

# Pesquisa Interactiva de Fotografias com Base na Regra dos Terços

Gabriel Barata Ricardo Dias Sandra Gama Manuel J. Fonseca Daniel Gonçalves

Departamento de Engenharia Informática  
INESC-ID / IST / Universidade Técnica de Lisboa  
R. Alves Redol, 9, 1000-029 Lisboa, Portugal

[gabriel.barata, ricardo.dias, sandra.gama}@ist.utl.pt, [mjf, daniel.goncalves}@inesc-id.pt

---

## Resumo

*Hoje em dia, geramos uma quantidade cada vez maior de fotografias digitais. Estas encontram-se, frequentemente, desorganizadas ou organizadas em hierarquias complexas. A pesquisa por determinadas imagens específicas torna-se, assim, difícil. Neste artigo apresentamos uma solução para a pesquisa interactiva de fotografias que combina técnicas de visualização interactiva com algoritmos de recuperação de imagem. Para a visualização utilizamos histogramas multifacetados e introduzimos uma nova técnica de apresentação dos resultados chamada Linha de Similaridade. Adicionalmente, comparamos duas formas de descrever o conteúdo das fotografias usando histogramas locais; uma que divide a fotografia em nove regiões iguais e outra que usa a regra dos terços para determinar as regiões. Finalmente, incluímos também um algoritmo para contar o número de faces presentes nas fotografias. Testes preliminares mostram que a utilização da regra dos terços para calcular os histogramas locais apresenta resultados comparáveis com a solução tradicional, com a vantagem de usar menos histogramas.*

## Palavras-Chave

*Recuperação de Fotografias, Visualização, Interface Utilizador, Interação Pessoa-Máquina*

---

## 1. INTRODUÇÃO

As câmaras digitais estão presentes no nosso quotidiano, sob a forma de máquinas fotográficas digitais comuns e de telemóveis. Tiramos uma quantidade cada vez maior de fotografias, descarregando-as para o sistema de ficheiros, muitas vezes sem grande preocupação quanto à organização. Como consequência, o processo de encontrar uma fotografia específica é cada vez mais complexo, o que faz com que o desafio de recuperar fotografias digitais tenha vindo a ganhar uma relevância significativa.

De facto, o armazenamento utilizando um sistema de ficheiros implica uma classificação hierárquica que, além de exigir um elevado esforço para manter alguma organização, não tira partido da informação contida nas próprias imagens [Lee 09]. Além disso, a pobre interactividade que este método apresenta leva o utilizador a um elevado esforço para recuperar as suas fotografias.

A extracção de características relativas às imagens (como texturas, cores dominantes e formas, entre outros), permite aplicar estratégias de classificação aos dados [Stojanovic 07] e promover a recuperação de imagens baseada no seu conteúdo, potencializando o desenvolvimento de ferramentas interactivas para pesquisa. Ferramentas como o VisualSEEK [Smith 97], que se baseia nas cores dominantes e usa histogramas HSV, têm sido usadas

como ponto de partida para outras investigações. Esta aplicação permite ao utilizador desenhar um conjunto de regiões (em que especifica a posição, a dimensão e a cor), devolvendo as imagens que melhor correspondem ao desenho. Contudo, além da cor, pode ser interessante considerar múltiplos parâmetros para pesquisa interactiva de fotografias. A pesquisa facetada, que dá liberdade aos utilizadores para explorar relações na informação, tem vindo a ser explorada nos últimos tempos.

Um exemplo desse tipo de abordagem é o FacetLens [Lee 09], um sistema de visualização interactivo, que suporta a exploração e compreensão de informação, utilizando conjuntos de dados facetados.

A solução que apresentamos aplica este conceito à imagem, fornecendo uma aplicação para pesquisa interactiva de fotografias. A interface que criámos baseia-se em duas técnicas principais de visualização que se complementam mutuamente: histogramas com diversas facetes e linhas de similaridade. A informação relativa aos resultados da pesquisa efectuada pelo utilizador é apresentada de forma completa, mas ao mesmo tempo simples e directa.

A abordagem que seguimos para a recuperação de fotografias baseia-se na criação de histogramas locais da imagem e na extracção de rostos humanos. Adicionalmente, usamos a Regra dos Terços para especificar as regiões da imagem

utilizadas para construir os histogramas locais.

Na secção 2 apresentamos os algoritmos usados para descrever o conteúdo e recuperar fotografias. Na secção 3 descrevemos a interface utilizador. A secção 4 refere-se aos resultados dos testes relativos à recuperação de fotografias, e finalmente, na secção 5 apresentamos as conclusões e trabalho futuro a realizar.

## 2. ALGORITMOS DE RECUPERAÇÃO

A nossa aplicação para recuperação interactiva de fotografias baseia-se na criação de histogramas locais da imagem e na extracção de rostos.

### 2.1. Histogramas Locais

Wyszecki e Stiles [Wyszecki 82] investigaram a percepção de cor pelo olho humano e concluíram que o nosso sistema visual consegue distinguir 13 cores básicas: vermelho, laranja, amarelo, verde, azul-esverdeado, azul claro, azul, roxo, rosa, castanho, branco, cinzento e preto. Tal como Belongie et al. [Belongie 01], utilizamos esta teoria.

A primeira abordagem para construção de histogramas locais consistiu em dividir a imagem em nove porções de igual dimensão. Numa segunda aproximação, decidimos adoptar a regra dos terços [Rul10], que tem sido amplamente usada em artes visuais para criar imagens equilibradas e com maior impacto. Tal como ilustrado na Figura 1, esta regra consiste na divisão da fotografia em nove partes iguais, usando duas linhas horizontais e duas linhas verticais igualmente espaçadas, de forma a que os elementos de interesse fiquem posicionados sobre as linhas e nas intersecções entre estas.



Figura 1. Regra dos terços

Considerámos cinco regiões para a construção de histogramas: as quatro áreas sobre as linhas que subdividem a imagem e uma região central (para maior tolerância relativamente a fotografias que não sigam a regra dos terços), como ilustrado na Figura 2.

### 2.2. Número de Faces

Além da criação de histogramas, é considerado o número de rostos numa imagem como parâmetro para recuperação de fotografias.

A extracção do número de faces da imagem é feita através de um classificador de Haar em cascata, que se baseia na

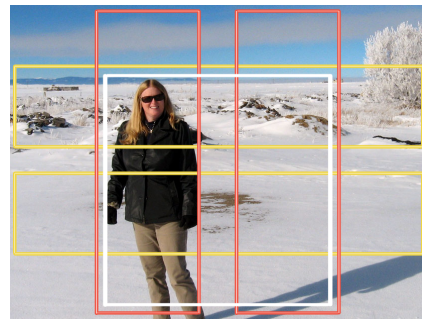


Figura 2. Regiões de interesse: áreas em redor das linhas de separação utilizando a regra dos terços e zona central da imagem.

investigação de Viola e Jones [Viola 01]. Este método consiste em distinguir padrões faciais de padrões não faciais através da subdivisão de uma imagem em janelas de menor dimensão e em submeter cada uma das regiões resultantes a um conjunto de testes em cascata em que serão aceites caso correspondam a padrões faciais e rejeitadas caso contrário.

### 2.3. Indexação

Dois mecanismos de indexação são empregues: um mecanismo de indexação multidimensional, de nome NBTree [Fonseca 03], que permite efectuar pesquisas pelos  $k$  vizinhos mais próximos de forma eficiente, e o tradicional índice invertido.

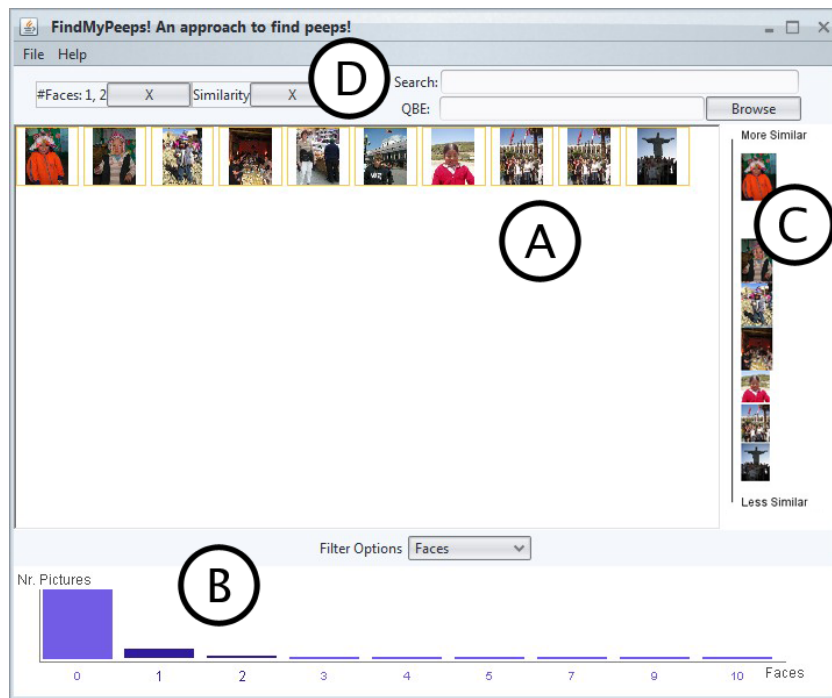
Na estrutura de indexação multidimensional guardamos os descritores de dimensão 14, obtidos pela concatenação do histograma (de dimensão 13, uma para cada cor) e do número de faces. Os descritores das várias regiões da fotografia são guardados na mesma estrutura de indexação.

Os dados referentes à data, orientação, nome, extensão, etc., são guardados em índices invertidos (um para cada tipo de dados).

Quanto às interrogações, o sistema suporta interrogações por exemplo, ou usando as características específicas. Na pesquisa por exemplo, o utilizador fornece uma imagem, a partir da qual se extrai um descritor que é usado para realizar uma interrogação KNN (*K Nearest Neighbours*) na estrutura multidimensional. Os descritores devolvidos correspondem às regiões mais semelhantes. Quantas mais regiões uma imagem tiver na lista devolvida, mais semelhante é a interrogação.

No caso dos outros dados, o utilizador usa o valor de uma propriedade para interrogar o índice invertido correspondente.

É de salientar que as interrogações podem ser armazenadas e combinadas de forma a gerar outras mais complexas. Por exemplo, podemos procurar todas as fotografias semelhantes a uma dada paisagem de montanha, tiradas em 2005, cuja orientação é a de paisagem e que mostre três pessoas.



**Figura 3. Interface Utilizador, mostrando o histograma multifacetado (B) e a linha de similaridade (C).**

### 3. INTERACÇÃO E VISUALIZAÇÃO

#### 3.1. Interface Utilizador

A interface que propomos consiste num mecanismo interactivo para recuperação de fotografias baseado em duas técnicas de visualização principais: histogramas multifacetados e linha de similaridade.

A interface está subdividida em quatro áreas principais, como ilustrado na Figura 3.

Na zona (A) são representadas, inicialmente, todas as imagens disponíveis. Quando é feita uma pesquisa, é nesta área que os resultados são apresentados. Esta interface, simples e natural, é semelhante aos esquemas de exploração de informação existentes nas interfaces gráficas dos sistemas operativos convencionais.

No entanto, quando uma fotografia é seleccionada na zona de navegação (A), sofre uma ampliação, inspirada na técnica *fisheye view*, apresentando dois botões, que permitem, respectivamente, procurar fotografias semelhantes (interrogação através de exemplo) e abrir a fotografia.

Na base desta área existe uma ComboBox que permite aplicar diferentes filtros (número de rostos, orientação, ano, extensão do ficheiro e palavras-chave), permitindo representar diferentes facetadas relativas às imagens.

A região (B) consiste numa das técnicas de visualização que adoptámos: histogramas multifacetados. Quando um filtro é seleccionado, na ComboBox da área (A), referida acima, o histograma é actualizado de acordo com o parâmetro correspondente. Por exemplo, se seleccionar-

mos o filtro "Ano", o histograma passará a mostrar essa faceta, isto é, a distribuição de fotografias por ano.

No painel (C) é representada a segunda técnica de visualização utilizada: linha de similaridade, permitindo a consulta simples e rápida do grau de semelhança entre a fotografia de interrogação e os resultados.

Finalmente, a área (D), no topo, permite uma pesquisa textual convencional, procurando fotografias que contenham, no seu nome, as palavras-chave introduzidas na caixa de texto. Além disso, a interface possibilita ao utilizador a selecção de uma qualquer fotografia existente no seu sistema de ficheiros para uma interrogação por exemplo.

#### 3.2. Técnicas de Visualização

Os histogramas multifacetados permitem visualizar informação relativa aos vários parâmetros disponíveis na filtragem, ou seja, ano, número de rostos, extensão do ficheiro e orientação (retrato ou paisagem). Optámos por manter o histograma bastante simples, contendo apenas informação essencial. No entanto, quando o rato é posicionado sobre uma barra do histograma, o valor correspondente é mostrado. A selecção de múltiplos valores relativos a uma faceta, que correspondem às várias barras do histograma, é suportada, quer para valores consecutivos, quer para valores alternados.

A linha de similaridade consiste num ordenamento e distribuição dos resultados de uma interrogação, que permite aos utilizadores ter uma perspectiva visual da semelhança entre a interrogação e as fotografias recupe-

radas, e destas entre si (Figura 3, área (C)).

Para tal, usamos a medida de semelhança entre a interrogação e os resultados, sendo estes apresentados ao longo de uma linha vertical com valor decrescente de similaridade. Assim, as fotografias que apresentam maior semelhança são apresentadas no topo, em tamanho maior, sendo que, à medida que a similaridade diminui, as restantes fotografias vão-se aproximando progressivamente do fundo da linha de similaridade. Esta técnica de visualização, que dispensa a comparação de valores numéricos de similaridade, complementa a interface ao fornecer uma perspectiva visual desta medida.

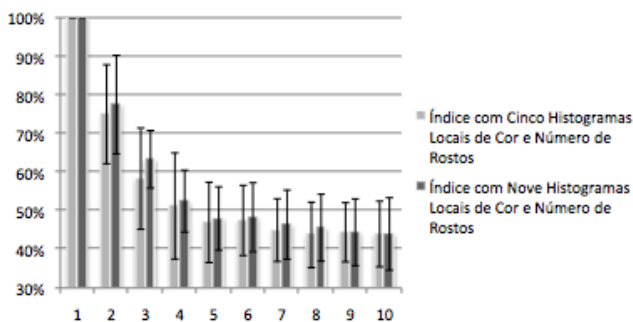
#### 4. TESTES

Neste momentos realizamos apenas testes para comparar as várias abordagens relativas ao algoritmo de recuperação de fotografias. Para isso utilizamos uma base de dados com 1000 fotografias e a R-Precision como medida de desempenho. Esta medida calcula a precisão para cada resultado considerado relevante.

Seleccionamos de forma aleatória 20 imagens da base de dados para usar como interrogações e depois submetemos aos dois algoritmos em avaliação e calculamos a R-Precision média para cada um dos 10 valores mais relevantes (Figura 4).

Da Figura 4 podemos ver que o algoritmo baseado na regra dos terços apresenta resultados semelhantes aos do algoritmo tradicional, com a vantagem de utilizar menos histogramas (cinco contra nove). Isto reflecte-se quer no espaço de armazenamento necessário quer no tempo de processamento das interrogações.

Pretendemos, num futuro próximo, comparar com o histograma global e ainda utilizar uma base de dados de fotografias onde a regra dos terços seja respeitada em todas as fotografias para verificar se os resultados são afectados.



**Figura 4. R-Precision média para os dois algoritmos de recuperação**

#### 5. CONCLUSÕES E TRABALHO FUTURO

A solução que apresentamos consiste numa ferramenta para pesquisa interactiva de fotografias que utiliza histogramas multifacetados e linhas de similaridade como técnicas de visualização. Fizemos ainda um estudo comparativo da utilização de histogramas locais para recuperação de fotografias e obtivemos indícios de que a regra dos terços será uma mais valia na recuperação de imagem. Depois de validados os algoritmos de recuperação, pretendemos efectuar testes com utilizadores para promover um desenho interactivo e a validação da interface.

#### Referências

- [Belongie 01] S. Belongie, C. Carson, H. Greenspan, e J. Malik. Color-and texture-based image segmentation using em and its application to content-based image retrieval. Em *Int. Conf. on Computer Vision (JCCV'01)*, 2001.
- [Fonseca 03] Manuel J. Fonseca e Joaquim A. Jorge. Indexing high-dimensional data for content-based retrieval in large databases. Em *Int. Conf. on database systems for advanced applications (DASFAA' 03)*, 2003.
- [Lee 09] B. Lee, G. Smith, G. Robertson, M. Czerwinski, e D. S. Tan. Facetlens: exposing trends and relationships to support sensemaking within faceted datasets. Em *Int. Conf. on Human factors in computing systems (CHI '09)*, 2009.
- [Rul10] Rule of thirds. <http://digital-photography-school.com/rule-of-thirds>, Accessed June 2010.
- [Smith 97] J. R. Smith e S.-F. Chang. *Querying by color regions using the VisualSEEK content-based visual query system*. AAAI Press, 1997.
- [Stojanovic 07] I. Stojanovic, S. Bogdanova, e M. Bogdanov. Content-based image retrieving improved by pixel-based search. Em *International Conf. on Systems Signals and Image Processing IWSSIP*, 2007.
- [Viola 01] P. Viola e M. Jones. Rapid object detection using boosted cascade of simple features. Em *Conf. on Computer Vision and Pattern Recognition (CVPR '01)*, 2001.
- [Wyszecki 82] G. Wyszecki e W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley, 1982.