

Comparação de Descritores de Formas no Reconhecimento de Objetos

Jonathan de Andrade Silva, Wesley Nunes Gonçalves, Bruno Brandoli Machado,
Hemerson Pistori, Albert Schiaveto de Souza
UCDB - Universidade Católica Dom Bosco
Grupo de Pesquisa em Engenharia e Computação
Av. Tamandaré, 6000, Jardim Seminário, 79117-900, Campo Grande, MS, Brasil
{jsilva, wnunes, bmachado}@acad.ucdb.br, {pistori,albert}@ucdb.br

Resumo

Os descritores de formas são amplamente utilizados para descrever ou representar objetos em imagens. Neste trabalho é realizado um estudo comparativo entre os descritores de formas baseados em contorno e região. Na abordagem por contorno, foram avaliadas as cadeias de códigos, k-curvatura e descritores de Fourier, e na abordagem por região utilizou-se momentos de imagem. A classificação é baseada pela extração de características das formas. Estes descritores foram aplicados ao problema de identificação do comportamento de camundongos. A avaliação foi realizada através da variação dos parâmetros de cada técnica e na acurácia dos classificadores, alcançando 92.6% de acerto a técnica momentos de imagem.

1. Introdução

O reconhecimento de padrões é um processo para tomada de decisão com o objetivo de classificar os objetos em um número de categorias ou classes. A capacidade de reconhecer padrões em uma imagem depende muito da quantidade de informações que se conhece de cada objeto [8]. Com isso, extrair características para descrever e representar os objetos é uma etapa fundamental em qualquer processo de reconhecimento de objetos. Há diferentes métodos para extrair informações de objetos em imagens, como os descritores de formas, que utilizam o contorno ou a região do objeto como atributo principal.

Um importante problema que vem sendo estudado por diversos grupos de pesquisa em visão computacional é a capacidade de reconhecer através de imagens os comportamentos de animais em experimentos na área da saúde. A observação dos comportamentos dos animais é importante para descobrir informações relacionadas, como a trajetória e o tipo dos movimentos realizados. Através dessas informações é possível identificar o comportamento desses

animais quando aplicada alguma droga ou realizada alguma modificação genética. Um experimento muito utilizado é o teste do campo aberto [11]. Este experimento é geralmente realizado com ratos e camundongos, que são colocados em uma arena e seu comportamento é analisado para que possam ser identificadas mudanças significativas.

Neste trabalho, foram utilizados dois principais grupos de técnicas para a descrição de formas, o primeiro baseado nas características das regiões e o segundo nas características do contorno do objeto. Entre as técnicas de descrição por região foi avaliada a técnica de momentos de imagem, enquanto na abordagem por representação do contorno do objeto foram realizados experimentos com os algoritmos de cadeia de códigos, k-curvatura e descritores de Fourier. As comparações foram realizadas com base em imagens reais de três comportamentos básicos dos camundongos, comportamento de auto limpeza, exploração vertical e horizontal. Para o auxílio na classificação dos comportamentos realizados pelos camundongos foi utilizado o *software* Weka¹. Para cada técnica foi escolhida a melhor configuração de parâmetros de acordo com a porcentagem de acerto dos classificadores.

Este artigo está estruturado em cinco seções. A Seção 2 cita os trabalhos correlatos de descritores de formas. Na Seção 3, são definidos os descritores de formas. Na próxima seção, são discutidas as ferramentas utilizadas e a realização dos experimentos. Na Seção 5, são analisados os resultados das técnicas implementadas. Finalmente, são apresentados a conclusão e os trabalhos futuros.

2. Trabalhos Correlatos

A representação de objetos é um importante passo para o reconhecimento de padrões. Diversas características e algoritmos de representação têm sido analisados e desenvol-

¹ O Weka é uma ferramenta livre que possui um conjunto de algoritmos de aprendizagem de máquina. <http://www.cs.waikato.ac.nz/ml/weka/>

vidos nos últimos anos, sendo o contorno uma das características mais utilizadas na representação dos objetos. Em [9] foi desenvolvido um método para o reconhecimento de objetos irregulares baseado em seu contorno. O contorno do objeto é descrito por cadeia de códigos baseada em histograma, sendo invariante a rotação e translação. O método foi avaliado no reconhecimento de letras e em objetos irregulares com superfícies defeituosas.

Wu [15] descreve um método de representação utilizando os pontos predominantes de cada objeto. Os pontos predominantes são aqueles onde a curvatura é extrema. Segundo Wu, a detecção desses pontos é uma importante etapa para representação de formas. O método apresentado utiliza cadeia de códigos para representar o contorno e o algoritmo K-cosseno como medida de curvatura. O método foi avaliado em contornos de folhas utilizando cinco critérios de avaliação: número de pontos predominantes, taxa de compressão, erro médio da área, erro médio quadrado e tempo de processamento.

Em [14] o algoritmo k-curvatura foi utilizado para identificar os dedos das mãos para o reconhecimento de gestos e outras aplicações de interação homem-máquina. Para avaliar o método, foi considerada a porcentagem de acertos de dedos identificados e do reconhecimento de 4 gestos (mão fechada, mão aberta, mão com um dedo estendido e mão com dois dedos estendidos). O resultado do reconhecimento de dedos foi em média de 97.6% em 6971 imagens. Para o reconhecimento de gestos o resultado foi ainda melhor, em média 99.39% de acertos avaliados sobre 4745 imagens.

Uma outra representação extensamente empregada são os descritores de Fourier. Em [16] os descritores de Fourier com diferentes representações de formas foram utilizados para recuperação de imagens através do contorno. A avaliação foi realizada através da taxa de formas relevantes retornadas pelo número de formas buscadas no banco de dados. Com isso, o melhor resultado foi alcançado utilizando os descritores de Fourier derivados da distância entre cada ponto e o centro do objeto. Já em [13], os descritores de Fourier, juntamente com redes neurais, foram utilizados na inspeção industrial. Os experimentos foram realizados em dois produtos e o treinamento foi realizado sobre 480 amostras com produtos aprovados e reprovados na inspeção. Na fase de reconhecimento foram utilizadas 200 amostras. Em ambos os experimentos com os produtos, a porcentagem de acertos foi máxima, utilizando 15 descritores para a representação da forma dos objetos.

A modelagem da variação da forma é uma importante característica para o reconhecimento de objetos. Liu [10] apresentou um método para modelar as mudanças nas formas 2D dos objetos. O método representa a forma do objeto utilizando descritores de Fourier e as mudanças utilizando modelos de auto-regressão. O método foi avaliado em três dife-

rentes aplicações. Na primeira aplicação, o método é avaliado na detecção de contornos de fogo, alcançando uma taxa de 99.9% de acerto. A segunda aplicação demonstra que a variação da forma sozinha é uma boa característica para discriminação. Os experimentos são realizados utilizando o método *cross-validation* em sete seqüências de pessoas correndo, bandeiras em movimento e seqüências de fogo. A terceira e última aplicação é utilizada em rastreamento de contornos. Uma limitação apresentada pelo método é que o modelo auto-regressivo (AR) é um sistema linear e pode ocorrer casos onde a linearidade não é suficiente para representar a variação da forma de um objeto deformável.

Em [1] a técnica de cadeia de códigos é utilizada para representar as formas geométricas de quadrados e retângulos em uma imagem, utilizando o conceito de conectividade-de-4 para representar quadrados e conectividade-de-8 para representar triângulos. Em [7] classifica-se os caracteres escritos à mão baseados em uma decomposição direcional da representação correspondente da cadeia de códigos de conectividade-de-8. O classificador é treinado pela cadeia de códigos que representa o contorno do caractere.

3. Descritores de Formas

As imagens são constituídas de diversos objetos, pertencentes ao plano de fundo e ao plano de interesse. Os objetos do plano de interesse são, geralmente, os que possuem maior importância para os sistemas de visão computacional. Esses objetos são constituídos de um conjunto de *pixels* representando a sua forma, cor e textura.

Os descritores de formas são métodos matemáticos que representam a forma do objeto ou a região. Os descritores são divididos em duas categorias, baseados no contorno do objeto e outros baseados em regiões. Os descritores baseados em contorno descrevem a forma do objeto considerando suas formas mais externas, seu contorno. Descritores baseados em região descrevem as formas detalhadas do objeto considerando suas formas internas. Os descritores ideais devem possuir invariança a translação, rotação, escala e a ponto de início.

3.1. Momentos de Imagens

Os momentos [2] de ordem $p + q$ sobre uma função contínua 2D são calculados de acordo com a Equação 1.

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (1)$$

Uma imagem pode ser visualizada como uma função 2D discreta I , onde a intensidade de cada *pixel* é indexada como $I(x, y)$. A Equação 2 é a derivação da Equação 1 de momentos aplicada a imagens.

$$M_{pq} = \sum_{x=1}^{width} \sum_{y=1}^{height} x^p y^q I(x, y) \quad (2)$$

Para a representação de um objeto através de momentos são considerados apenas *pixels* pertencentes ao objeto. Através dos momentos de imagens, aplicados a um objeto, é possível derivar algumas propriedades, como por exemplo, a área do objeto, que é representada pelo momento M_{00} e o centro de massa \bar{x} e \bar{y} representado pelos momentos $\frac{M_{10}}{M_{00}}$, $\frac{M_{01}}{M_{00}}$, respectivamente.

Além dos momentos regulares, é possível derivar os momentos centrais para imagens, definido na Equação 3.

$$u_{pq} = \sum_{x=1}^{width} \sum_{y=1}^{height} (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (3)$$

As propriedades derivadas dos momentos centrais neste trabalho são a variância em X e Y (Equação 4), a direção do objeto (Equação 5) e a excentricidade (Equação 6).

$$\sigma_x^2 = \frac{u_{20}}{m_{00}}, \quad \sigma_y^2 = \frac{u_{02}}{m_{00}} \quad (4)$$

$$\theta = \frac{u_{02} - u_{20} - 2u_{11} + \lambda}{u_{02} - u_{20} + 2u_{11} - \lambda} \quad (5)$$

$$e^2 = \frac{u_{20} + u_{02} + \lambda}{u_{20} + u_{02} - \lambda} \quad (6)$$

$$\lambda = \sqrt{(u_{20} - u_{02})^2 + 4u_{11}^2} \quad (7)$$

Na Figura 1 é ilustrada a aplicação dos momentos em uma imagem contendo um camundongo. A imagem foi dividida em quatro quadrantes com a mesma área, em seguida, para cada quadrante são extraídos o centro de massa e o desvio, ilustrados como eixos nas cores azul, laranja e verde. Além desses parâmetros, para a imagem como um todo, é extraída a elipse que representa o objeto.

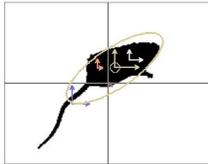
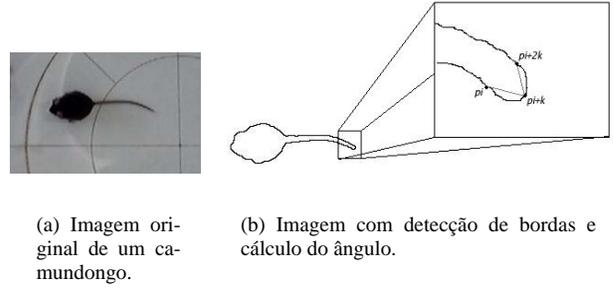


Figura 1. Exemplo da aplicação dos momentos de imagem com quatro quadrantes, duas linhas e duas colunas.



(a) Imagem original de um camundongo.

(b) Imagem com detecção de bordas e cálculo do ângulo.

Figura 2. Passos realizados para a extração de atributos com k-curvatura.

3.2. K-Curvatura

O extrator k-curvatura representa o contorno de um objeto através da relação do ângulo formado entre dois vetores. Com a aplicação de um detector de bordas na imagem segmentada é possível adquirir um conjunto de pontos (Equação 8) que representam o contorno, como mostra Figura 2(b). Esse conjunto é adquirido percorrendo-se os pontos em um sentido (e.g. horário e anti-horário) a partir de um ponto inicial qualquer.

$$P = \{p_0, p_1, \dots, p_n\} \quad \text{com} \quad p_i = (x_i, y_i) \quad (8)$$

Para eliminação de ruídos do contorno, não são considerados todos os pontos extraídos do objeto [14], sendo escolhidos pontos após um espaçamento de k valores. O cálculo do ângulo é realizado entre três pontos p_i , p_{i+k} e p_{i+2k} como mostra a Figura 2(b). O vetor v é formado pelos pontos p_i e p_{i+k} , enquanto o vetor w é formado pelos pontos p_{i+k} e p_{i+2k} . A Equação 9 mostra o cálculo do ângulo entre dois vetores.

$$\theta = \cos^{-1} \frac{v \cdot w}{|v||w|} \quad (9)$$

onde $v \cdot w$ é o produto escalar entre dois vetores (Equação 10) e $|v|$ é a norma de um vetor (Equação 11).

$$v \cdot w = v_1 w_1 + v_2 w_2 + \dots + v_n w_n \quad (10)$$

$$|v| = \sqrt{v \cdot v} \quad (11)$$

Após os cálculos de todos os ângulos do contorno é construído um histograma, onde cada posição i desse histograma corresponde à frequência de uma faixa de ângulos encontrado no contorno. Através dessa composição por histograma, o extrator k-curvatura é invariante à rotação e translação. Para que o k-curvatura seja invariante à escala é necessário que o k seja atualizado de acordo com a escala aplicada.

Como exemplo, considere o contorno de um objeto constituído pelos ângulos $0^\circ, 82^\circ, 33^\circ, 160^\circ, 102^\circ, 170^\circ, 179^\circ, 131^\circ, 20^\circ, 49^\circ, 51^\circ$ entre os pontos com cada faixa representado um intervalo de 50° para o histograma. O histograma que representa o exemplo acima é ilustrado na Figura 3.

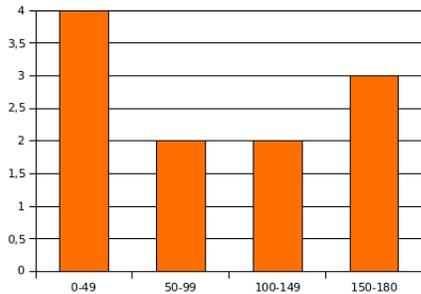


Figura 3. Histograma de ângulos extraídos pelo k-curvatura.

3.3. Cadeias de Códigos

As cadeias de códigos (*chain codes*) foram inicialmente introduzidas por Freeman [4, 12]. Este método de extração de atributos é utilizado para representar o contorno de objetos em uma imagem através de uma seqüência de símbolos com determinado tamanho e direção [3]. As primitivas de conectividade-de-4 ou de-8 segmentos são, geralmente, utilizadas como símbolos e direção. As primitivas são ilustradas na Figura 4.

A principal vantagem da cadeia de códigos é a flexibilidade em descrever um conjunto amplo de objetos. Em contrapartida, qualquer pequena perturbação ao longo do contorno, devido a ruído ou mudanças causadas pela segmentação defeituosa, pode não ser corretamente relacionada à forma do objeto.

A partir da cadeia de códigos é possível construir um histograma de cadeia de códigos (*Chain Code Histogram - CCH*) [9]. O CHH é descrito pela Função discreta 12.

$$h(k) = \frac{n_k}{n} \quad (12)$$

onde n_k é o número de códigos de direção k e n é o tamanho da cadeia de códigos. Com isso, o CCH mostra a porcentagem das diferentes direções presentes no contorno do objeto. O exemplo de um objeto, cadeia de códigos e CHH é ilustrado na Figura 5.

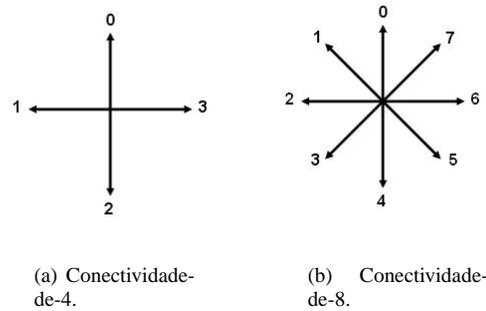


Figura 4. Primitivas utilizadas nas cadeias de códigos.

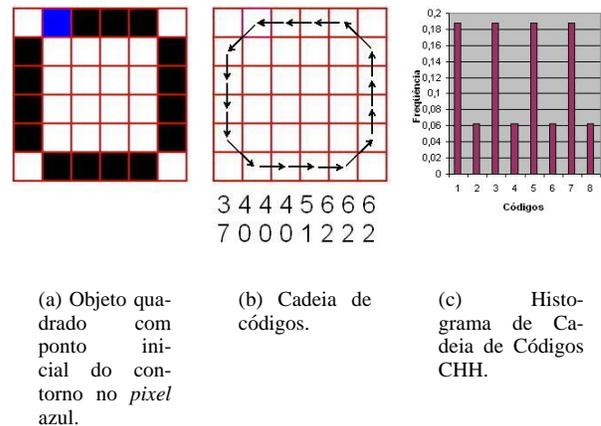


Figura 5. Exemplo da aplicação da cadeia de códigos em um objeto.

3.4. Descritores de Fourier

Os descritores de Fourier são muito utilizados para descrever os objetos através do contorno. Os descritores de Fourier aproximam os pontos dos contornos P (Equação 8) através da soma de k , com $k \leq \frac{n}{2}$, termos senos e cossenos [5] através da Equação 13.

$$p_l = \bar{p} + \sum_{k=1}^K a_k \cos\left(\frac{2\pi kl}{n}\right) + b_k \sin\left(\frac{2\pi kl}{n}\right), \quad l = 1, \dots, n \quad (13)$$

sendo \bar{p} o centro de massa do objeto, a_k e b_k descritos nas Equações 14 e 15.

$$a_k = \frac{2}{n} \sum_{l=1}^K p_l \cos\left(\frac{2\pi kl}{n}\right) \quad (14)$$

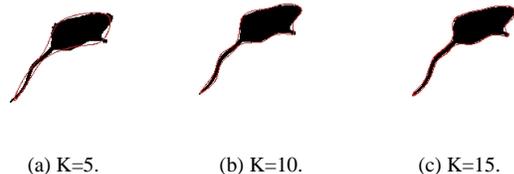


Figura 6. Aplicação dos descritores de Fourier para diferentes valores de K.

$$b_k = \frac{2}{n} \sum_{l=1}^K p_l \text{sen}\left(\frac{2\pi kl}{n}\right) \quad (15)$$

Para valores K iguais à $\frac{n}{2}$ a curva interpolada do contorno passa por todos os pontos originais. Para valores de K menores que $\frac{n}{2}$ a curva interpolada é uma aproximação do contorno original. As aproximações podem ser visualizadas na Figura 6 para o contorno de uma camundongo para diferentes valores de K .

Essas aproximações são importantes para eliminar pequenos ruídos fornecidos, por exemplo, pelo processo de segmentação. Além disso, através desses descritores é possível reconstruir o contorno original e manter a invariância com respeito à translação [5].

4. Experimentos

Os experimentos foram realizados com base em um banco composto de imagens de camundongos separadas em três comportamentos descritos como auto limpeza, exploração vertical e horizontal. Para cada comportamento foi utilizado 60 imagens de camundongos com iluminação artificial e fundo simples. As imagens contidas nesse banco foram convertidas em tons de cinza e segmentadas através da técnica de subtração de fundo [6].

O ambiente computacional utilizado nos experimentos foi um computador com processador Intel®Pentium 4™ com frequência de *clock* de 3.00 GHz, 512 MBytes de memória, com sistema operacional Ubuntu e a linguagem de programação Java versão 1.5. Para as implementações das técnicas foram utilizadas algumas funcionalidades das bibliotecas JMF², para manipulação de vídeos e imagens, ImageJ³, para o processamento das imagens e o Weka, que possui uma coleção de algoritmos de aprendizagem de máquina, aplicada na classificação dos comportamentos dos camundongos.

² JMF - <http://java.sun.com/products/java-media/jmf/>

³ ImageJ - rsb.info.nih.gov/ij/

Para cada técnica testada foram escolhidos os melhores parâmetros baseados na porcentagem de classificação. Na identificação dos comportamentos foram utilizados os classificadores de árvore de decisão J48, árvore de modelos logísticos LMT, máquina de vetor de suporte SMO e os classificadores IBK e Kstar implementados no Weka.

Para a técnica de momentos de imagens foi variado o número de quadrantes que dividem as imagens, representados por linhas e colunas, variando de duas a vinte divisões como mostra a Figura 1. Na técnica cadeia de códigos foi variado o parâmetro faixa de cadeias do histograma com intervalos de 1 a 8. O algoritmo de k-curvatura obteve a variação de 1 a 180 intervalos para a faixa de ângulos e o parâmetro k variou de 1 a 50. Para a técnica de descritores de Fourier o parâmetro variado foi o número de descritores iniciado em 5 até 30.

Os códigos desenvolvidos dos extratores estão disponíveis livremente na internet através de um sistema de controle de versão, o Subversion, integrado com o Trac, um sistema baseado em *web* para o gerenciamento de projetos utilizados pelo GPEC - Grupo de Pesquisa em Engenharia e Computação⁴. Estes descritores de formas implementados foram integrados ao projeto SIGUS⁵.

5. Análise e Resultados

O gráfico da Figura 7 apresenta a porcentagem de acerto de cada técnica implementada na identificação dos comportamentos de camundongos. Nota-se que a técnica de cadeia de códigos utilizando o classificador IBK atingiu a melhor porcentagem de acerto, com 96.3%, em relação as demais técnicas. Entretanto, a técnica de momentos de imagens manteve uma menor variação na porcentagem de acerto, alcançando uma média de 92.6% entre todos os classificadores, conforme mostrado na Tabela 1. O algoritmo de momentos de imagens obteve um melhor desempenho pelo fato de considerar as informações da área do objeto e não do seu contorno, onde o contorno é muito susceptível a ruídos e variante com imagens apresentando o mesmo comportamento.

6. Conclusão e Trabalhos Futuros

A principal contribuição deste trabalho foi a comparação destes descritores de formas para auxiliar no sistema de classificação dos comportamentos dos camundongos para o sistema Topolino [11]. Este sistema foi desenvolvido para auxiliar pesquisadores em experimentos com camundongos de forma automática na área da saúde.

⁴ <http://www.gpec.ucdb.br/>

⁵ SIGUS - Plataforma de Apoio ao Desenvolvimento de Sistemas para Inclusão Digital de Pessoas com Necessidades Especiais <http://www.gpec.ucdb.br/projetos/sigus/>

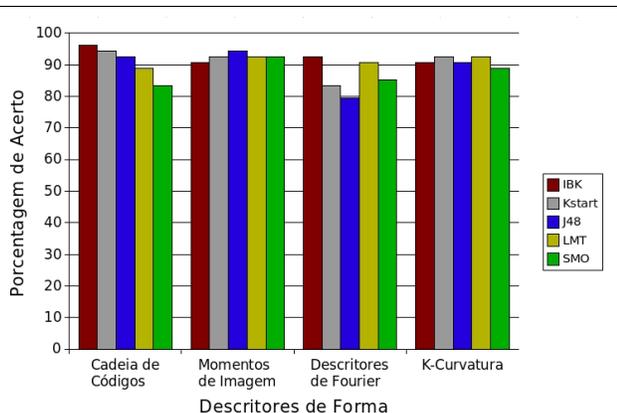


Figura 7. Gráfico dos extratores em relação aos classificadores utilizados.

Algoritmo	Média (%)
Momentos de Imagem	92.6
Cadeia de Códigos	91.1
K-Curvatura	91.1
Descritores de Fourier	86.3

Tabela 1. Média da porcentagem de acerto dos algoritmos.

Como trabalhos futuros, é proposto o desenvolvimento de outras técnicas de descritores de formas baseadas em região. Com isso, será possível realizar novas comparações com as técnicas de descritores baseados em contorno que são sensíveis a ruídos.

Agradecimentos

Este trabalho recebeu apoio financeiro da Universidade Católica Dom Bosco, UCDB, da Agência Financiadora de Estudos e Projetos, FINEP e da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, FUNDECT. Os autores também contaram com bolsas do Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPQ, nas modalidades PIBIC e Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora.

Referências

[1] E. R. Costa, A. R. Hirakawa, and J. J. Neto. An adaptive alternative for syntactic pattern recognition. In *Proceeding of 3rd International Symposium on Robotics and Automation, ISRA*, pages 409–413, Toluca, Mexico, September 2002.

[2] K. P. de Souza and H. Pistori. Implementação de um extrator de características baseado em momentos da imagem. In *Proceedings of SIBGRAPI 2005 - XVIII Brazilian Symposium on Computer Graphics and Image Processing*, Natal, RN, Brazil, October 2005. SBC, IEEE Press.

[3] P. Dukkupati and L. Brown. Improving the recognition of geometrical shapes in road signs by augmenting the database. *ICCSA International Conference on Computer Science and its Applications*, V2, June 2005.

[4] H. Freeman. Computer processing of line-drawing images. *ACM Computing Surveys*, 6(1):57–97, March 1974.

[5] C. A. Glasbey and G. W. Horgan. *Image analysis for the biological sciences*. John Wiley & Sons, Inc., New York, NY, USA, 1995.

[6] W. N. Gonçalves, V. A. Saueia, B. B. Machado, J. de Andrade Silva, K. P. de Souza, and H. Pistori. Técnicas de segmentação baseadas em subtração de fundo e modelos de cores: Um estudo comparativo. *XXVIII CILAMCE - Iberian Latin American Congress on Computational Methods in Engineering*, June 2007.

[7] S. Hoque, K. Sirlantzis, and M. C. Fairhurst. A new chain-code quantization approach enabling high performance. In *ICIDAR*. IEEE Computer Society, 2003.

[8] J. Iivarinen, M. Peura, J. Särelä, and A. Visa. Comparison of combined shape descriptors for irregular objects. In A. F. Clark, editor, *8th British Machine Vision Conference, BMVC'97*, volume 2, pages 430–439, september 1997.

[9] J. Iivarinen and A. Visa. Shape recognition of irregular objects. In D. P. Casasent, editor, *Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling, Proc. SPIE 2904*, pages 25–32, 1996.

[10] C.-B. Liu and N. Ahuja. A model for dynamic shape and its applications. *CVPR*, 02:129–134, 2004.

[11] B. B. Machado, J. de Andrade Silva, W. N. Gonçalves, H. Pistori, and A. S. de Souza. Topolino: Software livre para automatização do experimento do campo aberto. *XV Seminário de Computação*, pages 19–28, November 2006.

[12] A.-B. M. Salem, A. A. Sewisy, and U. A. Elyan. A vertex chain code approach for image recognition. *ICGST International Journal on Graphics, Vision and Image Processing*, 05, 2005.

[13] M. E. Stivanello and P. C. R. Gomes. Inspeção visual industrial automatizada por análise de forma com descritores de fourier e redes neurais artificiais. *XV Seminário de Computação*, 2006.

[14] M. A. Q. Truyenque. Uma aplicação de visão computacional que utiliza gestos da mão para interagir com o computador. Master's thesis, PUC-Rio, 2005.

[15] W.-Y. Wu. A dynamic method for dominant point detection. *Graph. Models*, 64(5):304–315, 2002.

[16] D. S. Zhang and G. Lu. A comparative study on shape retrieval using fourier descriptors with different shape signatures. *International Conference on Intelligent Multimedia and Distance Education*, pages 1–9, 2001.