



FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

DEPARTAMENTO DE
ENGENHARIA MECÂNICA

Sistema de classificação de tecidos para a indústria automóvel, uma abordagem baseada em aprendizagem

Dissertação apresentada para a obtenção do grau de Mestre em Engenharia E Gestão Industrial

Autor

Carlos Diogo Pires Carvalho Gomes

Orientador

Germano Manuel Correia dos Santos Veiga

Júri

Presidente Professor Doutor Cristóvão Silva
Professor Auxiliar da Universidade de Coimbra

Vogais Professor Doutor Jorge Manuel Miranda Dias
Professor Associado c/Agreg. da Universidade de Coimbra

Coimbra, Setembro, 2012

“Os computadores são incrivelmente rápidos, precisos e burros;
os homens são incrivelmente lentos, imprecisos e brilhantes;
juntos, o seu poder ultrapassa os limites da imaginação.”

Albert Einstein

Agradecimentos

O trabalho que aqui se apresenta só foi possível graças à colaboração e apoio de algumas pessoas, às quais não posso deixar de prestar o meu reconhecimento.

Gostaria de começar por agradecer ao Prof. Dr. Cristóvão Silva pelo seu apoio ao longo de todo o curso, ao Prof. Dr. Germano Veiga pela ajuda e orientação deste trabalho e ao Pedro Malaca pela sua constante disponibilidade.

Não quero perder esta oportunidade para agradecer às pessoas mais importantes da minha vida, às quais nunca consegui agradecer por tudo e, muito provavelmente, nunca conseguirei.

Agradeço aos meus amigos que, directa ou indirectamente, contribuíram para esta tese e, acima de tudo, por poder contar com a sua preciosa amizade.

À minha família, em particular à minha família mais próxima, aos meus pais e ao meu irmão. Aos meus pais, um agradecimento muito especial, pela educação que me deram, que é, sem qualquer margem de dúvida, a minha maior força e virtude.

À minha mulher gostaria de agradecer o apoio inigualável e a compreensão pelos períodos de ausência e pelas noites de trabalho.

Resumo

O objectivo deste trabalho é a classificação, em tempo-real, de tecidos da indústria automóvel sob condições de iluminação não-controlada.

Foram estudadas e analisadas diversas técnicas de pré-processamento e de escolha das características de textura mais adequadas para o problema. Ao nível do pré-processamento, a solução obtida utiliza processos de equalização do histograma, de aplicação de filtros de Laws e Sobel, e de análise multi-escala. O vector final de características das imagens compreende características estatísticas relativas aos histogramas dos 3 canais de cor e relativas às matrizes de co-ocorrência das imagens pré-processadas. A classificação foi realizada usando o algoritmo de Máquinas de Suporte Vectorial (SVM – Support Vector Machines).

Alguns processos industriais não possibilitam o controlo eficaz da iluminação para os seus sistemas de visão, prejudicando muito a eficácia dos mesmos. A possibilidade de ultrapassar esses problemas com técnicas adequadas de pré-processamento e de escolha de características permitirá aumentar a eficácia dos sistemas de visão industrial, com claros ganhos produtivos e, conseqüentemente, económicos.

Foram testadas um conjunto de amostras com grande variabilidade de condições de luz, de forma a simular o ambiente industrial com condições não ideais de luz. A solução encontrada apresentou taxas de sucesso de classificação de 100%, usando um algoritmo muito rápido e robusto.

Palavras-chave: Visão computacional, classificação, análise de textura, matrizes de co-ocorrência, processamento de imagens digitais, iluminação não-controlada.

Abstract

The aim of this work is the classification, in real-time, of fabric texture for the auto industry, under uncontrolled illumination.

It was studied and analyzed various preprocessing techniques and the choice of the most appropriate texture characteristics for our case study. In terms of preprocessing, the obtained solution uses histogram equalization, Laws and Sobel filters, and multi-scale analysis. The final characteristics' vector of the images includes statistics' features of the 3 color channels' histograms and of the co-occurrence matrices taken from the preprocessed images. The classification was carried out using the algorithm Support Vector Machines (SVM).

Some industrial processes do not permit effective control of illumination for their vision systems, hindering their effectiveness. The ability to overcome these problems with suitable techniques of preprocessing and choice of characteristics will increase the efficiency of industrial vision systems, with obvious production gains and thus economical.

It was tested a set of samples with great variability of illumination conditions in order to simulate the industrial environment with non-ideal illumination conditions. The solution presented classification's success rates of 100% using a very fast and robust algorithm.

Keywords Computer vision, classification, texture analysis, co-occurrence matrices, digital image processing, uncontrolled illumination.

Índice

Índice de Figuras	vi
Índice de Tabelas	vii
Simbologia e Siglas	viii
Simbologia.....	viii
Siglas	viii
1. INTRODUÇÃO.....	1
1.1. Motivação	1
1.2. Objectivos	1
1.3. Organização do trabalho	2
2. Processamento de Imagens Digitais	4
2.1. Imagens Digitais	4
2.1.1. Modelação Matemática.....	4
2.1.2. Modelo de Cores.....	5
2.1.3. Ruído	7
2.1.4. Histograma	7
2.1.5. Imagens em Pirâmide	8
2.2. Processamento de Imagens	8
2.2.1. Transformação de Cor	8
2.2.2. Equalização do Histograma	9
2.2.3. Filtros.....	10
3. Análise de Texturas	14
3.1. Abordagem Estatística	14
3.1.1. Medidas Baseadas na Distribuição de Níveis de Cinza.....	14
3.1.2. Matriz de Co-ocorrência.....	16
3.2. Abordagem Estrutural.....	19
3.3. Abordagem Espectral.....	20
3.4. Outras Abordagens.....	20
4. Classificação de Imagens.....	21
4.1. Introdução	21
4.2. SVM - Máquinas de Suporte Vectorial.....	21
4.3. Matriz de Confusão.....	23
5. Caso Prático.....	25
5.1. Introdução	25
5.2. Metodologia de Ensaio	26
5.2.1. Amostras de Imagens	26
5.2.2. Software Halcon	28
5.2.3. Classificação.....	28
5.2.4. Processamento das Imagens	29
5.3. Resultados	29
5.3.1. Modelos de Cores e Seus Histogramas	29
5.3.2. Filtros e Matriz de Co-Ocorrência.....	30
5.3.3. Conversão Escala de Cinzas e Filtro de Sobel	31

5.3.4.	Análise Piramidal	32
5.3.5.	Uso da Matriz de Confusão	33
5.3.6.	Representatividade da Amostragem	34
5.3.7.	Tempos de Simulação.....	34
6.	Conclusões.....	35
7.	Referências bibliográficas	37

ÍNDICE DE FIGURAS

Figura 1. Exemplos de Imagens e Sua Representação Digital	5
Figura 2. Representação modelo RGB (Schwartz e Pedrini, 2007)	6
Figura 3. Imagens Coloridas e Sua Representação Digital no modelo RGB	6
Figura 4. Representação modelo HSI (Schwartz e Pedrini, 2007)	7
Figura 5. Análise Piramidal (Gonzalez e Woods, 2002)	8
Figura 6. Exemplo de equalização do histograma. (Schwartz e Pedrini, 2007)	10
Figura 7. 9 Máscaras 3x3 de Laws	12
Figura 7. Exemplo da aplicação do Filtro de Sobel	13
Figura 8. Exemplo do cálculo de GLMC.....	17
Figura 9. Representação da GLMC das 4 direcções para uma amostra	17
Figura 9. Transformação espacial do SVM (Halcon, Solution Guide II-D).....	22
Figura 10. Vectores de Suporte do SVM (Halcon, Solution Guide II-D)	23
Figura 11. Exemplo de uma matriz de confusão	23
Figura 12. Classes de tecidos do caso de estudo	25
Figura 13. Variabilidade das Amostras	27
Figura 14. Exemplo de utilização da matriz de confusão	33

ÍNDICE DE TABELAS

Tabela 1. Conjuntos de amostras por cada classe do caso prático	26
Tabela 2. Resultados do primeiro grupo de simulações	30
Tabela 3. Resultado óptimo do segundo grupo de simulações.....	30
Tabela 4. Resultados do terceiro grupo de simulações.....	31
Tabela 5. Resultados do quarto grupo de simulações.....	32
Tabela 6. Resultados do quinto grupo de simulações.....	33
Tabela 7. Tempos de Simulação.....	34

SIMBOLOGIA E SIGLAS

Simbologia

μ – Média

σ – Desvio-padrão

ani – Anisotropia

con – Contraste

cor – Correlação

ene – Energia

ent – Entropia

hom – Homogeneidade

k – Curtose

s – Grau de assimetria

L_{\max} – nível máximo de cinza possível na imagem

Siglas

FL – Filtro de Laws

FS – Filtro de Sobel

GLCM – Grey Level Co-Occurrence Matrix – Matriz de Co-Ocorrência de Níveis de Cinza

JIT – Just in Time

LBP – Local Binary Patterns – Padrão Binário Local

MLP – Multi-Layer Perceptrons – Redes Neurais Artificiais

SVM – Support Vector Machines – Máquinas de Suporte Vectorial

1. INTRODUÇÃO

1.1. Motivação

Análise de texturas é uma área de estudo importante e útil em sistemas de visão computacional. Um sistema de visão computacional captura imagens através de um dispositivo ou sensor e analisa-as de forma a calcular valores que definem as mais diversas características dessas imagens. Uma aplicação típica de um sistema de visão computacional é na área de produção industrial, seja para inspeção visual automática, quer seja para automação da linha de produção (Theodoris, S. e Koutroumbas, K., 2003). Na inspeção, por exemplo, a visão computacional permite verificar defeitos em objectos com maior rapidez e, muitas vezes, com maior rigor do que a inspeção humana.

A crescente necessidade de aumento de produtividade e poupança de custos tem dinamizado muito a área de conhecimento e aplicação dos sistemas de visão computacionais na área de produção industrial.

Cartografia, medicina, microscopia, reconhecimento da fala, identificação de impressões digitais, leitura de códigos de barra, pagamento via-verde, etc são somente alguns exemplos de áreas e aplicações que utilizam sistemas de visão.

1.2. Objectivos

O objectivo principal deste trabalho é a concepção de um algoritmo de classificação para um caso de estudo real da indústria automóvel. Os tecidos encontram-se inseridos em portas de automóveis (Figura 1), as quais, pela sua dimensão e disposição da linha de montagem, não permitem ter condições de iluminação controlada para as imagens capturadas pelo sistema de visão. O caso prático de estudo exige a obtenção de um algoritmo que permita uma taxa de classificação elevada mas que seja suficientemente rápido no processo de classificação em tempo real na linha de montagem.



Figura 1. Porta de automóvel com uns das classes de tecido

Esse objectivo implicou o estudo, aplicação e análise de resultados dos vários parâmetros envolvidos no processo de classificação. Numa primeira fase, a aplicação de diversas técnicas de pré-processamento de imagens, de modo a reduzir os problemas de iluminação não controlada e a realçar as características texturais. Numa segunda fase, a análise e selecção das características que melhor definem (e distinguem) as diversas texturas é um ponto crítico do sucesso do algoritmo de classificação.

Além de encontrar uma solução para o caso prático de estudo, outro objectivo, sempre presente ao longo do desenvolvimento deste trabalho, passou por conseguir uma solução ajustável a outras situações com problemáticas semelhantes. Com os conceitos e metodologia de resolução deste caso prático, descritos ao longo deste trabalho, ambiciona-se conseguir tal objectivo.

1.3. Organização do trabalho

Para simplificação e limitação da extensão deste trabalho, a introdução teórica dos conceitos e algoritmos serão restringidos àqueles que foram utilizados no processo de obtenção da solução para o caso de estudo deste trabalho.

No capítulo 2 apresentam-se os conceitos básicos de imagens digitais e os métodos de processamento usados neste trabalho.

No capítulo 3 introduz-se os conceitos e problemáticas da análise de texturas. As principais abordagens são descritas sumariamente e os métodos de extracção de características de textura utilizados neste trabalho são apresentados com maior detalhe.

Uma explicação do processo de classificação e, em particular, o método da Máquinas de Suporte Vectorial (SVM – Support Vector Machines) são explicados no capítulo 4. Alguns métodos de optimização e ajuste dos algoritmos de classificação são também descritos neste capítulo.

No capítulo 5, o caso de estudo é explicado com detalhe, bem como toda a metodologia de estudo e análise do problema. Apresentam-se os principais resultados das simulações efectuadas e as suas conclusões.

Finalmente, no capítulo 6, indicam-se as principais conclusões deste trabalho.

2. PROCESSAMENTO DE IMAGENS DIGITAIS

2.1. Imagens Digitais

2.1.1. Modelação Matemática

Do ponto de vista matemático, uma imagem é considerada uma função bidimensional $f(x,y)$ onde x e y são coordenadas planas, e a amplitude de f em qualquer par de coordenadas (x,y) é chamada de intensidade ou nível de cinza da imagem no referido ponto. Quando (x,y) e a amplitude de f fazem parte de um conjunto de valores finitos, ou discretos, a imagem é chamada de imagem digital (Gonzalez e Woods, 2002), onde cada elemento da função $f(x,y)$ é denominado de pixel.

A função $f(x,y)$ pode ser expressa pelo produto entre dois componentes, a quantidade de luz incidente na cena e a quantidade de luz reflectida pelos objectos presentes na cena. Esses componentes são chamados de iluminância e reflectância, respectivamente, e são representados por $i(x,y)$ e $r(x,y)$. A natureza de $i(x,y)$ é determinada pela fonte de luz, enquanto $r(x,y)$ é determinada pelas características dos objectos na cena. (Schwartz e Pedrini, 2007).

Assim, a função $f(x,y)$ pode ser representada como:

$$f(x,y) = i(x,y) \cdot r(x,y) \quad (1)$$

Em que:

$$0 < i(x,y) < \infty \text{ e } 0 < r(x,y) < 1 \quad (2)$$

Gonzalez e Woods (2002) indicam alguns valores médios para $i(x,y)$. Num dia claro, a iluminância pode atingir os 90000 lúmen/m², enquanto que num dia nublado esse valor é inferior a 10000 lúmen/m². O nível de iluminância típico num escritório é de aproximadamente 1000 lúmen/m². Numa noite clara de lua cheia teremos cerca de 0.1 lúmen/m².

Na prática estipula-se um intervalo $[0, L-1]$ para os valores de intensidade da função $f(x,y)$, chamado de intervalo de valores de cinza (ou intensidade), onde o valor 0

representa o preto, L-1 representa o branco, e os restantes valores do intervalo representam tons de cinza.

Tipicamente considera-se que L, o número de níveis de intensidade possíveis numa imagem digital, é igual a 2^k . Assim, com k igual a 1, teremos uma imagem binária, e com k igual a 8, teremos 256 níveis de intensidade (ou cinza) possíveis na imagem (Figura 2).

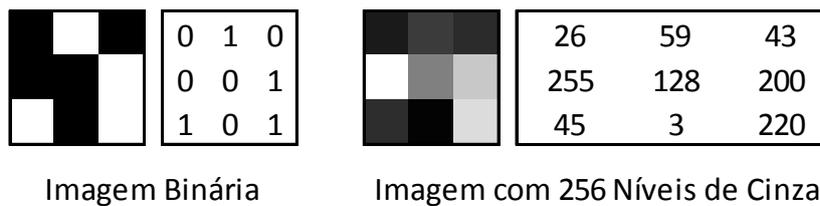


Imagem Binária

Imagem com 256 Níveis de Cinza

Figura 2. Exemplos de Imagens e Sua Representação Digital

2.1.2. Modelo de Cores

Dois conceitos são importantes para o entendimento do conceito de percepção de cor: luminância e a crominância. A luminância contém a informação da quantidade das cores pretas e brancas presentes na imagem. O cérebro interpreta essa informação como a quantidade de cinza presente na cor (ou brilho). A crominância informa a respeito da tonalidade de uma cor, é a frequência dominante do raio de luz. A combinação destes dois conceitos, em diferentes proporções, permite ao cérebro perceber o espectro de cores visível numa determinada imagem ou cena. Para representar as cores existem modelos ou sistemas de cores como, por exemplo, o RGB (*Red, Green, Blue*), o CMY (*Cyan, Magenta, Yellow*), o HSI (*hue, saturation, intensity*), o HSV (*hue, saturation, value*), entre muitos outros.

O sistema RGB é um sistema de representação de cor aditivo e baseia-se na teoria dos três estímulos proposta por Young-Helmholtz. Segundo esta teoria, o olho humano percebe a cor através do estímulo de três pigmentos visuais presentes nos cones da retina, que possuem sensibilidades para alguns comprimentos de onda. Em 1931, a Comissão Internacional de Iluminação atribuiu os valores de comprimento de onda 700 nanómetro para o vermelho, 546.1 nm para o verde e 435.8 nm para o azul.

No modelo RGB, cada componente de cor é representada pelos seus componentes espectrais primários, Vermelho, Verde e Azul. Este modelo é baseado num sistema de coordenadas cartesiano, representado na Figura 3, onde as cores primárias estão

representadas em 3 vértices, as cores secundárias em 3 outros vértices, a cor preta na origem e a branca no vértice mais afastado da origem. Neste modelo os pontos com cor cinzenta (pontos com igual intensidade nos canais R, G e B) estão representados por uma linha que une os pontos que representam as cores preta e branca (Gonzalez e Woods, 2002).

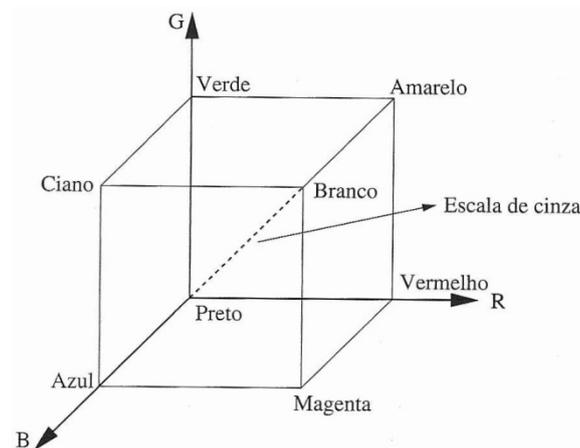


Figura 3. Representação modelo RGB (Schwartz e Pedrini, 2007)

O modelo RGB é indicado para implementações de hardware, como monitores e impressoras, mas não descreve a cor de forma prática para a interpretação humana. O exemplo da Figura 4 mostra um exemplo de representação de uma imagem colorida no sistema RGB, onde se pode verificar as cores a sua representação no modelo RGB.

	26,59,43	115,60,176	128,165,226
	255,128,200	210,32,240	34,93,126
	45,3,220	42,195,211	136,148,229

Imagem Colorida (RGB) com 256 Níveis de Cinza

Figura 4. Imagens Coloridas e Sua Representação Digital no modelo RGB

O modelo HSI, representado na Figura 5, corresponde praticamente à forma como os humanos descrevem e interpretam a cor. Além disso, este modelo apresenta também a vantagem de separar a informação de cor (canais de Matiz e Saturação) da informação de intensidade (canal de Intensidade) de uma imagem (Gonzalez e Woods, 2002), tornando este modelo teoricamente mais vantajoso para análise de imagens em condições de iluminação não controlada.

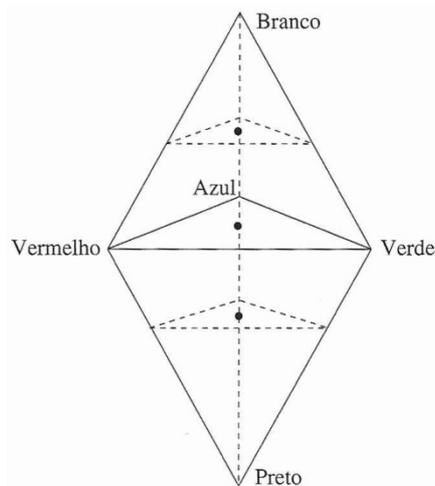


Figura 5. Representação modelo HSI (Schwartz e Pedrini, 2007)

2.1.3. Ruído

Imagens reais sofrem degradações durante o processo de aquisição, transmissão ou processamento. Essa degradação é normalmente chamada de ruído. O ruído pode ser considerado uma variável aleatória z , caracterizada por uma função densidade de probabilidade $p(z)$. Os tipos de ruído mais comumente modelados são o ruído impulsivo, Gaussiano, uniforme, Erlang, exponencial, Rayleigh e Poisson (Schwartz e Pedrini, 2007).

Este ruído pode ser reduzido por técnicas de filtragem, algumas delas apresentadas em 2.2.3.

2.1.4. Histograma

O Histograma de valores absolutos, $H(i)$, apresenta o número de pixel da imagem que têm o nível de cinza igual a i . É um vector com a dimensão igual ao número de níveis de cinza possíveis na imagem.

O Histograma de valores relativos, $h(i)$, obtém-se dividindo o histograma de valores absolutos pelo número de pixels da imagem, correspondendo a uma função de massa de probabilidade.

Os histogramas perdem a informação espacial do pixel e a relação com a sua vizinhança, mas são muito utilizados na caracterização de imagens, devido ao seu custo computacional baixo, à sua simplicidade de visualização e compreensão.

2.1.5. Imagens em Pirâmide

A técnica das imagens piramidais, ou multi-escala, é uma técnica simples, proposta por Burt and Adelson (1983), para analisar imagens em diferentes resoluções. Consiste no conjunto de imagens com resolução decrescente organizadas em forma de pirâmide. A base da pirâmide contém a imagem original, com a máxima resolução, a imagens seguintes terão sucessivamente menor resolução e tamanho (Figura 6).

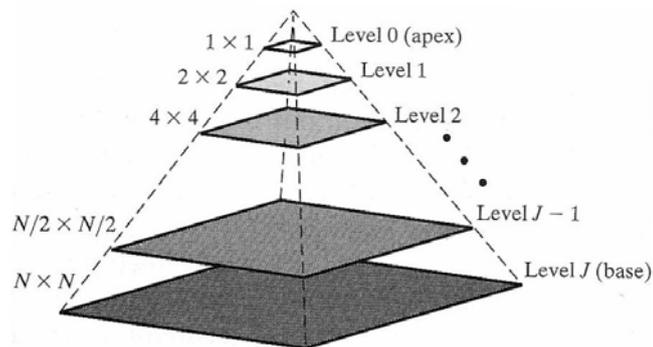


Figura 6. Análise Piramidal (Gonzalez e Woods, 2002)

Este tipo de técnica permite capturar características das texturas em diferentes níveis de resolução. Isso poderá permitir capturar e definir correctamente algumas características texturais num determinado nível que não terão sido capturadas noutros níveis.

2.2. Processamento de Imagens

Existem muitos métodos de processamento e transformação de imagens, com diversos objectivos e áreas de aplicação. A aplicação correcta destas técnicas podem permitir a diminuição da variabilidade nas imagens tiradas em condições de iluminação não ideais, realçar ou suavizar os detalhes da imagem, destacar arestas, etc

2.2.1. Transformação de Cor

Com uma imagem cromática é possível realizar diversas transformações ao nível da cor. As mais comuns, em sistemas de visão, são as conversões para outros modelos de cor e a transformação para uma imagem monocromática.

A transformação do modelo de cor de RGB para HSI é muito comum entre os especialistas em visão computacional, pois, devido às razões descritas em 2.1.2., este modelo separa a informação de cor da intensidade.

Muitas texturas são distinguíveis pelo sistema visual humano sem a informação de cor. A transformação de uma imagem cromática para uma imagem monocromática revela-se bastante útil, quer pela simplificação da análise dos resultados, quer pela poupança ao nível computacional, pois os métodos subsequentes terão uma imagem com um canal a processar, ao invés dos 3 canais da imagem original. Uma imagem no modelo de cor RGB é transformada, usando a equação (3), para uma imagem monocromática:

$$f_{x,y}^{\text{mono}} = \sum_{x=1}^M \sum_{y=1}^N [0.299 \cdot f_{x,y}^{\text{red}} + 0.587 \cdot f_{x,y}^{\text{green}} + 0.114 \cdot f_{x,y}^{\text{blue}}] \quad (3)$$

2.2.2. Equalização do Histograma

O objectivo da equalização do histograma é obter uma distribuição mais uniforme dos níveis de cinza da imagem (Jain, 1989), ou seja, os níveis de cinza devem aparecer na imagem aproximadamente com a mesma frequência. Esta técnica tem a grande vantagem de ser totalmente automática, isto é, sem parametrização ou ajuste manuais. No entanto, há situações em que esta equalização do histograma pode degradar uma imagem (Schwartz e Pedrini, 2007).

Na equalização do histograma foi utilizada a função de distribuição acumulada:

$$g_k = \sum_{i=1}^k h_i \quad \text{com } k = 0, 1, \dots, L-1 \quad (4)$$

Onde h_i representa o histograma relativo dos níveis de cinza da imagem a ser equalizada.

Mostra-se o exemplo da Figura 7, onde as alíneas a) e b) representam a imagem original e respectivo histograma, e as alíneas c) e d) a imagem equalizada e respectivo histograma.

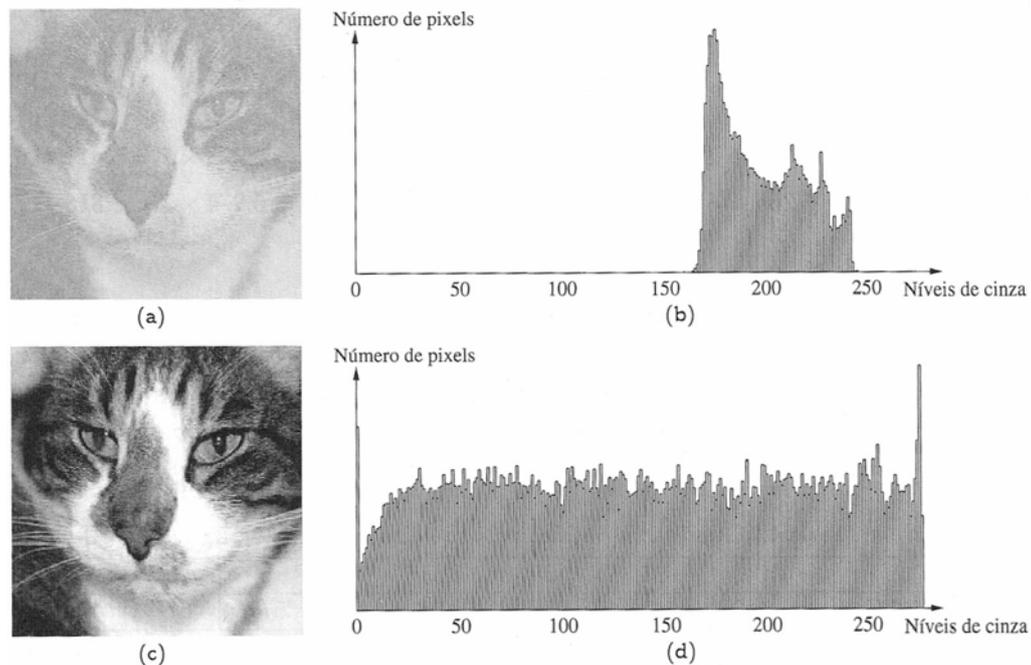


Figura 7. Exemplo de equalização do histograma. (Schwartz e Pedrini, 2007)

2.2.3. Filtros

As operações de filtragem podem ser realizadas, tanto no domínio do espaço, como no domínio da frequência. Neste trabalho apenas foram utilizados filtros espaciais.

Os filtros são normalmente classificados em 3 categorias: passa-baixas, passa-altas e passa-faixa. O filtro passa-baixas atenua as altas frequências, que estão relacionadas com a informação de detalhes na imagem, e tende também a diminuir o efeito do ruído nas imagens. O filtro passa-altas, pelo contrário, realça os detalhes na imagem. Um filtro passa-faixa realça um determinado intervalo de frequências do nível de cinza.

As técnicas de filtragem no domínio espacial atuam diretamente sobre a matriz de pixels da imagem digitalizada, logo, as funções de processamento de imagens no domínio espacial podem ser expressas como:

$$g(x,y) = T[f(x,y)] \quad (5)$$

onde $g(x,y)$ é a imagem processada, $f(x,y)$ é a imagem original e T é um operador em f , definido em uma certa vizinhança de (x,y) . A vizinhança definida ao redor de (x,y) poderá ser, por exemplo, uma região 3×3 pixels, denominada máscara, na qual o pixel central é o de referência. O centro dessa região é movido pixel a pixel, percorrendo a

imagem e aplicando à localidade o operador T para calcular um novo valor g naquele ponto.

2.2.3.1. Filtro da Média

Os filtros da média, do tipo passa-baixas, substituem o valor do pixel central pela média dos valores dos pixels na máscara. O exemplo seguinte mostra a máscara de 3x3 pixel de um filtro da média:

$$FM_{3 \times 3} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6)$$

2.2.3.2. Filtro Gaussiano

Nos filtros Gaussianos, do tipo passa-baixas, os coeficientes da máscara são derivados a partir de uma função Gaussiana bidimensional. Estes filtros apresentam diversas características que os tornam particularmente úteis em processamento de imagens. Para uma descrição mais pormenorizada, recomenda-se a leitura de Schwartz e Pedrini (2007). Ilustra-se o exemplo da máscara de 4x4 pixel de um filtro gaussiano:

$$FG_{4 \times 4} = \frac{1}{64} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 3 & 9 & 9 & 3 \\ 3 & 9 & 9 & 3 \\ 1 & 3 & 3 & 1 \end{bmatrix} \quad (7)$$

2.2.3.3. Filtro de Laws

As medidas de energia de textura desenvolvidas por Laws (1980) têm sido vastamente utilizadas em análise de texturas. Estas medidas foram definidas a partir de 3 simples vectores com 3 elementos:

$$\begin{aligned} L &= [1 \quad 2 \quad 1] \\ E &= [-1 \quad 0 \quad 1] \\ S &= [-1 \quad 2 \quad -1] \end{aligned} \quad (8)$$

Os quais representam um filtro de passa-baixas com ponderação no centro (l – *level*), um filtro de detecção de borda (e – *edge*) e um filtro de detecção de ponto (s – *spot*).

Multiplicando a transposta de um destes vectores por outro destes vectores obtemos as 9 máscaras de 3x3 de Laws, representados na Figura 8.

	L	E	S
L	$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix}$
E	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix}$
S	$\begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$

Figura 8. 9 Máscaras 3x3 de Laws

Mais tarde, baseados nos vectores originiais de Laws, foram propostos novos vectores com dimensões maiores, com os quais se calcularam máscaras de maiores dimensões.

Veja-se o exemplo da aplicação deste filtro do tipo “LS”, 3x3, a uma das amostras deste trabalho, do lado direito da imagem, e a original do lado esquerdo:

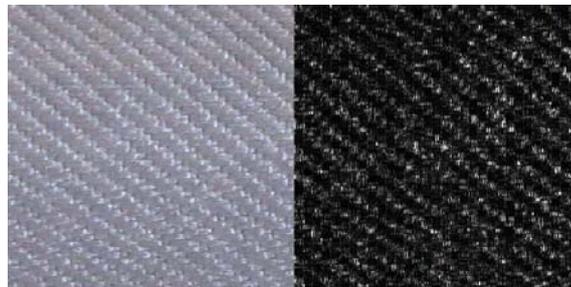


Figura 9. Exemplo aplicação filtro laws

2.2.3.4. Filtro de Sobel

Os filtros de Sobel permitem destacar as linhas das imagens. Actuam como um operador de gradiente que aproxima a magnitude do gradiente como a diferença de valores ponderados dos níveis de cinza da imagem como:

$$G_x(3x3) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y(3x3) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (9)$$

Estas máscaras podem ser aplicadas separadamente, em que a máscara G_x destacará as bordas perpendiculares ao eixo x, e a máscara G_y as bordas perpendiculares ao eixo y. Poder-se-á combinar os resultados da aplicação de ambas as máscaras de forma a obter uma imagem com todas as bordas destacadas

Veja-se o exemplo da aplicação deste filtro, com duas máscaras de tamanhos diferentes, a uma das amostras deste trabalho:

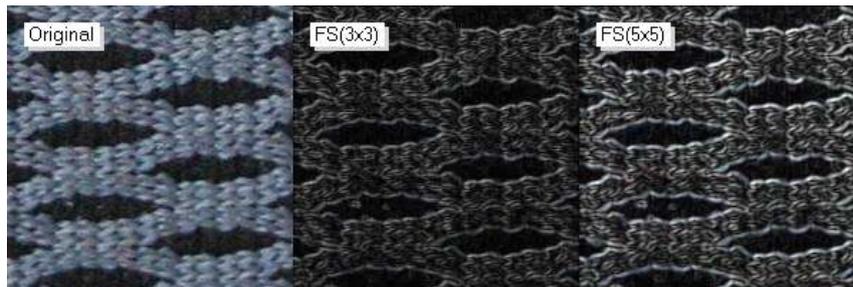


Figura 10. Exemplo da aplicação do Filtro de Sobel

3. ANÁLISE DE TEXTURAS

A definição de uma textura é simples para a visão humana, mas é bastante complexa de definir objectivamente. Prova disso é quantidade de definições de textura existentes na literatura, originando artigos que apenas abordam esta temática. Não existe nenhuma definição matemática rigorosa que seja amplamente aceite pelos autores (Vácha et al., 2011)

As propriedades associadas às texturas são consideradas uma fonte importante de dados usado na identificação de objectos ou regiões de interesse numa imagem, contribuindo para uma melhoria na precisão do processo de classificação.

Os principais métodos de extracção de características de textura podem ser classificados como abordagens estatísticas, estruturais, e espectrais (Palm 1994).

3.1. Abordagem Estatística

As abordagens estatísticas não procuram compreender explicitamente a estrutura hierárquica da textura, mas tentam representar a textura indirectamente por propriedades não-determinísticas, que definem distribuições entre os níveis de cinza dos pixéis pertencentes a uma imagem (Schwartz e Pedrini, 2007).

3.1.1. Medidas Baseadas na Distribuição de Níveis de Cinza

Esta abordagem tem como grande vantagem o custo computacional bastante baixo e como principais desvantagens a perda da informação espacial e da relação de intensidade entre pixéis vizinhos. Como só têm em consideração a intensidade do nível de cinza, as características extraídas com esta abordagem são denominadas de estatísticas de primeira-ordem

Sendo g_i o nível de cinza do pixel i e n o número de pixéis existentes na imagem, definem-se, respectivamente, a média (μ) e desvio-padrão (σ) dos níveis de cinza de uma imagem como:

$$\mu = \frac{1}{n} \sum_{i=1}^n g_i \quad (10)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (g_i - \mu)^2} \quad (11)$$

O grau de assimetria (s) da distribuição (12) dos níveis de cinza de uma imagem é um indicador da concentração de valores em relação à mediana. Um grau de assimetria negativa equivale a concentrações de valores à esquerda da mediana, e vice-versa. Por outro lado, a curtose (k), calculada através da equação (13), indica o achatamento da função de distribuição. Uma curtose negativa indica uma função mais achatada que a Gaussiana, e vice-versa.

$$s = \frac{1}{n \cdot \sigma^3} \sum_{i=1}^n (g_i - \mu)^3 \quad (12)$$

$$k = \left(\frac{1}{n \cdot \sigma^4} \sum_{i=1}^n (g_i - \mu)^4 \right) - 3 \quad (13)$$

Outra abordagem aplicada na extração de medidas estatísticas está na utilização do histograma de níveis de cinza. O histograma é um vector com dimensão L , sendo L o nível de cinza máximo possível na imagem.

Com a equação seguinte, pode ser calculada a partir do histograma a energia (ene), cujo significado físico será explicado no ponto 3.1.2.2:

$$ene = \sum_{i=0}^{L_{\max}-1} h(i)^2 \quad (14)$$

A entropia (ent) representa o grau de incerteza ou de desordem da imagem, é uma medida positiva, que será igual a zero quando todos os pixéis da imagem têm a mesma intensidade, pelo contrário será máxima quando uma imagem contém a mesma quantidade de pixéis para todas as intensidades:

$$ent = - \sum_{i=0}^{L_{\max}-1} h(i) \cdot \log_2(h(i)) \quad (15)$$

A anisotropia (*ani*) representa uma medida de simetria dos valores da intensidade.

$$ani = \frac{\sum_{i=0}^k h(i) \cdot \log_2(h(i))}{ent} \quad (16)$$

Com k igual ao valor mais baixo tal que: $\sum_{i=0}^k h(i) \geq 0.5$

3.1.2. Matriz de Co-ocorrência

3.1.2.1. Definição e Cálculo

A Matriz de Co-ocorrência de Níveis de Cinza (Gray Level Co-Occurrence Matrix – GLCM) considera a relação entre dois pixéis, um chamado de *pixel referência* e o outro de *pixel vizinho*, por este motivo, as características extraídas desta matriz são designadas de estatísticas de segunda-ordem. O *pixel vizinho* escolhido pode ser vizinho em qualquer direção: por exemplo à direita, à esquerda, acima, abaixo, ou na diagonal de cada pixel referência. Também a vizinhança não precisa ser exatamente de 1 pixel, pode ser de 2, 3, ou qualquer valor. Cada pixel dentro da imagem torna-se o pixel referência, iniciando num canto e percorrendo a imagem toda.

A co-ocorrência, na sua forma geral, pode ser especificada por uma matriz de frequências absolutas $P(i,j,d,q)$, na qual dois elementos de textura vizinhos, separados por uma distância d em uma orientação q ocorrem na imagem, um com propriedade i e o outro com propriedade j . Por exemplo, para um relacionamento com um ângulo q igual a 0° , e vizinhança considerada de 1 pixel, tem-se: $P(i,j,1,0^\circ)$.

Noutras palavras, a GLCM pode ser definida como $P(i,j,d,q)$ onde: i e j são pixéis vizinhos de acordo com a distância d . Esta distância entre os pixéis i e j é analisada de acordo com uma direção q . Normalmente são usadas as direções $0, 45^\circ, 90$ e 135° , e também, a matriz obtida da média das matrizes das 4 direções anteriores.

Desta forma, a GLCM, obrigatoriamente quadrada, com dimensão igual ao L_{\max} , representa em cada elemento a_{ij} o número de vezes que ocorreu uma transição do

nível de cinza L_i para L_j , considerando a distância d entre os pixéis vizinhos i e j na direcção q .

Estas matrizes são normalmente normalizadas através da divisão de cada entrada da matriz pelo somatório de todos os elementos da matriz $P(i,j,d,q)$, obtendo-se a matriz normalizada $p(i,j,d,q)$.

Para melhor ilustrar o exposto, apresenta-se um exemplo simples do cálculo de uma GLCM, não-normalizada para melhor compreensão. Considerando uma imagem com 3x3 pixel e quatro níveis de cinza [0, 1, 2, 3], calcula-se a GLCM com distância igual a 1 pixel e para as direcções de 0° e 90° :

<table style="border-collapse: collapse;"> <tr><td>0</td><td>1</td><td>2</td></tr> <tr><td>3</td><td>3</td><td>0</td></tr> <tr><td>2</td><td>2</td><td>1</td></tr> </table>	0	1	2	3	3	0	2	2	1		<table style="border-collapse: collapse;"> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td><td>2</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>2</td><td>2</td><td>0</td></tr> <tr><td>3</td><td>1</td><td>0</td><td>0</td><td>2</td></tr> </table>		0	1	2	3	0	0	1	0	1	1	1	0	2	0	2	0	2	2	0	3	1	0	0	2		<table style="border-collapse: collapse;"> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>2</td><td>1</td><td>0</td><td>0</td><td>2</td></tr> <tr><td>3</td><td>1</td><td>1</td><td>2</td><td>0</td></tr> </table>		0	1	2	3	0	0	1	1	1	1	1	0	0	1	2	1	0	0	2	3	1	1	2	0
0	1	2																																																													
3	3	0																																																													
2	2	1																																																													
	0	1	2	3																																																											
0	0	1	0	1																																																											
1	1	0	2	0																																																											
2	0	2	2	0																																																											
3	1	0	0	2																																																											
	0	1	2	3																																																											
0	0	1	1	1																																																											
1	1	0	0	1																																																											
2	1	0	0	2																																																											
3	1	1	2	0																																																											
Imagem		$P(i, j, 1, 0^\circ)$		$P(i, j, 1, 90^\circ)$																																																											

Figura 11. Exemplo do cálculo de GLMC

Representa-se a seguir, a GLCM, calculada para as 4 direcções, de uma amostra do caso de estudo deste trabalho. Inverteu-se a imagem, isto é, os pontos brancos correspondem a zeros e os pontos pretos a uns, para facilitar a visualização e impressão das imagens:

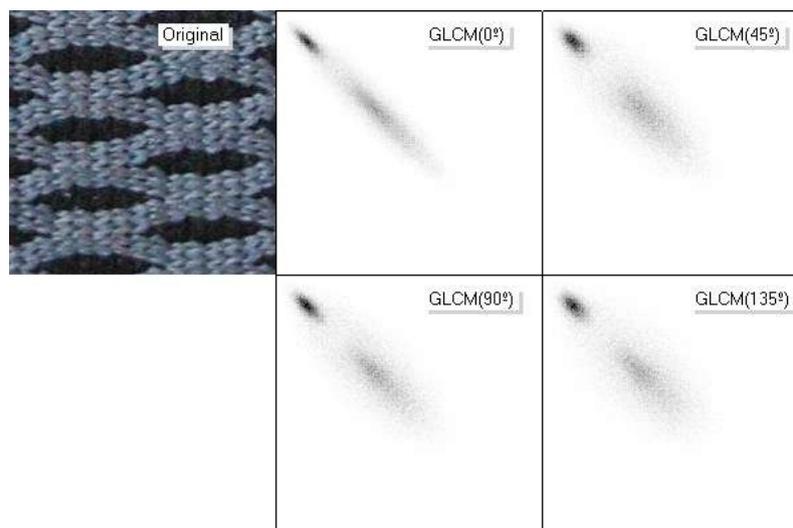


Figura 12. Representação da GLMC das 4 direcções para uma amostra

3.1.2.2. Características Extraídas da GLCM

Haralick et al. (1973) propuseram 14 descritores de textura baseados na GLCM. Segundo diversos autores, 6 destas medidas apresentam maior relevância na descrição da textura: energia, correlação, homogeneidade, contraste, entropia e variância. Neste trabalho usaram-se, a entropia e anisotropia, definidas no ponto 3.1.1, e as 4 medidas definidas a seguir.

A energia (*ene*) expressa a uniformidade de uma textura, isto é, reflecte a quantidade de repetições de pares de pixéis. Apresenta valores elevados em imagens com texturas muito ordenadas, nas quais, a energia, calculada através da equação (17), apresenta valores próximos de um:

$$0 < ene = \sum_{i=0}^{L_{\max}-1} \sum_{j=0}^{L_{\max}-1} p_{i,j}^2 < 1 \quad (17)$$

O contraste (*con*) caracteriza a diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza localizados numa região contígua da imagem. Em termos da GLCM, valores baixo de contraste, calculado através da equação (18), ocorrem quando há grande concentração de elementos que apresentam valores significativos em torno da diagonal principal.

$$con = \sum_{i=0}^{L_{\max}-1} \sum_{j=0}^{L_{\max}-1} (i - j)^2 \cdot p_{i,j} \quad (18)$$

A correlação (*cor*) mede a dependência linear entre os tons de cinza vizinhos presentes numa imagem. Valores altos calculados pela equação (20), próximos de 1, indicam a existência de relação linear entre os pares de níveis de cinza. Nessa equação, σ_x e σ_y indicam o desvio padrão das distribuições marginais obtidas com base nas equações (19), assim como μ_x e μ_y indicam o valor médio dessas distribuições.

$$var_i = \sum_{i=0}^{L_{max}-1} \sum_{j=0}^{L_{max}-1} (i - \mu_i)^2 \cdot p_{i,j} \quad (19)$$

$$var_j = \sum_{i=0}^{L_{max}-1} \sum_{j=0}^{L_{max}-1} (j - \mu_j)^2 \cdot p_{i,j}$$

$$-1 < cor = \frac{1}{\sigma_x \cdot \sigma_y} \sum_{i=0}^{L_{max}-1} \sum_{j=0}^{L_{max}-1} (i - \mu_i) \cdot (j - \mu_j) \cdot p_{i,j} < 1 \quad (20)$$

A homogeneidade (*hom*), calculada de acordo com a equação (21), apresenta valores elevados quando a textura apresenta pequenas variações de níveis de cinza entre pares de pixéis. Tal medida apresenta correlação inversa com as medidas de contraste e de energia (Ito et al, 2009).

$$hom = \sum_{i=0}^{L_{max}-1} \sum_{j=0}^{L_{max}-1} \frac{1}{1 + (i - j)^2} \cdot p_{i,j} \quad (21)$$

Limitações GLCM

As matrizes de co-ocorrência apresentam algumas limitações. Não existe nenhum método automático para a escolha da distância *d* e o cálculo das matrizes de co-ocorrência para diferentes valores de *d* tem um custo computacional elevado.

3.2. Abordagem Estrutural

A abordagem estrutural representa textura como sendo formada pela repetição de padrões que obedecem a alguma regra de posicionamento para a sua geração (Ballard e Brown, 1982). As abordagens estruturais são consideradas apropriadas quando a primitiva de textura é grande o suficiente para ser individualmente segmentada e descrita.

3.3. Abordagem Espectral

A abordagem espectral diz respeito a propriedades baseadas no domínio das frequências.

No espectro de Fourier, por exemplo, transforma-se a imagem do domínio espacial para o domínio da frequência, onde pode ser detectada a existência de padrões periódicos ou semi-periódicos (Kulkarni, 1995).

As transformadas *Wavelet* apresentam propriedades que proporcionam uma representação com informações espaciais e de frequência, além do aspecto de multi-resolução, que permite a extracção de descritores em diversas escalas.

Das transformadas de *Wavelet* resultam subimagens de onde poderão ser extraídas características texturais, permitem a análise multi-escala e apresentam uma ampla escolha de funções *wavelet*.

3.4. Outras Abordagens

Nos últimos anos surgiram diversas abordagens com resultados comprovados na descrição de texturas, destacando-se a LBP – *Local Binary Pattern*, e suas extensões (LBP/C, OCLBP,...) e os filtros de *Gabor*.

4. CLASSIFICAÇÃO DE IMAGENS

4.1. Introdução

Genericamente, o objectivo da classificação é seleccionar a classe mais apropriada, de entre um conjunto de classes conhecidas, para um objecto desconhecido. Como um sistema de classificação perfeita é por vezes impossível, a classificação pode também ser efectuada com base na probabilidade do objecto desconhecido pertencer a cada uma das classes possíveis. Outra abordagem é usar uma classe adicional, uma classe de rejeição, para objectos que não sejam possíveis de classificar com um grau de certeza suficiente.

Os objectos são processados com o objectivo de calcular as características que melhor definem as texturas em estudo. Essas características, representadas por números, reais em geral, são introduzidas num vector, denominado por vector de características.

Basicamente, o processo de classificação é realizado em duas fases: a fase de treino e a fase de classificação.

A fase de treino consiste em apresentar imagens de texturas cuja classe é conhecida. O classificador vai calcular as características de cada imagem e associar essas características à classe estipulada para essa imagem.

Depois da fase de treino, o classificador recebe imagens com classes desconhecidas. O classificador, com base nas características obtidas na fase de treino para cada classe, e nas características obtidas para a nova imagem, procede à classificação da imagem usando o método de classificação escolhido.

4.2. SVM - Máquinas de Suporte Vectorial

Com base nas características dos algoritmos disponíveis no software Halcon, usado para a aplicação do caso prático deste trabalho, optou-se por usar o algoritmo de Máquinas de Suporte Vectorial (SVM – Support Vector Machines). De acordo com as referências do Halcon, o algoritmo do SVM em relação ao algoritmo das Redes Neurais Artificiais (MLP – Multi-Layer Perceptrons) permite obter taxas de classificação

ligeiramente superiores, sendo mais rápido na fase de treino mas mais lento na fase de classificação.

Este algoritmo permite lidar com classes que não são linearmente separáveis. Os vectores de características são transformados num espaço de maior dimensão, obtendo-se hiper-planos não-lineares, por forma a conseguir separar linearmente as classes. Com esta transformação, o vector de características já é possível ser classificado usando um classificador linear.

Ilustra-se o exposto na Figura 13, com um exemplo simples, no qual 2 classes são definidas por um vector com 2 características. No espaço 2D não é possível definir uma linha recta que separe as duas classes. Adicionando uma terceira dimensão, as classes tornam-se separáveis por um plano.

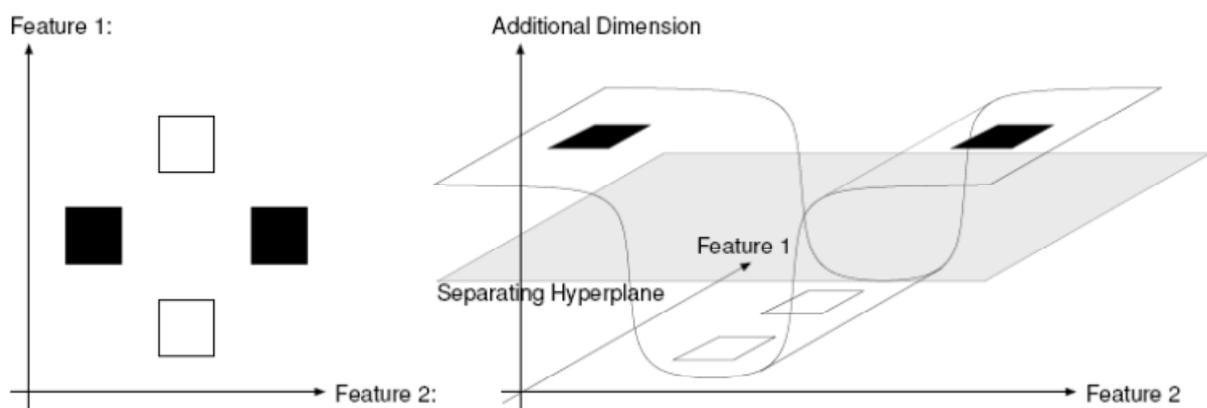


Figura 13. Transformação espacial do SVM (Halcon, Solution Guide II-D)

O hiper-plano de separação entre duas classes é construído de forma a maximizar a margem entre as duas classes. A margem é definida como a distância mais curta entre o hiper-plano de separação e qualquer amostra de treino. As amostras de treino que têm a distância mais curta até ao hiper-plano de separação são chamadas de vectores de suporte, como ilustrado na Figura 14:

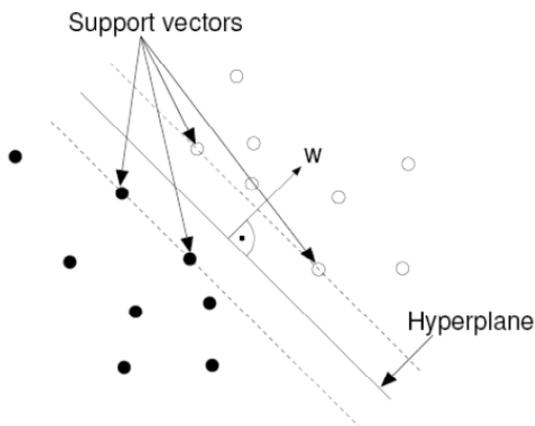


Figura 14. Vetores de Suporte do SVM (Halcon, Solution Guide II-D)

4.3. Matriz de Confusão

A matriz de confusão é uma técnica simples e bastante prática para analisar os resultados do algoritmo de classificação e ajudar na escolha ou rejeição de características das imagens para o processo de classificação, durante a fase de construção e otimização do algoritmo.

Consiste numa matriz quadrada com dimensão igual ao número de classes. No início do processo de classificação será constituída por zeros. Cada imagem classificada irá entrar nesta matriz na posição (Classe Correcta, Classe Classificada) ou vice-versa.

Desta forma, ter-se-á um algoritmo de classificação óptimo quando se obtiver uma matriz diagonal, ou o mais próximo disso possível.

Visualmente, será possível verificar as principais classes que estão a ser mal classificadas, se existem correlações entre 2 classes mal classificadas, etc

Veja-se o seguinte exemplo:

	1	2	3	4	5	6	7	8
1	136	0	0	0	0	0	0	0
2	0	133	0	4	2	4	1	0
3	3	6	105	16	14	2	12	0
4	0	1	0	118	0	2	0	2
5	0	0	2	1	121	0	2	0
6	0	0	1	0	0	130	0	0
7	1	0	32	1	3	2	125	0
8	0	0	0	0	0	0	0	138

Figura 15. Exemplo de uma matriz de confusão

O somatório das colunas (ou linhas, dependendo da forma de construir a matriz) será igual ao número de amostras classificadas em cada classe, neste caso, 140 por cada classe.

Nesta simulação, verifica-se rapidamente que a classe 3 é a classe mais problemática e que há uma correlação na má classificação da classe 3 com a classe 7. Com esta informação pode analisar-se as classes 3 e 7, definir as principais características diferenciadoras entre elas e introduzi-las, tanto quanto possível, no vector de características da próxima simulação, com vista à melhoria das taxas de classificação.

5. CASO PRÁTICO

5.1. Introdução

Pretende-se conceber um algoritmo de classificação de 8 tipos de peças da indústria automóvel. Estas peças são componentes de portas de automóvel. A dimensão das portas e a disposição da linha de montagem não permitem ter condições de luz controlada na captura de imagens das peças a classificar. A classificação é feita em tempo real na linha de montagem da fábrica, logo o algoritmo a ser implementado terá de ser rápido, na sua fase de classificação.

Os 8 tecidos a serem classificados são apresentados na Figura 16:



Figura 16. Classes de tecidos do caso de estudo

A empresa é responsável pelo fabrico da porta completa que contém as peças de tecido a classificar. Com base nessa classificação, entre outras características, a porta é marcada com uma referência. É usado o sistema de produção *just in time* (JIT), fornecendo a porta completa ao cliente, Grupo PSA em Vigo. Este grupo recebe, de outros fornecedores, outras peças, algumas das quais também contêm os mesmos tecidos aqui analisados. Na linha de montagem deste grupo, a montagem de todas as peças é feita com base nas referências de cada uma delas. Um erro causado, por exemplo, pela

incorrecta classificação do tecido, implicaria a montagem de peças com tecidos diferentes no mesmo veículo, acarretando perdas significativas de produtividade.

Neste momento, o sistema de visão e classificação encontra-se implementado nesta fábrica, mas apresenta taxas de classificação insatisfatórias.

5.2. Metodologia de Ensaio

5.2.1. Amostras de Imagens

Foram tiradas 70 imagens de cada tipo de tecido, com uma máquina fotográfica EOS 300D, com tripé, mantendo-se constantes a distância focal da máquina e distância entre a câmara e as amostras.

Estas 70 imagens foram tiradas em 6 condições diferentes, conforme descrito na Tabela 1, com o objectivo de introduzir a variabilidade da intensidade e direcção da luz incidente nas imagens, tal como acontece em ambiente industrial.

De cada uma dessas imagens foram criadas 10 amostras com 200x200 pixéis, perfazendo assim 700 amostras de cada tipo de tecido.

#	Nº de Amostras	Descrição
01	200	Dia N, 12h
02	100	Dia N+1, 12h
03	100	Dia N+1, 19h
04	100	Dia N+2, 10h
05	100	Dia N+3, 10h Dentro deste conjunto de imagens variou-se a distância (0-4 metros) e orientação da peça à entrada de luz natural na divisão.
06	100	Dia N+4, 12h Neste conjunto de imagens utilizou-se luz artificial e variou-se a quantidade de luz natural incidente nas peças.

Tabela 1. Conjuntos de amostras por cada classe do caso prático

A variabilidade das imagens deve ser o mais representativa possível, isso é essencial para o sucesso do classificador. Na Figura 17 apresentam-se alguns exemplos dessa variabilidade nas amostras utilizadas nesta simulação.



Figura 17. Variabilidade das Amostras

Nesta figura verifica-se a influência das condições de luz na captação da imagem. A maioria das imagens das 2 últimas colunas pertence aos dois últimos conjuntos de amostras.

No conjunto 5 aumentou-se progressivamente a distância da amostra à janela, bem como a orientação da peça relativamente à mesma. No conjunto 6 utilizou-se luz artificial e variou-se a luz natural para valores mínimos.

Tentou-se, desta forma, simular as variabilidades de iluminação da fábrica, a qual opera 24 horas por dia.

5.2.2. Software Halcon

Usou-se o software comercial Halcon da MVTec, para realizar o estudo e conceber o algoritmo de classificação. Este software tem disponível quase todos os algoritmos de visão computacional e permite também a criação de rotinas por parte do usuário.

A utilização de um software comercial tem como principais vantagens a rapidez e estabilidade de implementação dos algoritmos. A opção por este tipo de plataforma deveu-se não só a requisitos temporais, como ao carácter essencialmente prático da presente tese, que se enquadra no domínio da engenharia e gestão industrial.

5.2.3. Classificação

Foi utilizado o método de classificação SVM.

Para testar e otimizar os parâmetros da simulação desenvolveu-se uma rotina de classificação de teste. Nesta rotina, dividiram-se as imagens em 5 conjuntos de igual tamanho. A classificação foi realizada em 5 ciclos. Em cada ciclo, de forma rotacional, 4 conjuntos de imagens são utilizados para treinar e o outro conjunto é usada para testar a classificação. Desta forma, as imagens a serem classificadas não entram na fase de treino. Usando as 700 amostras por cada tipo de classe, cada simulação realizada usa 22 400 amostras (2 800 por cada classe) para a fase de treino e classifica 5600 amostras (700 por cada classe).

O algoritmo de classificação SVM do Halcon tem dois parâmetros: tipo de núcleo e um parâmetro que regulariza a separação das classes. Para determinar o par óptimo destes valores foi feita uma simulação de classificação com dois ciclos exteriores, nos quais se variaram estes dois parâmetros dentro do intervalo recomendando pelo Halcon.

5.2.4. Processamento das Imagens

Tal como já explanado no ponto 2.2, existem muitas alternativas de processamento de imagens de forma a, por um lado, reduzir os efeitos da iluminação não controlada e, por outro, para realçar as características texturais das imagens. Além da aplicação isolada de cada técnica de pré-processamento, temos ainda a possibilidade da aplicação sequencial de 2 ou mais técnicas de manipulação de imagens. Isso torna o processo de parametrização e optimização do pré-processamento difícil, pois o número de combinações possíveis é bastante elevado.

5.3. Resultados

Neste capítulo apresentam-se a descrição e resultados das principais simulações realizadas. Ao longo deste trabalho, para além de se procurar obter a solução concreta para este problema, esteve sempre presente seguir uma metodologia aplicável a outros problemas semelhantes a este caso prático. Por esse motivo, a metodologia utilizada será explicada da forma mais simples e prática possível.

Para melhor leitura e comparação dos resultados usa-se a simbologia de cada característica, como descrito anteriormente. Relativamente às matrizes de co-ocorrência, adopta-se a notação de GLCM para a GLCM média (média das 4 direcções), de GLCM(0°) para a direcção horizontal, etc

Usa-se GR como simbologia para indicar a imagem monocromática convertida a partir da imagem clorida em RBG.

5.3.1. Modelos de Cores e Seus Histogramas

Começou-se por testar as características mais simples, relativas aos histogramas dos canais de cor, em 3 modelos de cor: RGB, HSI e HSV. A Tabela 2 descreve cada simulação e o respectivo resultado:

#	Descrição	% Sucesso
1	Total de 18 Características: 3 Histogramas RGB: $\mu, \sigma, s, k, ene, ent$	87.1%
2	Total de 18 Características: 3 Histogramas HSI: $\mu, \sigma, s, k, ene, ent$	83.6%
3	Total de 18 Características: 3 Histogramas HSV: $\mu, \sigma, s, k, ene, ent$	82.4%

Tabela 2. Resultados do primeiro grupo de simulações

Em cada simulação separaram-se os 3 canais de cor. Nas 2ª e 3ª simulações, converteu-se o modelo de cor RGB para HSI e HSV, respectivamente, e calculou-se o histograma de cada um dos canais e as medidas estatísticas da média, desvio padrão, assimetria, curtose, energia e entropia de cada um dos histogramas.

Muitos investigadores da área de visão computacional recomendam o uso do modelo de cores HSI, pela separabilidade entre a cor e a intensidade (ver 2.1.2), no entanto, neste caso prático isso não se verificou. As taxas de sucesso foram superiores usando o modelo RGB. Por esse motivo, decidiu-se usar este modelo nas simulações subsequentes.

5.3.2. Filtros e Matriz de Co-Ocorrência

Nesta fase introduziram-se os filtros de Laws e as respectivas matrizes de co-ocorrência e calcularam-se 6 medidas estatísticas por cada GLCM.

#	Descrição	% Sucesso
4	Total de 30 Características: 3 Histogramas RGB: $\mu, \sigma, s, k, e, ent$ FL(le)+GLCM: <i>ene, cor, hom, con, ent, ani</i> FL(ls)+GLCM: <i>ene, cor, hom, con, ent, ani</i>	91.3%

Tabela 3. Resultado óptimo do segundo grupo de simulações

Os filtros, por exemplo, têm vários tamanhos de máscara, vários tipos de máscaras e possibilitam a aplicação de diversas combinações de filtros de forma sequencial. Simularam-se vários tipos de filtros e várias sequências de filtros para se

encontrarem os melhores filtros para o nosso caso de estudo. Numa primeira fase, aplicaram-se todos os tipos de filtros e variaram-se os tamanhos das máscaras, concluindo-se que o tamanho óptimo da máscara é de 3x3. Numa segunda fase fixou-se esse tamanho da máscara e eliminaram-se, sucessivamente, os filtros que não contribuíam positivamente para o sucesso da classificação. Com base nestas simulações decidiu-se usar uma máscara com o tamanho de 3x3 pixel e os filtros do tipo “le” e “ls”.

Como se pode observar na Tabela 3, os resultados melhoraram consideravelmente com a introdução destes filtros e das matrizes de co-ocorrência.

No final desta fase optou-se, também, por retirar as características de assimetria (*s*), curtose (*k*), energia (*ene*) e entropia (*ent*) do histograma, dada a morosidade no cálculo destas relativamente ao benefício da sua incorporação na classificação. Este cálculo foi efectuado numa rotina criada de raiz, encontrando-se, nesta fase, pouco optimizada.

5.3.3. Conversão Escala de Cinzas e Filtro de Sobel

Usando o mesmo procedimento adotado em Ojala et al (1996), converteu-se a imagem inicial em RGB para uma imagem de níveis de intensidade de cinza (monocromática). Isto permitiu uma simplificação do processamento posterior da imagem e da análise dos seus resultados. Além disso, introduziu-se a equalização do histograma obtido e aplicou-se um filtro gaussiano, na fase de pré-processamento.

#	Descrição	% Sucesso
5	<p>Total de 38 Características:</p> <p>Pré-processamento: conversão p/imagem monocromática (GR), equalização e filtro gaussiano.</p> <p>Histograma (GR): μ, σ</p> <p>FL(le)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FL(ls)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (0°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (45°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (90°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (135°): <i>ene, cor, hom, con, ent, ani</i></p>	98.9%

Tabela 4. Resultados do terceiro grupo de simulações

Acrescentou-se o Filtro de Sobel (FS) para realçar as linhas e detalhes das texturas, aplicando-se a GLCM em cada uma das 4 direcções à imagem obtida.

Como se pode observar na Tabela 4, a redução dos 3 canais de cor para uma imagem com um canal de níveis de cinza e a utilização do Filtro de Sobel permitiram um aumento significativo nos resultados.

5.3.4. Análise Piramidal

Nesta fase introduziu-se a análise piramidal, descrita no ponto 2.1.5., em que o conjunto de características é calculada para cada uma dessas imagens e adicionado ao vector de características. Testaram-se vários níveis de análise piramidal e obteve-se o valor óptimo de 4 níveis. De notar que apesar do vector de características quadruplicar de tamanho em relação ao anterior, o classificador mantém-se bastante rápido, classificando cada imagem na ordem das décimas de segundo.

#	Descrição	% Sucesso
6	<p>Total de 152 Características,</p> <p>38 características por cada um dos 4 níveis piramidais:</p> <p>Pré-processamento: conversão p/imagem monocromática (GR), equalização e filtro gaussiano.</p> <p>Histograma (GR): μ, σ</p> <p>FL(le)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FL(ls)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (0°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (45°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (90°): <i>ene, cor, hom, con, ent, ani</i></p> <p>FS+GLCM (135°): <i>ene, cor, hom, con, ent, ani</i></p>	99.7%

Tabela 5. Resultados do quarto grupo de simulações

5.3.5. Uso da Matriz de Confusão

Neste ponto pretende-se dar um exemplo prático do uso da matriz de confusão.

A matriz de confusão da última simulação apresentada (simulação #6) foi a seguinte:

	1	2	3	4	5	6	7	8
1	700	0	0	0	0	0	0	0
2	0	700	0	0	0	0	0	0
3	0	0	700	0	0	0	0	0
4	0	0	0	700	0	0	0	0
5	0	0	0	0	700	0	0	5
6	0	0	0	0	0	700	0	0
7	0	0	0	0	0	0	700	0
8	0	0	0	0	0	0	0	695

Figura 18. Exemplo de utilização da matriz de confusão

Os 5 resultados incorrectos foram todos da classe 8 e todos “confundidos” com a classe 5. Com a introdução dos parâmetros da média e desvio-padrão dos 3 canais de cor RGB, conseguiu-se uma taxa de classificação de 100%.

Durante esta ronda de corridas, experimentou-se com sucesso, usar a entropia e anisotropia para a imagem filtrada com FS, retirando-se o cálculo das mesmas em relação às GLCM para as 4 direcções.

#	Descrição	% Sucesso
7	<p>Total de 152 Características,</p> <p>38 características por cada um dos 4 níveis piramidais:</p> <p>3 Histogramas RGB: μ, σ</p> <p>Pré-processamento: conversão p/imagem monocromática (GR), equalização e filtro gaussiano.</p> <p>Histograma (GR): μ, σ</p> <p>FL(le)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FL(ls)+GLCM: <i>ene, cor, hom, con, ent, ani</i></p> <p>FS: <i>ent, ani</i></p> <p>FS+GLCM (0°): <i>ene, cor, hom, con</i></p> <p>FS+GLCM (45°): <i>ene, cor, hom, con</i></p> <p>FS+GLCM (90°): <i>ene, cor, hom, con</i></p> <p>FS+GLCM (135°): <i>ene, cor, hom, con</i></p>	100%

Tabela 6. Resultados do quinto grupo de simulações

5.3.6. Representatividade da Amostragem

Realizaram-se algumas simulações onde se utilizaram, como imagens de treino amostras dos primeiros 5 conjuntos (ver 5.2.1) e classificando o conjunto 6. Os resultados baixaram para menos de 90% de taxa de eficácia. Este resultado demonstra a importância de treinar o classificador com uma amostra de imagens o mais representativa possível das variabilidades possíveis de cada classe.

5.3.7. Tempos de Simulação

Na tabela seguinte apresentam-se os tempos de simulação para duas simulações, uma só com 6 características e outra, a solução final deste caso de estudo, com 152 características utilizadas no classificador:

	Tempo de Treino com 22400 amostras	Tempo de Classificação por cada imagem
6 Características	2.64 min	0.017 s
152 Características	18.8 min	0.046 s

Tabela 7. Tempos de Simulação

A simulação com 6 características necessita de bastante menos tempo, relativamente à simulação com 152 características, principalmente ao nível do tempo de treino. A proporção entre o tempo necessário e o número de características não é proporcional.

A principal restrição do caso de estudo é a exigência da rapidez na fase de classificação, pois a fase de treino poderá ser feita *offline*, isto é, sem a linha de montagem parar. O treino pode ser feito, guardado em disco ou memória, para ser lido posteriormente, para a fase de classificação.

O algoritmo de classificação da solução final, com 152 características, necessita de somente 0.046 segundos para processar e classificar uma imagem, tempo esse que é manifestamente menor do que o tempo exigido pela linha de montagem.

6. CONCLUSÕES

Resumem-se as principais conclusões obtidas com este trabalho:

- **Classificador SVM**

Provou-se a rapidez e eficácia deste algoritmo. Neste trabalho não foi necessário o estudo e comparação com outros algoritmos de classificação. No entanto, caso o tempo de classificação não fosse suficientemente baixo, estudar-se-ia também o algoritmo MLP, muito usado actualmente nesta área.

- **Seleção de características de textura**

Em primeiro lugar, constatou-se que usando somente as características básicas dos histogramas dos 3 canais RGB, conseguiu-se uma taxa de classificação de quase 90%.

Acrescentando os filtros de Laws e Sobel e calculando as respectivas matrizes de co-ocorrência, foi possível atingir os 100% de sucesso. Comprovou-se a grande eficácia, por um lado, dos filtros para realçar características texturais, e, por outro, da utilização das matrizes de co-ocorrência e a extracção das suas principais características.

- **Metodologia**

Utilizou-se uma metodologia não-exaustiva, mas simples, com o objectivo de apresentar um meio para resolver o caso prático deste trabalho, mas também tendo em vista a sua aplicação a outros casos semelhantes.

- **Representatividade das amostras**

Uma boa representatividade das amostras é essencial para o sucesso do algoritmo de classificação. Treinando com os 5 primeiros conjuntos de amostras e classificando o 6º conjunto, os resultados baixaram para os 90%.

- **Software comercial - Halcon**

Provou-se ser uma ferramenta muito prática e robusta. Tendo em vista o carácter da aplicação industrial deste trabalho, da complexidade da área de análise de texturas e da limitação de tempo, esta ferramenta foi essencial para atingir os objectivos propostos para este trabalho.

Pensando no segundo objectivo para este trabalho, a metodologia proposta poderia ser desenvolvida com a utilização de outras abordagens de análise de texturas.

A utilização de abordagens como o LBP e suas extensões, a utilização de transformadas de Fourier, Filtros de *Gabor*, transformadas de *Wavelet* e matrizes de co-ocorrência tridimensionais poderiam ser estudados e utilizados nas situações em que os métodos apresentados neste trabalho não atinjam resultados satisfatórios. Estes métodos têm sido muito estudados e utilizados com sucesso na área de classificação e segmentação de texturas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

Ballard, D. H. e Brown, C. M. (1982), “Computer Vision”, New Jersey, Prentice Hall.

Bur, P. J. e Adelson, E. H. (1983), “The Laplacian Pyramid as a Compact Image Code”, IEEE Trans. Commun., vol. COM-31, n.4.

Gonzalez, R. C. e Woods, R. E. (2002), “Digital Image Processing”, 3ª Edição, New Jersey, Prentice Hall.

Halcon (2010) “Solution Guide II-D: Classification”, MVTek Software GmbH, Germany.

Haralick, R. M. e Shanmugam, K. e Dinstein, I. (1973) “Textural Features for Image Classification”, IEEE Transactions on Systems, Man and Cybernetics.

Ito, R.H., Kim, H.Y., Salcedo, W. J. (2009), “Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência,” 8th International Information and Telecommunication Technologies Symposium.

Jain, A. K. (1989) “Fundamentals of Digital Image Processing”, New Jersey, Prentice-Hall.

Laws, K. I. (1980) “Textured image segmentation”, Report 940, Image Processing Institute, Univ. of Southern California.

Ojala, T., Pietikäinen, M. e Nisula, J. (1996), “Determining composition of grain mixtures by texture classification based on feature distributions”, International Journal of Pattern Recognition and Artificial Intelligence.

Palm, C. (1994), “Color Texture Classification by Integrative Co-occurrence Matrices”, Patter Recognition.

Pedrini, H. e Schwartz, W. R (2008), “Análise de imagens digitais: princípios, algoritmos e aplicações”, São Paulo, Thomson Learning.

Kulkarni, A. D. (1995), “Artificial Neural Networks for Image Understanding”, New York, Van Nostrand Reinhold.

Schwartz, W. R e Pedrini, H. (2007), “Color Textured Image Segmentation Based on Spatial Dependence Using 3D Co-occurrence Matrices and Markov Random Fields”, International Conference in Central Europe on Computer Graphics,

Visualization and Computer Vision'2007.

Silva, R. D. (2008), “Segmentação de Imagens Baseada em Descritores de Textura Extraídos por Transformadas Wavelets”, Curitiba.

Theodoridis, S. e Koutroumbas, K. (2003), “Pattern Recognition”, 2nd Edition, Elsevier Academic Press.

Vácha, P., Haindl, M., Suk, T. (2011), “Colour and rotation invariante textural features based on Markov random fields”, Pattern Recognition, Elsevier.