

# Uma Abordagem para Indexação e Recuperação Automática de Imagens Baseada em Vetor de Comportamento em Sistemas de Inspeção e Montagem

MARCELO RUDEK

*Laboratório de Automação e Sistemas – LAS  
Pontifícia Universidade Católica do Paraná PUC PR  
R. Imaculada Conceição, 1155 Prado Velho - Curitiba PR  
e-mail: rudek@rla01.pucpr.br*

**Resumo** – Um processo de indexação automática de imagens, associado aos procedimentos modernos de manufatura, deve permitir a classificação de imagens obtidas de células de trabalho, e prover o armazenamento de informações relevantes ao processo produtivo. Como também, permitir a recuperação de imagens de um banco de dados, por meio de consultas baseadas em exemplos, que possam suprir requerimentos de sistemas robóticos. Ao contrário dos métodos baseados em reconhecimento de objetos onde busca-se um componente específico da imagem, esta proposta apresentada uma forma de codificar a interpretação visual de um cenário, ou de uma peça, para a indexação e recuperação da imagem. Este artigo apresenta a proposta de um método de indexação automática e uma aplicação protótipo, fundamentada no conceito de análise do comportamento apresentado em (Rudek, 99).

**Abstract** – An automatic images indexing process, assigned to modern manufacturing requirements, must allow the images classification for work cells and provides the storage of important informations for productive process. Also, It must allow recovering images from an image data base, by examples queries, to supply the robotics systems. In contrast with the traditional object recognition methods, this propose presents a way to codify the visual scenary interpretation, or part of one product, by indexing and recovering the image. This paper presents an automatic indexing method based in the behavior analysis concepts, present in (Rudek,99).

**Key Words** – Computer Vision, Image Index, Robotics, Manufacture System.

## 1 Introdução

O tratamento adequado de uma imagem, através de um sistema automático de reconhecimento e indexação, pode fornecer um nível de organização de imagens, de forma mais natural e/ou conveniente para dispositivos que operam em sistemas de manufatura.

A utilização de vetores de atributos formados pela extração de informações, tais como cores, texturas e regiões importantes, é uma das soluções que pode ser utilizada para a representação de uma imagem em um banco de dados com imagens. Desta forma os sistemas de busca podem operar em um espaço reduzido de informações, usando os próprios atributos como índices para as imagens. Trabalhando em um espaço reduzido, sobre um vetor de atributos, tem-se uma diminuição da complexidade computacional, mas dependendo da profundidade da análise feita sobre a imagem, existe a desvantagem de se ter um vetor de atributos muito grande, o que pode tornar o processo de indexação muito lento. Com o uso de histogramas pode-se obter a distribuição de cores da imagem, mas contudo, não é possível obter a localização da cor. Segundo (YIHONG, 1996) é possível dividir a imagem em várias sub-regiões e criar um histograma para cada porção da imagem. Assim, pode-se localizar a maior distribuição de cor (cores mais significativas) para cada região, e usar estes histogramas para indexar uma imagem.

A informação semântica da imagem, ou seja, a forma que as pessoas interpretam ou entendem uma imagem é o objetivo deste estudo. Neste artigo é

proposto uma forma de organizar a distribuição de cores de uma imagem de maneira que se possa expressar o seu comportamento. Como apresentado na figura 1, o processamento de imagens realiza a interface entre os dispositivos de aquisição e armazenamento/repositório de informações. Este trabalho apresenta as definições e resultados de simulações, sobre os processamentos realizados pelos métodos representados dentro da região pontilhada da figura 1.

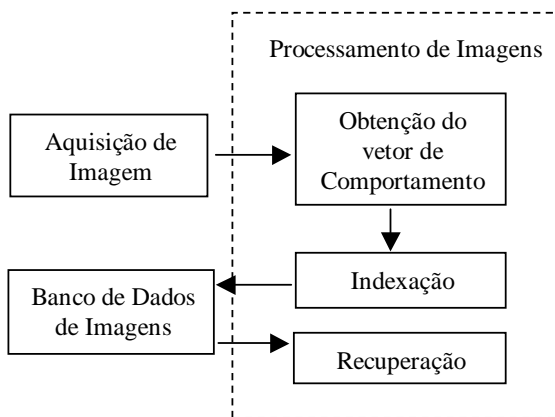


Figura 1. Visão geral do processo de indexação e recuperação de imagem.

## 2 Definição do Vetor de Comportamento

Uma imagem pode ser dividida na forma de uma matriz quadrada de dimensões  $M \times N$ . Por exemplo, se  $M=N=3$ , teremos 9 sub-áreas de tamanhos iguais, que podem ser ordenadas da maneira apresentada pela figura 2. A

distribuição dada pela matriz, gera o vetor de comportamento, que é uma estrutura que armazena a forma pela qual as sub-áreas de uma imagem se relacionam. A maneira como este vetor é composto, determina um comportamento específico para um grupo de imagens. Para obter o comportamento descrito por uma imagem, conforme apresentado na figura 2, deve-se extrair para cada região a cor mais significativa, ou seja, aquela que tem a maior distribuição dentro de cada sub-área.

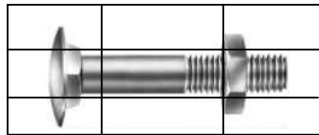


Figura 2. Divisão da imagem em 9 sub-áreas

Uma vez extraída a maior distribuição de cor para cada parte da imagem, dada pelo valor de pico do histograma, pode-se organizar as distribuições na forma de um vetor, denominado por (Rudek,99) vetor de comportamento.

O vetor de comportamento, pode ser obtido da forma apresentada na figura 3, agrupando as sub-regiões da esquerda para a direita e de cima para baixo, conforme representado na figura 4.

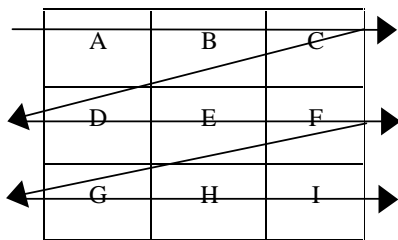


Figura 3. Descrição da seqüência de agrupamento das subáreas para formar o vetor de comportamento.

Pode-se observar pelo apresentado na figura 3, a configuração de um arranjo do relacionamento para cada parte da imagem, na forma de um vetor, como mostra a figura 4.

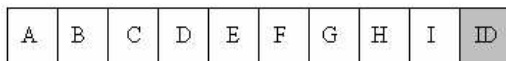


Figura 4. Formato do vetor de comportamento obtido da distribuição de regiões. Um elemento adicional é inserido no final do vetor para diferenciar vetores iguais.

Para cada região é realizada a contagem dos pixels. A distribuição dos pixels é classificada de acordo com a distribuição de frequência, conforme mostrado na tabela 1. Cada *pixel* da imagem é classificado de acordo com o intervalo de distribuição da tabela 1.

A classe que detém a maior quantidade de pixels por região, é associada ao vetor de comportamento. Conseqüentemente, para a imagem da figura 2, cada elemento do vetor representa uma classe de distribuição de cores, obtendo-se o vetor de comportamento dado por [FF FE58FFF][0].

Tabela 1. Descrição das classes e intervalo de distribuição de cores.

Classe	Intervalo de Distribuição
1	000-016
2	017-033
3	034-050
4	051-067
5	068-084
6	085-101
7	102-118
8	119-135
9	136-152
A	153-169
B	170-186
C	187-203
D	204-220
E	221-237
F	238-255

### 3 Indexação

O vetor de comportamento representa um subespaço de informações que caracteriza uma imagem. Cada imagem armazenada pode ser associada por um vetor de comportamento correspondente (que é o índice para cada uma). O vetor é gerado a medida que a imagem é entrada no sistema.

Após a criação do vetor, a imagem pode ser armazenada numa área de disco, referenciada pelo seu índice. Assim, todas as operações de consulta são realizadas sobre os índices da imagem, o que torna o processo de recuperação da informação robusto e rápido. O procedimento de indexação é dado pela:

- i. Criação do índice (vetor de comportamento);
- ii. Associação do índice à imagem;
- iii. Armazenamento ou busca da imagem no banco de imagens.

A indexação deve ser associada a um método de busca. Através da análise de comportamento entre uma imagem exemplo (amostra) e os índices do banco de dados, é possível encontrar a imagem procurada, ou um grupo de imagens semelhantes, que possuem o mesmo comportamento. Graficamente, o comportamento da imagem pode ser representado pela figura 5, onde a curva descrita entre os valores das classes de distribuição (eixo da abcissa) e o vetor de comportamento (eixo da ordenada) caracteriza o comportamento da imagem.

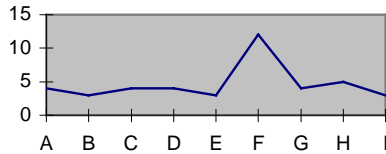


Figura 5. Representação gráfica do comportamento da imagem. Onde o eixo das abscissas corresponde às classes de distribuição e o eixo das ordenadas às posições do vetor de comportamento.

A partir do gráfico da figura 5 que descreve o comportamento da imagem, pode-se construir um descritor simbólico da imagem, como mostra a figura 6.

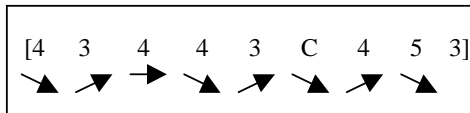


Figura 6. Descritor simbólico do vetor de comportamento.

Com base no descritor simbólico, é possível realizar as comparações entre o elemento de busca e às imagens armazenadas no banco.

#### 4 Recuperação da Informação

Um procedimento de busca deve estar associado a indexação. A forma como os índices são procurados podem influenciar a velocidade de resposta do sistema. No sistema protótipo utilizado para a validação desta proposta, é realizada a comparação de cada elemento do vetor de comportamento (índice) com o descritor simbólico da imagem de exemplo. A razão de igualdade entre os elementos do vetor define a precisão do sistema. Esta precisão é dada pelo nível de significância, que corresponde a um valor numérico entre 1 e N, onde N é o número de elementos do vetor. Quanto maior o nível de significância usado num processo de busca, mais precisa será a resposta, quanto menor, menos precisa a resposta, entretanto pode ocorrer uma melhora no agrupamento de imagens. Um agrupamento de imagens (imagens de mesmo conteúdo semântico) pode ser obtido, diminuindo-se o nível de significância a critério do usuário.

Em síntese, são possíveis três tipos de resposta, que são:

- Existência de total coincidência com apenas uma alternativa de resposta.
- Não ocorrência de total coincidência de valores na comparação dos vetores.
- Existência de total coincidência com mais de uma alternativa de resposta.

#### 5 Sistema Protótipo

A função do sistema protótipo, conforme apresentado na figura 7, é avaliar o método proposto, através do armazenamento e da recuperação de imagens em um

banco de dados de forma automática, sem a utilização de identificação textual. A medida que a varredura é executada sobre a imagem, o sistema identifica qual sub-área da imagem está sendo analisada. Cada sub-área que é processada gera o valor da distribuição de frequência, e passa a compor, campo a campo, o vetor de comportamento. Uma vez criado este vetor que representa um elemento de indexação, a imagem pode ser armazenada, ou servir de elemento de busca. Para a recuperação da imagem no banco, dependendo-se do que se busca, não basta recuperar apenas a imagem original, mas sim obter um grupo de imagens similares como resposta. Para isto pode-se determinar um padrão de semelhança, variando a precisão de busca na comparação dos índices da imagem. Esta precisão é determinada através de um nível de significância. Aumentando o nível de significância, espera-se uma maior precisão da resposta, onde as imagens encontradas devem ser iguais a imagem de exemplo. Ao diminuir o valor do nível de significância, espera-se encontrar um maior número de respostas que possa satisfazer ao critério de agrupamento da busca. A diferença de escala não influi na geração do vetor de comportamento, pois para a distribuição de frequência ocorre diferença apenas da quantidade de *pixels* a ser analisada, mas proporcionalmente não altera sua cor ou forma.

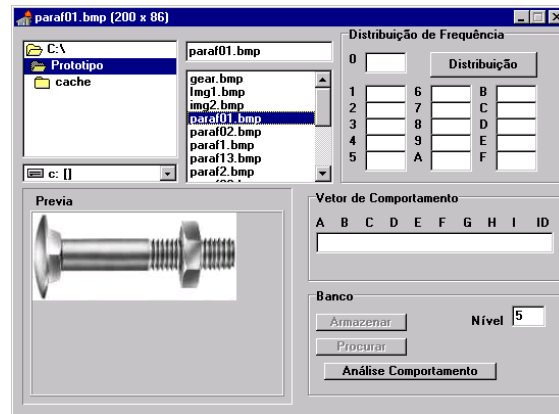


Figura 7. Tela do sistema protótipo

Outro fator em questão, é conhecer qual distribuição de cor é mais significativa. Na distribuição de frequência podem ocorrer empates de valores. Para cada região da imagem, pode ocorrer de não se ter um valor de distribuição bem definido. Então, defrontam-se duas situações. A primeira é considerar o maior valor de distribuição como sendo o valor real, e desconsiderar a ocorrência de empates. Se numa contagem de distribuição ocorrer por exemplo, 350 ocorrências para um índice de cor A, e 351 ocorrências para o índice de cor B, o vetor de comportamento assume o valor B como resposta. Porém, se ambos os valores forem iguais, não há como determinar, qual distribuição vai compor o vetor de comportamento. A segunda situação é definir um limite em que as distribuições podem ser consideradas como de igual valor (empate ajustado). Dois ou mais valores

podem ser considerados iguais, se forem idênticos, ou satisfizerem um limite percentual de proximidade. Na situação dos índices de exemplo a cor A e a cor B, podem ser consideradas empatadas, pois não existe uma diferença significativa entre eles (350 e 351). Para determinar de quanto pode ser o limite de erro para o empate, pode-se verificar a proporcionalidade de distribuição em relação a quantidade de *pixels* de cada região da imagem. O percentual de proximidade pode ser um valor arbitrário, obtido através de diversas simulações, ou pode ser estimado através do seguinte cálculo:

$$Limiar = \frac{(H \times W)}{Nr \times Qd \times 10} \quad (1)$$

onde, H = altura da imagem  
W = largura da imagem  
Nr = número de regiões  
Qd = quantidade de intervalos

Para uma imagem exemplo de 128 X 128 pixels, aplicando a equação (1), tem-se:

$$limiar = (128 \times 128) / (9 \times 15 \times 10) = 12,1 \quad (2)$$

Este limiar é equivalente a 10% do valor médio para cada intervalo de distribuição. Assim, para a situação apresentada, as distribuições que estiverem entre 339 e 351 são dadas como empate. Nesta situação, o elemento '0' é adicionado ao vetor, indicando uma situação de empate. É importante ressaltar que somente é considerado empate, se ele ocorrer entre as classes de distribuição. Dentro da mesma classe não existe empate já que os valores de distribuição se somam, originando o total da classe. Considerando a segunda situação, pode-se observar uma mudança da geração dos vetores de comportamento de acordo com a figura 8.

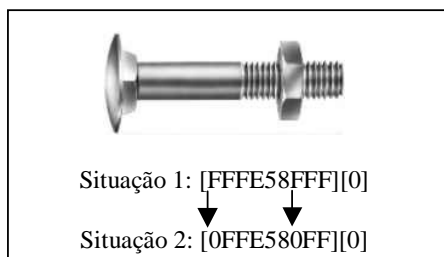


Figura 8. Diferença na geração do vetor de comportamento entre a situação 1, desconsiderando o empate e a situação 2, considerando a possibilidade de empate, dentro do limiar de 10%.

Os testes realizados com alteração do valor do limiar, demonstram que o vetor sofre atenuação, primeiramente em seus vales (pontos de mínimo) e posteriormente, com o aumento do limiar, a atenuação é nos picos (pontos de máximo). Desta forma, concluiu-se que um valor adequado para o limiar, deve ser de 10% do valor médio para cada intervalo, pois

permite identificar casos de empate da distribuição de cores, sem descaracterizar o vetor de comportamento para cada imagem.

## 6 Resultados Obtidos

Para a validação do método apresentado, foi utilizado um banco de teste, com cerca de 350 imagens de propósito geral extraídas da *internet*. Dois aspectos interessantes puderam ser observados. Um diz respeito à presença de agrupamento de imagens em relação a seu significado, como grupo de imagens de pessoas, grupo de imagens de paisagens, grupo de imagens de parafusos, etc. Neste caso ocorreu um total de 70% das respostas satisfatórias. O outro, em relação à semelhança de contexto, pelo posicionamento de seus componentes como cor e textura, complementando os 30% restantes. A partir da imagem de exemplo da figura 9, foi obtido um agrupamento de respostas, como mostra a figura 10.



Figura 9. Imagem de exemplo aplicada ao sistema protótipo.

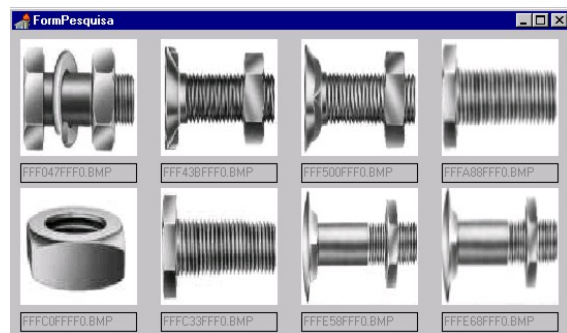


Figura 10. Imagens de resposta fornecidas pelo sistema protótipo, a partir da imagem de exemplo, com nível de significância igual a 5.

Apesar de não se ter no sistema nenhum processo de reconhecimento de objetos, imagens semelhantes podem ser agrupadas com boa precisão. Através dos testes efetuados, observou-se que o vetor de comportamento descreve melhor imagens que possuem grandes distribuições de cor originadas do plano de fundo da imagem (*background*), onde a quantidade de objetos importantes é menor em relação ao fundo, do que às de primeiro plano (*foreground*).

## 7 Aplicações em Automação

Do ponto de vista da automação industrial, principalmente em ambientes de manufatura, os sistemas robóticos que incluem capacidades de operação usando a visão artificial estão cada vez mais presentes. Dependendo do estágio que se encontra a confecção de um produto, (ou parte dele) como soldagem, pintura, montagem, empacotamento e expedição, são requeridos atributos visuais específicos para cada situação.

Na proposta apresentada neste artigo, uma peça pode ser classificada (indexada) para ser armazenada ou comparada com informações presentes em um banco de dados de imagem. A partir da obtenção da imagem, o sistema avalia o comportamento da imagem e busca os grupos de imagens similares no banco. Se a imagem existe no banco, o sistema pode obter por exemplo, as informações construtivas daquela peça, ou se a imagem é algo diferente do que existe no banco de dados, estas características podem ser percebidas pelo *software* que pode criar um nova classe de peças (defeituosas ou novos formatos) e indexá-las no banco para tratamento posterior. A aplicação da técnica do vetor de comportamento, está mais relacionada com processos de classificação de peças e de montagem, onde é necessário uma interação, por exemplo entre um banco de dados e o manipulador de um robô. Do ponto de vista do trabalho realizado por um operador humano em um processo de inspeção, requer um alto nível de concentração, que é obtido apenas na primeira meia hora de trabalho (BERGASA, 2000). Pela complexidade de uma tarefa e requerimentos de operação em tempo real, o uso de técnicas de processamento de imagens podem ser mais eficientes.

## 8 Conclusão

Para sistemas de indexação e recuperação, que manipulam coleções de imagens, a principal meta é fornecer respostas relevantes a partir de uma pesquisa visual, através de uma imagem de exemplo. Isto caracteriza os sistemas de indexação automática, em relação aos sistemas de indexação que necessitam de interação de um usuário ou mesmo dos que se baseiam em descrição textual. Conforme descreve (QIN-HE, 1996), é realmente difícil desenvolver um sistema de indexação e recuperação de imagens, de propósito geral, para grandes bancos de imagens. Também é difícil projetar, desenvolver e avaliar um sistema, sem relacionar com sua área de aplicação. A razão é que para diversos usuários ou propósitos, os requerimentos de indexação e recuperação são muito diferentes. Por isto, um sistema não pode ser usado para todas as aplicações. O mais interessante em termos práticos é se obter um sistema adaptável, capaz de dar resultados depois de um certo treinamento, e também capaz de improvisar sua performance. Sem compreender uma imagem, um sistema automático não pode fornecer um nível de organização de imagens de forma mais natural ou conveniente para as pessoas e aplicações com imagens.

Dentre os vários métodos de indexação, um dos objetivos dos sistemas de recuperação de imagens, é extrair imagens relevantes a partir de uma consulta visual. De acordo com o estudo de (RAVELA, 1995), a relevância está associada com a semântica. De modo geral, a semântica define o significado de uma imagem de acordo com o seu contexto, em termos de

seus atributos (cores, texturas, objetos), ou seja, a forma como o usuário vê e interpreta esta imagem. Decodificar a informação semântica em um sistema geral de indexação, envolve problemas como segmentação, reconhecimento e extração automática de atributos; - e isto ainda é um problema de difícil solução. De acordo com todas as referências bibliográficas utilizadas neste trabalho, e pelos testes realizados com a utilização do método do vetor de comportamento, pode-se concluir que nenhum sistema ainda é capaz de se comportar como o sistema de visão humano para a identificação e classificação de imagens de acordo com seu significado semântico.

A partir do contexto definido neste artigo, pode-se ter como perspectiva futura, a união da técnica de vetor de comportamento com novos elementos computacionais, como algoritmos genéticos (WANG et al., 2001), wavelets (QUAK & WEYRICH, 1994), rede neurais (AUGUSTEIJN & CLEMENS, 1996), sistemas nebulosos (CHENG & XU, 2000) entre outros, para resolver situações que ainda geram complexidade para manipulação de imagens, sobretudo problemas de ruído, condições adversas de iluminação e posicionamento de peças.

## Referências Bibliográficas

- AUGUSTEIJN, M. F; CLEMENS, L.E. A neural-network approach to the detection of texture boundaries; Engineering Applications of Artificial Intelligence, vol 9, n.1, 1996, pp.75-81
- BERGASA, L; DUFFY, N; LACEY, G; MAZO, M; Industrial Inspection Using Gaussian Functions in a Colour Space. Image and Vision Computing 18 (2000), pp. 951-957.
- CHENG, H.D. ; XU, H, X; A novel fuzzy logic approach to contrast enhancement; Pattern Recognition, v. 33, 2000, pp.809-819.
- QIN-HE; World of OAK : A Test of Image Indexing and Retrieval. Graduate School of Library and Information Science - University of Illinois at Urbana-Champaign Fall, 1996.
- QUAK, E.;WEYRICH, N. Decomposition and Reconstruction algorithms for spline wavelets on a bounded interval. Applied and Computational Harmonic Analysis, vol 1, n. 3, 1994, pp.217-231.
- RAVELA, R. MANMATHA , E. M. RISEMAN; Retrieval from Image Databases Using Scale-Space Matching. CmpSci Technical Report pp.95-104, UMASS, Amherst, MA, November 1995.
- RUDEK, M. Uma proposta para indexação e recuperação de cheques bancários através de vetor de comportamento. Dissertação de Mestrado, CEFET PR, 1999.
- WANG, R-Z; LIN, C.-F.; LIN, J.-C. Image hiding by optimal LSB substitution and genetic algorithm, Pattern Recognition, vol 34, 2001, pp. 671-683.
- YIHONG GONG; CHUA HOCK CHUAN; GUO XIAOYI; Image Index and Retrieval Based on Color Histograms; Multimedia Tools and Applications, 1996, vol-2, n.2, pp. 133-156.