A): vista da câmera. Em (B): vista lateral do tarugo.	p. 71
52	Comparação do efeito da iluminação estruturada nos dois planos de trabalho.	p. 71
53	Os dois planos de trabalho considerados. Em (A): plano aberto, para detecção da região dos números. Em (B): plano fechado para reconhecimento dos números.	p. 72
54	Plano de trabalho para a aplicação do método do decisor baseado em blocos para identificação da rotação do tarugo.	p. 73
55	Tarugo classificado incorretamente. A tinta provoca uma região homogênea próxima à margem direita da figura.	p. 74
56	Tarugo identificado corretamente mesmo com a presença de pintura. A tinta não contamina a margem inteira.	p. 74
57	Ajuste dos parâmetros para a correta identificação do tarugo. O aumento do afastamento da margem soluciona o problema.	p. 75
58	Análise do plano aberto da imagem, com detecção de contorno. Há presença de muito ruído.	p. 76
59	Análise das componentes de cor da imagem. Em (A): Componente vermelho. Em (B): Componente verde. Em (C): Componente azul.	p. 76
60	Evidenciação do tarugo. Em (A): Componente azul - Componente vermelho. Em (B): Imagem em (A) limiarizada pelo método de Otsu [41].	p. 77

61	Em (A): Imagem B da Figura 60 depois da aplicação do filtro de mediana. Em (B): Escolha do objeto oito conectado com maior área.	p. 77
62	Aplicação do método de segmentação do tarugo nas imagens coletadas.	p. 78

Lista de Tabelas

1	Resumo dos problemas e características dos posicionamentos da câmera na linha de inspeção, conforme Figura 7	p. 20
2	Resultado da correlação cruzada normalizada das imagens na Figura 12, com fundo de imagem = -1 e objeto = 1 . Os resultados são os mesmos para a situação inversa: fundo de imagem = 1 e objeto = -1	p. 39
3	Análise dos vetores formados seqüencialmente com as PLs de cada linha.	p. 48
4	Variáveis: 1 - Variância, 2 - Máximo, 3 - Mínimo, 4 - Média. O desempenho foi melhor considerando mais variáveis e sem a presença da média.	p. 49
5	Vetor classificado com erro. Todos os classificadores apontaram a seqüência 0110, seqüência consistente, mas errada.	p. 49
6	Resumo do desempenho dos métodos desenvolvidos.	p. 54
7	Valores da correlação normalizada para o grupo 1, conforme Figura 50. O número 1 identificado como 9 apresenta o menor valor de correlação cruzada normalizada.	p. 67
8	Análise do valor de correlação para os acertos das imagens do grupo 1.	p. 68
9	Índice de acerto nos 4 grupos de imagens.	p. 68
10	Tabela com os valores de correlação cruzada normalizada entre os padrões escolhidos.	p. 68
11	Resumo da aplicação do método “decisor baseado em blocos” nas imagens coletadas na linha de inspeção.	p. 75
12	Tabela de relação das OPs inspecionadas no dia 14/01/2005 com as bitolas (em mm).	p. 82

Resumo

Esta dissertação apresenta um estudo sobre as ferramentas de reconhecimento de caracteres para a aplicação em um problema real na Gerdau Açominas. O problema é a identificação de números de Ordem de Produção (OP) em peças de aço (tarugos). O reconhecimento visa a automatização completa da LIT (linha de inspeção de tarugos). Pela identificação da OP, o sistema modifica as variáveis paramétricas da linha de inspeção e separa os tarugos para cada cliente, evitando-se problemas como mistura de peças e inspeções com parâmetros errados. O estudo é realizado nas imagens coletadas, visando fornecer requisitos para a escolha dos algoritmos. O problema é dividido em etapas de segmentação, extração de características e classificação. São estudados e propostos métodos nestas áreas, levando em consideração a variação de escala e rotação dos caracteres. Primeiramente, são realizados testes em imagens coletadas em laboratório, sem condicionamento de escala ou rotação. A rotação é resolvida pela análise da característica de textura presente no tarugo, diminuindo o custo computacional e fornecendo rotação constante para a etapa de extração de características. Esta técnica forneceu 100% de acerto nas imagens coletadas em laboratório. Como as etapas de extração de características utilizam como base a imagem limiarizada, um estudo é realizado sobre os diferentes tipos de ferramentas propostas na literatura. Com a análise das imagens é proposto um método para a extração do caractere a partir da imagem em escala de cinza. Na etapa de reconhecimento foi testado o método de casamento de padrões, implementado através da correlação cruzada normalizada. Os resultados demonstram um desempenho de 98% para imagens com bom contraste e pouca deformação, indicando a necessidade de um maior condicionamento na iluminação e escala. A segunda fase demonstra resultados preliminares em imagens coletadas na linha de inspeção. A identificação da rotação foi correta em 94% das imagens testadas. As imagens apresentam um melhor condicionamento de iluminação e escala, fornecendo resultados promissores para a implementação do sistema.

Abstract

This dissertation presents an overview of character recognition methods for application in a real problem of Gerdau Açominas. The problem consists of the identification of OP (Production Order) numbers imprint on steel objects (billets). The recognition will complete the automation of the inspection line. The parameters for the inspection are defined by the OP number. After the inspection the billets are separated by client. The correct recognition prevent problems such as mixture of billets and the inspection of billets with wrong parameters. The study is carried through with the collected images, supplying requirements for the choice of the algorithms. The problem is divided in three stages: segmentation, feature extraction and classification. Methods are proposed in these areas, in two fronts. First, tests in images collected in laboratory are carried through, without conditioning the scale or rotation. The rotation is decided by the analysis of the texture present in the billets, diminishing the computational cost and supplying constant rotation for the stage of feature extraction. The method achieved a identification rate of 100% in images collected in laboratory. Most feature extraction methods are designed for solid binary images. A study is carried through on the different methods in literature. With the analysis of the collected images, a method for the extraction of the character from the gray level image is proposed. In the stage of recognition, the method of template matching was tested, implemented by a normalized cross correlation. The method provided a recognition rate of 98% in images with better contrast and low deformation, which indicates that better conditioning of the illumination and scale are necessary. The second front demonstrates the preliminary results in images collected in the inspection line. The identification of the rotation provided a rate of 94%. The images presents better conditioning of illumination and scale, supplying promising results for the implementation of the system.

1 *Introdução*

O trabalho descrito nesta dissertação faz parte do projeto de desenvolvimento e instalação de um ambiente computacional para o reconhecimento da ordem de produção (OP) em peças de aço chamadas “tarugos” da empresa Gerdau Aço Minas, conforme Figura 1¹. A Aço Minas, empresa sediada na cidade de Ouro Branco-MG, é forte no ramo da siderurgia. O Grupo Gerdau é acionista da Aço Minas desde 1997, tornando-se majoritário em 2002, após a aquisição das ações da empresa Natsteel. A Gerdau é uma das maiores produtoras de aço do continente americano, possuindo usinas em países como Estados Unidos, Canadá, Argentina e Brasil [2]. Dentre os produtos da Gerdau Aço Minas, os tarugos possuem uma larga aplicação em diversas áreas que, entre outras, se destacam [2]:

- fio-máquina para arames em geral;
- eletrodo de solda;
- estampagem a frio;
- molas;
- vergalhões.

Os tarugos são de fundamental importância na empresa, sendo, inclusive, um forte produto de exportação. Por este fator, a Aço Minas dispõe de um rígido controle de qualidade, atendendo às rigorosas especificações do cliente. Os tarugos, antes de serem encaminhados aos clientes, são inspecionados na LIT (**L**inha de **I**nspção de **T**arugos). Neste setor são realizados o controle de qualidade e a destinação da produção para o cliente.

Os tarugos são objetos de aço proveniente do processo de laminação de peças, denominadas “lingotes”, oriundas da Aciaria, com características próprias na composição

¹Figuras adaptadas do site da Aço Minas [1].

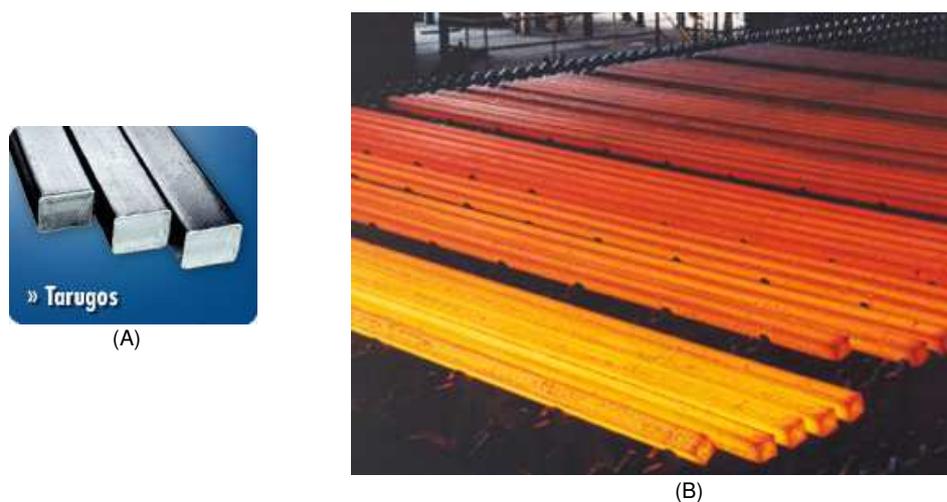


Figura 1: Figuras reais de tarugos. Em (A): Detalhe da seção reta do tarugo. Em (B): Detalhe da extensão do tarugo. Figuras adaptadas do site da Açominas [1].

química (especificadas pelo cliente). Estas peças passam por um processo de conformação mecânica denominado “laminação”, até atingirem a bitola requerida pelo cliente, apresentando, no final do processo, grandes extensões. Como o comprimento também é uma especificação de produção, estes tarugos são, então, cortados por um equipamento chamado “tesoura pendular”. Os tarugos são as peças no fim deste processo (Figura 2). Para haver controle da produção, os tarugos cortados são estampados com a ordem de produção (OP), número que designa as características químicas e as dimensões especificadas pelos clientes, gerada a partir da negociação entre o setor de vendas e o cliente. As estampas nos tarugos são marcados após o corte por equipamentos chamados “martelos”, que estampam no aço quente o número da OP (Figura 2). O primeiro e o último tarugo cortados recebem, além deste número, uma identificação chamada “localização”. A letra T (Top) designa o primeiro tarugo e a letra B (Bottom) o último. Estas estampas são realizadas por um segundo (letra T) e terceiro (letra B) martelos, que só são aplicados nestas duas ocasiões. Alguns clientes requisitam que a localização, além de ser estampada, seja pintada. Como não é um padrão para todas as peças, a tinta pode ou não aparecer nos tarugos, não podendo ser utilizado para identificação. Os tarugos intermediários não recebem nenhuma estampa adicional. A Figura 3 mostra dois tarugos estampados com OP e localização, com a presença da pintura em um deles.

Depois de produzidos e estampados, os tarugos são encaminhados para a etapa de resfriamento e estocagem, onde permanecem até serem destinados aos clientes. Os tarugos são conduzidos para a LIT (Figura 5), onde será realizado o controle de qualidade e a

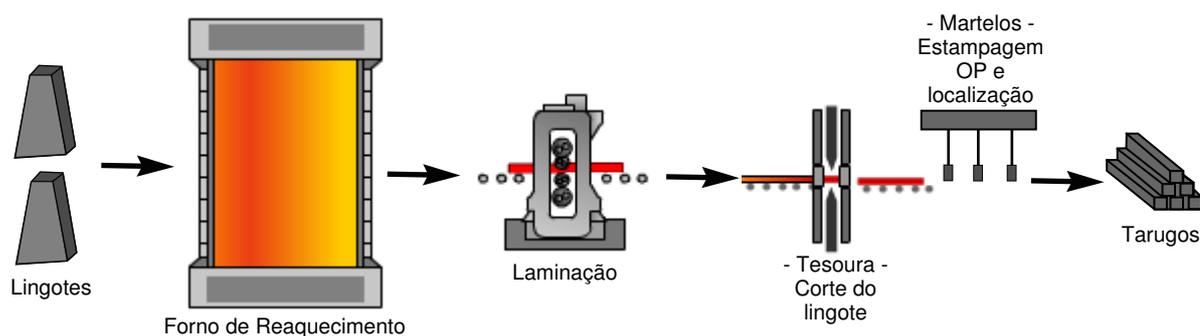


Figura 2: Processo de produção dos tarugos.

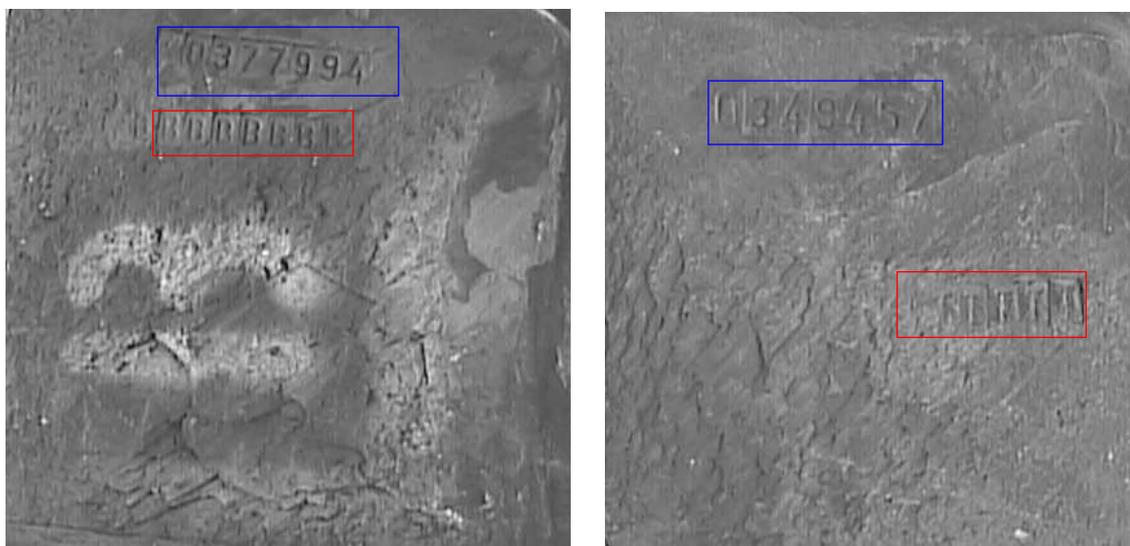


Figura 3: Exemplos de tarugos estampados com OP (retângulo azul) e localização (retângulo vermelho). A tinta está presente em um tarugo, realçando a informação da localização para o cliente.

correta separação e empacotamento dos tarugos. O controle de qualidade é realizado por inspeções internas e superficiais para detectar presença de defeitos na produção. Defeitos que podem ser reparados ainda na LIT, através de um equipamento chamado “esmerilhadeira” (para correção de defeitos superficiais) e através de cortes (para correção de defeitos internos). Os tarugos cortados recebem outra estampagem no final do processo. Os tarugos reformados são, então, conduzidos para a empacotadeira, onde se juntam aos tarugos aprovados na LIT, para serem encaminhados para o cliente.

Como cada conjunto de tarugo, especificados por OPs diferentes, detém composições químicas e dimensões diferenciadas, a inspeção também deve ser realizada de forma di-

ferente. Os parâmetros dos equipamentos que compõe a LIT e detectam os defeitos (superficiais e internos) são, por isso, determinados pela OP dos tarugos. Antes de o tarugo entrar efetivamente na LIT, é realizado o ajuste dos parâmetros através do repasse da informação da OP ao sistema nível 2 de automação instalado da LIT. Os níveis de automação fazem parte da arquitetura de um processo de automação industrial, e desempenham diferentes papéis, conforme Figura 4.

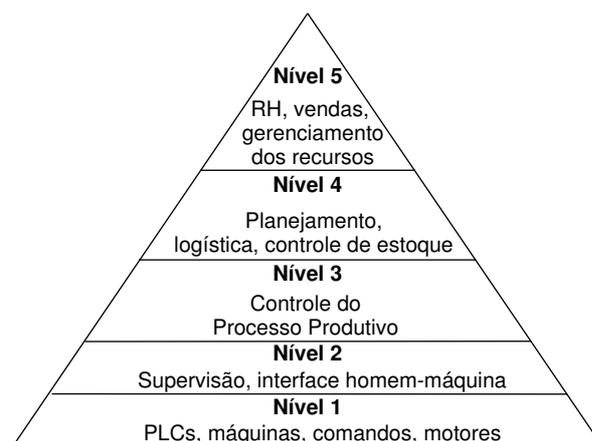


Figura 4: Arquitetura de um processo de automação industrial.

O sistema de nível 2 e o sistema supervisor são responsáveis pela interface homem-máquina, fazendo a supervisão e recebendo informações dos sistemas de nível 1, que atuam no chão de fábrica (máquinas, dispositivos, etc). O sistema de nível 3 desempenha um papel de controle do processo produtivo, dispondo de bancos de dados relevantes. O sistema de nível 2 da Açominas é responsável por analisar a ordem de produção e manter as características da inspeção até que todos os tarugos da mesma ordem de produção sejam inspecionados e estocados [39]. Quando recebe uma OP diferente, o sistema de nível 2 requisita ao sistema de nível 3 a parametrização dos equipamentos de inspeção e recebe o número de tarugos pertencentes àquela OP. O sistema de nível 2 só irá requisitar novas mudanças nos parâmetros dos equipamentos quando todos os tarugos da OP vigente forem processados. Portanto, se um tarugo de uma OP diferente aparecer na inspeção antes do final da OP atual, o sistema pede a retirada do mesmo. A leitura da OP é, portanto, parte fundamental neste processo e, atualmente, é realizada por funcionários da empresa que passam a informação diretamente para o sistema nível 2 da automação da LIT.

Para otimizar o processo de automação atualmente implantado, o sistema a ser implantado deverá repassar a informação do número da OP e da “localização” do tarugo

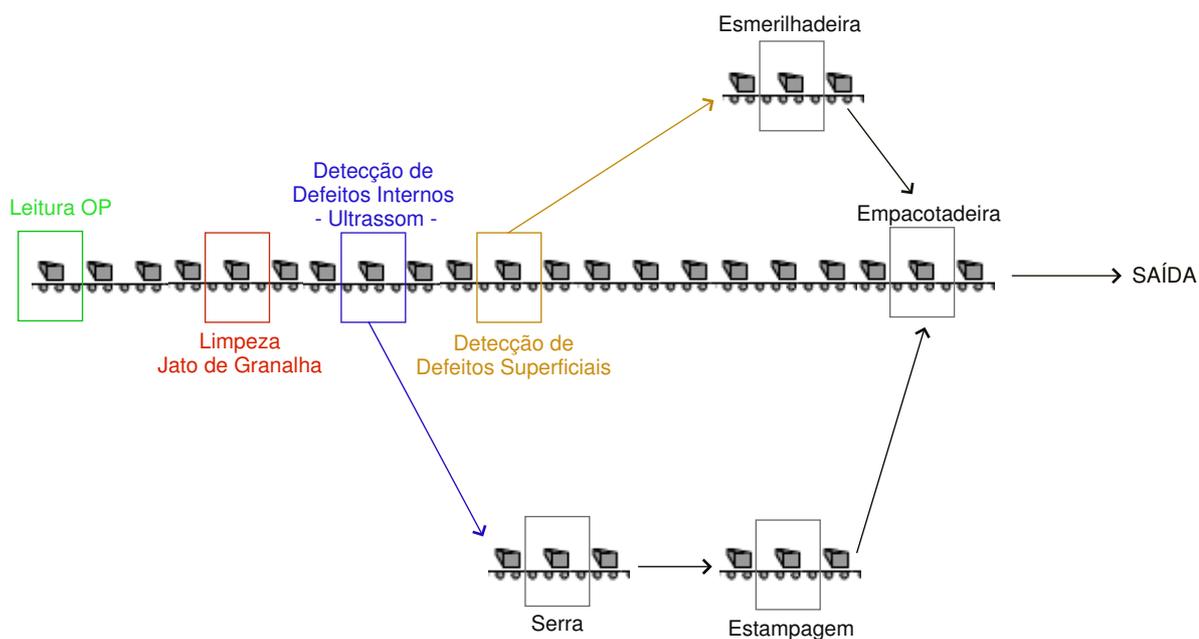


Figura 5: Sistema de inspeção dos tarugos. Os defeitos detectados são reparados. Os tarugos são então empacotados.

para o sistema de nível 2 (Figura 6). Isto agilizará o processo de inspeção, propiciando um maior escoamento da produção.

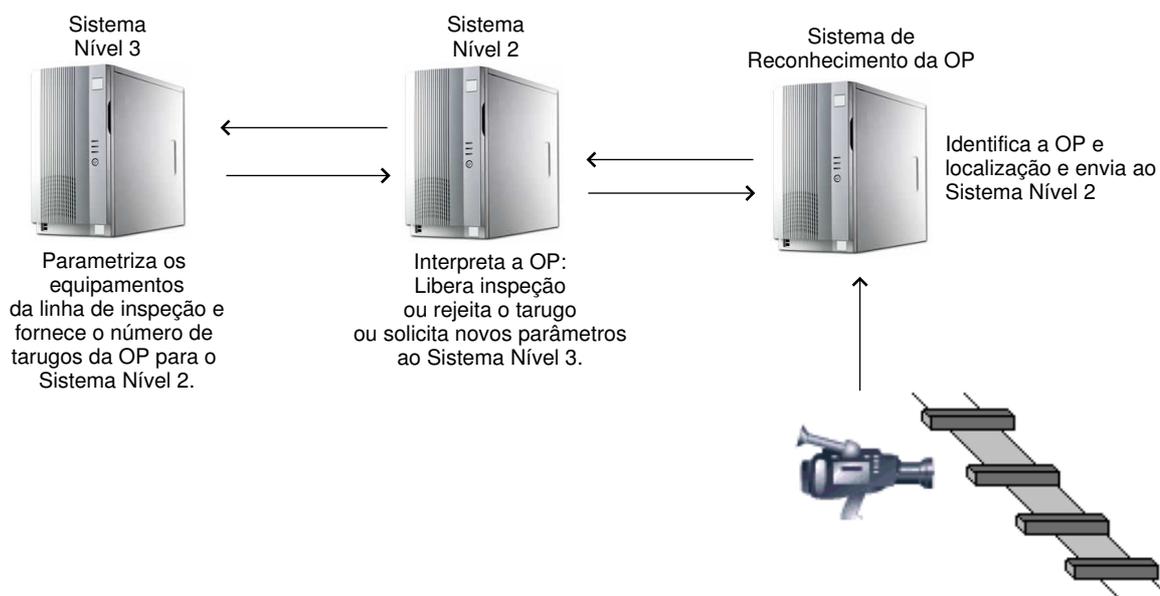


Figura 6: Integração do sistema de visão com os sistemas da automação.

1.1 Motivação do projeto

Apesar de existirem diversas ferramentas de análise de imagem, não foi encontrado nenhum produto industrial que identifique números estampados e forneça bom desempenho em um ambiente não controlado como o da linha de inspeção da Açominas. Este projeto é, portanto, importante para a Açominas e para a UFMG, que podem juntar forças e desenvolver uma solução, até certo ponto, inédita na indústria.

1.2 Especificações do Projeto

O índice de erro requerido pela Açominas é de 0%, uma vez que a identificação errada de uma OP acarreta conseqüências imediatas para o cliente final [39]. Há uma tolerância para o índice de acerto, que pode ficar em torno de 95%. O sistema deve ser capaz de identificar possíveis erros e repassar ao operador da LIT. O operador seria, então, notificado para solucionar a dúvida. Para obtenção deste índice, algumas características do processo de inspeção devem ser analisadas. O ambiente e as condições dos tarugos na LIT provocam algumas dificuldades para o sistema de identificação. São elas [39]:

- variações de distância do tarugo em relação a aproximação frontal da câmera (influenciaria no ajuste de foco);
- variações de posicionamento do tarugo no eixo lateral (podendo não parar no ponto exato do ângulo de visada da câmera);
- proximidade dos tarugos - os tarugos podem vir lado a lado, sem espaço entre eles, o que pode dificultar o processo de detecção do número da OP e identificação do mesmo;
- rotação dos tarugos, podendo deixar os números da OP em quatro posições diferentes;
- tempo reduzido de processamento;
- impressão sobre a parte rugosa.

Uma análise mais detalhada da LIT auxilia a compreensão de alguns destes conceitos. Na entrada da LIT, os tarugos são descarregados por guinchos com eletroímãs em uma esteira (Figura 7). O alinhamento dos tarugos nesta etapa não é controlado. A esteira

Tabela 1: Resumo dos problemas e características dos posicionamentos da câmera na linha de inspeção, conforme Figura 7

Posicionamento da Câmera *	Tempo de Processamento	Problemas		
		Variação da distância frontal para a câmera	Proximidade dos tarugos	Possibilidade de sair da visada da câmera
Posição 1	1min	sim	sim	sim
Posição 2	1min	sim	não	não
Posição 3	15s	não	não	não
* Conforme Figura 6				

conduz os tarugos para o braço mecânico, que posiciona o tarugo nos rolamentos que irão direcioná-lo para as etapas de inspeção. Ao depositar o tarugo na linha, o braço mecânico já prepara o próximo, que só é descarregado nos rolamentos depois que o tarugo inicial colidiu com o batente do final da linha, cerca de 1 min depois da entrada no rolamento.

Com base na Figura 7, três localizações podem ser propostas para o posicionamento da câmera: em frente da esteira, em frente do braço mecânico ou acima do rolamento que direciona o tarugo para a inspeção. O posicionamento na esteira apresenta três desvantagens: a de apresentar tarugos muito próximos, a de que os tarugos podem se apresentar fora do ângulo de visada da câmera e a presença de variações na distância dos tarugos para a câmera. Em frente ao braço mecânico dois destes problemas são contornados, uma vez que o tarugo fica isolado dos demais e não sai do ângulo de visada da câmera. No rolamento de inspeção, há a possibilidade de instalação de um dispositivo que interrompa momentaneamente o processo de encaminhamento do tarugo, solucionando também o problema de distância do tarugo para a câmera. Esta solução, proposta pela Açominas, contorna três dos itens descritos anteriormente, mas necessita de um tempo menor de processamento, cerca de 15", uma vez que interrompe o processo de inspeção, gerando um aumento do tempo de permanência do tarugo na LIT. Nos outros posicionamentos há um aproveitamento de um tempo inutilizado. Um resumo sobre as características e problemas dos posicionamentos está descrito na Tabela 1.

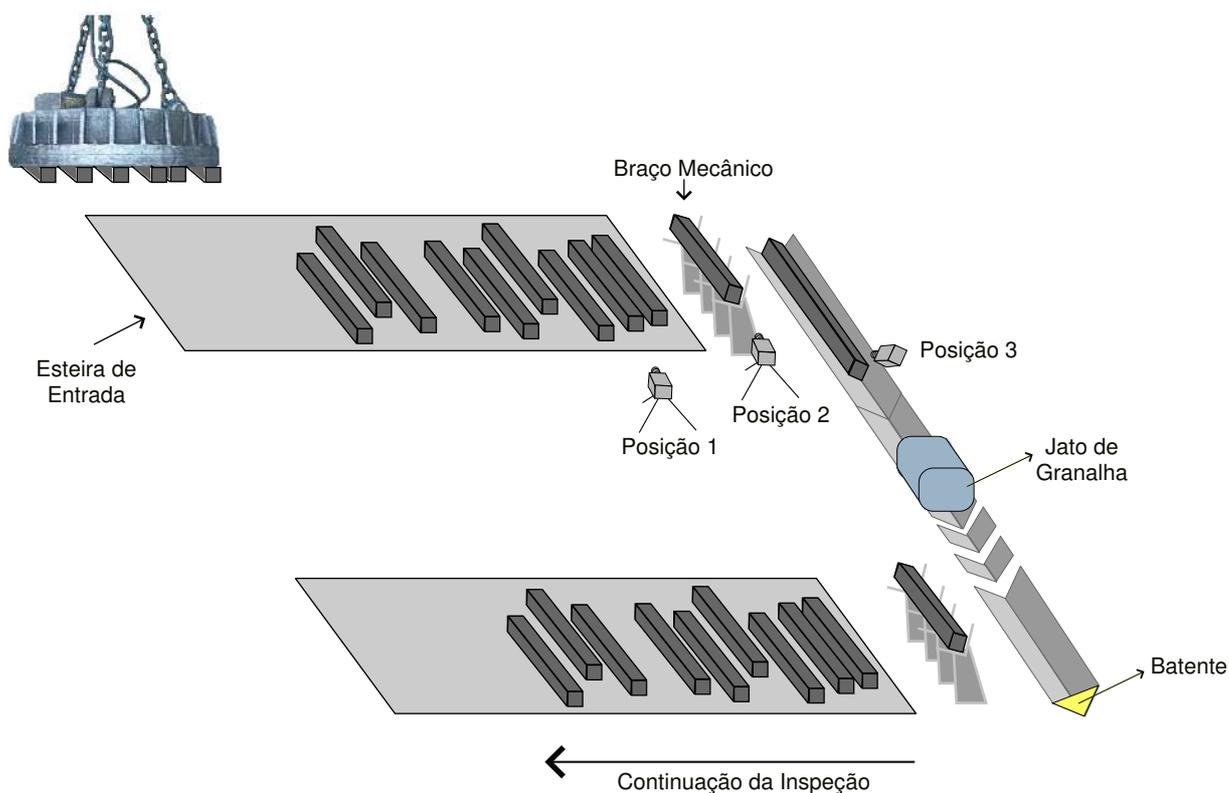


Figura 7: Descrição e definição de possíveis posicionamentos da câmera na LIT.

1.3 Contribuição da Dissertação

Esta dissertação visa pesquisar e desenvolver ferramentas para a análise da imagem do tarugo capturada. A parte de *software* desenvolvida leva em consideração as dificuldades e as características do processo de inspeção dos tarugos. Pelas especificações do projeto, o tarugo pode apresentar variações de rotação e distância. As duas características são muito importantes no desenvolvimento de algoritmos para reconhecimento de caracteres.

Com as imagens coletadas em laboratório, a identificação da rotação forneceu 100% de acerto e os dígitos foram corretamente reconhecidos em 98% das imagens com bom contraste e pouca deformação. Uma melhora na iluminação foi proposta nas imagens coletadas na LIT, fornecendo requisitos importantes para implementação do sistema.

1.4 Organização do Trabalho

Os capítulos irão mostrar o desenvolvimento do trabalho. No Capítulo 2 será feita uma definição das etapas do problema de reconhecimento de caracteres e uma revisão sobre as ferramentas utilizadas e os principais aspectos considerados para a escolha. No Capítulo 3, a metodologia empregada será detalhada, assim como a demonstração de resultados em imagens coletadas em laboratório. No Capítulo 4 serão demonstrados resultados com imagens coletadas na linha de inspeção. No Capítulo 5 será apresentada a conclusão do trabalho, discutindo os resultados e apresentando perspectivas de trabalho futuro.

2 *Revisão da Literatura*

2.1 Introdução

Neste capítulo será apresentada uma revisão sobre os métodos utilizados no reconhecimento de caracteres, com as diversas áreas envolvidas no processo. Primeiramente, será realizada a definição das etapas e em seguida serão apresentadas as técnicas utilizadas na literatura. Uma discussão sobre os métodos e sobre o problema encerra o capítulo.

2.2 Definição das Etapas do Problema

Na literatura existem várias técnicas desenvolvidas para a análise, reconhecimento e processamento de imagens, aplicadas a uma grande variedade de problemas. O sistema é usualmente dividido em etapas [18, 57, 37], para desenvolver uma solução completa e confiável. Isto organiza o trabalho e facilita o desenvolvimento e aplicação de métodos. A Figura 8 apresenta uma formulação utilizada na descrição das etapas em problemas de reconhecimento de caracteres, que se assemelha às configurações descritas por Gonzales [18], Taxt [57] e Morita [37]. As etapas serão descritas a partir desta figura.

A etapa de captação é responsável pela aquisição do sinal, fornecendo uma imagem digital para a etapa de pré-processamento. Fazem parte desta etapa itens como iluminação, câmera e placas de conversão A/D.

A primeira etapa, após a captação da imagem, é a fase de pré-processamento. O objetivo principal desta etapa é realçar e melhorar a condição das imagens, a fim de aumentar o desempenho das etapas posteriores [18]. Com o mesmo objetivo, as técnicas de pré-processamento podem ser utilizadas antes de outras etapas, como extração de características, conforme Figura 8.

A segunda etapa é a segmentação dos caracteres, ou seja, a isolação completa da sentença ou seqüência numérica em sub-imagens contendo apenas um caractere. O desen-

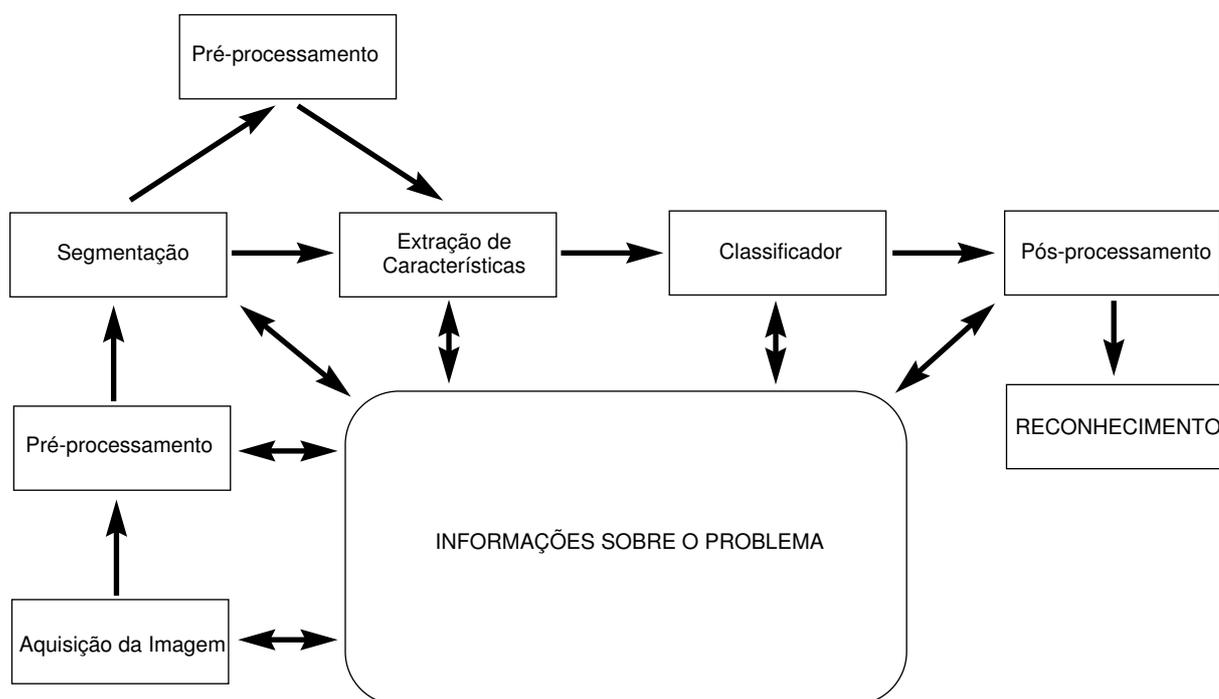


Figura 8: Etapas do problema de reconhecimento de caracteres, com base nas divisões propostas por Gonzales [18] e Taxt [57].

volvimento de um processo de segmentação automática eficiente é um dos grandes desafios na área de processamento de imagens [18]. Esta etapa influi diretamente no desempenho geral do processo, uma vez que erros neste estágio geram, necessariamente, falhas no reconhecimento.

Os caracteres segmentados são encaminhados para a etapa de extração de características que fornece variáveis relevantes para a etapa de classificação. O vetor de variáveis extraído da imagem pode apresentar invariabilidade a algumas distorções e variações presentes no problema [57], como escala e rotação, além de reduzir a dimensão do problema, o que agiliza e facilita a etapa de classificação. Além de ser uma etapa importante, a extração de características pode requerer uma etapa de pré-processamento, como indicado na Figura 8. Existem métodos de extração que trabalham com imagens em escala de cinza, com imagens limiarizadas, com o esqueleto das imagens (ou imagem vetorial ou “*wire frame*” da imagem) ou com o contorno do objeto. As imagens A, B, C e D, na Figura 9, representa os tipos de imagens citadas.

A próxima etapa, a de classificação, objetiva construir um procedimento para atribuir classes pré-determinadas ao vetor de variáveis extraídos da imagem do caractere [56]. A linha que separa as etapas de extração de características e classificação é tênue [14]. Uma

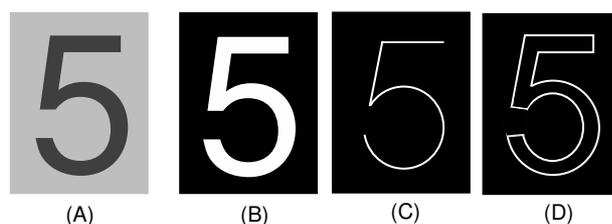


Figura 9: Tipos de figuras utilizadas em algoritmos de reconhecimento de caracteres. Em (A): Imagem em escala de cinza. Em (B): Imagem limiarizada. Em (C): Esqueleto da imagem. Em (D): Contorno da imagem.

etapa de extração ideal facilita a classificação e a recíproca é verdadeira. Mas, de um modo geral, a extração de características depende do problema [14] e a classificação do comportamento dos dados.

Após a etapa de classificação, uma forma de determinar e, se possível, corrigir eventuais erros é aplicação de pós-processamento. Nesta etapa, o contexto dos caracteres é analisado, ratificando ou corrigindo a resposta. A análise de contexto é largamente empregada em problemas de reconhecimento de letras, através da análise não só da palavra, como da sentença formada. Em problemas envolvendo reconhecimento de números e seqüências numéricas, na maioria das aplicações, não há restrições de contexto [37].

Todas as etapas descritas anteriormente podem ser modificadas e adaptadas a partir de um conhecimento prévio do problema. O bloco identificado como “Informações sobre o Problema” na Figura 8 fornece subsídios relevantes para as etapas, incorporando novos elementos à base de dados [18], a partir do processamento realizado nas etapas. Estas novas informações podem modificar parâmetros de etapas posteriores e anteriores, sendo possível gerar uma realimentação para a otimização da resposta [18]. Com a análise correta das informações sobre o problema é possível definir corretamente as etapas, restringir possibilidades de erros e agilizar o processo. Um exemplo do fornecimento de informações para etapas é a aplicação do pós-processamento, que necessita de um conhecimento prévio do problema para a formulação de uma análise de contexto eficiente.

Nas próximas seções serão discutidos as características de cada etapa, os trabalhos e técnicas propostas na literatura.

2.3 Pré-processamento

Nesta seção serão descritas as técnicas que objetivam preparar as imagens para a análise e medição de características e estruturas presentes [51]. A análise das imagens nas etapas de segmentação e extração de características geralmente necessitam que certas propriedades estejam bem definidas, como contorno, linhas, brilho, textura e possíveis combinação entre estes fatores [51]. O tipo de análise realizada nas etapas posteriores determinará o tipo de pré-processamento realizado [51]. Os métodos de detecção de contorno, limiarização, esqueletização, análise de textura e detecção de linhas serão discutidas nesta seção.

Nesta linha de pensamento, pode-se dividir os métodos desta etapa em duas propriedades dos *pixels* das imagens [18]:

- descontinuidade;
- similaridade.

A descontinuidade objetiva a detecção das variações que ocorrem entre *pixels* próximos, como as bordas de objetos e linhas. A similaridade agrega regiões que possuem características semelhantes, como as técnicas de limiarização. Os métodos de esqueletização são aplicados na imagem limiarizada e serão brevemente discutidos.

2.3.1 Limiarização

As técnicas de limiarização são utilizadas para realçar a informação do caractere em problemas de reconhecimento. Os métodos propõem a escolha de um limiar que forneça a máxima separação da informação do fundo de imagem para o caractere. A partir de uma imagem I com 256 níveis de cinza, a limiarização define um limiar L , formando uma imagem I_L segundo a relação:

$$I_L = \begin{cases} 1, & \text{se } I > L \\ 0, & \text{se } I \leq L \end{cases}$$

Os algoritmos de limiarização podem ser divididos em duas áreas: técnicas globais e locais. Os métodos globais definem um único limiar para a imagem, analisando, geralmente, a informação do histograma. Os métodos locais adaptam o limiar com informações locais,

dividindo e analisando partes da imagem original. Fator que aumenta a complexidade computacional [58, 59].

Os métodos de limiarização global, dos quais se destacam o algoritmo de Otsu [41], apresentam bons resultados em imagens com histogramas bimodais, nas quais há a presença notória de duas classes, conforme exemplificado na Figura 10. Contudo, em imagens complexas, sujeitas, por exemplo, a uma iluminação não uniforme, o desempenho decai. Os estudos realizados por Trier e Taxt [58, 59] avaliaram técnicas locais e globais de limiarização, com uma comparação em imagens com iluminação não uniforme. Os algoritmos locais obtiveram melhor desempenho, especialmente o algoritmo de Niblack [38], sujeito a um pós-processamento desenvolvido por Yannowitz e Bruckstein [62]. O método global com melhor desempenho foi o de Otsu [41]. A Figura 11 demonstra um exemplo deste tipo de imagem com os resultados comparativos da aplicação do método global de Otsu [41] e o local de Niblack [38].

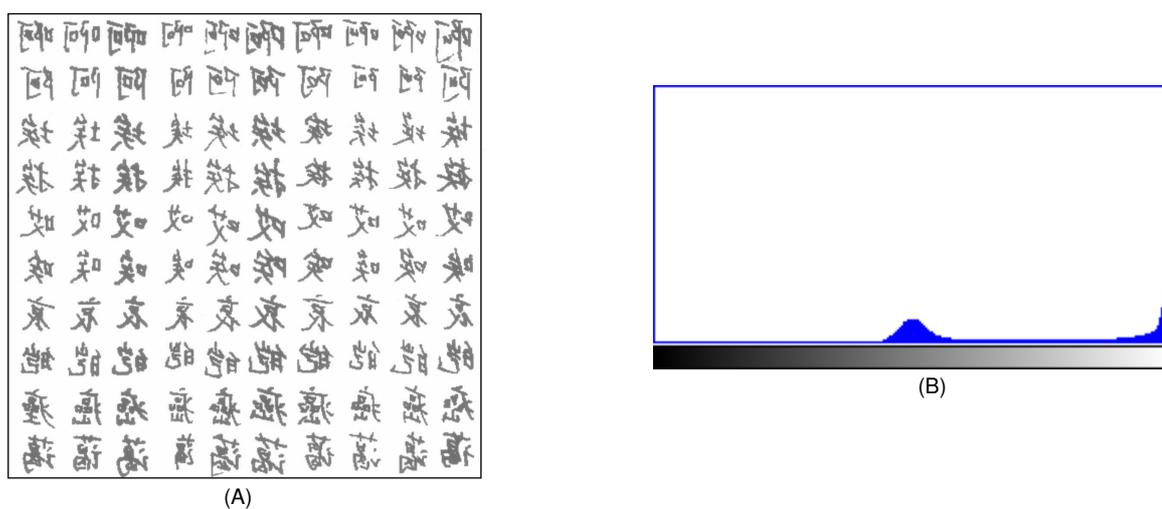


Figura 10: Exemplo de imagem com histograma bimodal. Em (A): Imagem de caracteres coreanos. Imagem adaptada da base de dados disponível em [27]. Em (B): Histograma da Imagem em (A).

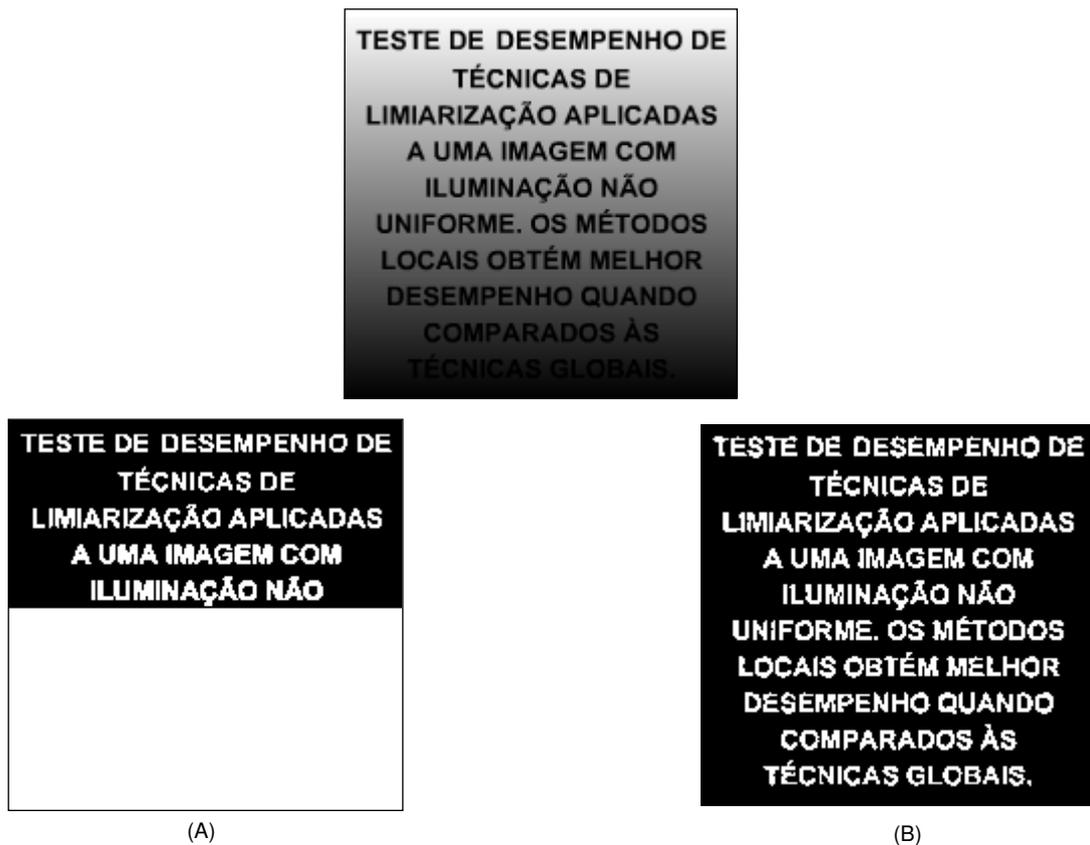


Figura 11: Exemplo de uma imagem com iluminação não uniforme. Em (A): Resultado da aplicação do método global de Otsu [41]. Em (B): Resultado da aplicação do método local de Niblack [38].

2.3.1.1 Método de Niblack

Este método foi proposto por Niblack em 1986 [38] e é largamente utilizado em algoritmos de limiarização local [58, 59]. Consiste em deslocar uma janela sobre a imagem original, calculando os limiares localmente. Considerando uma janela de dimensão $N \times N$ pixels e uma imagem I , o limiar é calculado segundo a equação:

$$L = \mu(i, j) + k\sigma(i, j)$$

$$i = 1, \dots, N$$

$$j = 1, \dots, N$$
(2.1)

$\mu(i, j)$ - Média da janela espacial com $i = 1, \dots, N$ e $j = 1, \dots, N$.

$\sigma(i, j)$ - Desvio padrão da janela espacial com $i = 1, \dots, N$ e $j = 1, \dots, N$.

k - Fator de peso, parâmetro do método.

Os parâmetros para esta técnica são o tamanho N , em *pixels*, da janela $N \times N$ e o peso k . O tamanho da janela deve ser uma solução de compromisso, uma vez que pequenas dimensões preservam detalhes enquanto maiores reduzem ruídos [59]. O valor do peso k atua principalmente nas bordas dos objetos, nas quais há *pixels* de transição, contendo valores intermediários entre objeto e o fundo da imagem [59]. O peso define quanto da borda do objeto será agregada à imagem limiarizada [59].

2.3.1.2 Método de Otsu

O método de Otsu é uma forma de limiarização baseada na análise do histograma da imagem. Uma das vantagens é apresentar característica não paramétrica [35]. O histograma da imagem é utilizado para estimar a função probabilidade dos valores de intensidade dos *pixels*. Em um problema de duas classes, caracterizado por um histograma bimodal, o objetivo é achar o limiar que minimize a variância intraclasse, resultando em regiões homogêneas [53].

A função probabilidade $p(i)$ é estimada pela frequência relativa dos valores dos *pixels*. Considerando uma imagem I de dimensão $M \times N$ *pixels* com 256 níveis de cinza, a função probabilidade pode ser estimada:

$$p(i) = \frac{Ni}{NT} \tag{2.2}$$

$$i = 0, \dots, 255$$

$$NT = M * N$$

Ni - Número de ocorrências do valor i na imagem.

NT - Número total de *pixels* na imagem. $NT = M \cdot N$.

A Equação 2.2 é o histograma da imagem I normalizada pelo total de amostras (*pixels*) $M \cdot N$.

Considerando um limiar L que separe a função probabilidade $p(i)$ em duas classes C_1 e C_2 , suas probabilidades podem ser definidas:

$$P_{C_1}(L) = P(0 \leq i \leq L) = \sum_{i=0}^L p(i)$$

$$P_{C_2}(L) = P(L+1 \leq i \leq 255) = \sum_{i=L+1}^{255} p(i) \quad (2.3)$$

$$P_{C_1}(L) + P_{C_2}(L) = 1$$

As variâncias e médias de cada classes podem ser calculadas:

$$\sigma_{C_1}^2(L) = E[(i - E(i|C_1))^2] = \sum_{i=0}^L [i - \mu_1(L)]^2 p(i|C_1) \quad (2.4)$$

$$\sigma_{C_2}^2(L) = E[(i - E(i|C_2))^2] = \sum_{i=L+1}^{255} [i - \mu_2(L)]^2 p(i|C_2)$$

$$\mu_{C_1} - \text{Média da classe 1. } \mu_{C_1} = E[i|C_1] = \sum_{i=0}^L ip(i|C_1)$$

$$\mu_{C_2} - \text{Média da classe 2. } \mu_{C_2} = E[i|C_2] = \sum_{i=L+1}^{255} ip(i|C_2)$$

$$p(i|C_1) = p(i)/P_{C_1}(L)$$

Para avaliar a escolha do limiar, pode-se definir três variâncias, a variância intraclasse (σ_{intra}), inter-classe (σ_{inter}) e a variância total (σ_{total}):

$$\sigma_{intra}^2 = P_{C_1}(L)\sigma_{i|C_1}(L) + P_{C_2}(L)\sigma_{i|C_2}(L)$$

$$\sigma_{inter}^2 = P_{C_1}(L)[\mu_{i|C_1} - \mu_i]^2 + P_{C_2}(L)[\mu_{i|C_2} - \mu_i]^2 \quad (2.5)$$

$$\sigma_{total}^2 = \sum_{i=0}^{255} (i - \mu_i)^2 p(i)$$

Pode ser provado [53] que há uma relação entre as variâncias descritas em 2.5:

$$\sigma_{total}^2 = \sigma_{intra}^2 + \sigma_{inter}^2 \quad (2.6)$$

A relação em 2.6 mostra que a variância total depende das variâncias intra e inter-classe. Mas não depende do limiar L , conforme Equação 2.5. Portanto, o limiar L selecionado para minimizar σ_{inter}^2 será o mesmo que irá maximizar σ_{intra}^2 [53]. A escolha do limiar é realizada segundo a fórmula [3, 35, 53]:

$$\sigma_{inter}^2(L^*) = \max_{0 \leq L \leq 255} [\sigma_{inter}^2(L)] \quad (2.7)$$

2.3.2 Detecção de Bordas

A borda de um objeto pode ser denominado como o conjunto de *pixels* que delimitam duas regiões com distintos valores de luminância [18]. Existem duas abordagens básicas para a detecção de contorno [43]: detecção por diferenciação e detecção baseada em modelos. A maioria dos métodos propostos e utilizados na literatura são compostos por operadores de diferenciação [18]. Por este motivo, o estudo conduzido nesta seção demonstrará os princípios e ferramentas desta abordagem.

A aplicação dos métodos de diferenciação é realizada pelo gradiente para derivadas primeiras e pelo laplaciano para derivadas segundas [18]. Contudo, o cálculo do gradiente fornece melhor desempenho, especialmente em imagens contendo ruído [51]. Por este motivo são propostos na literatura diversos métodos para a aproximação do gradiente discreto, através do cálculo das diferenças entre *pixels* adjacentes [51]. As técnicas definem janelas espaciais que são convoluídas com a imagem original, fornecendo uma aproximação do cálculo do gradiente para cada pixel.

Com relação à direção do cálculo do gradiente, as técnicas podem ser: ortogonais (realizam o cálculo nas direções verticais e horizontais) ou multi-derecionais.

As técnicas multi-derecionais calculam o gradiente em várias direções, escolhendo, em cada pixel, o maior valor. Supondo M direções, janelas H_m (para cada direção) e uma imagem I , o gradiente G é estimado [43]:

$$\begin{aligned} G &= \max[G_1, \dots, G_m] \\ G_m &= I * H_m \\ m &= 1, \dots, M \end{aligned} \tag{2.8}$$

Os métodos ortogonais combinam as duas estimativas:

$$G = \sqrt{G_x^2 + G_y^2} \tag{2.9}$$

Dentre as principais técnicas ortogonais, pode-se destacar os métodos [43]: Sobel [54], Roberts [48], Prewit [44] e Frei-Chen [16], que apresentam características particulares de desempenho:

Sobel Apresenta maior sensibilidade a contornos diagonais do que a contornos verticais e horizontais [43].

Prewitt Apresenta maior sensibilidade a contornos verticais e horizontais do que a contornos diagonais [43].

Frei-Chen O valor do gradiente é o mesmo para contornos verticais, horizontais e diagonais [43].

Roberts Maior sensibilidade a contornos diagonais, sendo mais sensível a ruídos do que os outros métodos [43].

Dentre as técnicas multidirecionais, pode-se destacar os métodos: Kirsch [25] e Robinson [49]. A definição das janelas espaciais propostas pelas técnicas descritas e um estudo mais detalhado sobre o desempenho dos métodos podem ser encontrados em [43].

2.3.3 Detecção de Linhas

Esta seção descreverá brevemente dois dos métodos mais utilizados no reconhecimento de linhas em uma imagem: Transformada de Hough [20] e RANSAC [15]. Maiores detalhes e explicações podem ser encontradas em [18, 43].

Os métodos descritos necessitam de etapas anteriores, que forneçam pontos relevantes para estimação do modelo. Os métodos trabalham com as dimensões x e y dos pontos fornecidos e não com seus valores. Pode ser considerado que estes métodos identificam a imagem de entrada como um conjunto bidimensional de dados. Estes pontos devem ser, portanto, estimados a partir da imagem em escala de cinza. Métodos como detecção de contorno e limiarização devem ser utilizados.

Transformada de Hough Método proposto e patenteado em 1962 por P. V. C Hough [20] para reconhecimento de linhas em uma imagem. Consiste em analisar a imagem pelo plano dos parâmetros, plano ab , analisando a equação da reta como: $-b = y - ax$. Primeiramente, são definidos os valores possíveis dos parâmetros a e b . Para cada ponto $I(x, y)$, no plano da imagem, são estimados os parâmetros b para cada a . Cada b_i estimado a partir de um determinado a_j modifica o valor do ponto $P(i, j)$ no plano ab : $P(i, j) = P(i, j) + 1$. No final do processo, o ponto $P(i, j)$ definirá quantos pontos $I(x, y)$ da imagem pertencem à reta de equação $a_j x + b_i = y$.

RANSAC O método RANSAC (*Random Sample Consensus*) foi proposto por Fischler e Bolles em 1981 [15], sendo utilizado para procurar modelos matematicamente definidos em um conjunto de dados. O método consiste em estimar os parâmetros de modelos sem os dados que normalmente contaminam as amostras (ruídos e erros de medição, por exemplo). Um exemplo ajudará a entender o conceito. A utilização do RANSAC para detectar linhas em um conjunto bidimensional de pontos (uma imagem limiarizada, por exemplo) consiste em achar o mínimo requisito para a formação da linha (dois pontos). O algoritmo seleciona aleatoriamente dois pontos (chamado de subconjunto do modelo), desenvolve o modelo da reta e agrega os pontos que estejam próximos ao modelo desenvolvido, o que descarta ruídos. Se houver pontos suficientes agregados ao modelo, eles são integrados à base de dados e os parâmetros são novamente estimados. Caso contrário, o modelo é descartado. Este algoritmo é muito utilizado para detectar linhas e elipses [63], e pode ser utilizado para encontrar qualquer modelo matematicamente definido (círculos e arcos, por exemplo).

2.3.4 Esqueletização

Os algoritmos de esqueletização atuam em imagens limiarizadas e objetivam reduzir a forma dos objetos às mínimas estruturas capazes de identificá-los [18]. Por isso são usualmente chamados de “esqueletos” da imagem. Este afinamento da imagem é extremamente sensível, podendo causar deformações não presentes na imagem limiarizada e causar erros de classificação [57]. A aplicação em imagens degradadas, com ruídos e incompletas pode comprometer a solução do problema. Um estudo sobre as diversas técnicas, com comparação de desempenho, pode ser encontrado em [31].

2.4 Segmentação

A segmentação de caracteres tem como objetivo o particionamento da imagem original em sub-imagens contendo somente a informação de caracteres [37]. A segmentação correta dos caracteres é extremamente difícil, principalmente em caracteres cursivos [37]. Para caracteres isolados, algumas técnicas simples obtêm um bom desempenho. São os casos das projeções de histograma nos dois planos da imagem [10]. Se não houver ruídos, a separação ocorre sem erros. Contudo, em imagens complexas, com fundo de imagem não uniforme, a segmentação não é uma tarefa simples. Por isso alguns autores utilizam técnicas mais robustas, como análise de textura para segmentar textos em imagens [12,

60, 32, 61].

Uma outra forma de contornar o problema de segmentação é dividir em sub-imagens e, em cada uma, aplicar as etapas seguintes. Segundo Lecoliet [10], esta solução pode ser vista como reconhecimento sem segmentação. Como não há segmentação explícita, os erros desse método são exclusivos das etapas de extração de características e classificação. A resposta do sistema de reconhecimento fica mais confiável [37], mas necessita de um esforço computacional, muitas vezes, inviável [37, 10].

2.4.1 Análise de Textura

A análise de textura é um campo importante na área de processamento de imagens com aplicações em diversas áreas [46], como segmentação de imagens médicas e aéreas [42]. A aplicação destas ferramentas apresenta grande importância na área de reconhecimento de caracteres, uma vez que regiões contendo textos na imagem podem ser consideradas um tipo de textura [46]. O reconhecimento de textura, portanto, é utilizado para delimitar regiões na imagem original que contenham textos [60, 32, 61, 12].

A análise de textura pode ser considerada uma etapa de extração de características quando aplicadas a classificação de texturas. O método foi alocado na etapa de segmentação porque, em problemas de reconhecimento de caracteres, não se procura classificar texturas mas segmentar possíveis regiões na imagem que contenham textos.

Várias abordagens são utilizadas para identificação de textura. Destas, duas podem ser destacadas [46, 42, 47, 11]:

Estatística Método de co-ocorrência [45], aplicação de ferramentas estatísticas de variância [51] e momentos de ordem superior [6].

Aplicação de técnicas de filtragem Aplicação de filtros ou banco de filtros espaciais [46]. Nesta área encontram-se muitos trabalhos desenvolvidos, com diversas configurações de filtros. Um trabalho pioneiro em filtros para análise de textura são as máscaras desenvolvidas por Laws [28]. Dentre as técnicas mais utilizadas estão: *wavelets* [33, 12], filtros de Gabor [22], filtros de Laws, etc.

Os diferentes tipos de filtros da área de processamento de sinais objetivam realçar características particulares das imagens, como detalhes horizontais e verticais (filtros de Laws [28]), diferentes orientações (filtros de Gabor [22]), sendo que sua escolha dependerá

do problema. Uma referência sobre a aplicação de processamento de sinais pode ser encontrada em [47, 46].

2.5 Extração de Características

A imagem, na etapa de extração de características, é convertida em vetores de características que deverão apresentar, idealmente, máxima diferença extraclasse e mínima variabilidade intraclasse [57].

Esta etapa está intrinsecamente ligada ao comportamento do problema [14], uma vez que existem diferentes tipos de características em diferentes tipos de problemas. Métodos utilizados na classificação de impressões digitais, por exemplo, não são aplicáveis em problemas de classificação de voz. O tipo de problema e o seu comportamento devem ser, portanto, cuidadosamente analisados.

Em reconhecimento de caracteres há vários métodos para extração de características. O comportamento do problema irá novamente auxiliar na escolha, já que os métodos apresentam diferentes propriedades, como invariância a rotação, escala, variabilidade dos caracteres e nível de degradação. A escolha de um método que seja invariante à rotação, por exemplo, não distinguiria o número “6” do número “9”, nem as letras “u” e “n” [57]. Além destes fatores, é necessário considerar a complexidade computacional e a integração com as outras etapas. O método de extração irá determinar e definir parâmetros nas etapas de pré-processamento e classificação.

Como mencionado na Seção 2.2, são quatro tipos de imagens utilizadas nos métodos de extração: imagem em escala de cinza, imagem limiarizada, “esqueleto” da imagem e contorno da imagem. A utilização de imagens em escala de cinza requer, de um modo geral, um algoritmo de extração invariante à iluminação e à diferença de contraste [57]. As imagens limiarizadas requerem essa propriedade nos algoritmos de transformação.

Com base nesses tópicos, pode-se listar algumas das principais técnicas utilizadas, explicitando o tipo de imagem utilizada e a aplicação (uma completa referência sobre as técnicas pode ser encontrada em [57]):

Casamento de Padrões Utilizado em problemas com a presença de uma única fonte de impressão. Trabalha com a imagem limiarizada e o resultado é sensível a variações de escala e rotação. A característica extraída é uma medida de similaridade, geralmente correlação, entre a imagem original e os padrões propostos. É possível normalizar a

resposta para que o resultado esteja entre -1 e 1 , possibilitando testar a consistência da resposta. O classificador desenvolvido para esse método é a escolha do maior valor de medição de similaridade acima de limiar definido.

Momentos de Zernike Melhor método de momento pelo estudo de Belkasim [4]. Utilizados em problemas com variabilidade de caracteres. Trabalha com imagem limiarizada e o resultado é sensível a escala, mas a amplitude dos momentos é invariante à rotação. Os momentos são números complexos extraídos a partir de projeções da imagem de entrada em espaços ortogonais [57]. A técnica diminui, portanto, a redundância presente na imagem, diminuindo a quantidade de variáveis disponíveis na etapa de classificação. Apresenta a vantagem de reconstruir a imagem original a partir dos momentos extraídos. Com isso, é possível estimar a quantidade de momentos para uma determinada aplicação [57]. Mas para algumas aplicações pode ser necessário um número extenso de momentos, o que inviabiliza o método.

Muitos problemas apresentam propriedades singulares, sendo difícil encontrar um método ótimo para a extração. Segundo Trier e Taxt [57] e Jain [23], é comum combinar métodos para alcançar melhor desempenho na separação das classes. Alguns autores combinam classificadores sujeitos a diferentes algoritmos para extração de características [52]. Esta característica reforça a idéia utilizada em desenvolvimento de sistemas em engenharia de confiabilidade, ramo que está crescendo nas indústrias nos últimos anos. Com dois ou mais algoritmos de extração de características diferentes, com desempenhos e sensibilidades diferentes, a resposta do sistema fica mais confiável, evitando modos de erro comum. O termo modos de erro comum é largamente utilizado na área de engenharia de confiabilidade, no desenvolvimento de sistemas confiáveis [24]. Um exemplo definirá melhor o conceito. Em um avião, o trem de pouso é um sistema essencial e deve funcionar com a maior confiabilidade. O sistema principal é acionado eletronicamente e o sistema reserva, manualmente. Se os dois sistemas fossem eletrônicos, poderia haver uma única fonte de problema que atacaria os dois sistemas ao mesmo tempo (como uma descarga atmosférica). Com o desenvolvimento de sistemas que tendem a apresentar erros em situações diferentes, a confiabilidade aumenta.

2.5.1 Casamento de Padrões

Casamento de Padrões é uma técnica clássica e um dos métodos mais utilizados no reconhecimento de padrões. Consiste em comparar um modelo pré-determinado a uma

imagem [43]. Um critério de similaridade é utilizado e o objeto na imagem é reconhecido como o modelo se este critério atender aos requisitos mínimos.

Uma medida de similaridade muito utilizada é a correlação cruzada [51, 18, 43, 14]. Neste método, o padrão é deslocado pela imagem e, em cada posição, é multiplicado pelos valores da imagem nesta região. O somatório desta multiplicação é armazenado em cada posição de deslocamento. Considerando uma imagem I de dimensão $M \times N$ *pixels* e um padrão P de dimensão $K \times J$ *pixels*, a correlação cruzada é definida como:

$$R_{IP}(m, n) = \sum_k \sum_j I(i, j)P(k - m, j - n)$$

$$\begin{aligned} m &= 0, 1, \dots, M - 1 \\ n &= 0, 1, \dots, N - 1 \\ j &= 1, 2, \dots, J \\ k &= 1, 2, \dots, K \end{aligned} \tag{2.10}$$

Na Equação 2.10, o somatório é realizado na região da imagem onde o padrão está localizado. A vantagem deste método é a implementação com menor custo computacional por meio da transformada rápida de Fourier (FFT). Pelo teorema da correlação [18, 9, 55, 51], o cálculo da correlação cruzada pode ser realizado no domínio da frequência:

$$P(x, y) \circ I(x, y) \Leftrightarrow \mathcal{P}^*(u, v)\mathcal{I}(u, v) \tag{2.11}$$

O cálculo da correlação cruzada, contudo, apresenta sensibilidade a regiões de brilho intenso na imagem. O resultado fica polarizado, apresentando altos valores mesmo em regiões diferentes do padrão [43, 14, 51, 18, 43]. Para contornar este fator, Duda e Hart [14] e Pratt [43] propõem a correlação cruzada normalizada pela autocorrelação da imagem na região do padrão:

$$N(m, n) = \frac{\sum_i \sum_j I(k, j)P(k - m, j - n)}{\sum_k \sum_j I^2(k, j)} \tag{2.12}$$

$\sum_k \sum_j I(k, j)^2$ - Cálculo da autocorrelação da imagem, na região do padrão.

Esta normalização compensa regiões na imagem original em que há maior brilho, não havendo polarização na resposta. É realizada para verificar se um padrão pertence às imagens testadas [57]. Trier e Taxt [57] propõem uma normalização para encontrar o padrão que melhor se adapta a uma única imagem. A normalização é relativa à autocorrelação

do padrão:

$$N(m, n) = \frac{\sum_k \sum_j I(k, j)P(k - m, j - n)}{\sum_i \sum_j P^2(k - m, j - n)} \quad (2.13)$$

Na maioria dos problemas há várias imagens e vários padrões. A correlação cruzada é sensível a mudanças na amplitude tanto da imagem original quanto do padrão. Russ [51] e Yoshimura [64] define a correlação cruzada normalizada pela média geométrica das autocorrelações:

$$N(m, n) = \frac{\sum_k \sum_j I(k, j)P(k - m, j - n)}{\sqrt{\sum_k \sum_j I^2(k, j) \sum_k \sum_j P^2(k - m, j - n)}} \quad (2.14)$$

$\sum_k \sum_j I(k, j)^2$ - Cálculo da autocorrelação da imagem, na região do padrão.

Esta normalização é mais adequada e a Equação 2.14 será adotada como correlação cruzada normalizada neste trabalho. Na mesma linha de pensamento, Gonzales [18] propõem o cálculo da similaridade através do coeficiente de correlação:

$$\gamma(m, n) = \frac{\sum_k \sum_j [I(k, j) - \overline{I(k, j)}][P(k - m, j - n) - \overline{P}]}{\sqrt{\sum_k \sum_j [I(k, j) - \overline{I(k, j)}]^2 \sum_k \sum_j [P(k - m, j - n) - \overline{P}]^2}} \quad (2.15)$$

$\overline{I(k, j)}$ - Média dos *pixels* da imagem na região do padrão.

\overline{P} - Média dos *pixels* do padrão.

$\sum_i \sum_j [I(k, j) - \overline{I(k, j)}]^2$ - Cálculo da autocovariância da imagem, na região do padrão.

A facilidade de implementação pela FFT não se aplica diretamente ao coeficiente de correlação e nem à correlação cruzada normalizada (Equação 2.14), devido aos cálculos necessários das médias e das autocorrelações locais da imagem. A forma mais utilizada é a aplicação direta das Equações 2.15 e 2.14. Contudo, se considerarmos a imagem limiarizada, é possível realizar o cálculo rápido. Modificando fundo de imagem como -1 e o objeto como 1 ou o objeto como -1 e o fundo de imagem como 1 , a normalização pela média geométrica utilizada na Equação 2.14 pode ser implementada. Considerando uma imagem I (de dimensão $M \times N$ *pixels*) e um padrão P (de dimensão $K \times J$ *pixels*), formados com os valores 1 e -1 , a média geométrica fica:

$$\sqrt{\sum_i \sum_j I^2(i, j) \sum_i \sum_j P^2(i - m, j - n)} = \sqrt{(K \cdot J)(K \cdot J)} = K \cdot J \quad (2.16)$$

A Equação 2.14 pode ser implementada de maneira rápida:

$$N = \frac{\mathcal{F}^{-1} \{ \mathcal{P}^* \cdot \mathcal{I} \}}{K \cdot J} \quad (2.17)$$

A aplicação da correlação cruzada normalizada nos três exemplos propostos na Figura 12 forneceu resultados consistentes. O máximo da correlação normalizada é 1, conforme Tabela 2. O número incompleto apresenta um valor próximo a 1 e o bloco apresenta valor menor do que 0.40.

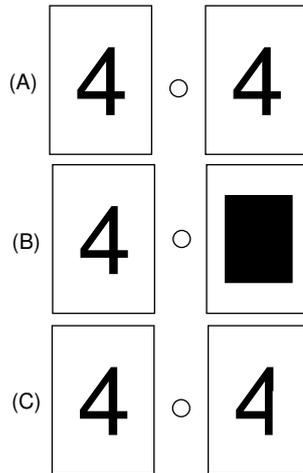


Figura 12: Exemplo de correlação cruzada normalizada entre três imagens. Os resultados estão na Tabela 2.

Tabela 2: Resultado da correlação cruzada normalizada das imagens na Figura 12, com fundo de imagem = -1 e objeto = 1. Os resultados são os mesmos para a situação inversa: fundo de imagem = 1 e objeto = -1

	Resultado da Correlação Cruzada Normalizada
(A)	1.000
(B)	0.398
(C)	0.966

2.6 Classificação

A etapa de classificação recebe as informações da etapa de extração de características e atribui a esses dados uma das classes pré-definidas [14].

O classificador procura particionar o espaço de características em regiões que definam as classes do problema [14]. O desenvolvimento do classificador é mais dependente do comportamento dos dados do que do tipo de problema. Após a etapa de extração, que depende do problema, as características dos dados devem ser cuidadosamente analisadas. O número de classes, as variabilidades intra e extraclasse, a dimensão do espaço de entrada são algumas características relevantes. Algumas topologias de redes neurais, por exemplo, são aplicadas somente a problemas que apresentam os dados de forma linearmente separável [7].

Redes Neurais Artificiais (RNA) são muito empregadas em problemas de classificação em geral e também são aplicadas a padrões em imagem [7, 19], servindo como solução ou complemento das técnicas clássicas [7]. Nos últimos anos, as RNAs têm sido muito utilizadas para reconhecimento de caracteres e dígitos [34, 30, 26, 29, 40]. As topologias mais utilizadas são as MLP (*Multi-Layer Perceptron*), com treinamento baseado em *backpropagation* [50]. Esse algoritmo de treinamento considera a minimização do erro quadrático na etapa de treinamento, podendo gerar problemas como *overfitting* ou *underfitting*. A utilização de algoritmos de treinamento multiobjetivos [13] fornece melhores resultados com a resposta otimizada entre o erro de treinamento e a norma do vetor de pesos da rede, relacionado com a capacidade de generalização da rede final [13]. A topologia MLP é muito utilizada pelo fato de resolver problemas não linearmente separáveis [7] e possuir a característica de aproximação de qualquer função contínua ou qualquer função matemática com a utilização de uma ou duas camadas intermediárias, respectivamente [7]. O conjunto de dados pode, também, polarizar e criar uma certa instabilidade na resposta dos sistemas. Técnicas como o método Bagging [8] manipulam o conjunto de treinamento a fim de apresentar uma solução mais confiável.

Classificadores estatísticos, como o de Bayes, também são utilizados [38]. O desempenho deste método é condicionado pela correta caracterização da função densidade de probabilidade a partir dos dados utilizados no treinamento [17]. Mas estas condições geralmente não são satisfeitas em problemas reais [52]. A maioria das aplicações em processamento de imagens utiliza a distribuição normal para caracterizar os dados, sendo que a estimação dos parâmetros é relativamente simples [38]. Segundo Niblack [38], a

aplicação de distribuições não gaussianas é mais difícil, tanto analiticamente quanto na implementação. A melhor forma de lidar com situações em que as variáveis não apresentam a distribuição gaussiana é dividi-las em classes menores, que possam ser aproximadas por uma gaussiana [38].

2.6.1 Classificador de Bayes

O classificador de Bayes é baseado na fórmula estatística de Bayes para cálculo das probabilidades *a posteriori* de eventos. O conhecimento prévio do comportamento estatístico das variáveis é essencial para o bom desempenho do classificador. Mas este não é o caso das maiorias das aplicações. A função densidade de probabilidade é estimada com dados de um conjunto de treinamento. Estes dados devem ser uma boa representação das classes para que a estimação seja bem sucedida.

Há dois tipos de classificadores que podem ser implementados a partir de uma base de treinamento. O primeiro seria estimar somente os parâmetros de uma distribuição, julgando que se pode afirmar o tipo de distribuição das variáveis consideradas. Este modo é chamado estimação paramétrica [23, 17]. O segundo método, não paramétrico, não define nem a distribuição nem seus parâmetros. Métodos são propostos para este cálculo, como o método de estimação de Parzen [17].

Aplicando a fórmula de Bayes para o problema de k classes, sujeita a um vetor com N variáveis:

$$p(C_i|x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N|C_i)p(C_i)}{p(x_1, \dots, x_N)}, i = 1, \dots, k. \quad (2.18)$$

$p(C_i|x_1, \dots, x_N)$ - Probabilidade do vetor de variáveis apresentadas x_1, \dots, x_N pertencerem a classe C_i . Esta é a probabilidade necessária para classificar o vetor de variáveis.

$p(x_1, \dots, x_N|C_i)$ - Função densidade de probabilidade condicional para a classe C_i , também conhecida como verossimilhança. Pode ser entendido como a probabilidade de um vetor de variáveis da classe C_i apresentar o valor x_1, \dots, x_N [38]. Esta função é estimada pelos dados de treinamento.

$p(C_i)$ - Probabilidade *a priori* de qualquer vetor de características pertencer a classe C_i .

$p(x_1, \dots, x_N)$ - A soma de todas as probabilidades $p(x_1, \dots, x_N|C_i)$ para todo i , conforme teorema de probabilidade total.

Considerando o teorema da probabilidade total [36]:

$$p(C_i|x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N|C_i)p(C_i)}{\sum_{i=1}^k p(x_1, \dots, x_N|C_i)p(C_i)}, i = 1, \dots, k. \quad (2.19)$$

Aplicando ao problema de duas classes, a equação pode ser simplificada:

$$p(C_i|x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N|C_i)p(C_i)}{p(x_1, \dots, x_N|C_1)p(C_1) + p(x_1, \dots, x_N|C_2)p(C_2)}, i = 1, 2. \quad (2.20)$$

Considerando a probabilidade *a priori* como sendo 0.5 para cada classe, pode-se cancelar estes termos da equação:

$$p(C_i|x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N|C_i)}{p(x_1, \dots, x_N|C_1) + p(x_1, \dots, x_N|C_2)}, i = 1, 2. \quad (2.21)$$

A equação fica somente restrita às probabilidades $p(x_1, \dots, x_N|C_i)$. Se considerarmos esta função densidade de probabilidade condicional como sendo gaussiana, a estimação a partir dos dados de treinamento fica relativo ao cálculo do vetor de média das variáveis (dimensão N) e da matriz de covariância (dimensão $N \times N$).

2.7 Conclusões

Este capítulo descreveu as etapas do problema de reconhecimento de caracteres, mostrando as características e as principais ferramentas e estudos realizados em cada área. O estudo desenvolvido auxiliou a desenvolvimento da metodologia proposta no próximo capítulo.

3 Metodologia

3.1 Introdução

Este capítulo demonstra a metodologia desenvolvida para o problema de identificação da ordem de produção OP estampados nos tarugos da Açominas e apresenta os resultados obtidos com imagens coletadas em laboratório, sem condicionamento de distância e iluminação. O problema foi dividido em etapas e estudado paralelamente. Primeiramente será descrita a divisão das etapas e em seguida serão apresentados a metodologia e os resultados obtidos. No final, é apresentada uma discussão sobre o capítulo.

3.2 Definição das Etapas

Como visto no Capítulo 2, o desenvolvimento de um método para reconhecimento de caracteres é extenso e formado por diversas etapas. As etapas devem ser formuladas com base em um conhecimento prévio do problema, para a correta aplicação e desenvolvimento de ferramentas.

As etapas propostas estão descritas na Figura 13. A rotação do tarugo é identificada, os caracteres são segmentados e direcionados para a etapa de extração de características e classificação.

Trabalhou-se com a imagem do tarugo enquadrado para definição da rotação dos números e paralelamente foram desenvolvidos algoritmos para o reconhecimento de caracteres isolados da imagem, utilizando a forma de segmentação explícita, ou seja, os caracteres são completamente segmentados e encaminhados para a etapa de extração de características e classificação. As Figuras 14 e 15 demonstram os planos de trabalho considerados.

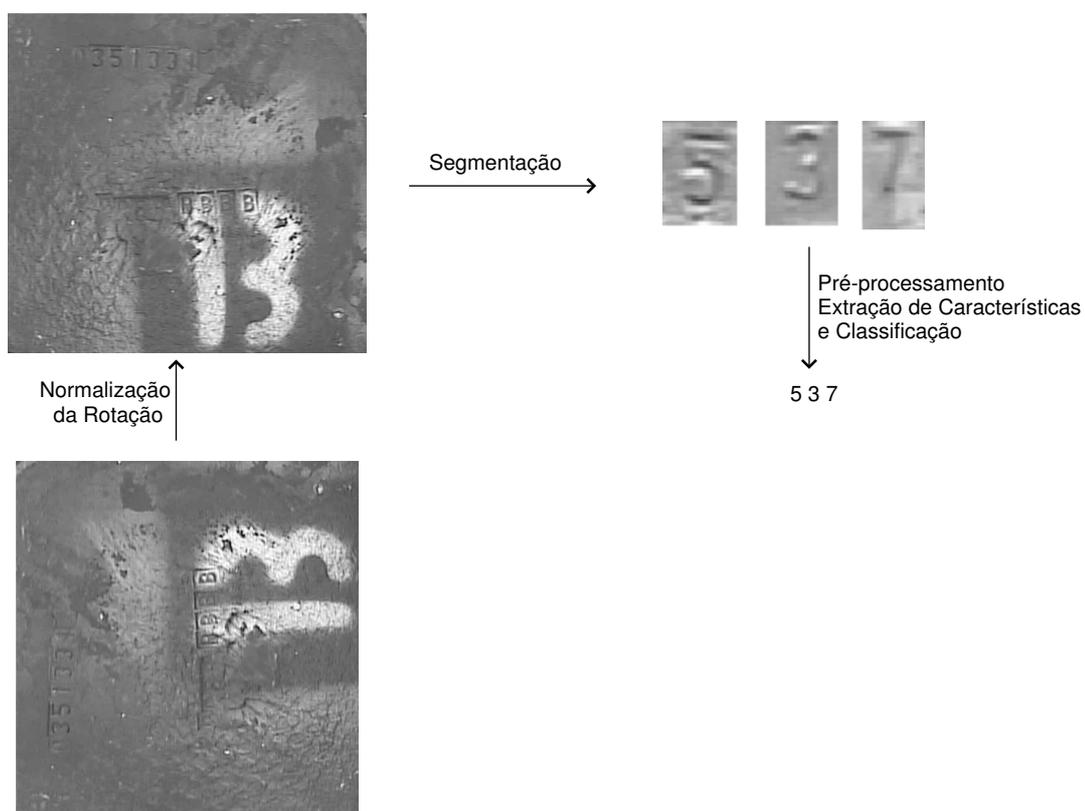


Figura 13: Divisão das etapas do problema, definindo, também, os possíveis planos de trabalho.



Figura 14: Plano de trabalho para reconhecimento da rugosidade e identificação da rotação.

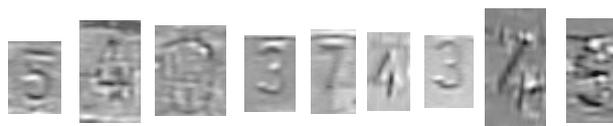


Figura 15: Plano de trabalho para extração de características e classificação.

As próximas seções apresentam as soluções desenvolvidas e os resultados com as imagens reais coletadas em laboratório.

3.3 Identificação da Rotação do Tarugo

A análise da textura da imagem, como mencionado no Capítulo 2, pode determinar a correta rotação do tarugo devido a uma característica particular do processo de corte do tarugo a partir da peça mãe. A peça mãe é produzida e, a partir dela, os tarugos filhos são cortados e estampados. O corte é realizado somente em uma parte, enquanto o restante é rasgado da peça mãe. Esta diferença produz duas texturas diferentes (Figura 16) no tarugo cortado, uma região lisa (efeito do corte) e uma região rugosa (efeito do rasgamento).

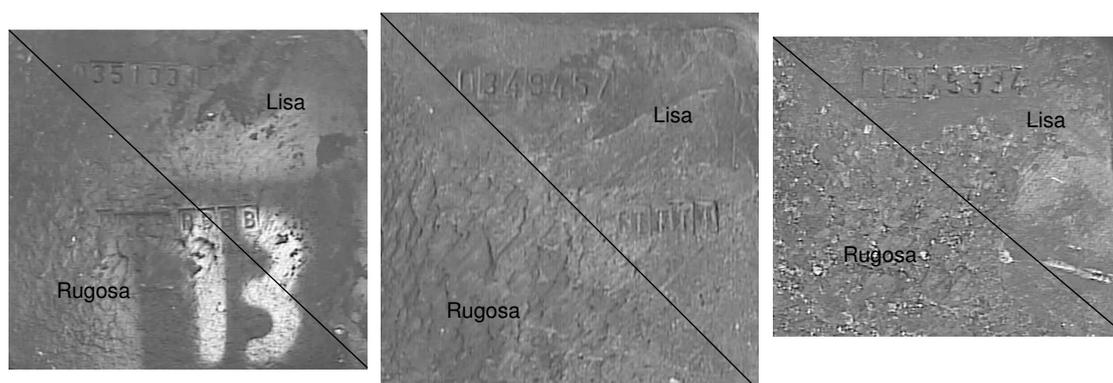


Figura 16: Demonstração das duas regiões distintas no tarugo.

Com a identificação das partes lisa e rugosa é possível determinar exatamente a orientação dos caracteres impressos, conforme Figura 17.

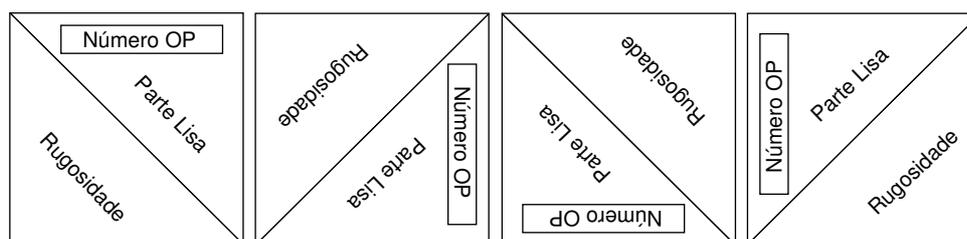


Figura 17: Definição da posição dos caracteres.

Os métodos utilizados na literatura para identificação de textura são aplicados na figura inteira, pois o objetivo é a segmentação completa da imagem [47], [11]. Neste problema é suficiente determinar a posição em que ela se encontra (Figura 17). Portanto, pode-se analisar partes das figuras, diminuindo a complexidade computacional. Este ponto é importante, pois o sistema trabalhará com um tempo fixo reduzido. É aconselhável reservar tempo de processamento para as etapas de extração e classificação. Foram analisados linhas e blocos das extremidades dos tarugos.

3.3.1 Análise das Linhas Extremas

A análise de linhas foi proposta por apresentar uma baixa complexidade computacional. O método consiste em analisar linhas próximas às bordas dos tarugos, a fim de identificar a localização da rugosidade.

Com relação às características analisadas, os métodos propostos levam em conta o fato da textura lisa e rugosa apresentar características básicas distintas. A informação da variância, que em muitos problemas é insuficiente para gerar uma separação eficiente [6], pode ser utilizado neste problema.

As linhas extremas do tarugo (verticais e horizontais) foram analisadas segundo variância, máximos e mínimos. Pelo padrão de corte, pode-se identificar duas linhas rugosas e duas lisas.

Estas informações foram encaminhadas a um classificador de Bayes gaussiano de duas classes, formulado segundo a Seção 2.6.1. O conjunto de treinamento servirá para estimar os parâmetros da distribuição, através do cálculo do vetor de médias e da matriz de covariância das variáveis.

Com o classificador de Bayes é calculada a probabilidade de a linha pertencer à classe

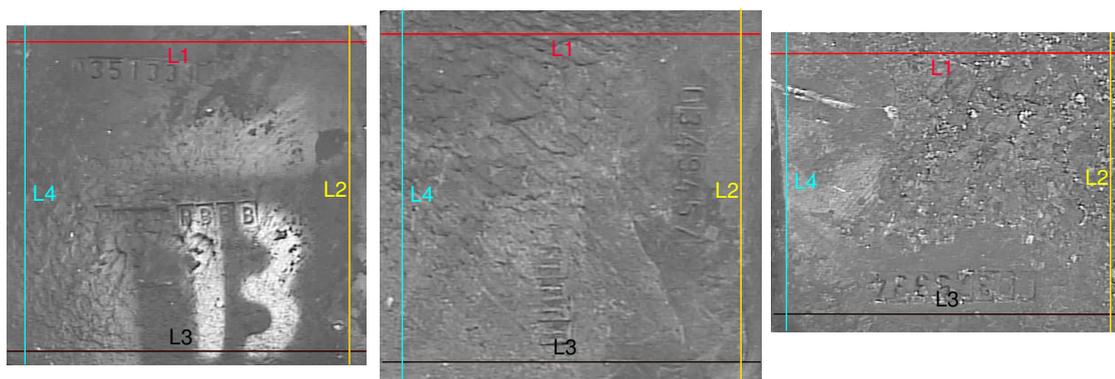


Figura 18: Disposição das linhas do tarugo.

lisa (será chamado de PL). O classificador é testado em imagens que não fazem parte do conjunto de treinamento.

- $PL = 1$ - Linha pertencente à parte lisa da imagem,
- $PL = 0$ - Linha pertencente à parte rugosa da imagem,
- $0 < PL < 1$ - Pode ser considerado algum limite para definição.

A análise isolada das linhas pode levar a erros indesejáveis. A informação adicional de contexto pode restringir o universo de soluções. Sabe-se que há duas linhas rugosas e duas linhas lisas e linhas com a mesma característica são adjacentes. Se três linhas forem classificadas como rugosas, por exemplo, há erro na classificação de uma delas. Da mesma maneira, duas linhas opostas não podem apresentar o mesmo comportamento. Não pode haver uma linha rugosa na parte superior do tarugo e outra na parte inferior.

A formulação destas regras pode ser realizada através de análise de um vetor com as PL s das linhas em uma seqüência determinada. A correta análise deste vetor indicará erros de classificação. Conforme a Figura 18, foi formado um vetor com a seguinte seqüência:

$$\begin{bmatrix} PL - LINHA1 \\ PL - LINHA2 \\ PL - LINHA3 \\ PL - LINHA4 \end{bmatrix}$$

A análise dos possíveis resultados dos vetores e sua correta interpretação estão na Tabela 3.

Tabela 3: Análise dos vetores formados sequencialmente com as PLs de cada linha.

Vetor	Descrição
0000	<i>Inconsistente</i> : todas as linhas são rugosas
0001	<i>Inconsistente</i> : Três linhas rugosas
0010	<i>Inconsistente</i> : Três linhas rugosas
0011	Seqüência Possível
0100	<i>Inconsistente</i> : Três linhas rugosas
0101	<i>Inconsistente</i> : Linhas opostas apresentam mesma característica
0110	Seqüência Possível
0111	<i>Inconsistente</i> : Três linhas lisas
1000	<i>Inconsistente</i> : Três linhas rugosas
1001	Seqüência Possível
1010	<i>Inconsistente</i> : Linhas opostas apresentam mesma característica
1011	<i>Inconsistente</i> : Três linhas lisas
1100	Seqüência Possível
1101	<i>Inconsistente</i> : Três linhas lisas
1110	<i>Inconsistente</i> : Três linhas lisas
1111	<i>Inconsistente</i> : Todas as linhas são lisas

Os vetores formados devem se aproximar de alguma destas soluções, limitando a decisão a quatro possibilidades consistentes. As outras apontam algum tipo de inconsistência, sugerindo que o método não foi capaz de identificar corretamente as linhas.

A saída do classificador produzirá a PL de cada linha. Ao colocá-las em uma seqüência, uma outra classificação será necessária. A identificação como alguma seqüência conhecida.

O método de classificação foi o cálculo das distâncias euclidianas do vetor de PLs para os vetores conhecidos. A menor distância aponta o melhor candidato.

Para otimizar o desempenho do classificador e eliminar chances de polarização e instabilidade na resposta [7], o método de Bagging foi implementado [8].

O classificador final é composto por cinco decisores, treinados com conjuntos de treinamentos diferentes. A decisão é realizada pela maioria absoluta. Os tarugos foram divididos para treinamento e validação:

- 26 tarugos fazem parte do conjunto de treinamento.
- 20 são escolhidos para o treinamento de cada classificador.
- A validação é feita em 10 tarugos, que não fazem parte do conjunto de treinamento.

A taxa representa uma média de 10 tentativas. Na primeira configuração, em 10

Tabela 4: Variáveis: 1 - Variância, 2 - Máximo, 3 - Mínimo, 4 - Média. O desempenho foi melhor considerando mais variáveis e sem a presença da média.

Classificadores	Taxa de acertos	Inconsistência	Erros	Variáveis
5	74%	25%	1%	1,2,3
5	60%	30%	10%	1,4

tentativas, apenas em uma, um único tarugo foi classificado errado.

A Tabela 5 apresenta a tabela de probabilidades de um vetor que foi classificado com erro.

Tabela 5: Vetor classificado com erro. Todos os classificadores apontaram a seqüência 0110, seqüência consistente, mas errada.

Linha	1	2	3	4
Prob. Lisa Classificador 1	0,0000	0,8965	0,834	0,0000
Prob. Lisa Classificador 2	0,0000	0,8944	0,916	0,0000
Prob. Lisa Classificador 3	0,0000	0,9358	0,881	0,0000
Prob. Lisa Classificador 4	0,0000	0,916	0,9358	0,0000
Prob. Lisa Classificador 5	0,0000	0,9196	0,9019	0,0000

3.3.2 Análise de Regiões

A análise das linhas não forneceu um bom precedente para a identificação da região de textura. A disposição bidimensional da textura aponta para uma análise de regiões.

O método proposto nesta seção consiste em analisar regiões da imagem (Figura 19), ao invés linhas.

A escolha das regiões foi realizada segundo a Figura 19, com quatro parâmetros: duas dimensões da região (a e b) e as distâncias da região para as margens (c e d). Os parâmetros são calculados de forma relativa, através do número de linhas e colunas da imagem do tarugo.

A escolha dos parâmetros influenciou extremamente nos resultados, fornecendo resultados conflitantes. Antes de utilizar algum classificador, como na Seção 3.3.1, é interessante analisar o comportamento das variáveis utilizadas. O classificador de Bayes depende deste comportamento e a análise pode fornecer informações adicionais.

Pela análise do gráfico na Figura 20 é possível separar as texturas com um perceptron simples. Além disto, a informação da variância é a grande responsável pela separação.

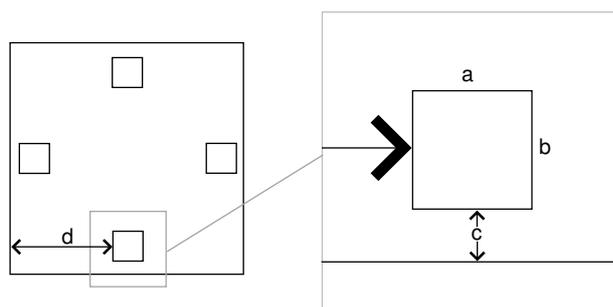


Figura 19: Regiões analisadas no método e os parâmetros considerados.

A análise individual dos tarugos na Figura 21, mostra que uma comparação direta entre as variâncias das margens opostas do mesmo tarugo gera uma separação eficiente. A hipótese, neste caso, é que a variância da parte rugosa é maior do que a da parte lisa, do mesmo tarugo. Com duas comparações, uma entre a margem esquerda e direita e outra entre a margem inferior e superior, a posição da rugosidade pode ser determinada. É a utilização da característica do tarugo: margens opostas pertencem a classes diferentes. Informação utilizada como contexto no método de linhas.

Na Figura 21 são demonstrados alguns confrontos entre margens opostas do mesmo tarugo. Os símbolos iguais definem regiões pertencentes ao mesmo tarugo. As cores verde e preto são as características das margens esquerda e direita respectivamente. As cores azul e vermelho são as características das margens inferior e superior respectivamente. Em todos os confrontos mostrados, as regiões das margens rugosas (nas cores verde e azul) possuem variância maior do que as regiões das margens lisas (nas cores vermelho e preto).

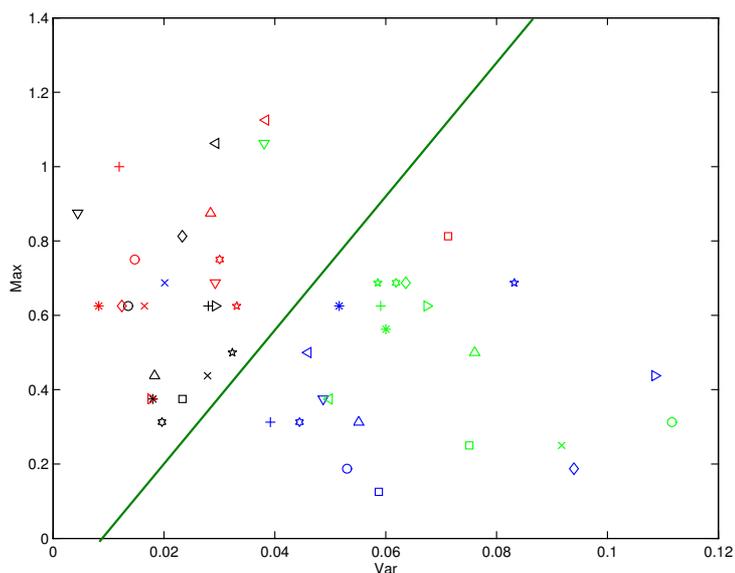


Figura 20: Separação conseguida pela análise das regiões. Azul e Verde = Rugoso. Vermelho e Preto = Liso.

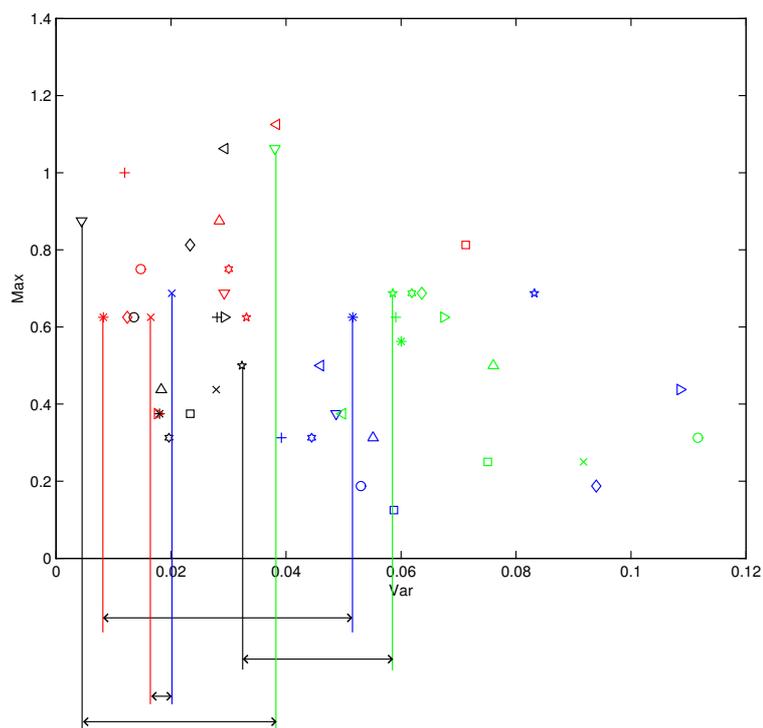


Figura 21: Comparação da variância de regiões em margens opostas do mesmo tarugo. Azul = abaixo, vermelho = acima, verde = esquerda e preto = direita. Símbolos iguais pertencem ao mesmo tarugo. As regiões rugosas (verde e azul) apresentam maior variância do que as regiões lisas (vermelho e preto).

3.3.3 Decisor Baseado em Blocos

A implementação de mais regiões, formando um conjunto de blocos, é direta a partir das conclusões da seção anterior. Aumenta a confiabilidade nos resultados sem comprometer a complexidade computacional.

A primeira configuração desenvolvida compara cada bloco ao bloco correspondente da margem oposta com relação à variância da região, segundo Figura 22.

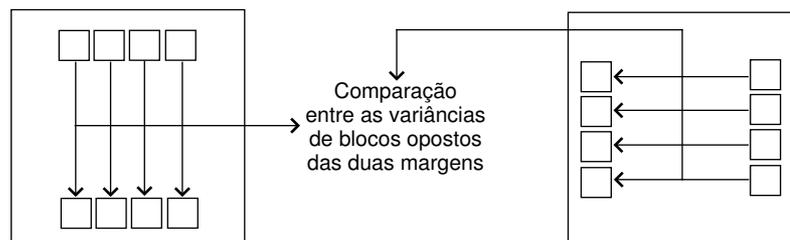


Figura 22: Configuração do reconhecimento da rotação pela variância de blocos.

Os parâmetros são diferentes nesta configuração: distância da margem mais próxima (a), dimensões dos blocos (b e c) e distância entre os blocos (d). O conjunto de blocos é posicionado de forma centralizada 23.

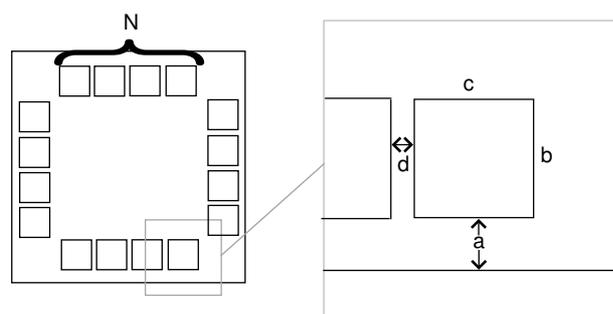


Figura 23: Parâmetros do decisor baseado em blocos.

A comparação de cada bloco com o seu oposto representa um decisor diferente. O classificador montado baseia-se na maioria absoluta.

Em uma segunda configuração, este processo foi otimizado (Figura 24). Não é preciso confrontar as variâncias dos blocos que estejam exatamente opostos. A suposição é que a maior variância da parte rugosa seja sempre maior do que a maior variância da parte lisa.

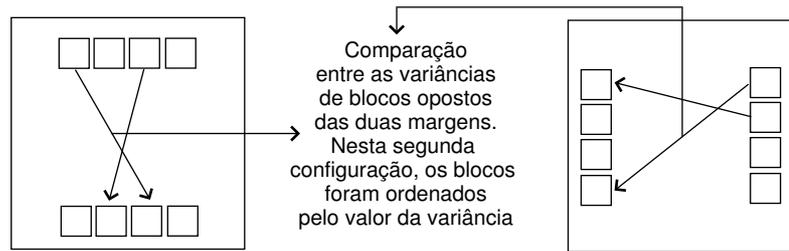


Figura 24: Otimiza o da configura o do reconhecimento da rota o pela vari ncia de blocos.

Com esta nova configura o, todos os resultados foram corretos. Os blocos superiores s o comparados aos inferiores e os blocos da direita com os da esquerda. O diagrama do classificador est  descrito na Figura 25.   composto por dois decisores, uma para cada localiza o. O primeiro decisor estabelece se a rugosidade est  na parte inferior ou superior do tarugo. O segundo se a rugosidade est  na parte direita ou esquerda. Com estas duas informa es a posi o da textura   identificada.

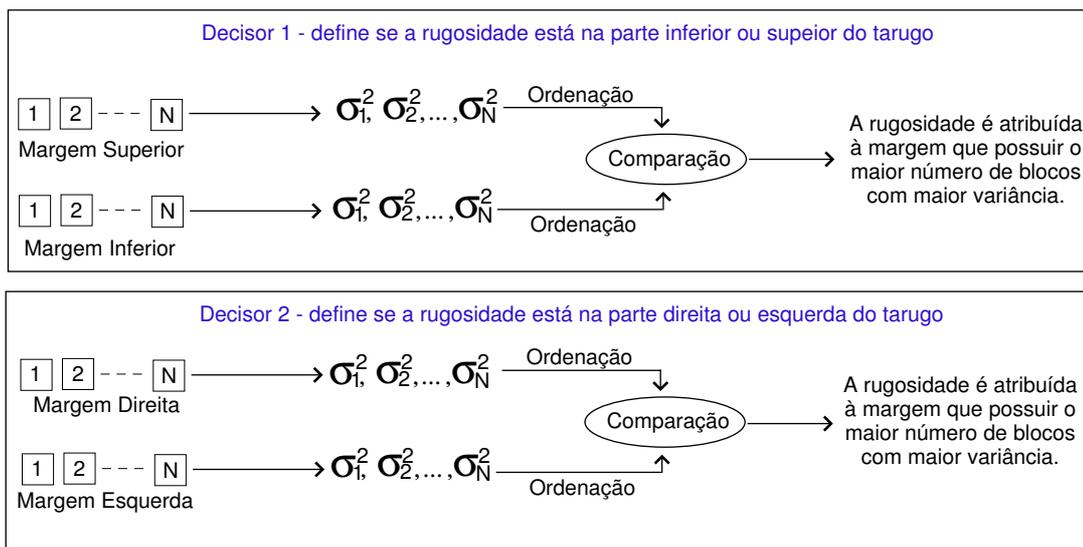


Figura 25: Montagem final do classificador. Dois decisores determinam a localiza o da rugosidade. Decisor 1: acima ou abaixo. Decisor 2: esquerda ou direita.

A solu o elaborada n o se apresentou muito sens vel   escolha de seus par metros. Com a configura o dos par metros, conforme Figura 23:

a - 3% de l ou c .

b - 3% de l ou c .

c - 3% de l ou c .

d - 0.

N - 5 blocos.

l - Número de linhas da imagem do tarugo.

c - Número de colunas da imagem do tarugo.

Todos os tarugos foram identificados corretamente em relação ao posicionamento da rugosidade. A Tabela 6 resume os resultados e características dos métodos de blocos propostos.

Tabela 6: Resumo do desempenho dos métodos desenvolvidos.

Método	Linhas Extremas	Regiões	Blocos
Classificador	Bayes	Maior Variância	Maioria Absoluta
Teste de Consistência	sim	não	não
% de Acerto	74,00%	91,67%	100,00%
% de Erro	1,00%	8,33%	0,00%
% de Inconsistência	25,00%	não se aplica	não se aplica
Sensibilidade aos Parâmetros	baixa	alta	baixa
Complexidade Computacional	baixa	baixa	baixa

3.4 Reconhecimento do Número da OP

Para o reconhecimento do número da OP, os caracteres foram segmentados explicitamente e entregues a etapa de extração de características (Figura 26). Para o método escolhido, deve se definir o tipo de imagem utilizada, conforme Seção 2.5. A maioria dos métodos de extração de características utiliza imagens limiarizadas. As formas de contorno e “esqueleto” da imagem também são utilizadas, mas são geradas a partir da imagem limiarizada.

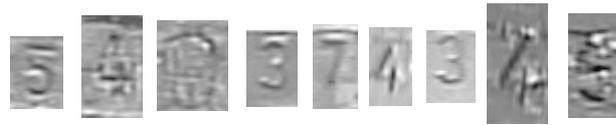


Figura 26: Imagens utilizadas na implementação de rotinas para identificação.

As imagens dos números isolados foram distribuídas em quatro grupos, segundo suas características:

- Grupo 1 - Imagens consideradas boas, com bom contraste e pouca deformação, Figura 27.
- Grupo 2 - Imagens sem deformação aparente, mas com baixo contraste, Figura 28.
- Grupo 3 - Imagens com deformações, Figura 29.
- Grupo 4 - Imagens que apresentam reflexos, Figura 30.

Esta divisão fornecerá maiores informações sobre o desempenho dos métodos, indicando a eficiência em condições diferentes.

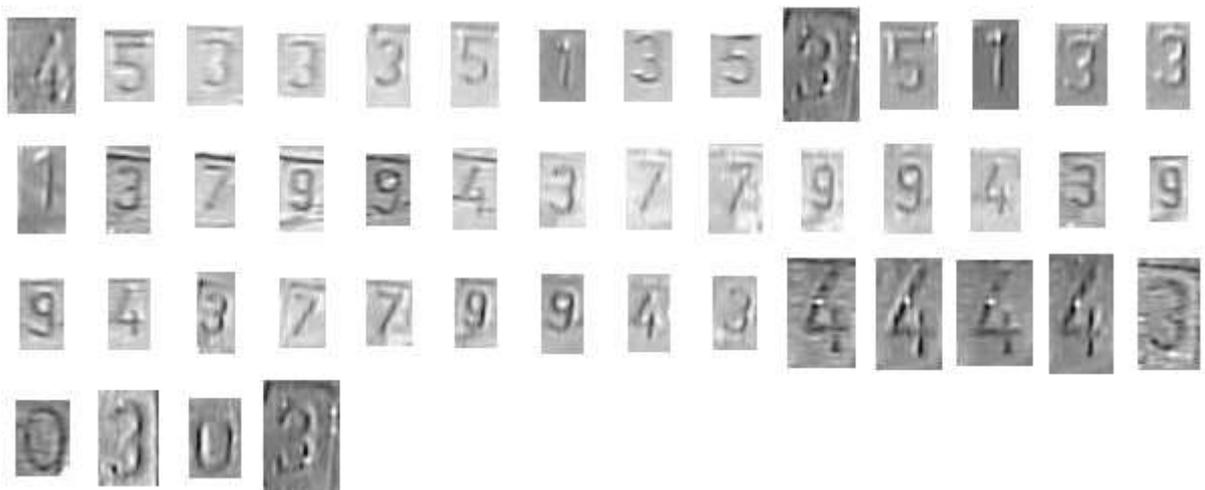


Figura 27: Grupo 1 - Imagens com melhor contraste e pouca deformação.

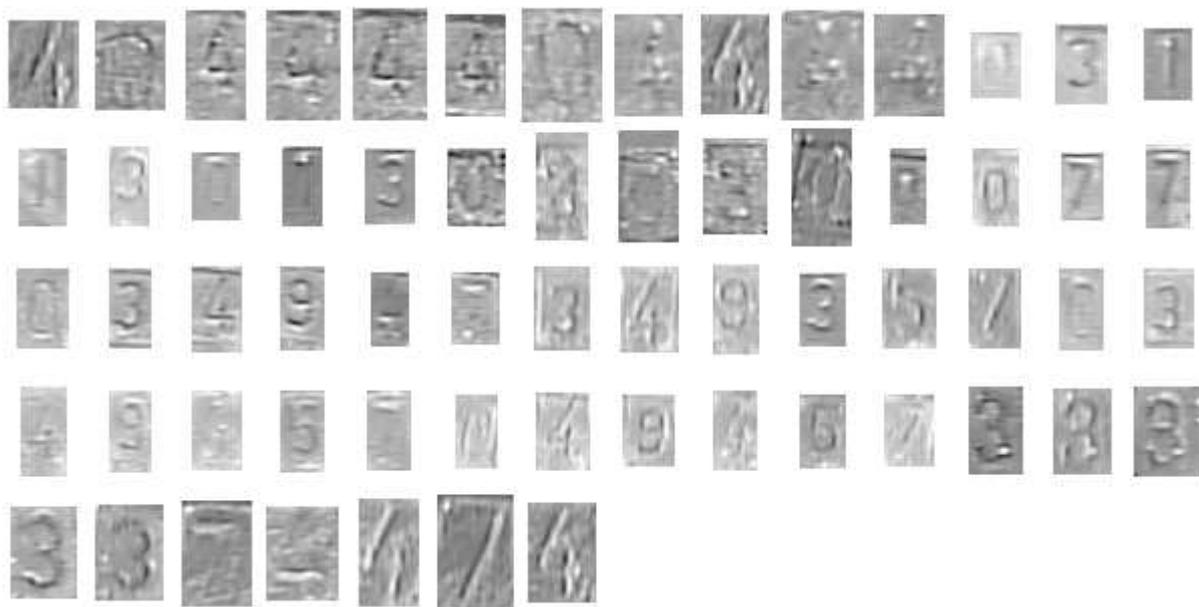


Figura 28: Grupo 2 - Imagens com baixo contraste.

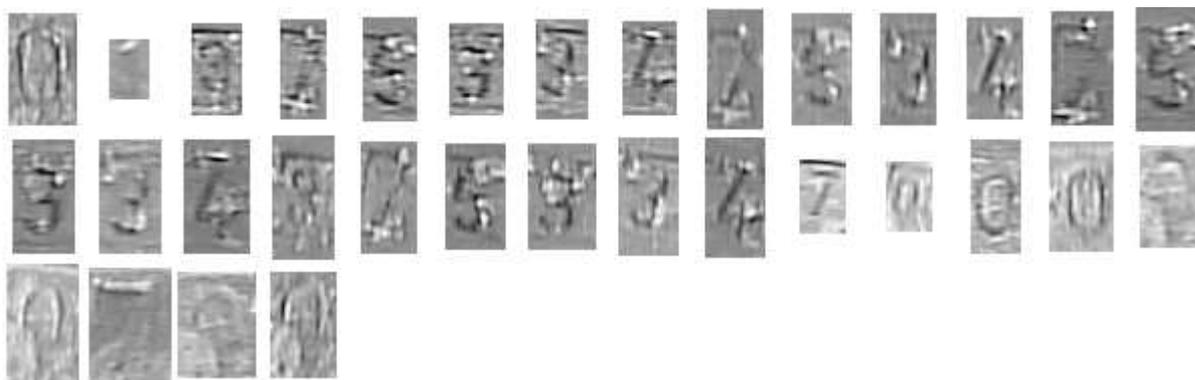


Figura 29: Grupo 3 - Imagens com deformações.

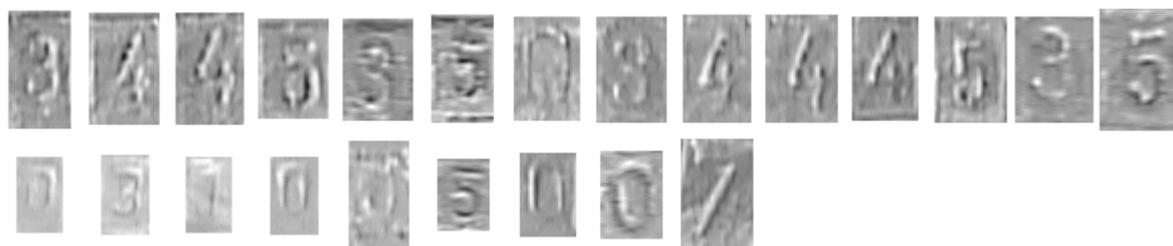


Figura 30: Grupo 4 - Imagens com reflexos.

3.4.1 Limiarização

O primeiro passo para a aplicação da maioria dos algoritmos de extração é a limiarização da imagem. Conforme a Seção 2.3.1, uma análise dos histogramas das imagens pode apontar qual método pode ser mais eficiente.

As imagens testadas apresentam pequenas variações de contraste entre o número e o fundo de imagem, apresentando um histograma praticamente unimodal, conforme Figuras 31, 32, 33 e 34.

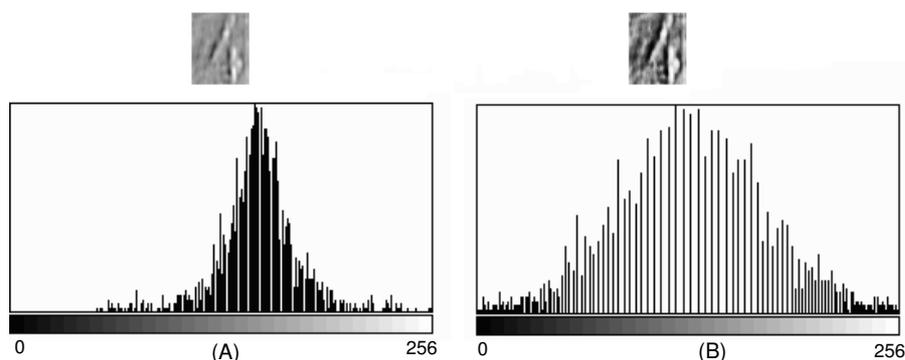


Figura 31: Imagem com baixo contraste. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.

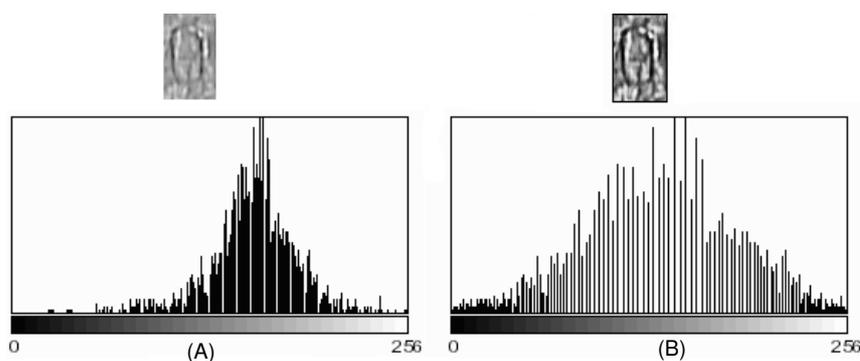


Figura 32: Imagem com deformação. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.

A utilização de métodos como o de Otsu [41] não é recomendada em imagens com este comportamento. A informação dos caracteres está sobreposta à informação do fundo de imagem no histograma.

Uma análise mais detalhada do histograma pode fornecer informações relevantes para a escolha do método. Pela análise da Figura 35, nota-se que o histograma do caractere

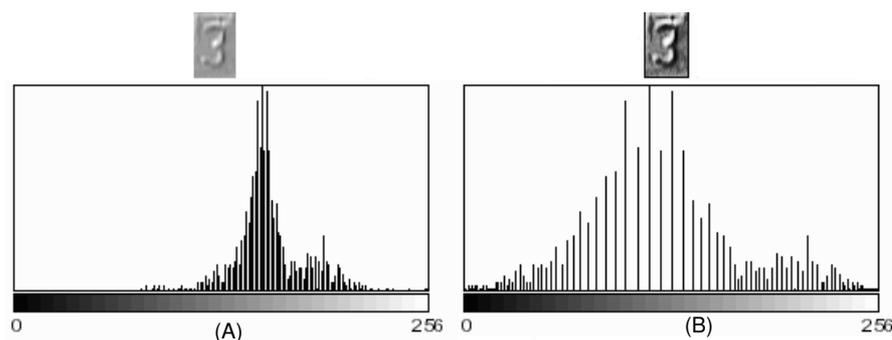


Figura 33: Imagem com bom contraste. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.

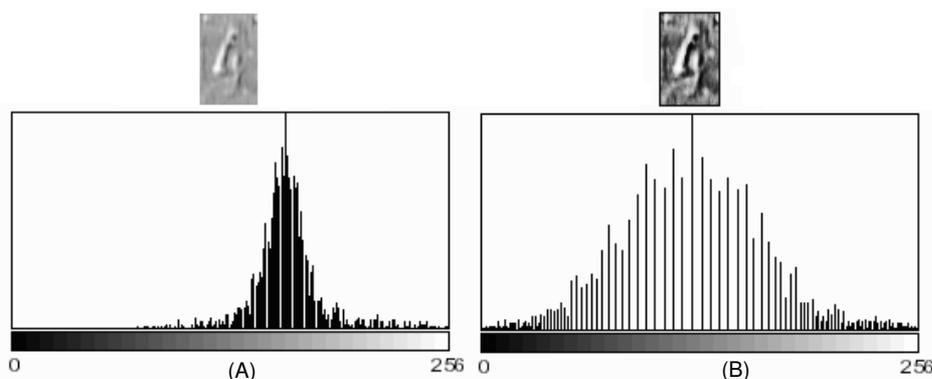


Figura 34: Imagem com reflexos. Em (A): imagem original e histograma. Em (B): imagem com equalização do histograma e histograma.

está completamente sobreposto ao histograma da imagem original. Não há uma separação aparente.

A Figura 36 compara os histogramas do caractere e da imagem original sujeita a equalização do histograma. A separação é mais evidente, sendo que um limiar ($\mu(i, j) - \sigma(i, j)$) pode ser considerado. Este limiar é o mesmo utilizado no método de Niblack [38] com $k = -1$, conforme Seção 2.3.1. Neste caso, entretanto, o método pode ser aplicado na imagem inteira, ao invés de em pequenas janelas. Como as imagens apresentam uma dimensão reduzida, cerca de 30×40 pixels, a iluminação não varia consideravelmente. Para efeito de comparação o limiar de Otsu [41] é aplicado, mas não apresenta um bom resultado.

Pela Figura 36, a imagem limiarizada apresenta ruídos indesejáveis. Este efeito pode ser amenizado com um filtro de médias (filtro passa-baixa) antes da equalização do histo-

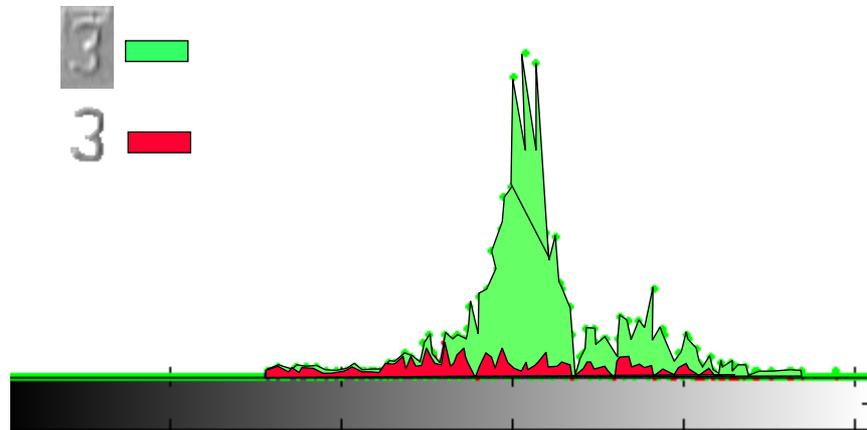


Figura 35: Comparação dos histogramas da imagem original e do caractere.

grama. Os ruídos esporádicos da imagem são suavizados, conforme Figura 37. Contudo o resultado também retirou alguma informação do número, o que pode ser prejudicial. Para o filtro de médias foi utilizada uma janela de 3×3 pixels.

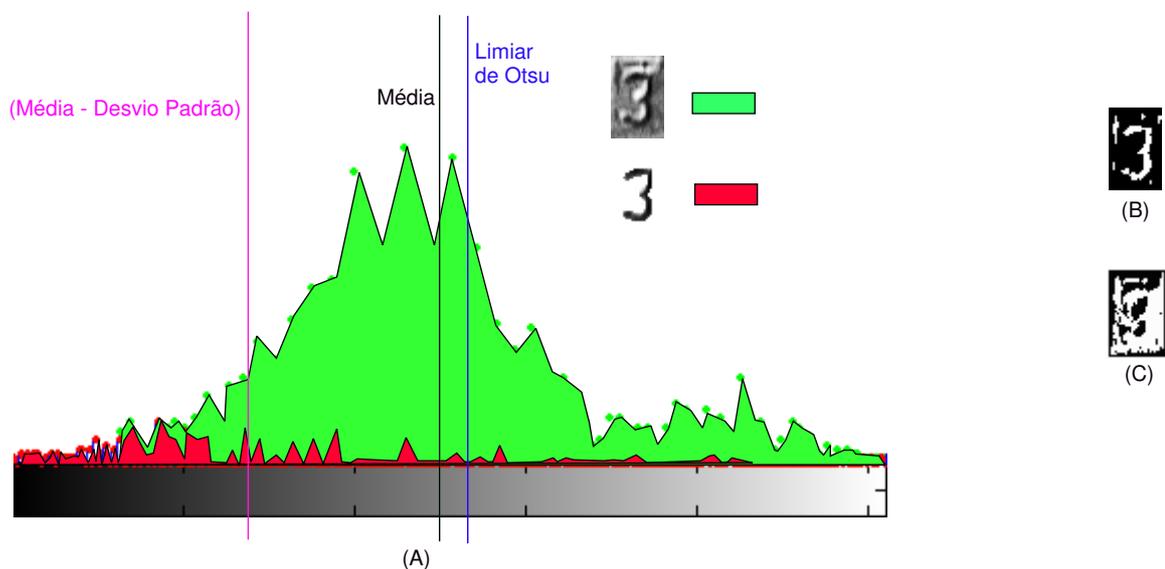


Figura 36: Comparação dos histogramas da imagem equalizada e do caractere. Em (A): comparação dos histogramas. Em (B): imagem limiarizada com o limiar $\mu(i, j) - \sigma(i, j)$. Em (C): imagem limiarizada com o limiar de Otsu.

O método foi aplicado nos grupos de imagens. A limiarização com filtragem antes da equalização provocou perda de informação em alguns números. A limiarização somente

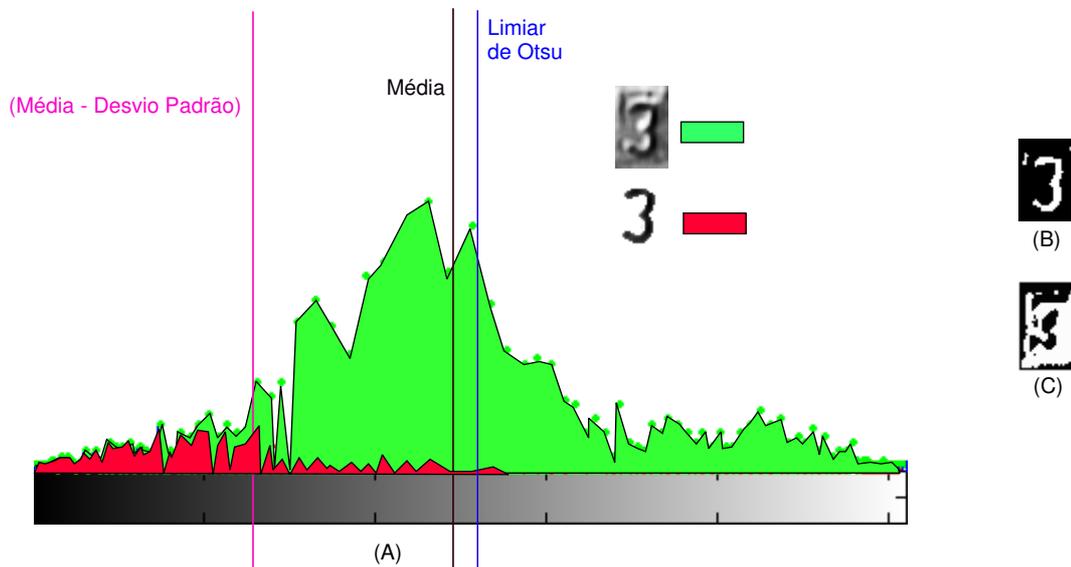


Figura 37: Comparação dos histogramas da imagem equalizada após a aplicação de um filtro de médias. Em (A): comparação dos histogramas. Em (B): imagem limiarizada com o limiar $\mu(i, j) - \sigma(i, j)$. Em (C): imagem limiarizada com o limiar de Otsu.

com a equalização do histograma preserva a informação, conforme Figuras 38 e 39. O método sem a etapa de filtragem foi aplicado nos outros grupos. Os resultados estão nas Figuras 40, 41 e 42.

A presença de ruído ainda é intensa e prejudica a visualização do número. A aplicação do filtro de médias após a etapa de limiarização limpa um pouco dos espúrios sem comprometer a formação do número, conforme Figuras 43, 44, 45 e 46.

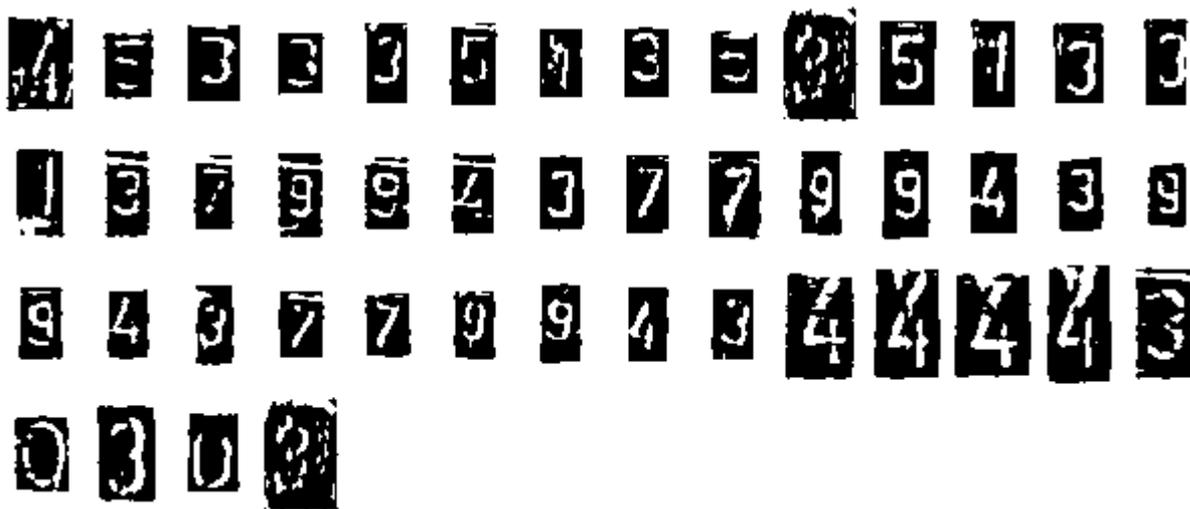


Figura 38: Imagens do grupo 1 limiarizadas pelo método de filtragem seguida de equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - \sigma(i, j)$.

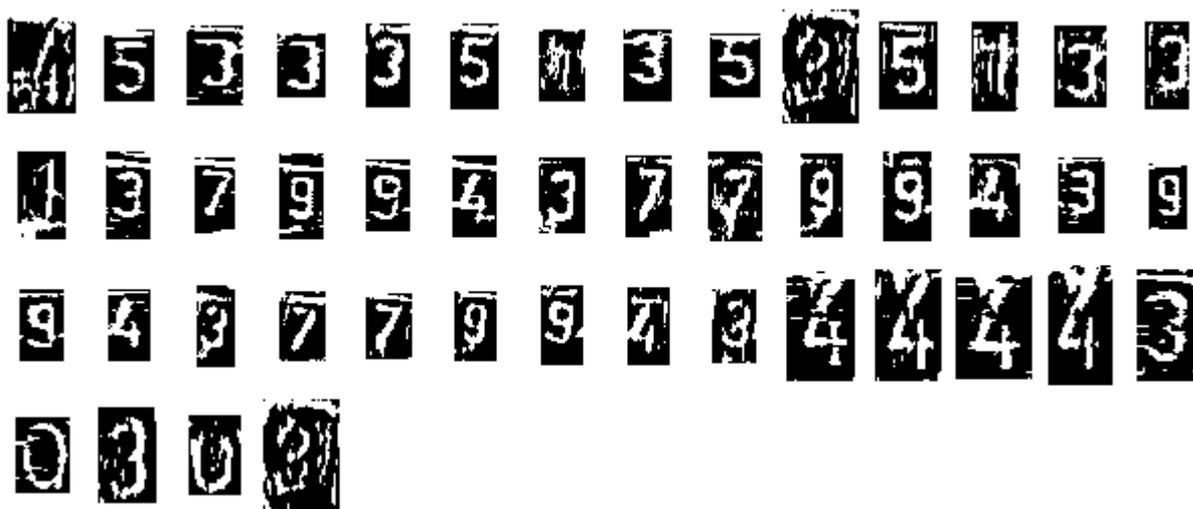


Figura 39: Imagens do grupo 1 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$.



Figura 40: Imagens do grupo 2 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$.



Figura 41: Imagens do grupo 3 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$.

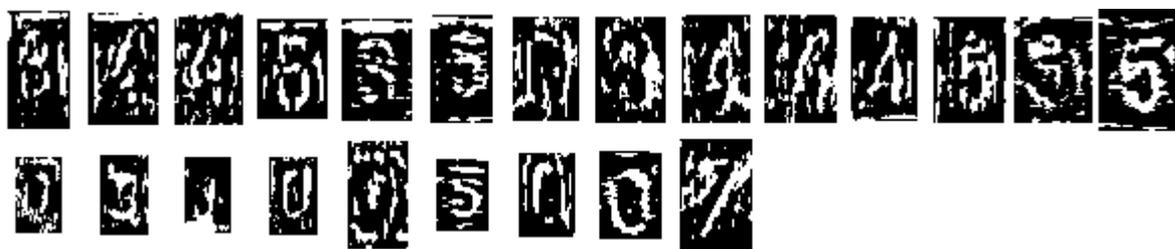


Figura 42: Imagens do grupo 4 limiarizadas após equalização de histograma. Imagens limiarizadas com o limiar $\mu(i, j) - 0,9 * \sigma(i, j)$.

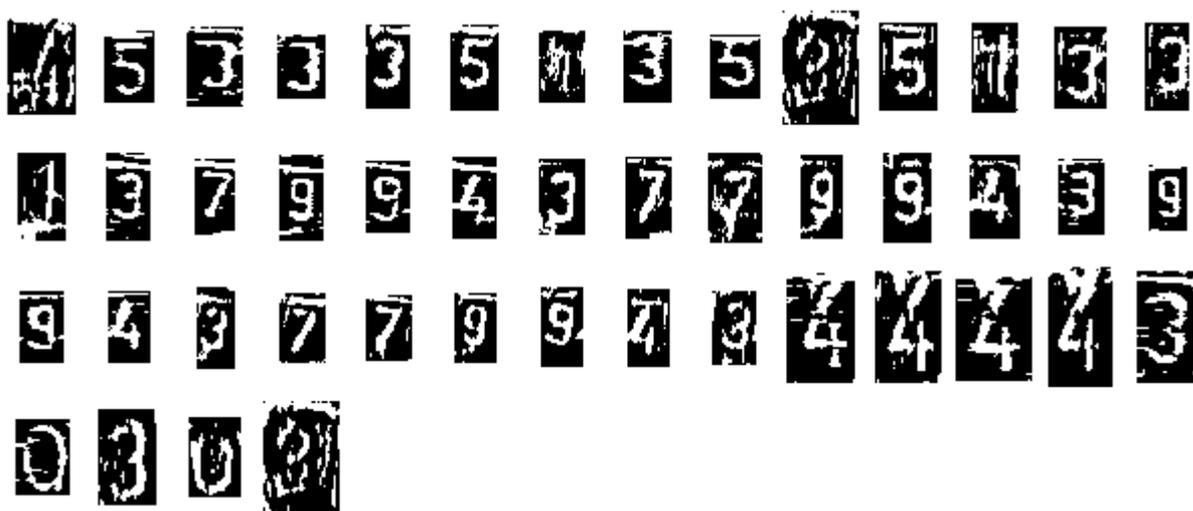


Figura 43: Imagens da Figura 39 após aplicação do filtro de médias de 3×3 pixels.

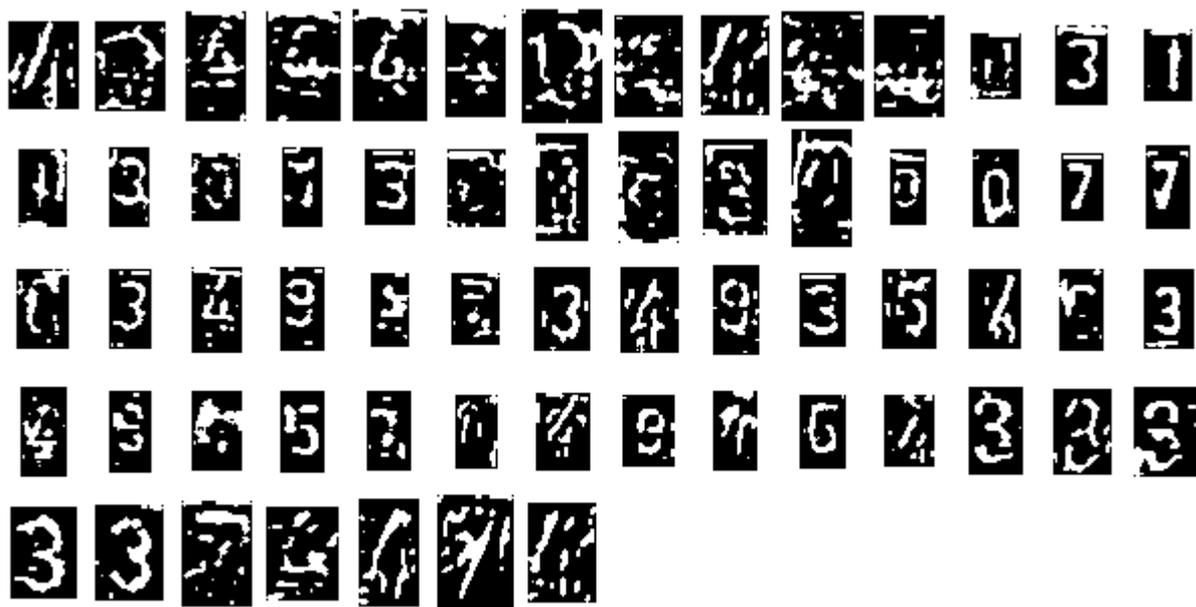


Figura 44: Imagens da Figura 40 após aplicação do filtro de médias de 3×3 pixels.

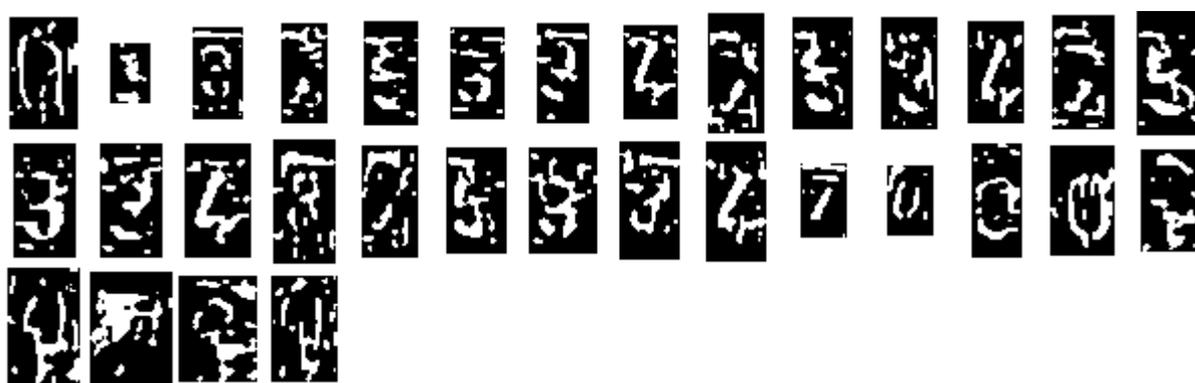


Figura 45: Imagens da Figura 41 após aplicação do filtro de médias de 3×3 pixels.



Figura 46: Imagens da Figura 42 após aplicação do filtro de médias de 3×3 pixels.

3.4.2 Casamento de Padrões

O método casamento de padrões é indicado em situações nas quais os padrões são constantes, como no caso de reconhecimento de caracteres a partir de documentos impressos. A utilização em problemas de caracteres cursivos, com grande variabilidade, não é recomendado [57].

O problema em questão apresenta somente uma fonte, que são as estampas dos números. É natural a consideração do método casamento de padrões para a resolução do problema. Além disto é possível obter resultados entre -1 e 1 , importante para uma análise de consistência.

A desvantagem do método está na sensibilidade a escala, uma vez que a rotação foi identificada com sucesso nas etapas anteriores. Para solucionar esta limitação, os padrões foram escalonados conforme Figura 48. Os números foram segmentados com uma margem de segurança, ou seja, não estão completamente enquadrados, fator importante para não haver segmentação errada. Com isto, a informação da escala do número não pode ser retirada dos dados da imagem isolada. A solução empregada redimensiona as imagens dos padrões em quatro escalas diferentes e correlaciona com a imagem do caractere.

Como a procura é realizada na imagem inteira, o caractere pode não estar com o número centralizado, conforme Figura 47.



Figura 47: O método utilizado permite que o número esteja deslocado do centro da figura. Em (A): Padrão. Em (B): Imagem que contém o número a ser identificado.

A imagem limiarizada do caractere é correlacionada com os 10 padrões de imagem (0, 1, 2, 3, 4, 5, 6, 7, 8, 9), em quatro escalas diferentes. O padrão que apresentar maior correlação é escolhido como resultado do algoritmo.

Os padrões foram escolhidos segundo as estampas dos números. Mas ainda outro ponto foi considerado: como o padrão foi segmentado com uma margem de segurança, o padrão deve conter informação do fundo de imagem. A informação do fundo de imagem presente no padrão procura na imagem um caractere que se assemelha ao padrão e que esteja isolado. Esta informação parece conflitante uma vez que o caractere já está

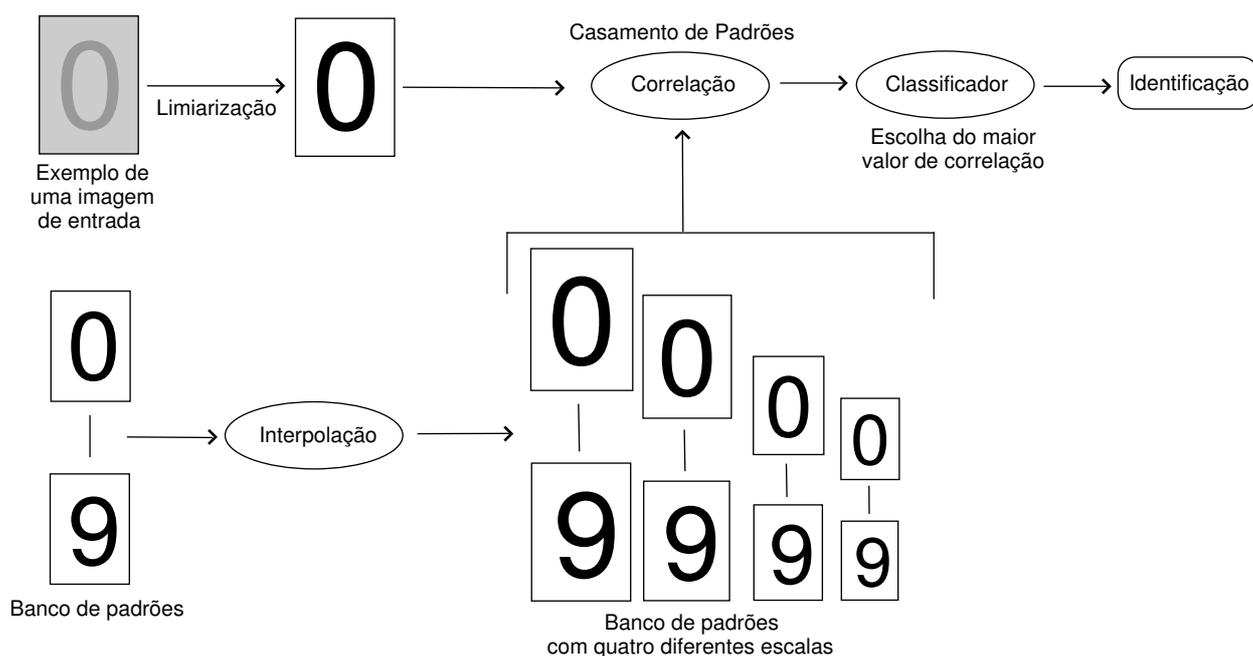


Figura 48: Esquema do algoritmo de identificação por correlação.

segmentado. Mas a presença de ruído e de escalas diferentes justificam essa escolha.



Figura 49: Banco de padrões utilizados na correlação, com a presença de fundo de imagem.

Os resultados para os padrões da Figura 49 para o grupo de imagens limiarizadas da Figura 43, estão na Figura 50 e os valores de correlação cruzada normalizada na Tabela 7.

Os resultados para imagens do Grupo 1, com bom contraste e pouca deformação apresentou 98% de acerto, segundo a Tabela 9. Os demais grupos apresentaram maiores taxas de erros. Entretanto, alguns erros podem ser atribuídos à similaridade dos padrões. A Tabela 10 mostra os valores de correlação cruzada normalizada para os padrões considerados. Há vários casos com valores acima de 0,8.

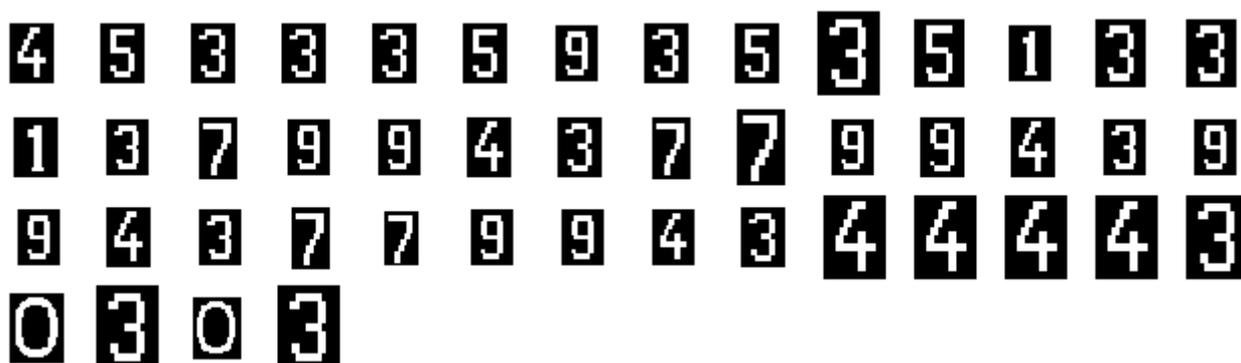


Figura 50: Resultado da correlação das imagens da Figura 43 com os padrões da Figura 49.

Tabela 7: Valores da correlação normalizada para o grupo 1, conforme Figura 50. O número 1 identificado como 9 apresenta o menor valor de correlação cruzada normalizada.

Número Real	Número Identificado	Correlação Normalizada	Número Real	Número Identificado	Correlação Normalizada
4	4	0,7000	9	9	0,7551
5	5	0,7424	9	9	0,7909
3	3	0,6545	4	4	0,6848
3	3	0,7273	3	3	0,8129
3	3	0,7727	9	9	0,6667
5	5	0,8121	9	9	0,8299
1	9	0,5986	4	4	0,7424
3	3	0,7212	3	3	0,6122
5	5	0,6848	7	7	0,6944
3	3	0,6656	7	7	0,7603
5	5	0,8165	9	9	0,6650
1	1	0,6964	9	9	0,7075
3	3	0,6988	4	4	0,6327
3	3	0,6306	3	3	0,7091
1	1	0,7875	4	4	0,8016
3	3	0,7891	4	4	0,7808
7	7	0,7131	4	4	0,7904
9	9	0,8129	4	4	0,7952
9	9	0,7891	3	3	0,7520
4	4	0,7727	0	0	0,6094
3	3	0,7515	3	3	0,7920
7	7	0,8065	0	0	0,7388
7	7	0,6493	3	3	0,6656

Tabela 8: Análise do valor de correlação para os acertos das imagens do grupo 1.

% do Total de Acertos	Valor da Correlação Cruzada Normalizada (C_n)
0%	$C_n < 0,60$
11%	$0,60 < C_n < 0,65$
22%	$0,65 < C_n < 0,70$
20%	$0,70 < C_n < 0,75$
31%	$0,75 < C_n < 0,80$
16%	$0,80 < C_n < 0,85$
100%	$C_n > 0,60$
89%	$C_n > 0,65$
67%	$C_n > 0,70$
47%	$C_n > 0,75$

Tabela 9: Índice de acerto nos 4 grupos de imagens.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
98%	73%	71%	68%

Tabela 10: Tabela com os valores de correlação cruzada normalizada entre os padrões escolhidos.

	0	1	2	3	4	5	6	7	8	9
0	1,0000	0,3683	0,3706	0,4749	0,2453	0,4014	0,4222	0,3293	0,3980	0,4601
1	0,3683	1,0000	0,4139	0,5153	0,4824	0,3878	0,4992	0,4300	0,4306	0,4793
2	0,3706	0,4139	1,0000	0,7288	0,4614	0,5405	0,5106	0,6211	0,6665	0,6196
3	0,4749	0,5153	0,7288	1,0000	0,3885	0,7230	0,6496	0,5130	0,8653	0,8683
4	0,2453	0,4824	0,4614	0,3885	1,0000	0,3731	0,4093	0,3977	0,4013	0,4680
5	0,4014	0,3878	0,5405	0,7230	0,3731	1,0000	0,8536	0,4690	0,7492	0,7616
6	0,4222	0,4992	0,5106	0,6496	0,4093	0,8536	1,0000	0,4517	0,8059	0,7172
7	0,3293	0,4300	0,6211	0,5130	0,3977	0,4690	0,4517	1,0000	0,4355	0,4626
8	0,3980	0,4306	0,6665	0,8653	0,4013	0,7492	0,8059	0,4355	1,0000	0,9114
9	0,4601	0,4793	0,6196	0,8683	0,4680	0,7616	0,7172	0,4626	0,9114	1,0000

3.5 Conclusões

A metodologia desenvolvida neste capítulo apresentou uma técnica eficiente de identificação da rotação do tarugo. Na etapa de reconhecimento dos caracteres, considerando as imagens segmentadas com uma margem de segurança, foram obtidos bons resultados para caracteres com bom contraste e pouca deformação. Um dos problemas encontrados nesta etapa foi a limiarização das imagens, sensíveis a condicionamento externo. A

deformação foi outro fator que influenciou o resultado do método casamento de padrões, apesar de se mostrar uma técnica viável em caracteres sem muita deformação e apresentar possibilidade de testar a consistência da resposta. Contudo, faltou uma comparação com outros extratores de características, como momentos de Zernicke, para uma análise de desempenho e também uma possível integração dos métodos. Os resultados para os outros grupos de imagem (com deformação, baixo contraste e reflexo) apontam para um condicionamento maior na captura da imagem (iluminação) que auxiliará a técnica de limiarização.

4 *Teste em Campo*

4.1 Introdução

Este capítulo descreve a aplicação dos métodos propostos em imagens coletadas na LIT, capturadas na posição 2 da Figura 7, demonstrando resultados preliminares. O condicionamento da iluminação foi proposto, apresentando resultados promissores. Foram aplicados os métodos de identificação de rotação e foi proposta uma técnica para isolamento do tarugo.

4.2 Condicionamento

Após a etapa de estudo em laboratório, foram coletadas imagens no local a ser instalado o sistema. A análise mostrou que era necessária uma iluminação mais específica para realçar a informação do baixo relevo.

Com auxílio da Açominas, alguns problemas que limitariam a confiabilidade e desempenho do sistema foram analisados e soluções viáveis foram propostas.

Um dos grandes problemas dos métodos testados em laboratório foram as condições das imagens testadas. A iluminação não condicionada, prejudicou a formulação de uma resposta geral para o problema. O contraste dos números para o restante do tarugo se mostrou tênue, provocando quedas de desempenho.

A melhoria na iluminação foi considerada, então, fator fundamental para a melhoria do desempenho e confiabilidade do sistema. A informação do relevo deve ser realçada com métodos de iluminação, principalmente com um ângulo próximo ao plano do tarugo (Figura 51).

O método apresentado na Figura 51 foi testado e os resultados foram extremamente relevantes. Pela comparação feita na Figura 52, fica claro que sem a iluminação, a implementação de uma solução eficiente ficaria difícil.

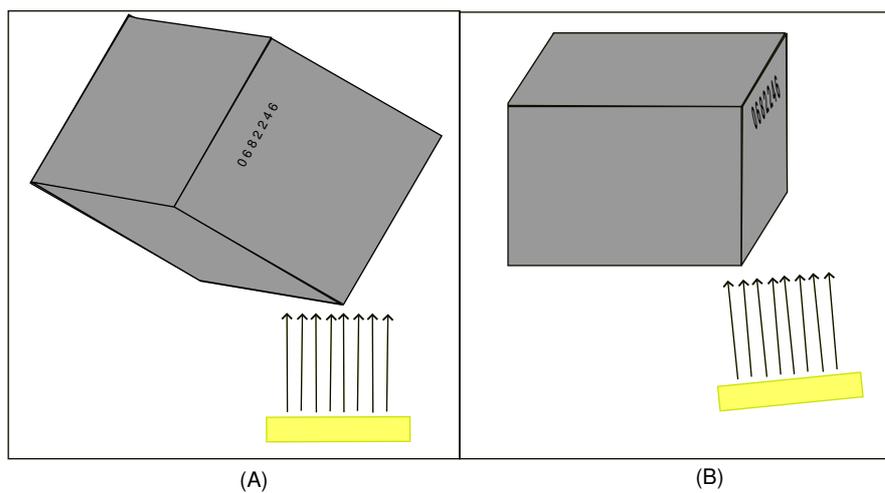


Figura 51: Iluminação para realçar detalhes no plano do tarugo. Em (A): vista da câmera. Em (B): vista lateral do tarugo.

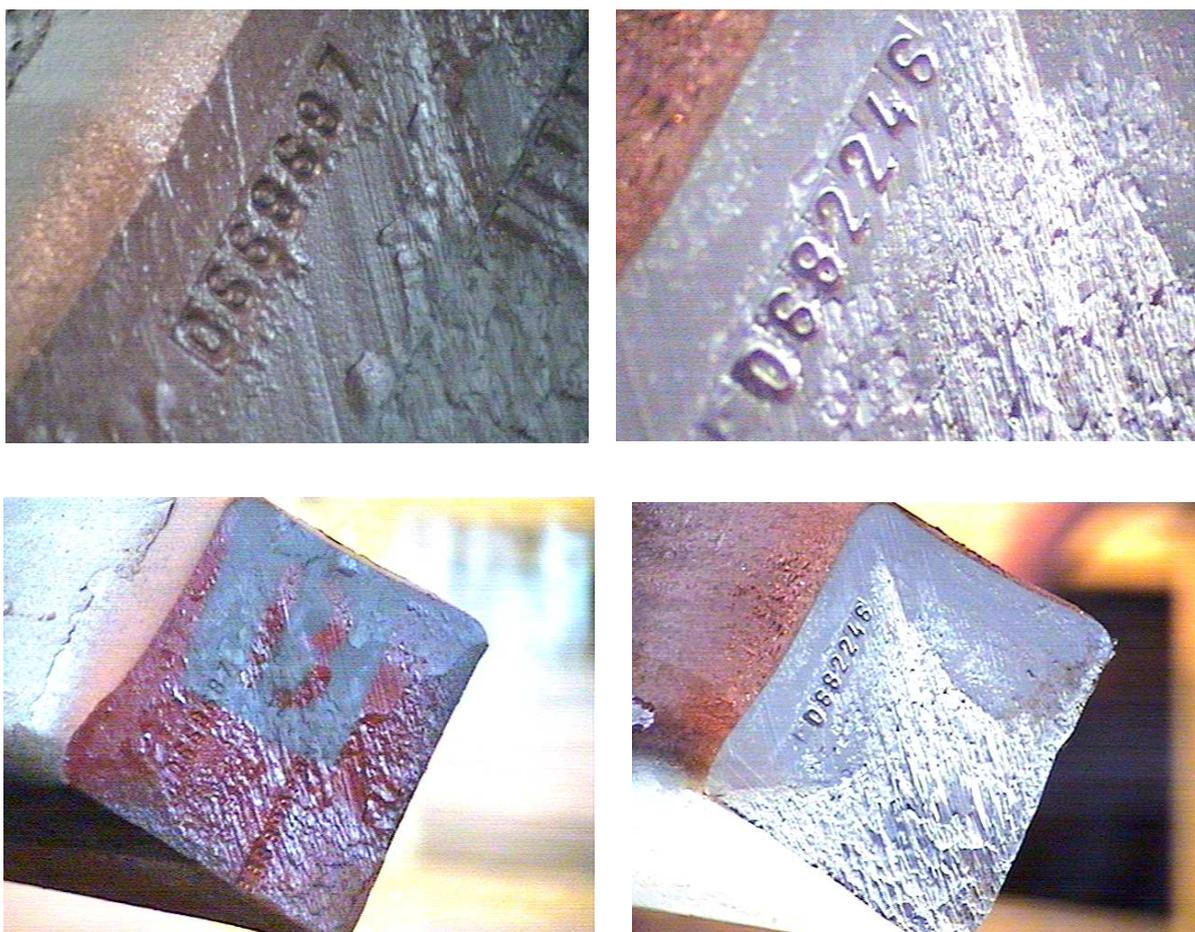


Figura 52: Comparação do efeito da iluminação estruturada nos dois planos de trabalho.

Além da iluminação, outro fator relevante para o sistema é a variação da distância na qual os tarugos são apresentados para a câmera. Além de influenciar vários métodos de extração de características e classificação com a variação na escala dos números, há problemas como foco automático. Este problema foi exposto para a Açominas que propôs, através da gerência do projeto, a normalização da distância a partir de sensores e a parada do tarugo na entrada da linha de inspeção sempre em um determinado ponto. Esta normalização será fundamental, não só para os algoritmos de extração e classificação, mas também para a própria iluminação, que não apresentará variações devido ao posicionamento do tarugo. Manter uma iluminação uniforme em uma distância relativamente grande (cerca de 1 metro), seria trabalhoso.

Outro grande problema nas etapas de laboratório foi considerar a segmentação e extração dos números diretamente da imagem original aberta. As informações dos números ficavam relativas somente a poucos *pixels*, comprometendo o sistema. Nesta etapa do projeto, trabalhou-se com um plano fechado nos números para uma melhor captação dos números. Um plano aberto seria utilizado para encontrar a região dos números do tarugo, como mostrado na Figura 53A. Essas informações da região seriam passadas para a câmera que faria um zoom na região escolhida. O plano fechado, conforme Figura 53B, contém maiores informações dos números, aumentando sua definição.

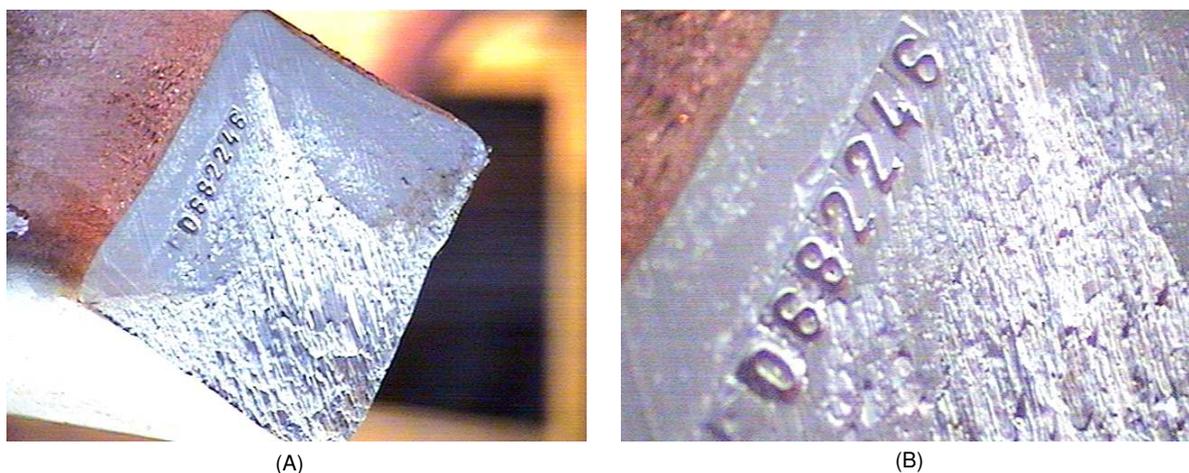


Figura 53: Os dois planos de trabalho considerados. Em (A): plano aberto, para detecção da região dos números. Em (B): plano fechado para reconhecimento dos números.

4.3 Identificação da Rotação do Tarugo

O método “decisor baseado em blocos”, definido no Capítulo 3, que obteve melhor desempenho, foi aplicado nas imagens coletadas na linha de inspeção. Com o tarugo enquadrado, conforme Figura 54, o método foi aplicado utilizando os parâmetros:

- a - 3%;
- b - 3%;
- c - 3%;
- d - 0;
- N - 7 blocos.

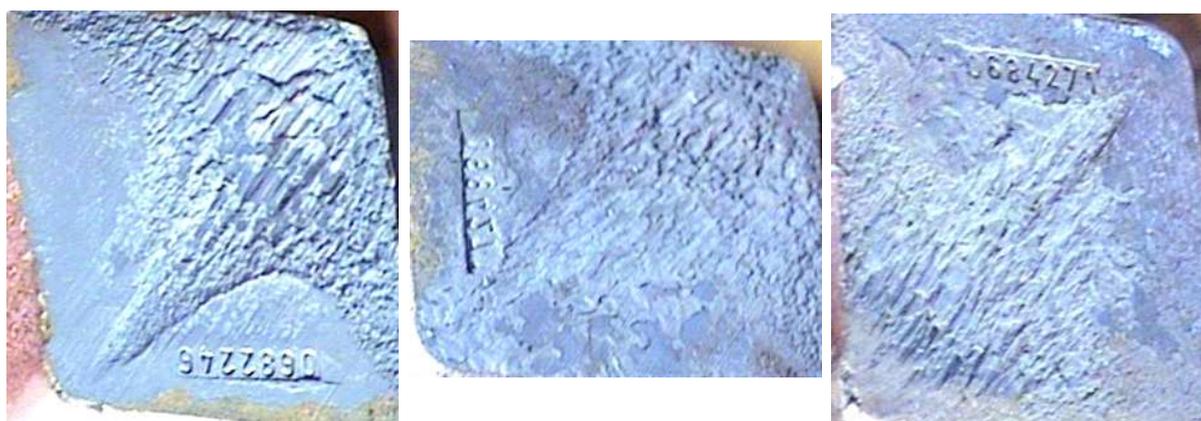


Figura 54: Plano de trabalho para a aplicação do método do decisor baseado em blocos para identificação da rotação do tarugo.

O método identificou a correta rotação em 16 de 17 tarugos. O tarugo identificado incorretamente, mostrado na Figura 55, apresenta uma tinta branca intensa na margem direita. A tinta está presente em outros tarugos, mas não há a contaminação da margem inteira, como neste caso. A tinta provoca uma região homogênea na imagem, fazendo com que a identificação fique incorreta. A Figura 56 demonstra uma figura classificada corretamente, mesmo apresentando partes pintadas.

Para solucionar o erro de classificação no tarugo, pode-se modificar os parâmetros do método, principalmente a distância dos blocos para a margem. Aumentando este fator, a contribuição da pintura iria diminuir consideravelmente. A Figura 57 demonstra o princípio. Utilizando os parâmetros:



Figura 55: Tarugo classificado incorretamente. A tinta provoca uma região homogênea próxima à margem direita da figura.



Figura 56: Tarugo identificado corretamente mesmo com a presença de pintura. A tinta não contamina a margem inteira.

- a - 3%;
- b - 3%;
- c - 10%;
- d - 0;
- N - 7 blocos.

O método foi capaz de identificar a correta rotação do tarugo. Contudo estes parâmetros não forneceram bons resultados nos outros tarugos. A tabela 11 resume a aplicação do método nas imagens coletadas na linha de inspeção.



Figura 57: Ajuste dos parâmetros para a correta identificação do tarugo. O aumento do afastamento da margem soluciona o problema.

Tabela 11: Resumo da aplicação do método “decisor baseado em blocos” nas imagens coletadas na linha de inspeção.

Parâmetros	Valor	
Largura do Bloco (a)	3%	3%
Comprimento do Bloco (b)	3%	3%
Afastamento da Margem (c)	3%	10%
Distância entre os Blocos (d)	0%	0%
Número de Blocos (N)	7	7
Número de Erros (em 17)	1*	4**
% de Acerto	94%	76%
* Erro causado pela interferência da pintura na margem		
** Configuração para correção do tarugo identificado incorretamente.		

4.4 Isolamento do Tarugo

Para que a etapa da identificação da rotação do tarugo funcione é necessário que o tarugo seja corretamente segmentado a partir da imagem no plano aberto. A análise em escala de cinza da imagem, considerando detecção de bordas, apresentou resultados promissores, conforme Figura 58. Mas ao analisar a imagem RGB, pode-se propor um método que forneça melhores resultados. Devido a característica da iluminação utilizada no tarugo, nota-se a tendência para uma coloração azulada. Analisando os componentes R, G e B da imagem original, conforme Figura 59, pode-se propor a operação $B - R$. Esta operação irá realçar a cor azul do tarugo. Aplicando o método de limiarização de Otsu [41], o tarugo é claramente evidenciado, conforme Figura 60.

As imagens podem apresentar ruídos não desejáveis, dificultando a resolução do pro-

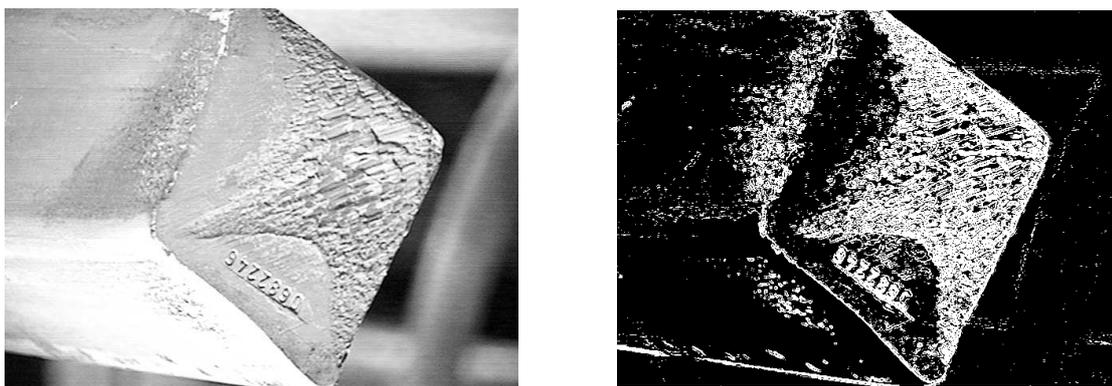


Figura 58: Análise do plano aberto da imagem, com detecção de contorno. Há presença de muito ruído.

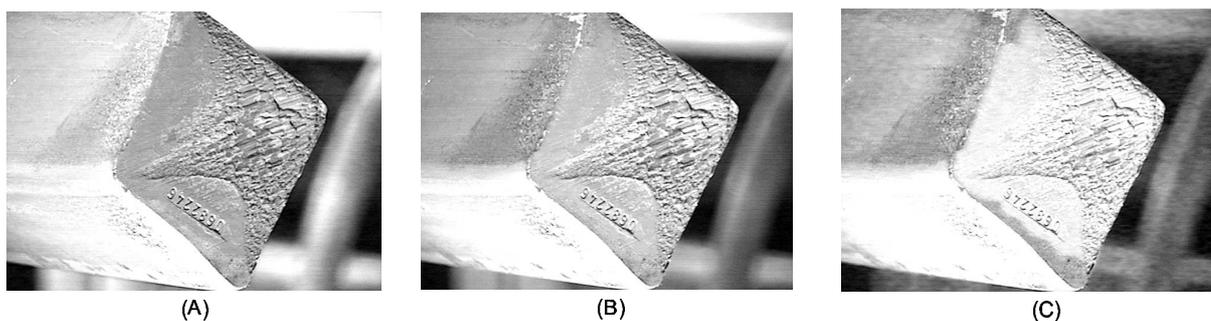
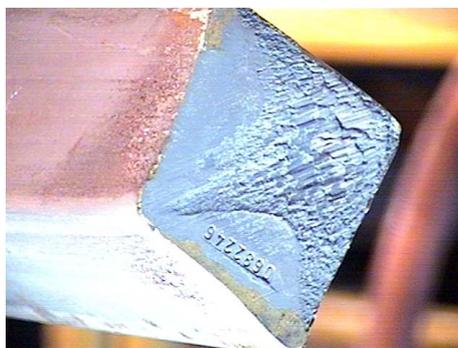


Figura 59: Análise das componentes de cor da imagem. Em (A): Componente vermelho. Em (B): Componente verde. Em (C): Componente azul.

blema. A solução seria escolher dentre os objetos na imagem limiarizada aquele que possuir maior área. A análise da imagem B da Figura 60 fornece 179 objetos oito-conectados. A aplicação de um filtro de mediana, conecta os objetos e ajuda a limpar espúrios. A aplicação do filtro de mediana na imagem B da Figura 60, fornece o resultado mostrado na Figura 61. Uma nova análise revela apenas 8 objetos oito-conectados. Isolando o objeto

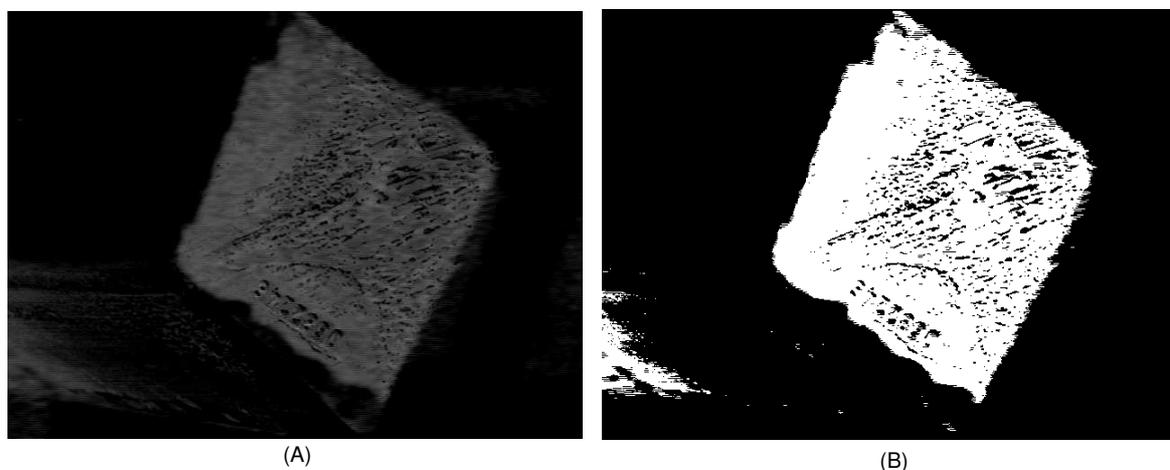


Figura 60: Evidenciação do tarugo. Em (A): Componente azul - Componente vermelho.
Em (B): Imagem em (A) limiarizada pelo método de Otsu [41].

com maior área, obtém-se o tarugo com uma boa definição das margens, conforme Figura 61.

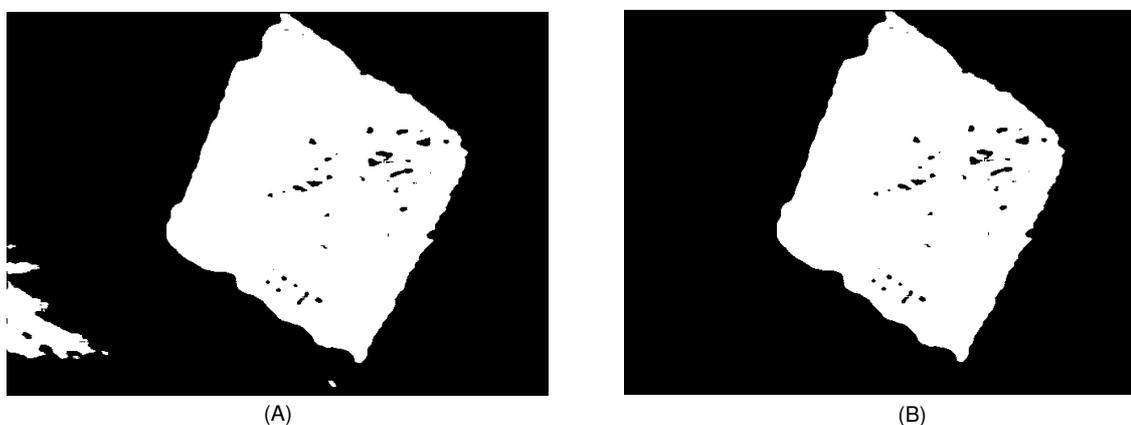


Figura 61: Em (A): Imagem B da Figura 60 depois da aplicação do filtro de mediana.
Em (B): Escolha do objeto oitto conectado com maior área.

As imagens coletadas apresentam uma rotação conhecida de 45 graus, devido ao braço que traciona as peças para a linha de inspeção. Pelo ângulo de visão da câmera, esta rotação foi estimada em 35 graus, para fornecer o tarugo na posição correta. Esta rotação foi aplicada nas 17 imagens, a Figura 62 demonstra alguns resultados.

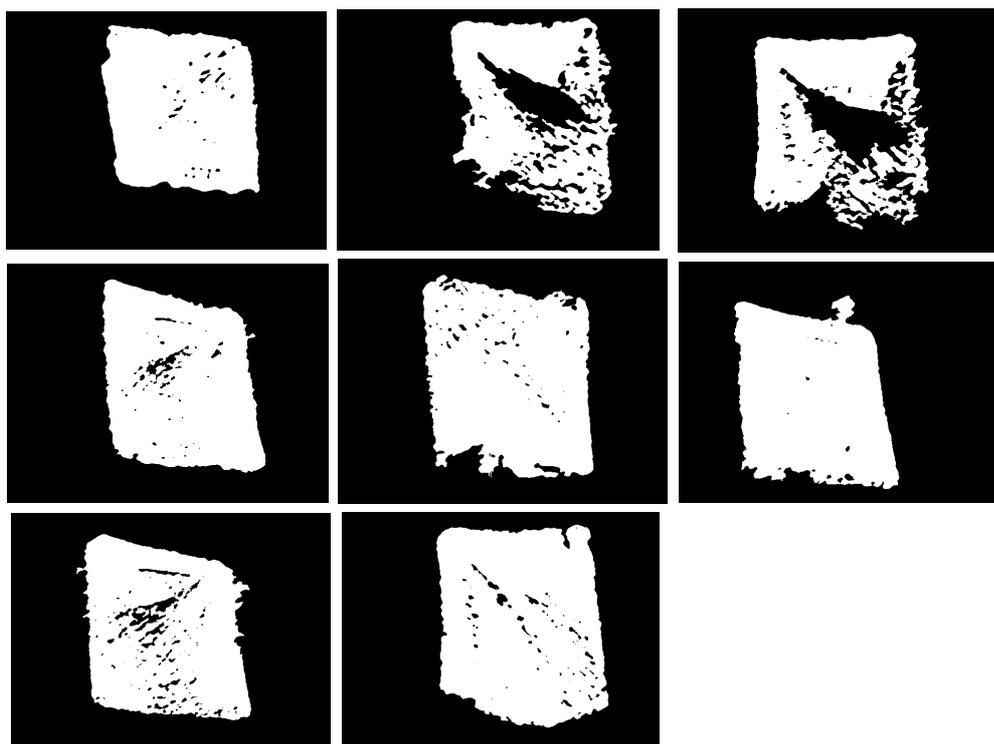


Figura 62: Aplicação do método de segmentação do tarugo nas imagens coletadas.

4.5 Conclusão

Este capítulo demonstrou o começo da aplicação dos métodos propostos no Capítulo de Metodologia. A aplicação do método “Decisor Baseado em Blocos” para identificação da rotação obteve somente um erro de classificação, explicado pela interferência de uma pintura na margem do tarugo. A alteração dos parâmetros do método classifica corretamente a imagem errada, mas gera erros em outros tarugos. O método proposto para isolamento do tarugo a partir da imagem RGB demonstrou resultados promissores, tanto no desempenho quanto na implementação.

5 *Conclusão*

Este trabalho estudou e testou ferramentas para utilização no reconhecimento da Ordem de Produção estampada em tarugos de aço da Açominas. O estudo realizou-se primeiramente em imagens coletadas em laboratório, depois em imagens coletadas na LIT (Linha de Inspeção de Tarugos).

Os resultados em laboratório indicaram uma boa técnica para a identificação da rotação do tarugo através da análise de textura. A metodologia proposta foi robusta, apresentando resultados corretos em todas as imagens testadas, mesmo com escala e iluminação não condicionadas. Os resultados preliminares com as imagens coletadas na LIT, apresentaram somente um erro, devido à interferência da pintura na margem. O método, contudo, apresenta a desvantagem de não apresentar teste de consistência. Não há como inferir sobre a confiabilidade do resultado. Este aspecto é consequência da consideração da maioria absoluta no confronto entre blocos de margens opostas. Uma maneira de contornar este problema seria a adoção de um mínimo de blocos de diferença entre a disputa. A solução desenvolvida foi baseada na informação de contexto: margens opostas pertencem a texturas diferentes. Outra forma de avaliar a consistência seria a adoção de mais uma informação de contexto, a de que margens adjacentes a uma determinada margem são de classes diferentes. Outro classificador pode ser desenvolvido, medindo a similaridade entre as margens adjacentes, fornecendo qual se aproxima mais da característica da margem em questão. Este classificador é independente e fornece um resultado para a localização de textura. Resultado que pode ser confrontado com o classificador desenvolvido, a fim de obter uma resposta coerente.

O reconhecimento baseado em casamento de padrões apresentou bons resultados em imagens com bom contraste e com pouca deformação. A correlação cruzada normalizada proposta pode ser implementada de maneira rápida, utilizando FFT e fornece resultados entre -1 e 1 , que podem ser utilizados para indicar consistência. Nas imagens com bom contraste e pouca deformação, grupo 1, o único erro apresentou o menor valor de correlação cruzada normalizada. Foram consideradas quatro escalas diferentes na resolução

do problema. Com certeza podem ser consideradas mais escalas, mas comprometeria o resultado. A presença de padrões com grande variação na escala pode polarizar o resultado para a menor escala, mesmo que a imagem original esteja em uma escala diferente. Por exemplo, o número 1 em uma escala menor, apresenta alto valor de correlação cruzada normalizada em vários números, mesmo que não estejam incompletos. A presença de escalas deve ser controlada, para que a diferença não seja grande. Uma maneira é estimar os valores possíveis de escala a partir das dimensões das imagens dos caracteres. Vale a pena ressaltar que a informação correta da escala da imagem pode ser retirada diretamente da imagem do caractere, mas é necessária uma segmentação extremamente confiável. Este problema não fornece requisitos para uma segmentação ótima, e mesmo nestes casos, é importante considerar uma margem de erro. Com isto, sempre será indicado considerar mais escalas na resolução do problema ou um extrator de características invariante à escala.

Estes requisitos foram levados em consideração na aplicação das ferramentas nas imagens reais. Como afirma Gonzales [18], uma aplicação industrial deve se beneficiar do condicionamento que lhe é permitido. Um bom condicionamento da iluminação e distância contribuem para um desempenho otimizado. O problema estudado apresenta várias características que, sem um condicionamento necessário, gerariam falhas na saída do sistema. Os caracteres são estampados em baixo relevo, a distinção pelo brilho dos pixels não se aplica diretamente. O que há, na verdade, é diferença de relevo e, por isto mesmo, os algoritmos baseados no brilho dos pixels não conseguem padronizar os resultados. Uma aproximação foi realizada, considerando que as diferenças de relevo provocariam, necessariamente, diferenças no brilho dos pixels. Esta aproximação é válida, uma vez que o baixo relevo pode apresentar sombras, gerando um brilho diferenciado. Mas a generalização dos resultados a partir desta idéia é difícil. Sem o condicionamento, há presença de reflexos e contraste tênue, como visto nas imagens coletadas em laboratório. O condicionamento da iluminação nestes casos é peça fundamental, sendo, em muitas ocasiões, aspecto principal na sua resolução [21].

Alguns estudos são realizados a partir de algoritmos que retiram a informação de sombras da imagem, chamados Shape from Shading [65]. Outros estudos propõem a análise da imagem iluminada por um padrão de luz conhecido [5]. Qualquer alteração deste padrão, na imagem coletada, significa informação de relevo. Para identificação de pequenas variações de relevo, é mais indicado a iluminação com feixes de laser [5]. No problema em questão, a proposição destas técnicas é complicada, devido aos pequenos relevos dos caracteres. Estes algoritmos são muito utilizados para reconhecer relevos significantes [65].

Contudo, o uso de iluminação estruturada é muito aplicado em soluções industriais. Mas uma solução que fizesse a varredura da imagem, gerando informação completa do relevo e com o nível de definição requerido, é inviável financeiramente. Contudo, a aplicação na medição da bitola do tarugo seria viável.

O condicionamento da iluminação foi fator decisivo para os bons resultados nos testes preliminares da linha de inspeção. O contraste entre os números e o restante do tarugo foi realçado e a presença de reflexos diminuída. Além destes aspectos, os caracteres estampados estavam com maior definição e com pouca deformação. A deterioração das estampas provocará uma queda imediata no desempenho do sistema. O tempo de troca das estampas deve ser analisado para o bom desempenho do sistema em longo prazo.

Na fase de implementação e testes *in locus* foi necessária a adoção de algumas medidas diferentes das realizadas em laboratório. As soluções desenvolvidas, contudo, podem ser empregadas. As modificações realizadas tendem a aumentar o desempenho das etapas, condicionando o ambiente e melhorando o contraste das imagens coletadas. A consideração de dois planos de trabalho também é importante, tendo a finalidade de aumentar a definição de caracteres para a fase de reconhecimento. Com a utilização de *zoom* ótico, a imagem é amostrada a uma frequência maior, fornecendo maiores informações de detalhes.

A resolução dos problemas foi adaptada às condições encontradas na LIT (linha de inspeção de tarugos). As adaptações são importantes, uma vez que as soluções industriais são, geralmente, realizadas sob medida. A aplicação em outras áreas é perfeitamente possível, mas ajustes serão necessários. A solução industrial deve ser personalizada, com foco em uma solução ótima para o problema, mesmo que não seja uma solução generalizada.

Trabalhos Futuros

Uma perspectiva de trabalho futuro é a aplicação do método de limiarização e reconhecimento nas imagens *in locus*, desenvolvidas no Capítulo 3. Além disto, outros extratores de características devem ser analisados para comparação de desempenho e integração com a técnica proposta.

A etapa de segmentação explícita dos caracteres deve ser mais bem formulada. Em testes preliminares, ainda não concluídos, o método para demarcação da região, utilizando análise de textura e a técnica proposta por Claudino [12] forneceu bons precedentes para sua utilização na segmentação. Com a aplicação deste método, os caracteres poderiam ser

Tabela 12: Tabela de relação das OPs inspecionadas no dia 14/01/2005 com as bitolas (em mm).

OP	Inspeção	Bitola	OP	Inspeção	Bitola
0659391	14/1/2005 04:17	150,00	0667671	14/1/2005 18:12	215,00
0675452	14/1/2005 04:18	160,00	0667679	14/1/2005 18:26	215,00
0675009	14/1/2005 04:21	160,00	0663019	14/1/2005 19:22	150,00
0675455	14/1/2005 04:29	160,00	0668879	14/1/2005 19:57	129,00
0651645	14/1/2005 09:49	100,00	0668880	14/1/2005 20:13	129,00
0659395	14/1/2005 09:58	150,00	0641021	14/1/2005 21:44	255,00
0667707	14/1/2005 10:28	215,00	0668162	14/1/2005 22:48	130,00
0669542	14/1/2005 11:09	129,00	0663339	14/1/2005 23:05	240,00
0668942	14/1/2005 11:10	129,00	0651641	14/1/2005 23:13	100,00
0675495	14/1/2005 15:47	160,00	0665454	14/1/2005 23:13	150,00
0675492	14/1/2005 15:47	160,00	0665450	14/1/2005 23:14	150,00
0651603	14/1/2005 15:51	100,00	0675510	14/1/2005 23:14	160,00
0675501	14/1/2005 15:56	160,00	0675504	14/1/2005 23:14	160,00
0675498	14/1/2005 15:56	160,00	0671968	14/1/2005 23:14	160,00
0675521	14/1/2005 16:14	160,00	0675507	14/1/2005 23:14	160,00
0675458	14/1/2005 16:20	160,00	0655183	14/1/2005 23:14	150,00
0675524	14/1/2005 16:23	160,00	0675513	14/1/2005 23:26	160,00
0662085	14/1/2005 16:39	130,00	0675516	14/1/2005 23:33	160,00
0663395	14/1/2005 16:52	240,00	0665446	14/1/2005 23:44	150,00

segmentados diretamente da região delimitada, considerando uma margem de segurança.

As etapas de pós-processamento também devem ser formuladas, analisando a lista das OPs diárias e também a bitola do tarugo. A bitola é a dimensão da seção reta do tarugo, relacionada com a OP. Há somente uma bitola para uma determinada OP, mas diversas OPs com mesma bitola. Contudo, esta análise pode diminuir ainda mais as possibilidades de erros, diminuindo o espaço de busca de soluções. A Tabela 12 apresenta a lista das OPs de tarugos inspecionados no dia 14/01/2005. Se a medição da bitola fosse realizada de maneira eficiente, o reconhecimento das OPs no dia ficariam sujeitas a, no máximo, 16 opções (bitola 160). Há casos, como a bitola 240, que apresentam somente duas opções.

Para implementação da medição da bitola, o método “Isolamento dos Tarugos”, apresentado no Capítulo 4 fornece bons requisitos para uma medição confiável.

Referências

- [1] Gerdau S. A. Gerdau Açominas - Produtos, [Online]. Disponível: <http://www.acominas.com.br/br/produtos/tarugos.asp>.
- [2] Gerdau S. A. A Gerdau, [Online]. Disponível: <http://www.gerdau.com.br/agerdau/>.
- [3] Lars Aurdal. Image Segmentation, thresholding - An Introduction. Notas de aula das disciplinas INF3300, INF4300, September 2004. Norwegian Computing Center.
- [4] S. O. Belkasim, M. Shridhar, and A. Ahmadi. Pattern recognition with moment invariants: A comparative study and new results. *Pattern Recognition*, 24:1117–1138, 1991.
- [5] J.-A. Beraldin, F. Blais, L. Cournoyer, G. Godin, and M. Rioux. Active 3d sensing. In *Modelli e Metodi per lo studio e la conservazione dell'architettura storica*, volume 10, pages 22–46, Pisa, April 2000. Scuola Normale Superiore.
- [6] Al Bovik. *Handbook of Image and Video Processing*. Communications, Networking, and Multimedia. Academic Press, 2000.
- [7] A. P. Braga, A. P. L. F. Carvalho, and T. B. Ludemir. *Redes Neurais Artificiais: teoria e aplicações*. Livros Técnicos e Científicos, Rio de Janeiro, RJ, 2000.
- [8] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
- [9] E. Oran Brigham. *The Fast Fourier Transform and Its Applications*. Prentice Hall, 1974.
- [10] R. Casey and E. Lecolinet. A survey of methods and strategies in character segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7), 1996.
- [11] C. H. Chen and L. F. Pau. *The Handbook of Pattern Recognition and Computer Vision*. World Scientific Publishing Co., 2nd edition, 1998.
- [12] Leonardo Max Batista Claudino. Segmentação de fragmentos de texto para a localização de placas de automóveis. Master's thesis, Universidade Federal de Minas Gerais, April 2005.
- [13] Roselito de A. Teixeira, Antonio de P. Braga, Ricardo H.C. Takahashi, and Rodney R. Saldanha. A multi-objective optimization approach for training artificial neural networks. *VI Brazilian Symposium on Neural Networks (SBRN'00)*, page 168, 2000.
- [14] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.

- [15] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.
- [16] W. Frei and C. Chen. Fast boundary detection: A generalization and a new algorithm. *IEEE Trans. Computers*, pages 988–998, 1977.
- [17] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition*. Computer Science and Scientific Computing. Academic Press, second edition, 1990.
- [18] Rafael C. Gonzales and Richard E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, second edition, 1992.
- [19] S. Haykin. *Neural Networks - A Comprehensive Foundation*. Prentice-Hall, 1994.
- [20] P. V. C. Hough. Methods and means for recognizing complex patterns. U. S. Patent 3069654, December 1962.
- [21] Inc. Illumination Technologies. Lighting Tutorial, [Online]. Disponível: <http://www.illuminationtech.com/seminars/index.htm>, 2001.
- [22] A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12):1.167–1.186, 1991.
- [23] Anil K. Jain, Robert P.W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and machine intelligence*, 22(1):4–37, January 2000.
- [24] D. Kececioglu. *Reliability Engineering Handbook*, volume 1. Prentice Hall, 1991.
- [25] R. Kirsch. Computer determination of the constituent structure of biomedical images. *Computers and Biomedical Research*, 4(3):315–328, 1971.
- [26] S. Knerr, L. Personnaz, and G. Dreyfus. Handwritten digit recognition by neural networks with single-layer training. *IEEE Transactions on Neural Networks*, 3(6):962–968, November 1992.
- [27] Kaist AI Lab. Popular image database, [Online]. Disponível: <http://ai.kaist.ac.kr/Resource/dbase/Image>
- [28] K. Laws. *Textured Image Segmentation*. PhD thesis, University of Southern California, January 1980.
- [29] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten Zip code recognition. *Neural Computation*, 1:541–551, 1989.
- [30] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel. Handwritten digit recognition with a back-propagation network. In D.S. Touretzky, editor, *Neural Information Processing Systems*, volume 2, pages 396–404, Denver 1989, 1990. Morgan Kaufmann, San Mateo.

- [31] S. W. Lee, L. Lam, and C. Y. Suen. Performance evaluation of skeletonizing algorithms. In *First International Conference on Document Analysis and Recognition*, pages 260–271, 1991.
- [32] H.P. Li and D. Doermann. Automatic text detection and tracking in digital video. *IEEE Trans. on Image Processing*, 9(1):147–156, January 2000.
- [33] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11:674–693, July 1989.
- [34] Gale L. Martin and James A. Pittman. Recognizing hand-printed letters and digits using backpropagation learning. *Neural Computation*, 3(2):258 – 267, 1991.
- [35] Paulo Roberto Martins. Segmentação de histogramas multimodais: Simulações, implementações e aplicação em cheques bancários brasileiros. Master’s thesis, Pontifícia Universidade Católica do Paraná, Curitiba, 2000.
- [36] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, March 1997.
- [37] Marisa Emika Morita. *Automatic Recognition of Handwritten Dates on Brazilian Bank Cheques*. PhD thesis, École de Technologie Supérieure Université du Québec, Montreal, June 2003.
- [38] W. Niblack. *An introduction to Digital Image Processing*. Prentice-Hall, 1986.
- [39] Douglas Rodrigues Oliveira. Especificações para identificações da OP nos tarugos. Relatório Técnico da Açominas, 2003.
- [40] L. S. Oliveira, R. Sabourin, F. Bortolozzi, and C. Y. Suen. High-level verification of handwritten numeral strings. In *Proceedings of the 14th Brazilian Symposium on Computer Graphics and Image Processing*, pages 36–43, Florianópolis-Brazil, 2001. IEEE Computer Society.
- [41] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.
- [42] N. P. Paragios. *Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision*. PhD thesis, INRIA Sophia Antipolis, France, January 2000.
- [43] William K. Pratt. *Digital Image Processing: PIKS Inside*. John Wiley & Sons, Inc., Los Altos, California, third edition, 2001.
- [44] J. M. S. Prewitt. Object enhancement and extraction. In B. S. Lipkin and A. Rosenfeld, editors, *Picture Processing and Psychopictorics*. Academic Press, New York, 1970.
- [45] K. Shanmugam R. Haralick and I. Dinstein. Textural features for image classification. *IEEE Trans. Systems, Man, Cybernetics*, 3:610–621, Novembro 1973.

- [46] T. Randen. *Filter and Filter Bank Design for Image Texture Recognition*. PhD thesis, Norwegian University of Science and Technology, November 1997.
- [47] T. Randen and J. H. Husoy. Filtering for texture classification: A comparative study. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(4):291–310, 1999.
- [48] L. G. Roberts. Machine perception of three-dimensional solids. In J. T. Tippett, editor, *Optical and Electro-Optical Information Processing*. MIT Press, 1965.
- [49] G. S. Robinson. Edge detection by compass gradient masks. *Computer Graphics and Image Processing*, 6(5):492–501, 1977.
- [50] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.
- [51] John C. Russ. *The Image Processing Handbook*. CRC Press LLC, third edition, 1999.
- [52] M. Sabourin, A. Mitiche, D. Thomas, and G. Nagy. Classifier combination for hand-printed digit recognition. In *Proceedings of the Second International Conference on Document Analysis and Recognition*, pages 163–166, Tsukuba Science City, Japan, October 1993. IEEE Computer Society Press.
- [53] Linda Shapiro and George Stockman. *Computer Vision*. Prentice Hall, March 2000.
- [54] I. Sobel. *Pattern Classification and Scene Analysis*, chapter 7 - Representation and Initial Simplifications, page 271. John Wiley & Sons, 1973.
- [55] Robert D. Strum and Donald E. Kirk. *First Principles of Discrete Systems and Digital Signal Processing*. Addison-Wesley Publishing Company, 1988.
- [56] J. G. Taylor. Spontaneous behavior in neural networks. *Journal of Theoretical Biology*, 36:513–528, 1972.
- [57] O. D. Trier, A. K. Jain, and T. Taxt. Feature extraction methods for character recognition - a survey. *Pattern Recognition*, 29(4):641–662, 1991.
- [58] O. D. Trier and T. Taxt. Evaluation of binarization methods for document images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17:312–315, 1995.
- [59] Oivind Due Trier and Anil K. Jain. Goal-directed evaluation of binarization methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1191–1201, 1995.
- [60] R. Manmatha V. Wu and E. M. Riseman. Finding text in images. In *2nd ACM Int. Conf. on Digital Libraries*, pages 3–12, Philadelphia, USA, July 1997. ACM.
- [61] A. Wernicke and R. Lienhart. On the segmentation of text in videos. In *IEEE Int. Conf. on Multimedia and Expo (ICME 2000)*, pages 1511–1514, New York, USA, July 2000. IEEE.
- [62] S. D. Yanowitz and A. M. Bruckstein. A new method for image segmentation. *Comput. Vis., Graph., Image Processing*, 46(1):82–95, 1989.

-
- [63] Y.C.Cheng and S.C.Levine. A new method for quadratic curve detection using k-ransac with acceleration techniques. *Pattern Recognition*, 16(9):663–682, 1995.
- [64] S. Yoshimura and Takeo Kanade. Fast template matching based on the normalized correlation by using multiresolution eigenimages. In *Proceedings of the 1994 IEEE/RSJ/GI International Conference on Intelligent Robots and Systems, Advanced Robotic Systems and the Real World (IROS '94)*, volume 3, pages 2086 – 2093, September 1994.
- [65] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690–706, August 1999.