Spatial color image retrieval without segmentation using thumbnails and the Earth Mover’s Distance

Thomas Hurtut, Haroldo Dalazoana, Yann Gousseau, Francis Schmitt

1. Introduction
2. Features
3. Distance
4. Matching criterion
5. Experimental evaluation
6. Conclusion
Introduction

Color image retrieval:

Without spatial information

With spatial information

Color-spatial image retrieval, many approaches:
- augmented features: e.g. chromatic histograms [Lambert et al. CGIV 2004]
- points of interest: e.g. [Heidemann 2004]
- region based image retrieval (RBIR): e.g. BlobWorld [Malik et al. 2002], [Rugna et al. CGIV 2002]

→ Unavoidable segmentation errors → false retrievals
Features

Coarsely sampled thumbnails, size $n$ pixels

$\rightarrow$ Signature: $n$ features $f^i$

$$f^i \equiv \{L^i, d^i, b^i, X^i, Y^i\}$$
Earth Mover Distance (EMD) is used [Rubner et al. 2000].

General case (transportation problem):

![Diagram showing transportation problem](image1)

Best solved with simplex algorithm (can be exponential with n)

Assignment problem case:

![Diagram showing assignment problem](image2)

Solved with Kuhn-Munkres (1955) algorithm for assignment problems in $O(n^3)$

Our features/segmentation step takes full advantage of EMD
An *a contrario* model gives an unsupervised answer: we do perceive events that happen rarely in a noisy situation.

→ A target image will be considered as similar to a query if it is closer to the query than an image of « noise » could be.
Unsupervised matching criterion

The noise model we consider is a dead leaves model [Matheron 1968], consisting in the superimposition of random objects, together with a power law distribution for the size of objects.

The distribution of the color of objects is learned from the color distribution of pixels from thumbnails of the database.
Experimental evaluation

Tests on a illuminated manuscripts database from IRHT gathering 1500 images

Blobworld + EMD
Our method
Experimental evaluation

Tests on a illuminated manuscripts database from IRHT gathering 1500 images
Experimental evaluation

Tests on a illuminated manuscripts database from IRHT gathering 1500 images
Experimental evaluation

Tests on a database gathering 25000 images from multiple Canadian museums
Experimental evaluation
Experimental evaluation

Tests on a database gathering 25000 images from multiple Canadian museums
Conclusion and future work

**Conclusion**
We propose a spatial color image retrieval method without initial segmentation, using thumbnails and EMD:
- robust description
- efficient use of the EMD distance

We also propose an unsupervised matching criterion relying on an a contrario method.

**Future work**
More elaborated features: local quantization, texture characteristics
Tests on bigger databases (> 100 000 images)
Recent EMD approximation, to speed up querying
Questions ?

More results are visible at: [http://www.tsi.enst.fr/recherche/cbir](http://www.tsi.enst.fr/recherche/cbir)

**Acknowledgments**

We thank Gilles Kagan from IRHT for providing us the IRHT database and for his helpful comments. We also thank Farida Cheriet from LIV4D Polytechnic School of Montreal for her suggestions. This work was supported by a CNRS ACI supervised by Anne-Marie Edde and Dominique Poirel from IRHT, and by a grant from the LIV4D laboratory.
Distance

\[
d(Q, T) = \min_{\phi} \sum_{i=1}^{n} d_e \left( f_Q^{\phi(i)}, f_T^i \right)
\]

\[
d_e(f_Q^i, f_T^i) = \alpha \left( 1 - e \left( \frac{(L_Q^i - L_T^i)^2 + (a_Q^i - a_T^i)^2 + (b_Q^i - b_T^i)^2}{\delta_C^2} \right) \right)
\]

\[
+ (1 - \alpha) \left( 1 - e \left( \frac{(x_Q^i - x_T^i)^2 + (y_Q^i - y_T^i)^2}{\delta_S^2} \right) \right)
\]

\[
\alpha \quad \text{balances color and spatial distribution}
\]

\[
\delta_C \quad \text{tune the color dynamic,}
\]

\[
\delta_S \quad \text{tune the spatial dynamic}
\]
Experimental evaluation

Tests on a illuminated manuscripts database gathering 1500 images

\( \alpha \) balances color and spatial distribution

\[ \alpha = 0.5 \]
(color and spatial)

\[ \alpha = 1 \]
(color only)
Unsupervised matching criterion

Given an image database \( \mathcal{B} = \{T_1, \ldots, T_m\} \)

We make a statistical test of two hypothesis against each other:

\[
H_1 = \{ T_i \text{ is similar to } Q \} \\
H_0 = \{ T_i \text{ follows the background model } \}
\]

And we will say that:

\[
T_i \text{ is an } \varepsilon\text{-meaningful match of } Q \text{ if } d(Q, T_i) \leq C_B
\]

Where \( C_B \) is such that:

The expected number of false matches is one