

A novel method for shape analysis based on statistics of Euclidean distances

Isaac Newton Melo Machado
Universidade Federal do Ceará
Campus de Sobral
Email: engisaacnewton@gmail.com

Francisco Evangelista Nobre Filho
Universidade Federal do Ceará
Campus de Sobral
Email: evan.filho06@gmail.com

Jarbas Joaci de Mesquita Sá Junior
Universidade Federal do Ceará
Campus de Sobral
Email: jarbas_joaci@yahoo.com.br

Abstract—Shape analysis is an important field in computer vision and, throughout the years, many methods have been proposed for shape characterization. In order to contribute to this line of research, this paper presents a novel shape descriptor based on statistics of Euclidean distances. For this purpose, the technique computes Euclidean distances between points of a shape contour and stores them into a matrix, which in turn is subdivided into submatrices. For each submatrix, the algorithm computes values such as mean, standard deviation and minimum distance for each column. Finally, attributes are computed from these data in order to construct a shape signature. The proposed method presented high performance in two benchmark datasets, surpassing other methods found in the literature.

1. Introdução

Ao definirmos um objeto, consideramos algumas de suas características, como, cor, textura ou forma, sendo esta última apenas dependente da existência do objeto, ou seja, ela é invariante em relação a alterações na iluminação e textura, por exemplo. Tratando-se de identificação visual, a forma é um atributo que permite o reconhecimento dos objetos de forma fácil e objetiva.

Na área de visão computacional, os problemas envolvendo classificação e segmentação de formas são antigos e, ao longo dos anos, diversos métodos de análise foram propostos. Eles se dividem em três grupos: baseados em contorno, região ou esqueleto [1]. Basicamente, suas diferenças se encontram no local de extração de características [2]. Os métodos baseados em contorno consideram a imagem como um conjunto de pontos ordenados e conectados por segmentos de reta. Nesses métodos, como esperado, pontos não interligados podem gerar alterações nos resultados. Como exemplos dessa categoria, podemos citar os descritores de Fourier [3], [4] e *Curvature Scale Space* (CSS) [5]. Já os métodos baseados em região, que são amplamente utilizados em formas genéricas, utilizam agrupamentos de pixels semelhantes para gerar descritores [6]. Os momentos de Zernike [7], [8] são um exemplo. Por fim, as técnicas baseadas em esqueleto geralmente modelam as formas por meio de grafos ou dimensões fractais, utilizando as informações dos eixos médios do objeto. O método de

“similaridade de caminho” [9], [10] é um exemplo dessa categoria. As assinaturas desses métodos são adequadas para imagens que possuem oclusão ou articulação [1], [2].

Considerando as explicações acima, o presente trabalho propõe um descritor que utiliza características estatísticas (média, desvio padrão e distância mínima) das distâncias relativas dos pontos de contorno entre si. Dessa forma, as propriedades invariantes a escala, rotação e translação são mantidas. Além disso, haja vista a complexidade de alguns algoritmos, propomos um método de simples entendimento.

Este artigo é apresentado da seguinte forma: a Seção 2 apresenta o método desenvolvido, além de defini-lo matematicamente. A Seção 3 descreve as bases de imagens usadas e os experimentos realizados. A Seção 4 apresenta os resultados obtidos e uma discussão sobre os mesmos. Finalmente, a Seção 5 apresenta algumas considerações sobre este trabalho.

2. Método proposto

Seja $C_N = \{c_1, c_2, \dots, c_N\}$ o conjunto de todos os pontos do contorno de uma forma, em que $c = (x, y) \in R^2$ é a coordenada do ponto e D uma matriz de distâncias euclidianas. Cada linha w da matriz D representa a distância do ponto c para todos os outros pontos do conjunto C_N . Com a finalidade de tornar o método invariante à rotação e escala, as distâncias das linhas de $D_{N \times N}$ são ordenadas e divididas por N , sendo este o número de pontos do contorno. A propriedade de invariância à translação é respeitada, pois são tomados como referência os próprios pontos do contorno.

Após a transformação acima descrita, a matriz resultante é subdividida, de forma a obter sub-matrizes, com o intuito de extrair um maior número de características e fornecer um conjunto mais abrangente de estatísticas acerca dos valores das distâncias. Para isso, definimos pela Equação 1 um conjunto H_k de sub-matrizes que contém número igual de linhas, N , e suas colunas definidas por $N(t_k)$.

$$H_k = \{S_{N \times N(t_1)}^1, S_{N \times N(t_2)}^2, \dots, S_{N \times N(t_k)}^k\} \quad (1)$$

Desta forma, o conjunto de sub-matrizes depende diretamente do conjunto de parâmetros T_k , definido na Equação 2, ordenados de forma crescente e composto pelos coeficientes t_i , com $i \in \mathbb{N}^*$.

$$T_k = \{0, t_1, t_2, \dots, t_k\} \quad | \quad 0 < t_1 < t_2 < \dots < t_k \leq 1 \quad (2)$$

Esses coeficientes definem o intervalo de colunas da matriz resultante $D_{N \times N}$ pertencente a cada sub-matriz $S_{N \times N(t_i)}^i$, dada por $D[\alpha, \beta]$, com α e β definidos na Equação 3. Tais intervalos dizem respeito às colunas que variam de $\lfloor N \cdot (t_{i-1}) \rfloor + 1$ até $\lfloor N \cdot (t_i) \rfloor$. Portanto, T_k formará os argumentos do descritor proposto, que variam de forma a obter um melhor arranjo para a aplicação.

$$\alpha = \{1, 2, \dots, N\} \quad \beta = \{\lfloor N \cdot (t_{i-1}) \rfloor + 1, \dots, \lfloor N \cdot (t_i) \rfloor\} \quad (3)$$

Após a obtenção do conjunto H de sub-matrizes, extraem-se três vetores dessas, que caracterizam a média ($\vec{\tau}_i$), desvio padrão ($\vec{\Psi}_i$) e mínima distância ($\vec{\Upsilon}_i$) de cada coluna j das sub-matrizes. Esses vetores são definidos, respectivamente, nas Equações 4, 5 e 6.

$$\vec{\tau}_i = [A_1 \quad A_2 \quad A_3 \quad \dots \quad A_{N(t_i)}] \quad \text{em que} \quad (4)$$

$$A_j = \frac{1}{N} \sum_{w=1}^N S_{w,j}^i$$

$$\vec{\Psi}_i = [B_1 \quad B_2 \quad B_3 \quad \dots \quad B_{N(t_i)}] \quad \text{em que} \quad (5)$$

$$B_j = \sqrt{\frac{1}{N} \sum_{w=1}^N (S_{w,j}^i - A_j)^2}$$

$$\vec{\Upsilon}_i = [C_1 \quad C_2 \quad C_3 \quad \dots \quad C_{N(t_i)}] \quad \text{em que} \quad (6)$$

$$C_j = \min_{\forall w \in \alpha} S_{w,j}^i$$

A partir desse ponto, criam-se os novos vetores $\vec{\phi}_1, \vec{\phi}_2, \dots, \vec{\phi}_k$. Esses foram obtidos a partir da média μ e desvio padrão σ dos vetores definidos nas Equações 4, 5 e 6 extraídos de cada uma das sub-matrizes $S_{N \times N(t_i)}^i$. Desse modo, é definido $\vec{\phi}_i$, vetor de seis características da sub-matriz de índice i por

$$\vec{\phi}_i = [\mu(\vec{\tau}_i) \quad \sigma(\vec{\tau}_i) \quad \mu(\vec{\Psi}_i) \quad \sigma(\vec{\Psi}_i) \quad \mu(\vec{\Upsilon}_i) \quad \sigma(\vec{\Upsilon}_i)]. \quad (7)$$

Por fim, chegamos ao vetor de características $\vec{\Omega}$ (Equação 8) composto pelos vetores de características de todas as sub-matrizes S^i .

$$\vec{\Omega} = [\vec{\phi}_1 \quad \vec{\phi}_2 \quad \dots \quad \vec{\phi}_k] \quad (8)$$

3. Experimento

Na seção que segue será apresentado o experimento realizado, com o objetivo de analisar a eficiência do método descrito na Seção 2. Para isso, serão classificados dois bancos de imagens: (i) *generic shapes* e (ii) *fish*. O primeiro

é composto por 9 classes com 11 imagens cada, o que nos fornece um total de 99 imagens [11], [12]. Cada classe tem uma forma única e com diferenças significativas com relação às outras, como pode ser visto na Figura 1. Estão incluídas nessas diferenças oclusão, rotação e partes que faltam (Figura 2).

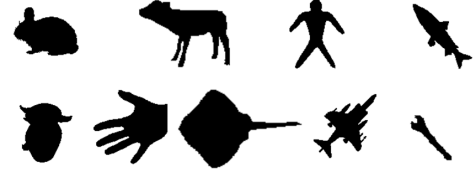


Figura 1. Exemplo das classes do banco de imagens *generic shapes* [11], [12].



Figura 2. Imagens do banco *generic shapes* com oclusão e rotação [11], [12].

O segundo banco de imagens apresenta formas de peixes [13]. Ele é dividido em 1.100 classes distintas, cada uma possuindo 10 imagens, o que totaliza 11.000 [17]. Cada imagem possui escala e rotação diferentes, o que permite a realização de um teste mais abrangente e preciso das propriedades de invariância a rotação e escala. Exemplos do banco de imagens (ii) são dados na Figura 3.

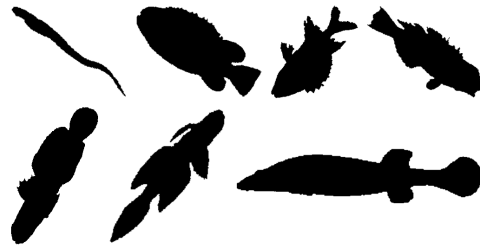


Figura 3. Exemplo das classes do banco de imagens *fish* [13].

O experimento foi realizado em 3 etapas: (i) segmentação, (ii) utilização do descritor proposto e (iii) validação. As imagens utilizadas como objeto de estudo deste trabalho foram adquiridas pré-processadas (binarizadas). Em seguida, com auxílio do *software* MATLAB, uma segmentação morfológica foi realizada em cada uma das imagens a fim de obter o contorno.

A partir das imagens segmentadas, o método descrito na Seção 2 é aplicado. As imagens passam pelo descritor proposto, no qual a matriz C é gerada e, consequentemente, o vetor de características $\vec{\Omega}$ também. Quando as características

são extraídas, a etapa (iii) inicia. A validação consiste em classificar e validar as informações contidas em $\vec{\Omega}$. Para o presente trabalho foram utilizadas duas técnicas: *Linear Discriminant Analysis* (LDA) e validação cruzada *leave-one-out*. Uma breve introdução sobre esses métodos será dada a seguir.

Para a classificação foi utilizado o LDA. Esse é um classificador amplamente utilizado e consiste em estimar um sub-espço linear, de forma que a variância entre as classes seja significativamente maior quando comparada à variância dentro das classes [14]. O classificador realiza esse processo definindo uma função densidade $f_i(x), i = 1, 2, \dots, z$, em que z é o número de classes que um banco de imagens possui e x é a amostra. Com esses dados, o classificador LDA obtém o valor da função densidade para cada classe i e classifica a amostra x para k , em que k é uma classe que representa o maior valor de $f_i(x)$. É importante ressaltar que essa função é definida através dos valores de média da classe i e da matriz de covariância de todas as classes, como mostra a Equação 9.

$$f_i(x) = -\frac{1}{2} \ln(|\Sigma|) - \frac{1}{2} (x - \mu_i)' \Sigma^{-1} (x - \mu_i), \quad (9)$$

Por fim, a validação cruzada *leave-one-out*. Ela é uma técnica que busca avaliar a confiabilidade da generalização de um modelo. De forma simples, ela estima o erro do nosso classificador. Esse processo é realizado utilizando uma amostra x como teste, enquanto as demais são utilizadas para um conjunto de treinamento. Esse processo continua até que todas as amostras x passem pelo processo de validação. Ao final, a taxa de acerto média é a medida de desempenho do classificador.

A fim de obter resultados aprimorados, a Subseção 4.2 mostra uma tabela comparando o método proposto com alguns tradicionais, como: descritores de Fourier [3], [4], momentos de Zernike [7], [8], descritores de Curvatura [15], *Multiscale Fractal Dimension* [16] e *Curvature Scale Space* (CSS) [5].

4. Resultados e discussões

4.1. Avaliação de parâmetros

Para obter o vetor de características $\vec{\Omega}$ a partir da matriz D é necessário um conjunto de parâmetros T . A fim de encontrar um resultado que pode ser usado em comparações futuras, é necessário definir a mesma configuração de parâmetros para as duas bases de imagens e assim configurar T de modo a encontrar a melhor descrição de características. As taxas de acerto do experimento definido na Seção 3 foram obtidas variando-se T , valores expostos na Tabela 1.

De acordo com a tabela foi constatado que há dificuldade em diferenciar valores de características similares e em grande quantidade. Por esse motivo, os conjuntos T que contêm cardinalidade elevada ou intervalos pequenos (0.01 - 0.04) têm suas taxas de acerto muito baixas, portanto foram retirados. Outro fato verificado diz respeito à primeira

metade da matriz D possuir informações mais relevantes, o que permite uma classificação mais precisa das imagens.

Após realizar a análise dos resultados expostos na Tabela 1, um vetor T de tamanho médio foi escolhido ($T = \{0.05, 0.1, 0.2, \dots, 0.8\}$). Sua função é servir como métrica de comparação dos resultados. Para essa configuração, é possível distinguir as classes com uma acurácia de 99.22% na base de imagens *fish* e também manter eficiência de 100% de acertos quando aplicada na base de imagens *generic shapes*.

TABELA 1. ACCURACY RATES FOR SOME VALUES OF $\vec{\Omega}$.

Parâmetro T	Nº Desc.	Acurácia(%)	
		<i>Generic shapes</i>	<i>Fish</i>
{0.5, 1}	12	97.97	98.00
{0.33, 0.66, 1}	18	100	98.49
{0.25, 0.5, 0.75, 1}	24	100	98.77
{0.2, 0.4, 0.6, 0.8, 1}	30	100	98.87
{0.1, 0.2, 0.4, 0.6, 0.8, 1}	36	98.98	99.11
{0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1}	42	97.97	99.15
{0.1, 0.2, ..., 0.5, 0.6, 0.8, 1}	48	97.97	99.17
{0.05, 0.1, 0.2, ..., 0.8}	54	100	99.22
{0.05, 0.1, 0.15, ..., 0.5}	54	93.93	98.85
{0.1, 0.2, ..., 0.8, 1}	60	96.96	99.24
{0.05, 0.1, ..., 0.4, 0.5, 0.75, 1}	66	98.98	99.30
{0.05, 0.1, ..., 0.5, 0.75, 1}	72	97.97	99.31
{0.05, 0.1, 0.15, ..., 0.7}	84	71.71	99.11

4.2. Comparação dos resultados

Após definidos os parâmetros que permitem que o descritor proposto tenha um melhor desempenho, foi realizada a comparação com alguns descritores tradicionais, citados ao final da Seção 2. Os resultados desses outros descritores foram obtidos nos trabalhos [2], [17].

A Tabela 2 mostra que o descritor proposto atinge uma taxa de acerto elevada, que é significativamente maior que os resultados dos descritores tradicionais encontrados na literatura.

A acurácia encontrada para o banco *generic shapes* se mostra superior às alcançadas em outros métodos. Isso indica que o método proposto é eficiente para identificar contornos com deformações ou oclusões. Apesar de o resultado escolhido ter um grande número de descritores em comparação aos outros métodos, a quantidade de classificações corretas se mantém relativamente alta para escolhas de parâmetros que geram menos descritores (12, 18 e 24), mantendo as taxas de acertos próximas à máxima possível.

Quanto ao resultado do segundo banco, *fish*, o descritor anteriormente proposto é eficiente para extrair características mantendo as propriedades de invariância a rotação e escala. Com exceção dos descritores de *Fourier* e do *Curvature*, nos quais os ganhos são de 0.15% e 1.67%, respectivamente, o descritor proposto nesse trabalho mostra um resultado,

no mínimo, 50% superior quando comparado aos demais. Esse resultado mostra que o método também é eficaz para distinguir formas similares. Outra característica relevante é a acurácia, que se mantém superior a 98%, independente da variação do número de características.

Quanto à complexidade do algoritmo, necessitamos calcular as distâncias de todos os pontos do contorno entre si, gerando N^2 distâncias. Assim, caracteriza-se um algoritmo $O(N^2)$. Essa complexidade é uma desvantagem em comparação aos outros métodos, como *Zernike* e *Fourier* que possuem complexidade $O(N \log N)$.

TABELA 2. AVALIAÇÃO DOS DESCRITORES. OS RESULTADOS DOS MÉTODOS DE COMPARAÇÃO FORAM OBTIDOS NOS TRABALHOS [2], [17].

Descritores	Acurácia(%)	
	<i>Generic shapes</i>	<i>Fish</i>
Descritor proposto	100	99.22
<i>Fourier</i>	83.84	99.07
<i>Zernike</i>	91.92	12.23
<i>Curvature</i>	76.77	97.55
<i>M. S. Fractal Dimension</i>	87.88	37.32
<i>Curvature Scale Space (CSS)</i>	71.72	43.52

5. Conclusão

No presente trabalho foi apresentado um descritor de formas que utiliza o contorno de simples entendimento e implementação. O descritor foi testado em dois bancos de imagens diferentes e obteve resultados superiores aos tradicionais. Podemos concluir pelo experimento que o método proposto é eficiente em classificar imagens com oclusão, articulação, partes faltantes, e é invariante a rotação, escala e translação.

Por fim, como produções futuras, métodos semelhantes que reduzam a complexidade da assinatura proposta, bem como trabalhos com otimização ou seleção de características, devido ao alto número parâmetros e de descritores, são exemplos de trabalhos que podem vir a melhorar os resultados expostos neste artigo.

Referências Bibliográficas

- [1] A. R. Backes and O. M. Bruno, Shape classification using complex network and Multi-scale Fractal Dimension, *Pattern Recognition Letters* 31(1): 44–51, 2010.
- [2] J. J. M. Sá Junior and A. R. Backes. Shape classification using line segment statistics. *Information Sciences*, 305: 349–356, 2015.
- [3] R. Gonzalez and R. E. Woods. *Processamento de Imagens Digitais*. Editora Edgard Blucher, São Paulo, 2000.
- [4] W. P. Lee, M. A. Osman, A. Z. Talib, K. Yahya, J. C. Burie, J. M. Ogier, and J. Mennesson, Recognition of fish based on generalized color Fourier descriptor, In *Science and Information Conference (SAI)*, London, pages 680–686, 2015. Disponível em: <http://ieeexplore.ieee.org/document/7237215/>
- [5] F. Mokhtarian and M. Bober, *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*, Kluwer Academic Publishers, 2003.
- [6] M.K. Hu, Visual pattern recognition by moment invariants, *IEEE Trans. Inform. Theory* 8(2): 179–187, 1962.
- [7] N. D. Mustaffa Kamal and N. Jalil, Fastener recognition using combination of colour features with shape-based features and Zernike moments, In *IEEE Student Conference on Research and Development (SCORED)*, Kuala Lumpur, pages 1–5, 2016. Disponível em: <http://ieeexplore.ieee.org/document/7810064/>
- [8] M. Zhenjiang, Zernike moment-based image shape analysis and its application, *Pattern Recogn. Lett* 21 (2): 169–177, 2000.
- [9] X. Bai, X. Yang, D. Yu, and L. J. Latecki, Skeleton-based shape classification using path similarity, *International Journal of Pattern Recognition and Artificial Intelligence*, 22(4): 733–746, 2008.
- [10] X. Bai and L. J. Latecki, Path similarity skeleton graph matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(7): 1282–1292, 2008.
- [11] D. Sharvit, J. Chan, H. Tek, and B. B. Kimia, Symmetry-based indexing of image databases, *Journal of Visual Communication and Image Representation*, 9(4): 366–380, 1998.
- [12] T. B. Sebastian, P. N. Klein, and B. B. Kimia, Recognition of shapes by editing their shock graphs, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5): 550–571, 2004.
- [13] S. Abbasi, F. Mokhtarian, and J. Kittler, *Search for similar shapes in the SQUID system: shape queries using image databases*, 2005. Disponível em: <http://www.ee.surrey.ac.uk/CVSSP/demos/css/demo.html>.
- [14] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, second ed., Academic Press, 1990.
- [15] W. Y. Wu and M. J. J. Wang, Detecting the dominant points by the curvature-based polygonal approximation, *Graph. Models Image Process.* 55(2): 79–88, 1993.
- [16] R. da S. Torres, A. X. Falcão and L. da F. Costa, A graph-based approach for multiscale shape analysis, *Pattern Recognition* 37(6): 1163–1174, 2003.
- [17] A. R. Backes, D. Casanova, and O. M. Bruno, A complex network-based approach for boundary shape analysis, *Pattern Recognition*, 42(1): 54–67, 2009.