

# CLASSIFICAÇÃO DE TEXTURAS POR MODELAGEM HMM DE CARACTERÍSTICAS AM-FM

E. O. T. Salles<sup>1</sup>, L. L. Lee<sup>2</sup>

<sup>1</sup>Departamento de Engenharia Elétrica – DEL, Universidade Federal do Espírito Santo – UFES  
Caixa Postal 01-9011, 29090-970 Vitória ES – Brasil, email: e.salles@ele.ufes.br

<sup>2</sup>Departamento de Comunicações–FEEC, Universidade Estadual de Campinas – UNICAMP  
Caixa Postal 6101, 13083-970 Campinas SP – Brasil, email: lee@decom.fee.unicamp.br

## RESUMO

Neste artigo é abordado o problema de classificação de texturas pelo uso da modelagem HMM de características AM-FM. Para tal fim, ditas características são extraídas empregando-se um banco de filtros com resolução espectral ao estilo das Wavelets de Gabor. O sistema foi modificado para também classificar texturas em presença de rotação, onde adaptou-se um procedimento de correção do giro à abordagem inicial. Como banco de dados de imagens, empregou-se texturas digitalizadas a partir do álbum de Phill Brodatz e os resultados obtidos foram comparados a outras abordagens encontradas na literatura.

## 1. INTRODUÇÃO

O estudo sobre imagens de texturas é quase tão antigo quanto a própria área de processamento de imagens mas, ainda hoje, não existe um modelo formal que as defina [10]. O alto grau de dificuldade envolvida em tal tarefa deve-se à variabilidade dos padrões e a forma pelo qual o ser humano os percebe, entre outros, o que resulta em diversas abordagens envolvendo áreas de conhecimentos distintas. Entretanto, devido à sua aplicação em determinadas atividades industriais, temas como análise, segmentação e classificação de texturas ainda são explorados na literatura. O objetivo deste artigo é avaliar a capacidade de características instantâneas AM-FM (amplitude e frequência moduladas), modeladas pela técnica das cadeias ocultas de Markov (do idioma inglês: *Hidden Markov Models*) na classificação de texturas em presença ou não de rotação

## 2. A MODELAGEM DAS TEXTURAS

Como mostrado em [2][3], a detecção de características AM-FM pode ser obtida empregando-se filtros localizados de Gabor, cada um deles apresentando uma resolução distinta no domínio espaço-frequência espacial. Este tipo de função, disposta segundo uma wavelet, pode aproximar uma textura  $f(x,y)$  por um modelo que apresenta as propriedades de frequência emergente e coerência localizada.

Para a aplicação da técnica, como passo inicial deve-se transformar  $f(x,y)$ ,  $f(\cdot) \in \mathfrak{R}$  em uma imagem complexa  $t(x,y)$ ,  $t(\cdot) \in \mathbb{C}$  a fim de evitar problemas com a representação de componentes instantâneas e, para este fim, a modelagem de uma textura por características AM-FM pode ser representada pelo modelo multicomponente a seguir [3]:

$$t(\mathbf{x}) = \sum_{k=1}^K a_k(\mathbf{x}) e^{-j\phi_k(\mathbf{x})}, \text{ for } t(\mathbf{x}) : \mathfrak{R} \rightarrow \mathbb{C}, \mathbf{x} \in \mathbb{Z}^2, \quad (1)$$

onde  $\mathbf{x}=[x,y]^T$ ,  $a_k(\mathbf{x})$  é a  $k^{\text{th}}$  componente instantânea AM,  $\phi_k(\mathbf{x})$  é a  $k^{\text{th}}$  componente de fase instantânea com uma frequência instantânea associada dada por  $\omega_k(\mathbf{x})=\nabla\phi_k(\mathbf{x})$ . Como apresentado em [3][4], o emprego da equação (1) passa pela definição das propriedades de frequência emergente e coerência localizada: texturas não homogêneas apresentam comportamento localizado homogêneo onde é possível assumir que as componentes AM e FM são descritas por funções suaves. Portanto, é preciso empregar algum procedimento localizado a fim de isolar tais características e a ferramenta é proporcionada pelos filtros localizados de Gabor, como descrito a seguir.

## 3. AS WAVELETS DE GABOR

A estrutura de multiresolução empregada, conhecida como Wavelet de Gabor [10], é definida a partir de uma gaussiana bidimensional modulando uma exponencial complexa que se propaga em direções distintas, conforme apresentado pela equação (2) abaixo ( $i = \sqrt{-1}$ ):

$$h_{r,j}(x, y) = g(x', y') \times \exp[i2\pi(Ux'+Vy')], \quad (2)$$

onde:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right], \quad (2.a)$$

$$\begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = y \cos \theta - x \sin \theta \end{cases}, \quad (3)$$

$\theta=j\pi/J$ ,  $j \in \{0, \dots, J-1\}$ ,  $r \in \{0, \dots, R-1\}$  e  $\omega_r=(U^2+V^2)^{0.5}$ ;

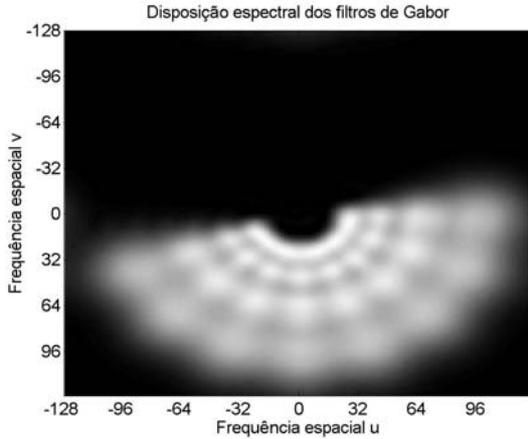
$(U, V)$  seleciona o posicionamento frequencial dos filtros,

$(\sigma_x, \sigma_y)$  determina a largura de banda de cada filtro  $h_{r,j}(x,y)$ ;

$J$  define o número de orientações discretas;

$R$  expressa a quantidade de frequências radiais  $\omega_r$ .

A transformação proposta em (3) garante que um dado filtro pode ser orientado em um determinado ângulo. A figura 1 apresenta a resposta espectral de todos os filtros que caracterizam o banco empregado, onde é possível visualizar o efeito das múltiplas resoluções. Observe que o ângulo  $\theta$  controla a orientação no domínio da frequência espacial



**Figura 1:** Resposta espectral do banco de filtros.

É importante salientar que para obter a propriedade de multiresolução é necessário controlar  $(U, V)$  e  $(\sigma_x, \sigma_y)$ . Novamente, os limites dos filtros realçam a curva de meia magnitude.

Assim sendo, a imagem original é convoluída com cada um dos filtros a fim de se isolar componentes que descrevam convenientemente uma textura. O resultado da operação  $y_{r,j}(\mathbf{x}) = t(\mathbf{x}) * h_{r,j}(\mathbf{x})$  (\* representa a convolução bidimensional), realizada sobre cada canal, é um mapa de características 2D que, de acordo com a aproximação das quase-autofunções, pode ser escrito pela equação (4) abaixo:

$$y_{r,j}(\mathbf{x}) \approx t(\mathbf{x}) H_{r,j}(\nabla \varphi(\mathbf{x})), \text{ for } t(\mathbf{x}) : \mathfrak{R} \rightarrow \mathbb{C}, \mathbf{x} \in \mathbb{Z}^2, \quad (4)$$

onde  $H_{r,j}(\nabla \varphi(\mathbf{x}))$  é a transformada de Fourier do filtro  $h_{r,j}(\mathbf{x})$ , avaliada em  $\nabla \varphi(\mathbf{x})$ . Novamente, tal aproximação só é possível de ser empregada devido às boas propriedades de localização dos filtros de Gabor.

Para a obtenção das componentes AM-FM, aplica-se um procedimento de demodulação sobre  $y_{r,j}(\mathbf{x})$  [4], obtendo-se os mapas instantâneos 2D de amplitude  $a(\mathbf{x})$  e de frequência  $\nabla \varphi(\mathbf{x})$ . No intuito de reduzir a dimensão das características extraídas, optou-se pela aplicação de um procedimento de redução de dimensionalidade obtendo-se o valor médio de cada  $y_{r,j}(\mathbf{x})$ . Assim, uma classe de textura pode ser representada apenas por um conjunto de vetores, cada um deles de baixa dimensão. O uso do valor médio mostrou ser representativo dado a capacidade do banco de filtro em detectar mapas instantâneos suaves. Outras características foram obtidas a partir de  $a(\mathbf{x})$  e  $\nabla \varphi(\mathbf{x})$ , perfazendo um total de cinco componentes (porém, nem todas são utilizadas simultaneamente). A saída de cada canal é representada por um vetor de dimensão reduzida, segundo a equação (5):

$$\mathbf{v}_{r,j} = [c_0(\omega_r, \theta_j), \dots, c_k(\omega_r, \theta_j), \dots, c_{K-1}(\omega_r, \theta_j)]^T, \quad (5)$$

onde  $k \in \{0, \dots, K-1\}$ , para  $K$  mapas instantâneos e  $c_k(\omega_r, \theta_j)$  expressa o resultado do valor médio.

## 4. O MODELAMENTO POR HMM

O processo de classificação escolhido se dá por extração de características que sejam representativas das classes de textura

que, como se vê em (5), gera uma sequência de vetores obtidos na saída de cada filtro de Gabor. Para diversos exemplares de uma única classe, o sistema proposto gera sequências de vetores de características representativos desta classe. Tais vetores traduzem como evolui as características AM-FM de uma dada textura. Ou seja, admite-se que uma boa representação é obtida assumindo-se um conjunto de vetores de observações que, como tal, podem ser modelados por alguma técnica estocástica.

Para lidar com este problema, neste artigo é empregado um procedimento de aproximação estocástica, por meio dos modelos ocultos de Markov (HMM) na classificação de texturas, baseando-se no desenvolvimento apresentado em [7][8]. Seu uso proporciona uma forma eficiente para lidar com as relações entre vetores de uma sequência de observações que sabidamente são correlatados, sem a necessidade de modelar toda as dependências existentes.

Uma cadeia de Markov assume que o processo a ser modelado se manifesta pela evolução estatística de um número finito de estados e a probabilidade de ocupação de um estado é determinado pela história recente de tais evoluções. Para o caso de modelos de Markov de primeira ordem, ditas probabilidades dependem, em um dado estado, tão somente do estado anterior. Isto posto, os HMM's se caracterizam pela evolução de dois processos estocásticos a saber:

- um processo não observável relativo à propriedade markoviana;
- um processo observável relativo à estatísticas da sequência de dados em sí.

A combinação desses dois processos confere o poder dos HMM's em capturar a evolução estatística de uma série estocástica. Seu treinamento é levado a cabo pelo algoritmo de Baum-Welch, estimando-se uma função de verossimilhança que melhor descreve os vetores de observações. Em geral, é desejável atribuir sentido físico aos estados de um HMM através da identificação de primitivas, que são dependentes fortemente do tipo de dado a ser descrito [1]. Entretanto, devido à falta de um modelo formal que descreva uma textura, tais primitivas são ainda vagas. Uma proposta para o problema, como apresentado em [7][8], é a escolha do modelo K-Means Segmental HMM, discutido inicialmente em [6], que também estima uma função de verossimilhança para cada classe porém fazendo uso do algoritmo de Viterbi.

Na presente abordagem, associa-se um determinado HMM para cada classe de textura. Durante a fase de classificação, emprega-se o algoritmo de Viterbi sobre todos os modelos treinados previamente e decide-se pela classe por aquele que apresente como resultado a máxima verossimilhança. Abaixo são mostrados as etapas de classificação de texturas não giradas.

### Procedimento de Treinamento

1. para uma determinada frequência radial  $\omega_r$  são realizadas  $K \times J$  medidas ( $K$  características para cada um dos  $J$  filtros orientados),
2. para  $k \in \{0, \dots, K-1\}$ ,  $r \in \{0, \dots, R-1\}$  e  $j \in \{0, \dots, J-1\}$ , monta-se o vetor de evolução das características  $\mathbf{v}_{r,j} = [c_0, c_1, \dots, c_{k-1}]^T$  e  $k \leq K$ ,
3. monta-se  $O_p \in \{v_{0,0}, v_{0,1}, \dots, v_{0,J-1}, v_{1,0}, \dots, v_{R-1,J-1}\}$ , sendo  $O_p$  uma sequência de vetores de observações para  $p \in \{1, \dots, T\}$ ,

- de posse das diversas seqüências  $O_p$ , treina-se um HMM para cada classe de textura.

**Procedimento de classificação de texturas:**

Para uma seqüência de testes  $O$ :

- executa-se os passos 1. 2. e 3. da etapa de treinamento,
- aplica-se o algoritmo de Viterbi e seleciona-se a classe vencedora.

## 5. O PROBLEMA DA ROTAÇÃO DE TEXTURAS

Para tratar o problema de rotação de texturas, adotou-se a técnica descrita na seção anterior, incorporando-se características invariantes à rotação. É importante ressaltar que dado um banco de filtros e dado uma frequência radial  $\omega_r = \omega_0$ , é possível obter  $v_{r,j}$  em diferentes direções. Neste caso, não é realizada a transformação analítica sobre a imagem inicial a fim de se aproveitar da propriedade do conjugado simétrico que aparece na transformação de imagens reais ( $f(\cdot) \in \mathfrak{R}$ ). Portanto para cada  $c_k(\omega_r, \theta_j)$ , avaliado para um par fixo  $(k, r)$  e  $j = \{0, \dots, J-1\}$ , pode-se interpretar tais medidas, indexadas por  $j$ , como um período de uma seqüência  $e$ , de acordo com a expansão em série de Fourier, é possível transformar esta seqüência em uma segunda, agora independente da rotação, conforme demonstrado em [11]. Assim, novamente, o treinamento e a classificação de amostras de texturas podem ser implementadas utilizando-se a técnica do HMM. Os passos a seguir ilustram o explicado:

- para uma mesma  $\omega_r$ ,  $c_k(\omega_r, \theta_j) \neq c_k(\omega_r, \theta_m)$  com  $j \neq m$ ,  $j, m \in \{0, \dots, J-1\}$ ,  $k$  indexando uma das características extraídas,
- obtem-se  $c_k(\omega_r, \theta_j)$  a partir de uma textura e extrai-se  $c_k(\omega_r, \theta_m)$  da mesma textura, agora girada de  $\Delta\theta = \theta_k - \theta_m$ , proporciona  $c_k(\omega_r, \theta_j) = c_k(\omega_r, \theta_m)$ ,
- para  $d_n = \frac{1}{J} \sum_{j=0}^{J-1} c_k(\omega_r, \theta_j) e^{-i2\pi jn/J}$ ,  $i = \sqrt{-1}$ ;  $|d_n| = |d_{J-n}|$ ,  $n \in \{0, \dots, J/2-1\}$ , toma-se o valor absoluto de  $d_n$ , o que significa retirar a dependência angular das características.

Isto posto, as etapas necessárias de treinamento e classificação de texturas giradas são apresentadas a seguir:

**Procedimento de Treinamento**

- para  $k \in \{0, \dots, K-1\}$ ,  $r \in \{0, \dots, R-1\}$  e  $j \in \{0, \dots, J-1\}$ , monta-se o vetor de evolução das características  $v_{r,j} = [c_0, c_1, \dots, c_{k-1}]^T$  e  $k \leq K$ ,
- uma vez obtidos  $v_{r,j}$ , toma-se todos os correspondentes  $c_k$ , para cada um dos  $J$  filtros, e corrige-se a rotação. Os novos vetores são  $r_{r,m}$ ,  $n \in \{0, \dots, J/2-1\}$  ( $J$  par),
- constrói-se a seqüência de observações  $O_p$  de tal forma que  $O_p = \{r_{0,0}, r_{0,1}, \dots, r_{1,J/2-1}, r_{1,1}, \dots, r_{R-1,J/2-1}\}$ ,  $p \in \{1, 2, \dots, T\}$ ,
- dada as diversas seqüências de treinamento  $O_p$ , treina-se um HMM para cada classe de textura.

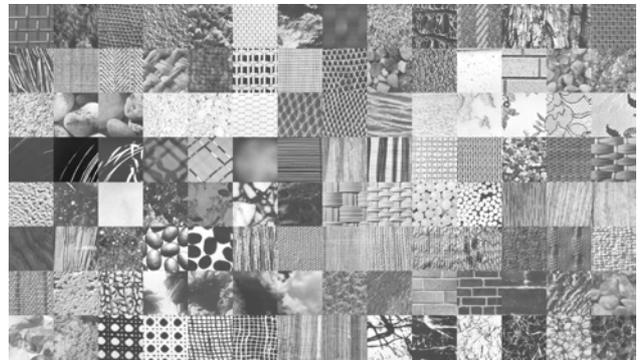
**Procedimento de classificação de texturas:**

Para uma seqüência de testes  $O$ :

- executa-se os passos 1. 2. 3. da etapa de treinamento,
- aplica-se o algoritmo de Viterbi e seleciona-se a classe.

## 6. RESULTADOS EXPERIMENTAIS

Todas as 112 imagens, obtidas a partir do álbum de P. Brodatz [5] e referenciadas de D1 a D112, foram digitalizadas em 512x512 pixels e 8 bits por pixel através de um scanner de 200 dpi. Cada imagem foi dividida em sub-imagens para a definição das amostras de treinamento e teste (amostras com valor médio igual a zero). Gerou-se sub-imagens com dimensões de ora 64x64 pixels (32 amostras para teste e 32 para avaliação) ora de 128x128 pixels (8 amostras para treinamento e 8 para avaliação). Como tomou-se tais amostras sem superposição, e a fim de manter uma boa quantidade de exemplos, dimensões maiores que 128x128 foram evitadas. Dimensões menores de 64x64 não foram empregadas pois para algumas classes não haveria uma boa representatividade da textura. Em geral, testes demonstraram melhores resultados para amostras de dimensões 128x128 pixels. As 112 texturas são mostradas na figura 2.



**Figura 2:** Um mosaico de todas as 112 texturas.

A fim de mostrar a robustez do sistema proposto, sobretudo no que tange à rotação de texturas, o sistema foi treinado apenas com texturas não giradas e testado tanto para amostras não giradas como giradas. Os ângulos de rotação empregados foram: 30°, 45°, 60°, 90°, 120° e 150°.

Com o objetivo de avaliar o desempenho da técnica, foram criados diversos subgrupos de amostras, obtidos a partir dos 112 exemplos de textura, subgrupos estes desenvolvidos segundo os grupos propostos em alguns trabalhos encontrados na literatura.

Subgrupo 1: todas as 112 texturas;

Subgrupo 2: D3, D4, D6, D9, D11, D16, D19, D21, D24, D29, D34, D36, D52, D53, D55, D57, D65, D68, D74, D77, D78, D79, D82, D83, D84, D92, D95, D102, D103 e D105;

Subgrupo 3: D9, D12, D15, D17, D19, D24, D29, D38, D68, D84, D92, D94 e D112; amostras girados no sentido anti-horário segundo os ângulos: 0° (somente para treinamento) 30°, 45°, 60°, 90°, 120°, 150°;

Subgrupo 4: D10, D12, D15, D17, D22, D24, D34, D37, D53, D56, D65, D68, D80, D82 e D84; amostras girados no sentido anti-horário segundo os ângulos: 30°, 45°, 60°, 90°, 120°, 150° (texturas anisotrópicas).

Na tabela 1 é sumarizado o desempenho do método proposto e de outras técnicas encontradas na literatura. A terceira coluna apresenta aspectos relativos às etapas de treinamento e teste.

Subgrupo	Referência	Comentários	Resultados
1	[9]	2 e 6	80,6%
	[este trabalho]	1 e 6	85,94%
	[este trabalho]	2 e 6	84%
2	[9]	2 e 6	100%
	[este trabalho]	2 e 6	99,17%
	[este trabalho]	1 e 6	100%
	[11]	1 e 6	99,35%
3	[9]	3 e 6	96,8%
	[este trabalho]	4 e 5	90,38%
4	[12]	3 e 6	89,3%
	[este trabalho]	4 e 5	88,33%

**Tabela 1:** Na coluna “Comentários”: 1: Sistema desenvolvido para texturas não giradas, 2: Sistema desenvolvido para texturas giradas, 3: Treinamento com texturas giradas, 4: Treinamento com texturas não giradas, 5: Amostras de imagens de 64×64 pixels, 6: Amostras de imagens de 128×128 pixels.

Para o subgrupo 1 de texturas a taxa obtida pela presente abordagem obteve um resultado superior a técnica apresentada em [9]. Apesar da forma de extração de característica em ambos os casos basear-se nas Wavelets de Gabor, a modelagem HMM de características AM–FM mostrou-se mais efetiva.

Para o subgrupo 2 observe que o emprego do procedimento para tratar texturas giradas na classificação de texturas não giradas gera uma pequena degradação na taxa alcançada (99,17%). Isto se deve ao procedimento de invariância à rotação que descartar informações relevantes para uma correta classificação.

Para o subgrupo 3, apesar do sistema ter sido treinado apenas com amostras de texturas não giradas, diferentemente do realizado em [9], a taxa de acertos final ficou comparável à obtida para o caso de treinamento com texturas giradas, demonstrando a robustez do método.

Finalmente, para o subgrupo 4 composto exclusivamente de texturas anisotrópicas, a taxa obtida foi comparável, ainda que menor, a aquela alcançada por [12]. Entretanto, obteve-se o valor de 88,33% com sistema treinado para texturas não giradas.

## 7. CONCLUSÕES

Neste artigo é proposta uma abordagem alternativa para o problema de classificação de texturas baseando-se no sequenciamento de características AM–FM. O emprego da técnica HMM mostrou-se eficiente também para o problema de classificação de texturas em presença de rotação. Especialmente neste caso, o presente método mesmo tendo sido treinado para texturas não giradas, proporcionou um bom desempenho, comparável a outros resultados apresentados na literatura onde o treinamento fora realizado com texturas giradas.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

[1] K. Aas, L. Eikvil, R. B. Huseby, “Applications of Hidden Markov Chains in Image Analysis”, *Pattern Recognition*, 32, (1999), pp. 703-713.  
[2] A. C. Bovik, N. Gopal, T. Emmoth, A. Respreto (Palacios), “Localized Measurement of Emergent Image Frequencies by

Gabor Wavelets”, *IEEE Trans. On Information Theory*, vol. 38, no.2 march 1992, pp. 644-664.

[3] J. P. Havlicek, A. C. Bovik, “AM–FM Models, the Analytic Image, and Nonlinear Demodulations Techniques”, *Center for Vision and Images Sciences*, CVIS –TR-95-001, march, 1995.

[4] J. P. Havlicek, D. S. Harding, A. C. Bovik, “Multidimensional Quasi-Eigenfunction Approximations and Multicomponent AM–FM Models”, *IEEE Trans. On Image Processing*, vol. 9, no. 2, february 2000, pp. 227-241.

[5] P. Brodatz, “*Textures: A Photographic Album for Artists and Designers*”, Mineola, NY: Dover Publications, 1966.

[6] B.-Hwang Juang, L. R. Rabiner, “The Segmental K-Means Algorithm for Estimating Parameters of Hidden Markov Models”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 9, september 1990, pp. 1639-1641.

[7] J.-L. Chen, A. Kundu, “Rotation and Gray Scale Transform Invariant Texture Identification Using Wavelet Decomposition and Hidden Markov Models”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, february 1994, pp. 208-214.

[8] J.-L. Chen, A. Kundu, “Unsupervised Texture Segmentation Using Multichannel Decomposition and Hidden Markov Models”, *IEEE Trans. On Image Processing*, vol. 4, no. 5, may 1995, pp. 603-619.

[9] G. M. Haley, B. S. Manjunath, “Rotation-Invariant Texture Classification Using a Complete Space-Frequency Model”, *IEEE Trans. on Image Processing*, vol. 8, n. 2, february 1999, pp. 255-269.

[10] S. Mallat, *A Wavelet Tour to Signal Processing*, Academic Press - 1<sup>a</sup> Edition, 1998.

[11] A. Mojsilovic, M. V. Ppopovic, D. M. Rackov, “On the Selection of an optimal Wavelet Basis for Texture Characterization”, *IEEE Trans. on Image Processing*, vol. 9, n. 12, december 2000, pp. 2043-2050.

[12] T. N. Tan, “Rotation-Invariant Texture Features and Their Use in Automatic Script Identification”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, july 1998, pp. 751-756.