

# ANÁLISE DE FORMAS PLANAS EM IMAGENS DIGITAIS

SOUZA, Gustavo Botelho de  
Universidade Estadual Paulista (UNESP)  
gustavo.botelho@gmail.com

MARANA, Aparecido Nilceu  
Universidade Estadual Paulista (UNESP)  
nilceu@fc.unesp.br

**RESUMO:** Com a difusão do uso dos computadores, o reconhecimento de padrões visuais tem sido automatizado em especial para poder tratar a enorme quantidade de imagens digitais disponíveis. Aplicações de diversas áreas utilizam técnicas de processamento de imagens bem como algoritmos de extração de características e reconhecimento de padrões visuais a fim de identificar pessoas, facilitar o diagnóstico de doenças, classificar objetos, etc. a partir de imagens digitais. Dentre as características que podem ser analisadas nas imagens encontra-se a forma de objetos ou regiões. Em alguns casos a forma é a única característica passível de análise com precisão. Este trabalho apresenta alguns dos mais importantes métodos de análise de formas descritos na literatura e compara seus desempenhos quando aplicados em três bases de dados públicas contendo imagens de formas. Por fim, propõe-se a criação de um novo descritor de formas baseado na Transformada de Hough.

**PALAVRAS-CHAVE:** análise de formas; análise de imagens; transformada de Hough.

**ABSTRACT:** *Given the widespread use of computers, the visual pattern recognition task has been automated in order to address the huge amount of available digital images. Many applications use image processing techniques as well as feature extraction and visual pattern recognition algorithms in order to identify people, to make the disease diagnosis process easier, to classify objects, etc. based on digital images. Among the features that can be extracted and analyzed from images is the shape of objects or regions. In some cases, shape is the unique feature that can be extracted with a relatively high accuracy from the image. In this work we present some of most important shape analysis methods and compare their performance when applied on three well-known shape image databases. Finally, we propose the development of a new shape descriptor based on the Hough Transform.*

**KEYWORDS:** *shape analysis; image analysis; Hough transform.*

## INTRODUÇÃO

A habilidade que o ser humano tem no trato com imagens é extremamente grande. Desde seu surgimento, o homem desenvolveu sistemas neurais altamente sofisticados voltados à tarefa de “processamento de imagens” e “reconhecimento de padrões visuais” (DUDA et al., 2000). Neste sentido, com o surgimento dos computadores, pesquisadores e empresas de todo o mundo voltaram suas expectativas à tentativa de mecanização do processo de captura, melhoramento, extração de características e reconhecimento de padrões a partir de imagens digitais dada a grande variedade de aplicações derivadas do trato automatizado das imagens.

Atualmente, diversas técnicas de análise de imagens e *hardwares* mais eficientes e baratos estão disponíveis para a construção de sistemas de visão computacional. Entretanto, os desempenhos dos sistemas computacionais ainda estão distantes dos apresentados pelos sistemas visuais biológicos (como o sistema visual humano, por exemplo) em relação ao tratamento de imagens. Isto pode ser explicado pelo fato de ainda não se ter compreendido perfeita e completamente o processo de inteligência humana e pelo uso de processamento primordialmente serial.

Neste contexto, novas técnicas de análise de imagens continuam surgindo aprimorando as já existentes ou tratando os problemas a partir de novas abordagens na

tentativa de automatizar o processo de análise de imagens tipicamente realizado pelos seres humanos. *Hardwares* mais rápidos e eficientes também estão sendo criados. Com tudo isso, as máquinas cada vez mais se assemelham aos seres humanos em se tratando do processo de reconhecimento de padrões visuais e assim se tornam mais úteis aos mesmos.

A análise de imagens pode se basear em várias características extraídas das mesmas a fim de obter informações sobre seus conteúdos. A forma de objetos ou regiões é uma das principais. Além de possuir grande poder de distinção entre elementos diferentes das imagens, em alguns casos a forma é a única passível de extração e análise com precisão. Deste modo, a análise de formas se configura como uma ferramenta essencial ao reconhecimento de padrões visuais nas imagens.

Neste artigo são apresentados alguns dos mais importantes métodos de análise de formas descritos na literatura especializada. Os desempenhos desses métodos são avaliados sobre três bases de dados contendo formas de objetos (silhuetas). Ao final, é proposta a criação de um novo descritor de formas (rápido e preciso) baseado na Transformada de Hough (DUDA e HART, 1972).

## 1. DEFINIÇÃO DE FORMA

Como dito, dentre as características que podem ser analisadas em uma imagem a fim de descobrir informações sobre seu conteúdo encontra-se a forma de objetos ou regiões da mesma.

Apesar de o ser humano lidar com formas (e seu reconhecimento) a todo instante, a definição em termos matemáticos do conceito “forma” (no caso deste trabalho, planas) não é uma tarefa fácil.

De acordo com Costa e Júnior (2000), uma forma pode ser entendida como qualquer entidade visual singular. Em outras palavras, quando se menciona “forma” faz-se alusão a um objeto como um todo, um conjunto de pontos conectados (quer no espaço discreto ou no contínuo).

Segundo Latecki e Lakämper (2000), independente do método usado

na análise de formas, ele deve funcionar como o nosso sistema de percepção visual. Isto significa que ele deve apresentar as seguintes características: (i) a medida de similaridade entre formas deve permitir o reconhecimento de objetos visualmente similares, mesmo que não matematicamente idênticos; (ii) deve abstrair distorções, como ruídos da digitalização e erros de segmentação (localização e demarcação de estruturas na imagem); (iii) deve dar atenção à partes visuais significantes dos objetos e (iv) não deve depender da escala, orientação e posição dos objetos.

## 2. Descritores de formas

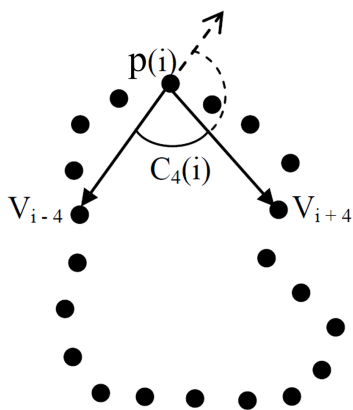
Os métodos de análise de formas apresentados na literatura se baseiam em diferentes propriedades a fim de representá-las e identificá-las. Dentre estas propriedades podemos citar: a área, a curvatura da borda (contorno), a quantidade de concavidades, a posição dos seus pontos de borda, dentre outras. A seguir são apresentados alguns descritores de formas tradicionais encontrados na literatura.

### 2.1. BAS (*Beam Angle Statistics*)

Um importante e eficiente descritor de formas proposto por Árica e Vural (2003) é o BAS (*Beam Angle Statistics*). O método BAS se baseia na ideia da representação do contorno de um objeto (que por natureza é uma função 2-D) através de uma função 1-D. Esta função 1-D deve representar todas as concavidades e convexidades do contorno do objeto com o mínimo de distorções.

Uma borda (contorno)  $B$  previamente segmentada, isto é, localizada e demarcada na imagem sob análise, é representada por uma sequência de pontos conectados  $p_i = (x_i, y_i)$ , onde  $i = 1, 2, \dots, N$ , sendo  $N$  o número de pontos da borda e  $p_i = p_{i+N}$ . Para cada ponto da borda  $p_i$ , os raios (*beams*) de  $p_i$  podem ser representados pelo conjunto de vetores  $L(p_i) = \{V_{i+j}^+, V_{i-j}^-\}$  onde  $V_{i+j}^+$  e  $V_{i-j}^-$  são os vetores posterior e anterior que conectam  $p_i$  aos pontos de borda  $p_{i+j}$  e  $p_{i-j}$ , respectivamente, na borda  $B$ , para  $j = 1, 2, \dots, N/2$ .

O ângulo entre o par de raios posterior ( $V_{i-j}$ ) e anterior ( $V_{i+j}$ ) de  $p_i$ , fixado um valor de  $j$  (valor inteiro entre 1 e  $N/2$ ), é definido por  $C_j(i)$ . A Figura 1 exhibe o ângulo  $C_4(i)$ , entre os raios de um ponto de borda  $p_i$ , obtido fixando-se o valor de  $j=4$ .



**Figura 1:** Ângulo  $C_4(i)$  entre o par de raios  $V_{i-4}$  e  $V_{i+4}$  (ou seja,  $j=4$ ) do ponto  $p(i)$  (ou ainda  $p_i$ ) (FALGUERA e MARANA, 2008).

O processo de encontrar os ângulos  $C_j(i)$  entre cada par de raios  $L(p_i)=\{V_{i+j}, V_{i-j}\}$  associados a  $p_i$  e com  $j=1, 2, \dots, N/2$  é realizado para todos os pontos da borda da forma. A partir do conjunto de ângulos encontrados e associados a cada ponto  $p_i$  da borda calcula-se os três primeiros momentos (média, desvio padrão e momento de ordem 3) dos valores dos mesmos e associa-se ao ponto  $p_i$ , não mais o conjunto de ângulos  $C_j(i)$ , mas sim estes três valores de momentos, calculados com base no conjunto.

Dado um ponto de referência da borda da forma e percorrendo-se a borda em sentido horário a partir deste ponto obtêm-se três funções 1-D com base nos três valores de momentos associados a cada ponto da borda (cada valor de momento gera uma função).

Dada uma forma de teste, suas três funções 1-D são encontradas e a partir delas um vetor de características de tamanho  $k$  é gerado para a forma ao se amostrar as três funções em um número  $k$  de posições. Deste modo, cada elemento do vetor de ca-

racterísticas da forma contém três valores (os três momentos amostrados das funções em uma certa posição).

Durante a fase de reconhecimento, o vetor de características da imagem de teste é comparado com os vetores das imagens da base de dados por meio do algoritmo OCS (*Optimal Correspondent Subsequence*) proposto por Wang e Pavlidis (1990). A cada forma da base de dados é associada uma distância em relação à forma de teste com base na comparação de seus vetores de características. A forma de teste é então classificada como sendo da mesma classe da imagem da base de dados que apresentou a menor distância, ou seja, da mesma classe da imagem da base de dados mais similar (com vetor de características mais parecido).

## 2.2. Dimensão Fractal Multiescala

Disseminada por Mandelbrot (1982), a dimensão fractal, que pode ser expressa por números fracionários, provê um eficiente meio de caracterizar a auto-similaridade de objetos abstratos e reais chamados fractais (TORRES et al., 2004).

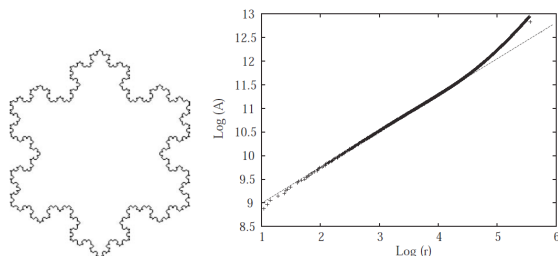
Dada uma forma, representada pelo conjunto  $S$  de todos seus pontos, e sendo  $S_r$  a dilatação exata (TORRES et al., 2004) da mesma por um disco de raio  $r$ , considerando-se  $A_r$  a área da versão dilatada da forma ( $S_r$ ). A dimensão fractal de Minkowski-Bouligand ( $F$ ), da forma operando-se em um espaço bidimensional, é definida por:

$$F = 2 - \lim_{r \rightarrow 0} \frac{\log(A_r)}{\log(r)} \quad (1)$$

A dimensão fractal é, portanto, um número no intervalo  $[0, 2]$ . Ela pode ser estimada por meio de uma interpolação linear da curva logarítmica da área  $A_r$  em função do raio de dilatação  $r$ . Após encontrar a reta que melhor se “aproxima” dos pontos da curva logarítmica, seu coeficiente angular  $A_r'$  pode ser utilizado para encontrar o valor de  $F$  aproximado por meio da equação:

$$F = 2 - A_r' \quad (2)$$

A Figura 2 ilustra esta interpolação para uma forma fractal e o cálculo do valor da dimensão fractal ( $F$ ) a partir da mesma.



**Figura 2:** Forma similar à estrela de Koch à esquerda e à direita os valores de  $\log(A_r)$  em função de  $\log(r)$  e a reta (tracejada) que melhor representa os pontos. A partir do coeficiente angular ( $A_r'$ ) desta reta pode-se obter o valor de  $F$  (TORRES et al., 2004).

A representação de formas por meio de um único valor (dimensão fractal tradicional) pode ser muito pobre e perder informações relevantes das formas uma vez que a curva  $\log\text{-}\log$  (extremamente complexa) é “aproximada” por uma simples reta. Para tratar este problema foi proposta a dimensão fractal multiescala. Diferente da dimensão fractal tradicional, a qual utiliza uma interpolação linear para estimar o coeficiente angular da curva  $\log\text{-}\log$ , esse método explora o limite infinitesimal da interpolação linear por meio da derivada. A partir da curva  $\log\text{-}\log$  obtém-se, por exemplo, uma curva polinomial por regressão que melhor represente os pontos da primeira e gera-se como que uma função baseada na derivada da função polinomial (adaptação da Equação 1), a qual é chamada de dimensão fractal multiescala da forma sendo analisada (TORRES et al., 2004).

Nesta técnica, para se gerar o vetor de características da forma em análise faz-se uma amostragem da função de dimensão fractal multiescala em  $n$  posições. Após, dados os vetores de características da forma sob análise e das formas da base de dados (extraídos das funções de dimensão fractal multiescala das mesmas), utiliza-se a métrica  $L_2$  para a comparação

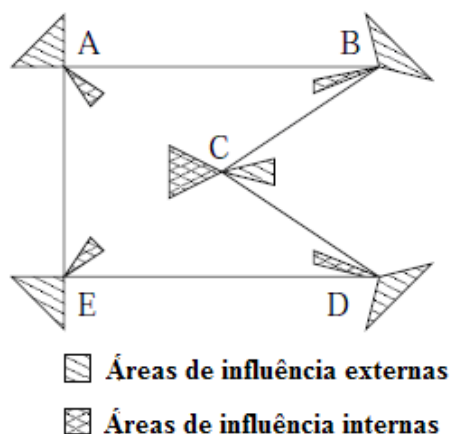
entre eles a fim de verificar a similaridade dos mesmos e permitir classificar a forma sob análise como sendo da mesma classe da forma da base de dados com menor distância, ou seja, mais similar.

### 2.3. Saliências do Contorno

Costa et al. (2001) propuseram o uso das saliências dos contornos das formas como meio de representá-las. As saliências do contorno de uma forma são definidas como as áreas máximas de influência de seus pontos (do contorno) de maiores curvaturas, denominados pontos de saliência, respeitando as regiões de Voronoi dos pontos e limitando-se as áreas a regiões próximas aos pontos apenas (TORRES e FALCÃO, 2007).

O algoritmo proposto por Costa et al. (2001), e adaptado por Torres e Falcão (2007) para determinar as saliências de uma forma opera de maneira que, dado um conjunto  $S$  de sementes (*pixels* da borda da forma), ele associa a cada *pixel*  $p$  da imagem um valor  $C(p)$  e um rótulo  $R(p)$  os quais são, respectivamente, a distância euclidiana entre  $p$  e a semente mais próxima e o rótulo (identificação) da semente mais próxima de  $p$ .

Após encontrar as áreas de influências internas e externas ao contorno da forma associadas aos *pixels* da borda (sementes), estes passam a ser representados pelos valores de suas respectivas áreas máximas. As áreas de influência de pontos de maiores curvaturas na borda, chamados de pontos de saliência, em geral, são maiores que as áreas de outros *pixels* da borda da forma e assim, para localizar estes na borda da forma basta aplicar um limiar nos pontos de borda selecionando os que apresentam áreas (valores de saliência) acima desse limiar. Com isto os pontos de saliência da borda da forma são encontrados (TORRES e FALCÃO, 2007). A Figura 3 ilustra as áreas de influências (internas e externas) de pontos de saliência de uma forma (polígono).



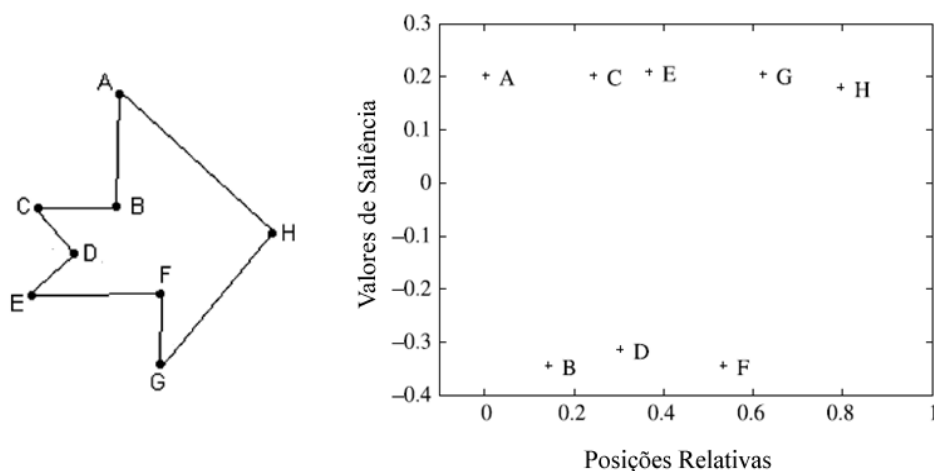
**Figura 3:** Áreas de influência internas e externas de pontos de saliência convexos (A, B, D e E), isto é, cujas áreas de influência externas são maiores que as internas, e côncavos (C), no caso oposto. As áreas são limitadas por um ângulo  $\theta$  (que respeita os limites das regiões de Voronoi dos pontos) e por uma distância  $r$  que garante que não se afaste muito dos pontos da borda (TORRES e FALCÃO, 2007).

Após encontrar os pontos de saliência do contorno da forma representa-se a mesma pelos valores de saliência (maiores áreas de influência) associados aos mesmos.

O descritor Saliências do Contorno (TORRES e FALCÃO, 2007) se baseia nesta representação para classificar as formas presentes em imagens digitais. Neste descritor, após determinar os pontos de saliência do contorno, pontos de saliência côncavos têm seus valores de saliência tomados como negativos e pontos de saliência convexos continuam com seus valores

de saliência positivos. Um ponto de saliência aleatório é tomado como referência e o método calcula as posições relativas de cada ponto de saliência do contorno em relação ao ponto de referência.

Então, os valores de saliências dos pontos de saliência do contorno e suas posições relativas formam dois vetores de características de mesmo tamanho, os quais são usados pelo descritor Saliências do Contorno na comparação entre formas. A Figura 4 ilustra estes vetores para um contorno poligonal.



**Figura 4:** Polígono (forma) e seus pontos de saliência à direita e os vetores de características da forma representados no gráfico à esquerda. O ponto de saliência de referência neste caso é o ponto A (TORRES e FALCÃO, 2007).

O descritor Saliências do Contorno usa um algoritmo de casamento heurístico de vetores de características que alinha os vetores usando os pontos de referência e calcula suas similaridades considerando a diferença de tamanhos entre eles. Ele é baseado no algoritmo proposto por Mokhtarian e Abbasi (2002). Após comparar a forma sob análise com as formas da base de dados, pode-se considerar a primeira como sendo da mesma classe que a forma mais parecida da base de dados.

#### 2.4. *Tensor Scale*

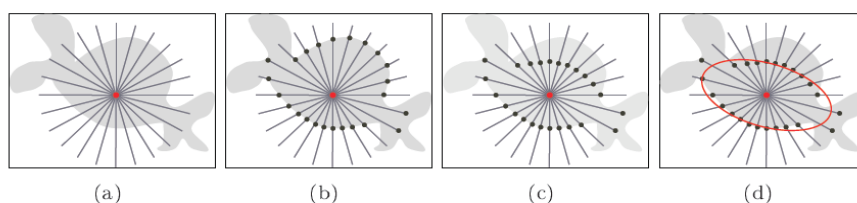
Saha (2005) introduz um método local para análise de imagens em tons de cinza, chamado *Tensor Scale*, o qual se resume em encontrar, para cada ponto  $p$  da forma, a maior elipse centrada em  $p$  dentro da região homogênea deste ponto na forma sendo analisada.

O primeiro passo para se encontrar a elipse centrada em um ponto  $p$  corresponde em traçar segmentos de reta opostos em relação ao ponto e iniciando nele, variando suas inclinações em relação ao eixo  $x$  no intervalo  $[0^\circ, 180^\circ)$ , conforme mostra a Figura 5(a). Um vetor unitário é usado para representar a direção de cada segmento, existindo assim  $m$  pares destes vetores (onde  $m$  não equivale a 180 necessariamente, isto é, pode-se “espaçar” mais ou menos os segmentos de reta em relação aos seus ângu-

los) (ANDALÓ et al., 2007).

Após traçar os segmentos a partir do ponto  $p$  verifica-se, para cada um deles, o ponto da borda da forma mais próximo de  $p$  por onde o segmento passa como mostrado na Figura 5(b). Para encontrar estes pontos a intensidade (nível de cinza) presente ao longo de cada segmento de reta é verificada. Como uma observação, caso a distância entre o ponto de borda a ser encontrado sobre o segmento e o ponto  $p$  ultrapasse o valor de  $L$  escolhido (valor máximo de distância), o ponto escolhido para o segmento é o que está à distância  $L$  de  $p$  no mesmo (último ponto no segmento). Isto ocorre com dois segmentos na Figura 5(b) na parte inferior direita do objeto sob análise.

Depois de encontrar o ponto da borda associado a cada segmento, em cada par de segmentos de reta (opostos), o ponto de borda com maior distância em relação ao ponto  $p$  é “deslocado” sobre o segmento de forma a ficar com a mesma distância em relação a  $p$  que o ponto de borda oposto (Figura 5(c)) a fim de garantir a simetria da elipse a ser gerada para o ponto  $p$ . Com os pontos todos rearranjados, uma elipse (que melhor se aproxima da posição dos mesmos) é obtida através da Análise do Componente Principal (*Principal Component Analysis – PCA*) como mostrado na Figura 5(d).



**Figura 5:** “Cálculo” da elipse associada ao ponto vermelho. Em (a) são traçados os segmentos de reta, em (b) os pontos de borda do objeto são localizados sobre os segmentos, em (c) os pontos são rearranjados e em (d) a elipse é traçada (SAHA, 2005).

Miranda et al. (2005) otimizam algumas partes deste método e o usam como descritor de formas. Este descritor atua de forma que a todo ponto da forma (da sua borda ou interno à mesma) realiza-se o processo de encontrar a elipse centrada nele. Após encontrar a elipse para um dado ponto, a orientação dela é associada ao mesmo.

A partir disto, o descritor *Tensor Scale* calcula o histograma das orientações locais dos pontos da forma sendo investigada (orientações das elipses centradas nos pontos da forma e encontradas como descrito anteriormente) e usa o mesmo como vetor de característica da forma.

O casamento de duas imagens pelo método *Tensor Scale* é realizado por meio da diferença absoluta das áreas de seus histogramas após corrigir eventuais deslocamentos causados por rotações no objeto. A forma sob análise pode ser classificada como pertencente à mesma classe da forma da base de dados mais parecida (menor distância).

### 3. Comparação dos métodos de análise de formas

Os métodos de análise de formas apresentados possuem princípios e estratégias de funcionamento bastante diferentes na extração de características, representação e classificação de formas. Com o objetivo de avaliar e comparar seus desempenhos, alguns experimentos foram realizados sobre três bases de imagens públicas contendo formas (silhuetas de objetos).

#### 3.1. Base Kimia-99

O primeiro experimento foi realizado usando-se a base de imagens Kimia-99 (SEBASTIAN et al., 2004). Nesta base existem 99 imagens de 9 classes de objetos (11 imagens por classe). Em cada imagem há a silhueta de um objeto em preto e um fundo branco. As dimensões das imagens não passam de 200 *pixels* e algumas das silhuetas estão rotacionadas, deformadas e transladadas.

Cada descritor (BAS, Dimensão Fractal Multiescada, Saliências do Contorno e *Tensor Scale*) foi testado em separado e para analisar seus desempenhos foram calculados os valores “Top n” dos mesmos da seguinte forma: dada forma de consulta, toda vez que ela e a imagem da base de dados mais parecida com ela (menor distância) segundo o descritor realmente eram da mesma classe, incrementava-se em uma unidade o valor “Top 1” do método. De forma similar, toda vez que as duas imagens da base de dados recuperadas como sendo as duas mais parecidas com a forma de consulta segundo o descritor eram da mesma classe desta, o valor de “Top 2” do descritor era incrementado em uma unidade. Processo análogo foi repetido para se calcular o valor “Top 3” para cada descritor.

As 99 imagens da base de dados foram tomadas, uma por vez, como imagens de consulta. Desta forma um descritor ideal (com 100% de acerto) teria os valores “Top 1”=99, “Top 2”=99 e “Top 3”=99. Os valores de “Top 1”, “Top 2” e “Top 3” dos descritores avaliados são mostrados na Tabela 1. O melhor desempenho foi obtido pelo método BAS e o pior pelo método Saliências do Contorno.

**Tabela 1.** Total de acertos operando-se com a base Kimia-99 (SEBASTIAN et al., 2004). Os métodos estão ordenados (nas linhas) pelos seus desempenhos, iniciando pelo melhor (BAS) e terminando pelo pior (Saliências do Contorno).

Descritor	Top 1	Top 2	Top 3
BAS (40 amostras)	97	95	93
Dimensão Fractal Multiescala	94	88	81
<i>Tensor Scale</i>	84	66	52
Saliências do Contorno	70	45	30

### 3.2. Base Kimia-216

A mesma metodologia de avaliação utilizada nos experimentos com a base Kimia-99 (SEBASTIAN et al., 2004) foi repetida com a base de imagens de formas Kimia-216 (SEBASTIAN et al., 2004), a qual é uma extensão da Kimia-99. Nesta base as imagens apresentam as mesmas características que as da base anterior, no entanto

existem 18 classes de objetos, com cada uma delas contendo 12 amostras (imagens de silhuetas), totalizando 216 imagens na base.

Os valores de “Top 1”, “Top 2” e “Top 3” dos descritores são mostrados na Tabela 2. Novamente os métodos BAS e Saliências do Contorno apresentaram o melhor e o pior resultados, respectivamente.

**Tabela 2.** Total de acertos operando-se sobre a base Kimia-216 (SEBASTIAN et al., 2004). Os métodos estão ordenados (nas linhas) pelo desempenho: iniciando pelo melhor (BAS) e terminando pelo pior (Saliências do Contorno).

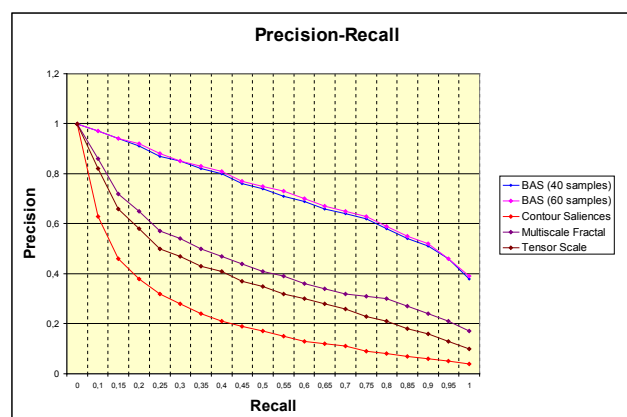
Descritor	Top 1	Top 2	Top 3
BAS (40 amostras)	212	207	200
Dimensão Fractal Multiescala	200	179	154
<i>Tensor Scale</i>	184	154	133
Saliências do Contorno	139	98	75

### 3.3. Base MPEG-7 (Part B)

O terceiro experimento para avaliação dos desempenhos dos descritores de formas foi realizado utilizando-se as imagens da base MPEG-7 Part B (*The MPEG Project*), que possui 1400 imagens de formas, sendo 20 amostras para cada uma das 70 classes de formas. As silhuetas dos objetos nas imagens desta base de dados estão em branco sobre um fundo em preto. Esta base apresenta mais rotações, translações e mudanças na escala, bem como deformações e ruídos nas silhuetas, do que as imagens das bases Kimia. Além disso, há grande variabilidade intraclasse e simi-

laridade interclasses, o que torna esta base bastante desafiadora aos métodos de análise de formas.

A metodologia de avaliação consistiu em calcular as curvas de Precisão-Revocação (*Precision-Recall*) para cada um dos descritores sendo analisados. Essas curvas são exibidas na Figura 6. Vale observar que quanto mais “alta” (mais próxima de “Precisão”=1 e “Revocação”=1) for a curva gerada por um descritor, melhor é o seu desempenho. Assim como nos experimentos realizados com as duas bases de dados Kimia, o BAS teve o melhor desempenho e o Saliências do Contorno o pior.



**Figura 6:** Curvas *Precision-Recall* dos descritores testados sobre a base MPEG-7 Part B.



#### 4. DISCUSSÃO

A partir dos resultados obtidos nos experimentos pode-se perceber que o método BAS apresenta melhor desempenho nos testes realizados com as três bases de imagens e há uma boa diferença entre as taxas de acerto deste em relação ao segundo melhor método, o Dimensão Fractal Multiescada. Mesmo tendo sido proposto em 2003, o método BAS ainda apresenta desempenho superior em relação a outros métodos mais novos. A ideia do funcionamento deste descritor é relativamente simples, mas ele é capaz de analisar e classificar as formas com eficiência.

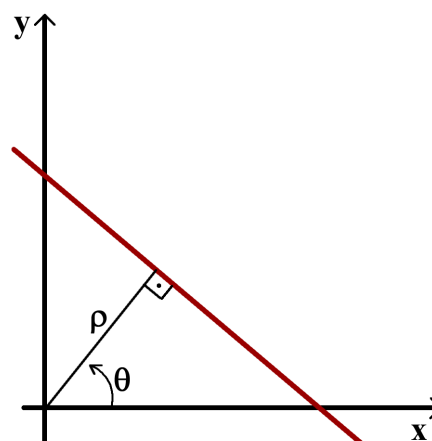
Apesar dos bons resultados do BAS, este descritor pode ser muito lento quando as imagens são grandes ou as bordas apresentam muitos pontos. A complexidade da fase de extração de características do BAS é  $\theta(n^2)$  uma vez que para gerar as funções 1-D associada a uma certa forma é necessário calcular, para cada um dos  $n$  pontos da borda do objeto,  $n/2$  ângulos. A fase de casamento do descritor BAS, a qual emprega o algoritmo OCS, também tem complexidade quadrática.

Com o aumento do poder de armazenamento e a evolução dos sensores de captura, as bases de imagens e as próprias imagens tendem a se tornar cada vez maiores. Desta forma, o BAS pode acabar não sendo indicado para ser utilizado em aplicações onde o baixo tempo de processamento é um dos principais requisitos.

Neste sentido, a criação de novos descritores de formas robustos, mas também ágeis, é um campo de pesquisa promissor. Uma abordagem ainda não utilizada nos trabalhos de análise de formas, mas que pode ser usada na descrição e classificação eficiente de formas corresponde à representação das bordas dos objetos por meio da Transformada de Hough (DUDA e HART, 1972).

Para detectar retas em imagens por meio da Transformada de Hough, Duda e Hart (1972) propuseram a utilização da equação da reta definida por coordenadas polares,  $\rho = x \cdot \cos(\theta) + y \cdot \sin(\theta)$ , uma vez que os parâmetros  $\theta$  e  $\rho$  desta equação são limitados e, portanto, mais indicados que os

parâmetros  $a$  e  $b$  da equação tradicional da reta  $y = a \cdot x + b$ . Como ilustrado na Figura 7, esta parametrização representa cada reta da imagem por meio do ângulo  $\theta$  de seu vetor normal e de sua distância  $\rho$  em relação à origem do sistema (origem da imagem).



**Figura 6:** Representação de retas através dos parâmetros  $\theta$  e  $\rho$ .

Se restringirmos  $\theta$  ao intervalo  $[0, \pi)$ , então os parâmetros  $\rho$  e  $\theta$  de uma dada reta são únicos. Com esta restrição, toda reta no plano  $x$ - $y$  corresponde a um único ponto no plano  $\theta$ - $\rho$  (espaço de parâmetros ou espaço de Hough).

Cada ponto da imagem com suas coordenadas  $(x_p, y_p)$  é representado por uma senóide (obtida a partir das coordenadas do ponto e da equação da reta citada) no espaço  $\theta$ - $\rho$ .

Uma importante propriedade da Transformada de Hough (geração do espaço  $\theta$ - $\rho$  com base no espaço da imagem) é que as senóides do espaço de parâmetros  $\theta$ - $\rho$  que representam pontos colineares da imagem têm um ponto de intersecção comum  $(\theta_0, \rho_0)$  o qual representa a reta da imagem à qual os pontos pertencem.

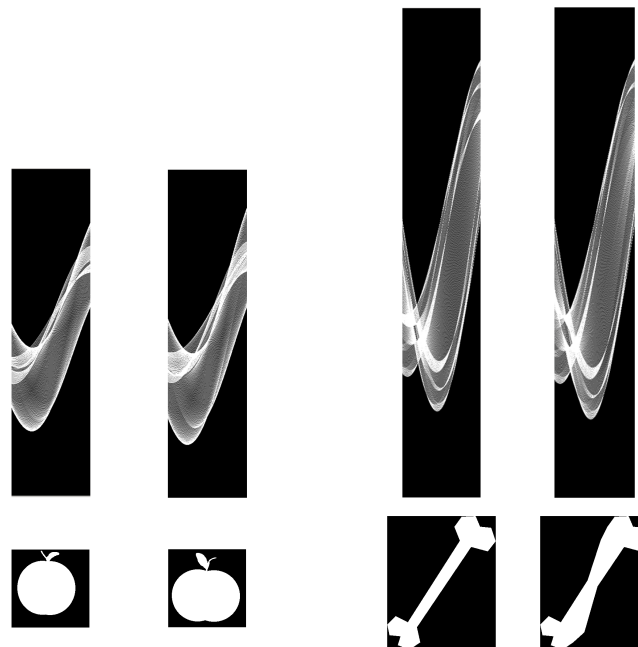
Desta forma, após gerar o espaço de Hough de uma dada forma, isto é, após representar todos seus pontos de borda no espaço de parâmetros  $\theta$ - $\rho$  por meio de senóides, pontos da borda pertencentes a segmentos de reta da borda terão senóides que passam por posições onde há grande quantidade de intersecções de senóides no espaço de parâmetros (posições que representam os segmentos de reta da borda

neste espaço). Por outro lado, pontos de regiões arredondadas da borda terão senóides que não passam por posições do espaço de parâmetros de alta concentração de senóides, isto é, o número de senóides que cruzam a curva senoidal do ponto no espaço  $\theta$ - $\rho$  estará bem distribuído ao longo da extensão da mesma.

Com isto diversos descritores de formas podem ser criados, por exemplo, baseando-se na disposição das intersecções das senóides no espaço de parâmetros ou então na quantidade e disposição de intersecções ao longo de cada senóide. Diversas outras medidas podem ser extraídas deste espaço o qual pode ser representado por uma matriz de inteiros, estrutura de fá-

cil armazenamento e acesso, viabilizando o desenvolvimento de descritores de formas rápidos e robustos, talvez até melhores que o BAS.

A capacidade que os espaços de Hough têm de armazenar informações a respeito de formas a fim de se poder, com base nestas, caracterizar e identificar objetos presentes em imagens é grande. A Figura 8 mostra os espaços de Hough gerados a partir dos pontos de borda de quatro silhuetas de objetos. Nota-se que objetos semelhantes apresentam espaços de Hough similares, enquanto que objetos distintos apresentam espaços de Hough também distintos.



**Figura 8:** Exemplos de espaços de Hough gerados a partir de silhuetas de maçãs (esquerda) e ossos (direita). Cada espaço de Hough foi gerado a partir da borda da respectiva forma. As regiões mais claras do espaço de Hough indicam posições de grande concentração de intersecções de senóides (as quais representam cada um dos pontos da borda da forma da imagem no espaço de parâmetros).

## CONSIDERAÇÕES FINAIS

Com base nos descritores apresentados e nos resultados dos experimentos realizados percebe-se que existem várias abordagens para a caracterização e reconhecimento de formas.

Dependendo da abordagem utilizada, um descritor de formas pode obter melhor ou pior desempenho dada uma base

de imagens. Nos experimentos realizados neste trabalho o descritor bas se configurou como o melhor método de análise de formas.

Apesar dos bons resultados o bas pode se tornar muito lento, em especial quando se trabalha com grandes bases de imagens ou com imagens de grandes dimensões uma vez que tanto para extrair as funções 1-d de uma borda quanto para ca-

sar vetores de características de duas bordas diferentes ele apresenta complexidade  $\theta(n^2)$ .

Desta forma, novos métodos de descrição e identificação de formas continuam surgindo. Independente do enfoque do método na análise de formas, ele deve, pelo menos, ser invariante à translação, rotação e escala.

Uma abordagem que pode ser empregada na análise de formas corresponde ao uso de espaços de hough na representação das mesmas. Estes espaços guardam grande quantidade de informações as quais podem ser extraídas a fim de identificar os objetos. Além disto, a criação do espaço de hough para uma dada forma é simples (usa-se uma matriz de inteiros) e o acesso ao mesmo é bastante rápido.

Desse modo, descritores de formas baseados na transformada de hough se apresentam como uma ideia inovadora e promissora. Extraíndo as medidas certas dos espaços de hough, excelentes resultados (até melhores que os do bas) poderão ser obtidos de maneira mais rápida.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANDALÓ, F. A.; MIRANDA, P. A. V.; TORRES, R. S.; FALCÃO, A. X. A New Shape Descriptor Based on Tensor Scale. **Proceedings of the 8th International Symposium on Mathematical Morphology**, v. 1, p. 141-152, 2007.
- ARICA, N.; VURAL, F. T. Y. BAS: a Perceptual Shape Descriptor Based on the Beam Angle Statistics. **Pattern Recognition Letters**, v. 24, n. 9-10, p. 1627-1639, 2003.
- COSTA, L. F.; JÚNIOR, R. M. C. **Shape Analysis and Classification** – Theory and Practice. Estados Unidos: CRC Press, 2000.
- COSTA, L. F.; CAMPOS, A.; MANOEL, E. An Integrated Approach to Shape Analysis: Results and Perspectives. **Proceedings of the International Conference on Quality Control by Artificial Vision**, p. 23-34, Maio, 2001.
- DUDA, R. O.; HART, P. E. Use of the Hough Transformation to Detect Lines and Curves in Pictures. **Communications of ACM**, n. 1, p. 1-15, 1972.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. Estados Unidos: Wiley-Interscience, 2000.
- FALGUERA, J. R.; MARANA, A. N. **Reconhecimento Automático de Sinus Frontais para Identificação Humana Forense Baseada na Transformada Imagem-Floresta e no Contexto da Forma**, 2008. Dissertação (Mestrado em Ciência da Computação) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista (UNESP), São José do Rio Preto. Disponível em: <<http://www.dcce.ibilce.unesp.br/ppgcc/dissert/Diss-03-Juan.pdf>>. Acesso em: 10 jun. 2012.
- LATECKI, L. J.; LAKÄMPER, R. Shape Similarity Measure Based on Correspondence of Visual Parts. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 10, p. 1185-1190, 2000.
- MANDELBROT, B. **The Fractal Geometry of Nature**. Estados Unidos: H. Freeman and Co., 1982.
- MIRANDA, P. A. V.; TORRES, R. S.; FALCÃO, A. X. TSD: a Shape Descriptor Based on a Distribution of Tensor Scale Local Orientation. **Proceedings of Brazilian Symposium on Computer Graphics and Image Processing**, p. 139-146, 2005.
- MOKHTARIAN, F.; ABBASI, S. Shape Similarity Retrieval under Affine Transforms. **Pattern Recognition**, n. 35, v. 1, p. 31-41, 2002.
- SAHA, P. K. Tensor Scale: a Local Morphometric Parameter with Applications to Computer Vision and Image Processing. **Computer Vision and Image Understanding**, n. 3, p. 384-413, 2005.
- SEBASTIAN, T.; KLEIN, P.; KIMIA, B. Recognition of Shapes by Editing their Shock Graphs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 26, p. 551-571, Maio, 2004.
- TORRES, R. S.; FALCÃO, A. X.; COSTA, L. F. A Graph-Based Approach for Multiscale Shape Analysis. **Pattern Recognition**, n. 37, p. 1163-1174, 2004.
- TORRES, R. S.; FALCÃO, A. X. Contour Saliency Descriptors for Effective Image Retrieval and Analysis. **Image and Video Computing**, n. 25, p. 3-13, 2007.
- The MPEG Project**. Disponível em: <<http://www.chiariglione.org/mpeg>>. Acesso em: 10 jun. 2012.
- WANG, Y. P.; PAVLIDIS, T. Optimal Correspondence of String Subsequences. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, n. 11, v. 12, p. 1080-1087, 1990.

**Gustavo Botelho de Souza** é técnico em informática pelo Colégio Técnico Industrial da Universidade Estadual Paulista (UNESP), campus de Bauru, e bacharel em Ciência da Computação pela mesma universidade. Atualmente é aluno de mestrado (em Ciência da Computação) da UNESP, campus de São José do Rio Preto. Tem interesse nas áreas de pesquisa: análise de imagens, visão computacional e biometria.

**Aparecido Nilceu Marana** é graduado em Matemática pelo Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto. Mestre em Ciência da Computação pelo Instituto de Computação da UNICAMP, Campinas, e doutor em Engenharia Elétrica pela Faculdade de Engenharia Elétrica e de Computação da UNICAMP, Campinas. Realizou pós-doutorado na Michigan State University na área de biometria. É livre-docente em Sistemas Biométricos pela Universidade Estadual Paulista. Também tem interesse nas áreas de pesquisa: análise de imagens, visão computacional e biometria.