

## *ImageSelector: um classificador de imagens*

Ana Cristina Nunes Soares  
Francisco do Nascimento Júnior  
Marinas Carrasco de Ribamar Dantas  
Patrícia Maforte dos Santos

Recife, Novembro de 2002

### **1 Resumo**

O objetivo deste projeto é o desenvolvimento de um classificador que permita separar imagens coletadas em um banco de dados em duas classes semânticas: imagens fotográficas e imagens gráficas. Imagens fotográficas incluem cenas naturais, como pessoas, faces, animais, flores, paisagens e cidades. Imagens gráficas são logotipos, desenhos, ícones, mapas e panos de fundo, geralmente gerados utilizando o computador.

Na maioria das vezes, não há problema para nós decidirmos se uma imagem é gráfica (Figura 1) ou fotográfica (Figura 2). Baseado nisso, o sistema tentará classificar da mesma maneira que os humanos fariam. No entanto, às vezes não fica claro que categoria pode ser aplicada: é o caso de imagens mistas ou de fotografias de desenhos. Figuras nestes casos não serão utilizadas no treinamento, nem na fase de testes.[1]

Figura 1: Exemplo de imagem gráfica

Figura 2: Exemplo de imagem fotográfica

Os classificadores podem ser utilizados em ferramentas comerciais de recuperação de imagens com base no conteúdo. Meios diferentes podem ser utilizados para coletar as imagens a serem classificadas. Apesar de utilizarmos uma base de dados, uma forma muito interessante e que iremos usar como motivação de projeto é a classificação de imagens da web.

## 2 Objetivos

O objetivo do classificador envolve as seguintes atividades:

- leitura da base de imagens do banco de dados;
- preparação das amostras de treinamento;
- estudo, definição, implementação e aplicação de métricas (número de cores, cor predominante, vizinho mais distante, saturação, histograma de cor, histograma do vizinho mais distante, razão de dimensão, dimensão menor) baseadas nas diferenças entre as duas classes em questão;
- disponibilização dos valores resultantes das métricas em tuplas de atributos com valores normalizados e utilização destas tuplas da amostra de treinamento na construção do classificador utilizando método supervisionado ID3 (Itemized Dichotomizer 3)[2].

Após a execução desses procedimentos, será criada uma árvore de decisão geradora de regras para a classificação.

Ao final da separação das imagens, o sistema permitirá a visualização gráfica das métricas para as figuras classificadas e ainda gráficos que exibirão os resultados obtidos (taxas de classificação e erros, desvio padrão, número médio de iterações).

### 3 Motivação

A Web é uma complexa e vasta fonte de informação multimídia. A cada dia, mais imagens digitais são acrescentadas e retiradas da web, tornando a necessidade de encontrar um determinado grupo de imagens uma tarefa cada vez mais difícil. Desta forma, para se fazer uso dessas informações de forma eficiente é necessário que estas estejam organizadas, de forma a permitir a pesquisa e a recuperação.[3]

A recuperação de imagens representa uma área muito ativa desde a década de 1970. Existem duas linhas de pesquisa que estudam a recuperação de imagens, que são o gerenciamento de banco de dados e a visão computacional. Estas duas áreas estudam a recuperação de imagens sob dois ângulos diferentes, um baseado em texto e outro baseado no conteúdo visual das imagens. A recuperação de imagens baseada em textos primeiro descreve as imagens de forma textual e depois utiliza um SGBD (Sistema de Gerenciamento de Banco de Dados) baseado em textos para a recuperação das imagens. Porém, existem algumas dificuldades neste tipo de recuperação. A primeira delas seria a quantidade de trabalho requerida na anotação manual das imagens. A segunda e mais importante seria o aspecto da subjetividade relacionada com a análise do conteúdo das imagens. Ou seja, para uma mesma imagem pessoas diferentes têm percepções diferentes, que ocasionariam anotações diferentes, podendo levar a erros irrecuperáveis no processo de recuperação.

Neste cenário, a organização e recuperação de imagens baseada no conteúdo tornou-se uma área muito importante da visão computacional e da multimídia. A necessidade de se encontrar uma imagem dentro de um conjunto de informações digitais é importante para diversos grupos de usuários, como jornalistas, engenheiros, historiadores, designers, professores, artistas e agências de propagandas, entre outros. A tecnologia que permite o acesso a essas informações vem mudando rapidamente, assim como a forma como as pessoas interagem com a informação visual.

O desenvolvimento de ferramentas que permitam o agrupamento de imagens em grupos semanticamente significativos tornou-se algo de extrema importância, representando uma solução para o desafio da recuperação de imagens baseada no conteúdo. Estas ferramentas tornarão possível a busca eficiente de imagens específicas em um grande banco de dados, pois estas imagens estarão esquematicamente categorizadas.

## 4 Cronograma

1. 11/11/02 – 25/11/02: Obtenção da Base de dados;  
Estudo e definição das métricas de classificação;  
Busca do algoritmo ID3 (Itemized Dichotomizer  
3).
2. 26/11/02 – 30/11/02: Estruturação da implementação.
3. 01/12/02 – 20/12/02: Implementação lógica.
4. 21/12/02 – 01/01/03: Recesso Acadêmico.
5. 02/01/03 – 15/01/03: Implementação lógica.
6. 16/01/03 – 08/02/03: Implementação gráfica.
7. 09/02/03 – 11/02/03: Preparação da documentação e apresentação do projeto.

## Referências

- [1] ATHITSOS, V., SWAIN, M. J. Distinguishing photographs and graphics on the world wide web. Chicago Illinois USA, Oct. 1.
- [2] OLIVEIRA, C. J. S., ARAÚJO, A., SEVERIANO JR, C. A., GOMES, D. R. Classifying images collected on the world wide web. *14th Brazilian Symposium on Computer Graphics and Image Processing*, Belo Horizonte MG Brasil, Oct. 2002.
- [3] SWAIN, M., FRANKEL, C., ATHITSOS, V. Webseer: An image search engine for the world wide web, 1997.