

Classificação Automática de Imagens do Acervo de Obras de Candido Portinari

Daniel Pacheco de Queiroz, Antonio da Luz Júnior e Arnaldo de Albuquerque Araújo
Universidade Federal de Minas Gerais - UFMG
Av. Antônio Carlos, 6627 - Pampulha - Belo Horizonte - MG
{pacheco, daluz, arnaldo}@dcc.ufmg.br

Resumo

Métodos eficientes de busca por imagens são cada vez mais necessários e importantes. Nos últimos anos, fatores como a Internet, a disseminação de câmeras digitais, o aumento da capacidade de armazenamento e a diminuição do preço deste armazenamento, fizeram com que a quantidade de informação visual disponível aumentasse consideravelmente. Com isto, surge a necessidade do desenvolvimento de métodos eficientes para recuperação de imagens em grandes bases de dados. Para alcançar este objetivo é necessário descrever o conteúdo das imagens de forma eficaz e de modo que gere dados que possam ser utilizados em uma classificação automática de imagens. Este objetivo é alcançado por meio de atributos relacionados com as características visuais das imagens. Este trabalho visa identificar alguns destes atributos e verificar a eficácia dos atributos selecionados utilizando um classificador automático. Como base foi utilizado o acervo de obras do artista Candido Portinari, cujas imagens digitais se encontram disponíveis para todos e já possui uma classificação manual, com a qual o método proposto pode ser confrontado.

1. Introdução

O desenvolvimento de sistemas capazes de realizar a tarefa de Recuperação de Imagens Baseado em Conteúdo (Content Based Image Retrieval - CBIR) é a solução ideal para o problema de busca por imagens de interesse em grandes bases de dados [4]. Com este tipo de abordagem é possível recuperar imagens utilizando uma descrição do conteúdo desejado, utilizando para isso um esboço (sketching) ou buscando imagens semelhantes às outras (Query By Example - QBE). O processo de CBIR pode ser dividido em duas etapas principais: a) construção da base de dados; b) recuperação de imagens. Idealmente essas etapas ocorrem em momentos distintos. Na etapa de construção da base de dados, todo conjunto de imagens disponível é processado de

forma a extrair informações, denominadas nesta área como atributos ou características visuais, capazes de identificar o conteúdo de cada imagem. Na segunda etapa, que consiste no processo de recuperação de imagens propriamente dito, é recebido como entrada um conjunto de atributos que descrevem o tipo de conteúdo que se espera encontrar nas imagens. Utilizando-se de técnicas de classificação e medidas de similaridade entre objetos, os dados recebidos são comparados com as imagens previamente obtidas no processamento de um conjunto inicial. Por fim, as imagens são organizadas de acordo com sua similaridade com os atributos de entrada.

A base de imagens das obras de Candido Portinari, que foi utilizada durante a elaboração deste trabalho, é composta por um conjunto de mais de 5.000 imagens [7]. Estas imagens foram previamente classificadas, em um esforço manual, de acordo com as técnicas e materiais utilizados em suas criações. Como durante a etapa de digitalização das imagens, não foram observados os critérios de qualidade adequados, este trabalho visa auxiliar na classificação das imagens caso seja realizada uma nova etapa de digitalização para corrigir as falhas apresentadas. Para isto é necessário desenvolver métodos capazes de realizar um processo automático de classificação destas novas imagens. Para tal, foi utilizado o conhecimento da categorização das imagens atuais no intuito de construir um classificador eficaz.

Outro objetivo do trabalho é constatar que com o uso de um algoritmo de seleção de atributos, podemos construir classificadores mais eficientes e eficazes, mostrando que nem sempre é melhor extrair o maior número de atributos possível.

2. Referencial Teórico

O primeiro sistema CBIR comercial foi o Query By Image Content (QBIC) [6], desenvolvido pela IBM. Sua análise de atributos é baseada em cor e textura. Alguns sistemas on-line utilizam a engine do QBIC como modelo para suas buscas.

Entre os algoritmos de classificação automática de dados não supervisionada, encontrados na literatura da área, temos o k-Means [11]. Este recebe como entrada uma matriz, as linhas são os itens a serem classificados e as colunas são atributos descritivos, e um número k de categorias. Já entre os algoritmos de classificação supervisionada, temos o Support Vector Machine (SVM) [2] e o k-Nearest Neighbors (k-NN) [1], que calcula a distância de uma imagem de entrada para todas as outras presentes na sua base, ordena as imagens da base por esta distância e define como a classe da imagem de entrada a classe mais freqüente nas k primeiras imagens.

Dentre as abordagens para seleção de atributos determinantes, existem soluções baseadas em algoritmos comerciais como o C4.5 [8] e outras, acadêmicas, como o algoritmo Statistical Association Rule Miner (StARMiner) [10]. O StARMiner foi originalmente utilizado para classificação de imagens médicas. O objetivo deste método é encontrar regras estatísticas envolvendo os atributos que melhor discriminam a imagem em suas categorias. Desta forma, alimentando o algoritmo com as imagens já representadas por seus atributos, ele irá separá-las em categorias e retornará os atributos que foram mais relevantes para a sua classificação.

3. Metodologia

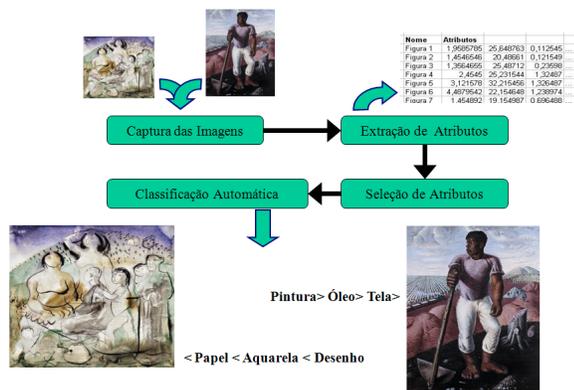


Figura 1. Metodologia

3.1. Captura das Imagens

Inicialmente foi feita a aquisição das imagens que irão compor o universo de pesquisa. Elas estão disponíveis no site do Projeto Portinari [7]. A captura foi realizada acessando-se cada uma das classes de imagens que estão classificadas por suas técnicas. Esta classificação

das obras já existente no site foi mantida, pois será utilizada como referência para quantificar a eficácia da classificação automática gerada neste trabalho.



Figura 2. Menina, pintura a aquarela e grafite sobre cartão, 1941

3.2. Métodos para a Extração dos Atributos

Após a etapa de aquisição de imagens, foi realizada uma pesquisa bibliográfica para identificar quais são os principais métodos utilizados para a extração de atributos visuais, considerando os aspectos de textura e cor.

3.2.1. Baseados em textura. A textura é uma das características mais importantes para a classificação de imagens, e pode ser caracterizada por variações locais nos valores dos pixels que se repetem de maneira regular ao longo da imagem. As texturas podem estar relacionadas a algumas características das imagens como uniformidade, densidade, aspereza, regularidade, intensidade, entre outras.

Como a análise de textura não depende de cor, as imagens foram convertidas para tons de cinza antes da aplicação dos métodos, mantendo-se uma cópia colorida do acervo para a extração dos atributos de cor. Foram selecionados os seguintes métodos baseados em textura:

- Matriz de co-ocorrência
- Transformada de Fourier
- Transformada discreta do cosseno

3.2.2. Baseados em cor

- Histograma normalizado
- Histograma métrico

- Vetor de coerência

3.3. Atributos

A partir da aplicação dos métodos descritos acima, foram extraídos atributos que, segundo a literatura pesquisada, estão associados a características visuais. Os atributos extraídos a partir de cada método estão relacionados abaixo:

3.3.1. Para a matriz de co-ocorrência. Em [5] é descrita uma metodologia de classificação de imagens a partir do uso da abordagem estatística, onde são definidas diversas características advindas do cálculo de matrizes de co-ocorrência, as selecionadas para este trabalho foram:

- Contraste: medida da quantidade de variação local de níveis de cinza em uma imagem.
- Correlação: representa uma idéia de linearidade de dependências de tons de cinza em uma imagem.
- Energia: medida da homogeneidade local dos níveis de cinza em uma imagem.
- Entropia: medida do grau de dispersão de ocorrências de níveis de cinza em uma imagem.
- Variância: medida da dispersão estatística, indicando quão longe, em geral, os valores se encontram do valor esperado (segundo momento central).
- Diferença entropia
- Diferença variância
- Medida correlação
- Máximo coeficiente de correlação
- Soma da entropia
- Soma da média
- Soma da variância
- Variância inversa

3.3.2. Para a transformada de Fourier e a transformada discreta do cosseno. Baseado em [9], extrairemos os seguintes atributos destas transformadas:

- Primeiro e segundo momentos angulares: o primeiro mede a probabilidade da distribuição dos pixels, enquanto o segundo corresponde à variância.

3.3.3. Para os histogramas. Segundo [3], os histogramas podem ser bem representados por estes atributos:

- Máximo: valor da cor com maior contagem (para cada uma das 3 camadas).
- Mínimo: valor da cor com menor contagem (para cada uma das 3 camadas).
- Média: média das contagens de todas as cores.
- Mediana
- Desvio padrão

3.4. Algoritmo para Seleção dos Atributos Relevantes

Dado o problema da alta-dimensionalidade de atributos a serem analisados, foi adotada uma abordagem para seleção de atributos mais relevantes. Neste trabalho, foi utilizado o algoritmo StARMiner [10] para a mineração de regras de associação estatísticas em imagens. O objetivo do algoritmo StARMiner é encontrar regras estatísticas envolvendo os atributos que melhor discriminam a imagem em suas categorias.

Seja uma base de imagens T , x uma categoria de imagens, $T_x \in T$ o subconjunto de imagens da base da categoria x e A um atributo da imagem, A será relevante se:

$$|\mu_A(T_x) - \mu_A(T - T_x)| \geq \Delta_{\mu min} \quad (1)$$

onde $\mu_{T_x}(Z)$ é a média dos valores do atributo A para o subconjunto de imagens Z e $\Delta_{\mu min}$ é um parâmetro de entrada que indica a diferença mínima permitida entre a média dos valores de A para as imagens da categoria x e a média dos valores de A para as imagens restantes.

3.5. Algoritmo para a Classificação Automática

Para a classificação automática das imagens, foi utilizado o algoritmo de classificação supervisionada k-Nearest Neighbor (k-NN). Este algoritmo realiza a classificação por similaridade. Para isso o vetor de atributos da imagem de entrada é comparado com os vetores de atributos das imagens contidas na base de conhecimento, as quais as categorias são conhecidas. Esta comparação é feita através de uma função de distância (utilizamos a distância euclidiana) que quantifica o quão perto (ou similar) estão cada par de vetores [1].

A próxima etapa é ordenar as imagens da base de conhecimento, baseando na distância para a imagem de entrada. Desta forma imagens semelhantes, que apresentam distâncias menores, aparecerão no início do vetor ordenado.

A categoria atribuída à imagem de entrada será a mais freqüente nas k primeiras posições deste vetor ordenado. A categoria original da imagem, já conhecida, é mantida para a quantização dos resultados posteriormente.

4. Resultados

Os testes foram realizados utilizando o classificador automático k-NN (com k=4), tendo como imagens de sua base metade das figuras de cada classe, escolhidas de uma forma aleatória, mas repetível, para que pudessem ser feitas as comparações entre os testes.

Inicialmente foi feita a classificação sem a seleção dos atributos e posteriormente com a seleção. Com a seleção

a classificação foi testada em três níveis, sendo o terceiro nível o de classificação mais profunda e difícil, pois leva em conta a técnica, o material de pintura e o suporte (ex.: pintura>óleo>tela). O segundo nível considera apenas os dois primeiros níveis, técnica e o material; já o primeiro nível classifica apenas quanto à técnica.

O índice de acerto foi calculado comparando-se a categoria dada pelo classificador automático, com a categoria já atribuída àquela imagem, caso elas sejam iguais configura-se um acerto. Os resultados são os seguintes:

| | Sem seleção de atributos | Com seleção de atributos (3 níveis) | Com seleção de atributos (2 níveis) | Com seleção de atributos (1 nível) |
|--|--------------------------|-------------------------------------|-------------------------------------|------------------------------------|
| Média de atributos por classe | 1.342 | 1.342 | 1.342 | 1.342 |
| Média de atributos relevantes por classe | 1.342 | 1.264 | 1.169 | 850 |
| Número de classes | 220 | 220 | 73 | 8 |
| Número de acertos | 370 | 714 | 960 | 2.094 |
| Índice de acerto | 14,6361 % | 28,5206 % | 38,5285% | 82,8323% |

Tabela 1. Resultados obtidos

5. Conclusões

A melhora observada com a seleção dos atributos determinantes se mostrou bem relevante, alcançando quase 100%. Contudo, analisando os resultados como os de uma máquina automática de classificação, eles estão aquém do que se esperaria.

Este baixo índice de acerto global pode ser justificado por vários fatores, sendo o principal fator a inexistência de um padrão no conteúdo das imagens, imagens com conteúdos completamente diferentes pertencem a uma mesma categoria. Também foram observados problemas com a aquisição das imagens, sendo que algumas foram feitas através de fotos, que englobam fundos e outros objetos que não pertencem às obras. O armazenamento das imagens, em formato JPEG, no qual há perda de qualidade e a geração de macroblocos, também prejudica muito a extração de atributos, principalmente de textura.

6. Agradecimentos

Agradecemos as entidades que apoiaram este projeto: CNPq, CAPES e FAPEMIG.



Figura 3. Exemplo da dissimilaridade de imagens de uma mesma categoria (Gravura > Litografia > Papel)

Referências

- [1] T. Cover and P. Hart. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1):21–27, 1967.
- [2] K.-S. Goh, E. Chang, and K.-T. Cheng. Support vector machine pairwise classifiers with error reduction for image classification. In *MULTIMEDIA '01: Proceedings of the 2001 ACM workshops on Multimedia*, pages 32–37, New York, NY, USA, 2001. ACM.
- [3] W. R. E. Gonzalez, Rafael C. *Digital Imaging processing*. Editora Edgar Blucher Ltda, São Paulo, 2007.
- [4] V. N. Gudivada and V. V. Raghavan. Content-based image retrieval systems. *Computer*, 28(9):18–22, 1995.
- [5] R. M. Haralick, Dinstein, and K. Shanmugam. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3:610–621, November 1973.
- [6] C. W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin. Qbic project: querying images by content, using color, texture, and shape. volume 1908, pages 173–187. SPIE, 1993.
- [7] C. Portinari. Projeto portinari. <http://www.portinari.org.br>.
- [8] J. R. Quinlan. *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [9] T. R. Reed and J. M. H. du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Underst.*, 57(3):359–372, 1993.
- [10] M. X. Ribeiro, J. Marques, A. J. M. Traina, and C. T. Jr. Statistical association rules and relevance feedback: Powerful allies to improve the retrieval of medical images. In *CBMS '06: Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems*, pages 887–892, Washington, DC, USA, 2006. IEEE Computer Society.
- [11] L. Zheng. *Automated feature extraction and content-based retrieval of pathology microscopic images using k-means clustering and code run-length probability distribution*. PhD thesis, Pittsburgh, PA, USA, 2005. Adviser-Paul Munro and Adviser-Michael Becich.