Adaptive image retrieval based on the spatial organization of colors

Thomas Hurtut, Yann Gousseau, Francis Schmitt

Adaptive image retrieval based on the spatial organization of colors

Thomas Hurtut [a,b], Yann Gousseau [a] , Francis Schmitt [a]

[a] *Telecom ParisTech, LTCI CNRS, Paris, France*

[b] *Ecole Polytechnique de Montréal, LIV4D, Montreal, Canada*

Abstract

This work proposes to compare the spatial organization of colors between images through a global optimization procedure relying on the Earth Mover's Distance. The resulting distance is applied to image retrieval. Unlike most region-based retrieval systems, no segmentation of images is needed for the query. We then address the decision stage of the retrieval, that is the problem of automatically deciding which images from a database match a query. To this aim, we make use of an *a contrario* method. Two images are matched if their proximity is unlikely to be due to chance; more precisely, a matching threshold on distances is computed by controlling the average number of false matchings in an unsupervised way. This threshold is adaptive, yielding different numbers of result images depending on the query and the database.

Key words: Spatial organization of colors; color image retrieval; *a contrario* method; earth mover's distance; optimal transport; dead leaves model; dead leaves model; image distance.

1. Introduction

Image retrieval is the process of finding an image or a group of images of interest in an image database. It usually relies on some distance or dissimilarity measure between image features that takes information about color, shape or texture into account. Since the seminal work of Swain and Ballard [53], many image retrieval systems have been strongly oriented toward representing the color content of images by using histograms. This is probably due to the ability of color histograms to concisely represent the global aspect of natural images and to the striking impression of global coherence given by series of images having similar color content. Tutorials about such methods can for instance be found in the monographs [5,37].

More recently, many works took interest in the spatial organization of colors and in dependence relationships between pixels. A first type of approach to this problem has been to select some spatial statistics between pixels color, such as correlograms or some filter responses, see [45,28,13,49,34]. A related category of methods makes use of points of interest within images in a way similar to classical object recognition methods in computer vision, see [26,54]. In both cases, the features of color images that are accounted for are various types of color transitions, usually using various scales and orientations.

A second type of approach aims at describing and comparing images with respect to their global spatial content, that is the coarse organization of colored regions in the image. Region-based approaches first perform a segmentation of images which is then used for image comparison. The *Blobworld* image representation, for instance, uses multi-dimensional Gaussian mixtures to represent the image [10]. Many other unsupervised segmentation methods have been used to index images, see e.g. [27,50,51,55,48,18,46,32,41]. Other methods rely on a small number of predefined image regions, see [52,40]. Several works specifically study the mutual arrangement of regions in images. Representation and comparison of these relationships can be symbolic [11] or nonsymbolic [4]. The goal of a global description is to be able to detect images that have a similar color layout, such as "a green stripe on the top of the image, above a red blob that is itself in between two blue blobs". The present work is concerned with the development of a method for image querying relying on the similarity of such a global, coarse organization of colors. In what follows, this aspect of image content will be denoted by *spatial organization of colors*.

The spatial organization of colors is an important aspect of the structure of an image and definitively one of the first to be perceived. It is an intermediate feature between low-level content such as color histograms and image semantics. In the context of image retrieval, various taxonomies of both image content and user-needs have recognized its importance. It is part of the *visual relationships* stressed by Burford et al. [7], which is characteristic of the spatial arrangement of the scene in two dimensions and of the *color* content. In the cultural heritage domain, it also concerns the *cultural abstraction* content since it models the artistic composition [20,12]. Such a feature can be used for illustration, generation of ideas, aesthetic decoration, and learning [14]. Within one semantic category reached with metadada for instance, a journalist may pursue its illustration needs according to composition clues. Image retrieval with such features may be a source of ideas for fines-arts students, professors, artists and curators. Image retrieval using the spatial organization of colors can also be serendipitous, meaning one can accidentally discover a fortunate association of semantically different images delivering the same visual impact, which is practically impossible using metadata or ontology-based description.

Our aim in this paper is twofold. First we introduce a robust method enabling the comparison of the spatial organization of color between images. This method does not require any segmentation step. By making use of the well-known Earth Mover's Distance (EMD), first introduced in [47] in the context of image querying, we show that it is unnecessary and sometime misleading to segment the image into regions. The comparison of the spatial organization of colors between two images is addressed as a global optimization procedure. To achieve reasonable computational costs, the EMD is computed between pixels of coarse thumbnail representations of images. Second we address the matching step of our querying system. In short, given a query image, we want to decide automatically which images from a database should be paired with the query. In particular, this yields a retrieval system in which the number of result images may vary from one query to another. This is what we refer to as *adaptive* image retrieval. This question has rarely been addressed by the image querying literature. It is however of primary concern since databases tend to become huge and knowing the number of images that are similar to a query gives a quick insight on the content of a database. To reach a decision, we follow an *a contrario* approach similar to the one in [44]. Roughly speaking, the system returns images for which the proximity to a query image is very unlikely to be due to chance. More precisely the method relies on a model of random generic color layouts to estimate the probability that two images are at a small distance apart by chance. To the best of our knowledge, this is the first proposition of an automatic matching threshold in the context of color image retrieval.

This article extends our conference paper [30], providing more detailed description and discussion. The plan of the paper is as follows. Section 2 introduces the distance used to compare the spatial organization of colors. In Section 3, an automatic matching threshold is derived in the context of image querying by example. Finally, the efficiency of the method is illustrated in Section 4 by querying two databases : a database of 1500 high quality illuminated manuscripts and a database of 15000 non-specialized digital photographs. This section includes an analysis of the behavior of the matching threshold depending on the database and the query. We give perspectives and conclude in Section 5.

2. Global comparison of the spatial organization of colors

In this paper, the problem of comparing the spatial organization of colors between two images is viewed as a global optimization procedure, in which the best transform matching an image to the other one is sought. The best transform refers to a set of pixels displacements that optimize some cost function relying on the spatial and color differences between pixels.

2.1. *Earth Mover's Distance*

Intuitively, if one is allowed to move the pixels of a first image and to change their color, one seeks the minimal amount of work needed to match a second image. The natural mathematical modeling of this approach relies on the theory of Optimal Transport, initiated by G. Monge at the end of the 18th century. The application of this theory to the problem of image matching and image querying is not new, and started with the work of Rubner et al. [47].

In [47], the Earth Mover's Distance (EMD) is introduced as the smallest amount of work needed to match a set of weighted features $(f_i, w_i)_{i=1,\ldots,n}$ to another one. The f_i's are features belonging to some space E on which a metric (i.e. a distance) d_e is defined, and the $w_i \geq 0$ are the weights. To illustrate this notion, the authors suggest an analogy between EMD and the minimal amount of work needed to put some mass of earth spread in space (the first set of features) into a collection of holes (the second one), which is precisely the problem addressed by G. Monge more than two centuries ago. In this context, the work corresponds to the quantity of earth that is displaced times the length of the displacement. Formally, the EMD between two sets of weighted features $f^1 = (f_i^1, w_i^1)_{i=1,\ldots,n_1}$ and $f^2 = (f_i^2, w_i^2)_{i=1,\ldots,n_2}$ having same total weight [1] $\sum_i w_i^1 = \sum_j w_j^2$ is defined in [47] as

$$d(f^1, f^2) = \min_{f_{ij}} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_e(f_i^1, f_j^2) x_{i,j}, \qquad (1)$$

where the *flow* $x_{i,j}$ is constrained by $x_{i,j} \geq 0$,

$$\forall i = 1, \ldots, n_1, \quad \sum_{j=1}^{n_2} x_{i,j} = w_i^1,$$

$$\forall j = 1, \ldots, n_2, \quad \sum_{i=1}^{n_1} x_{i,j} = w_j^2.$$

With the same analogy as before, the first equation means that all available earth is displaced, while the second one means that holes are completely filled. In the context of color image retrieval, features f_i either correspond to some color in a 3-dimensional space equipped with a perceptually adapted metric or to some point in a 5-dimensional space accounting for both color and spatial information. In the first case, the EMD compares the global color content of images. It is shown in [47] that this method is more satisfying than approaches relying on bin to bin distances on histograms that are very sensitive to the chosen quantization and to small color changes. In the second case (features in a 5-dimensional space), the (f_i, w_i) are clusters representing important color regions of the image. In [47], the clustering relies on the use of recursive splitting and k-d trees. This same approach can rely on any region-based segmentation method by assigning a position and various features to each region. This has been recently investigated

[1] It is possible to extend the definition to the case where total weights are not equal, see [47].

by various authors [19,32,41]. The *region integrated* method developed in [55] can also be considered as a simplification of the same idea.

2.2. *Applying EMD to pixels*

Instead of relying on a preliminary segmentation to perform region-based image comparison, we suggest using EMD directly at the smallest possible detail level. We thereafter explain why this approach is more robust than using segmentations and allows the same degree of efficiency. An image is associated with a set of features $(f_i, 1/n)_{i=1,\ldots,n}$ where n is the number of pixels in the image and f_i the color and position of the ith pixel. Each feature is given the same weight because pixels correspond to a fixed area in the original image. In this case (see [22,36]), the computation of the EMD is reduced to an assignment problem and is therefore greatly simplified:

if $f^1 = (f_i^1, 1/n)_{i=1,\ldots,n}$ *and* $f^2 = (f_i^2, 1/n)_{i=1,\ldots,n}$ *are two sets of identically weighted features in a metric space* (E, d_e)*, then*

$$d(f^1, f^2) = \min_{\phi} \sum_{i=1}^{n} d_e\left(f^1_{\phi(i)}, f^2_i\right), \qquad (2)$$

where ϕ *belongs to the set of permutations of* $\{1, \ldots, n\}$.

This means that when features all have the same weight, they are not split apart by EMD: each feature of the first image is assigned to exactly one feature of the second one. Each $x_{i,j}$ in Equation (1) is either one or zero. Computation of the EMD is therefore simplified and can be achieved by solving an optimal assignment problem [8,9]. In the absence of specific structure for the cost matrices involved in the computation of Equation (2) (in particular they are not sparse), we chose to solve the assignment problem using the classical Hungarian method, also called the Kuhn-Munkres algorithm [33], which has a complexity upper bound in $O(n^3)$. An alternative to this choice would be to bound the displacements of pixels, yielding sparse cost matrices and enabling faster computations, see [8,9].

2.3. *Using thumbnails*

Of course, even with this simplification, the spatial resolution of even a small image entails excessive computational time when querying large databases. Therefore, the method is applied to coarsely sampled thumbnails, see Figures 1 and 2. This is not a strong limitation, as we are interested in the global color content of images, and not in their small scale structure. This point will be further discussed in the next paragraph. The total number of pixels of these thumbnails is fixed to n, depending on the database content and on computing constraints. The influence of this important parameter of the method will be studied in Section 4.1. The typical order of magnitude of n is several hundreds. Image sub-sampling is achieved by averaging the image pixels. This is done first by gamma correcting the numerical RGB image components. The resulting tristimulus values can then be summed according to the Grassman's law of the colorimetry [56]. Assuming the image colorimetric space to be sRGB space, thumbnails are then converted into the psychometric CIELab color space [2] in order to enable the use of the Euclidean distance to compare color between their pixels. The signature for an image is therefore composed of the n thumbnail pixel features $f_i = \{L_i, a_i, b_i, X_i, Y_i\}$, where L, a, b are the color coordinates and X, Y the spatial positions.

2.4. *Thumbnails versus segmentation*

When using EMD with region based methods, each region is usually weighted according to its area. In this case, EMD breaks apart the weights of the features when optimizing the cost of the mapping. Therefore, the usefulness of a segmentation step before EMD is questionable since pixels that have been grouped into regions can be mapped separately to different target regions by EMD. Indeed no constraints are included in EMD that take into account the fact that regions represent groups of pixels. On the other hand, a segmentation method may fail and produce an under-segmentation which can yield regions that are not well represented by for instance their mean color, thus affecting the performances of the matching step. In short, using large regions with EMD is useless in favorable cases and misleading in unfavorable cases. Our approach is simpler, avoids the complex tunings of segmentation methods, is more robust and takes full advantage of EMD: we oversegment the image into small and regularly spaced regions (the thumbnail pixels), and then let the EMD distance resolve the matching problem. These regions are chosen to be of fixed size (using thumbnails) to reduce the computation of EMD to an assignment problem. Of course, using a fixed number of elementary regions (the thumbnail pixels) can be seen as a disadvantage. However, as stressed in the previous discussion, it is only necessary to use a number of pixels that is large enough to account for coarse regions in the image. The number of regions that are considered is only limited by computation constraints. We claim that it is generally enough to use an order of magnitude of several hundred pixels to achieve this. This fact is corroborated by the number of regions usually considered by Region Based Image Retrieval systems. The influence of the size of thumbnails will be further studied in the experimental Section 4.1.

Fig. 1. A coarsely sampled thumbnail. Left: source image (3200×2200 pixels). Right: the thumbnail (10×15 pixels).

2.5. *Distance*

The last tool that is needed in order to compare two images is a distance between two pixel features f_i^1 and f_j^2. We choose to use a weighted exponential distance:

$$d_e(f_i^1, f_j^2) = \alpha * \left(1 - e^{-\left(\frac{((L_i^1 - L_j^2)^2 + (a_i^1 - a_j^2)^2 + (b_i^1 - b_j^2)^2)^{1/2}}{\delta_c}\right)}\right) + (1-\alpha) * \left(1 - e^{-\left(\frac{((X_i^1 - X_j^2)^2 + (Y_i^1 - Y_j^2)^2)^{1/2}}{\delta_s}\right)}\right). \quad (3)$$

The parameter α balances color and spatial contributions. Parameters δ_c and δ_s are chosen according to the color and spatial dynamics. Typical values used in this paper are $\alpha = 0.5$, $\delta_c = 15$ and δ_s equal to a quarter of the thumbnails diagonal. The weighted exponential distance supports the idea that beyond a certain distance, two different images are just considered as not similar. This nonlinear distance is adapted to an image retrieval task, where a target image is close to a query, in which case the similarity is measured, or is simply different and penalized.

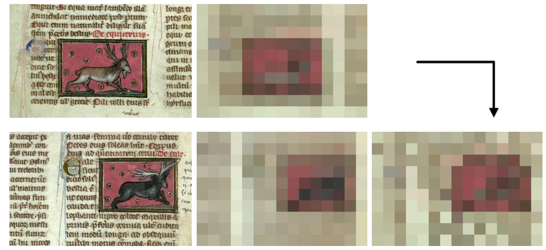


Fig. 2. Two images (left) and the associated thumbnails (middle), made of $n = 10 \times 15$ pixels. Each pixel of a thumbnail is associated to a feature in 5 dimensions (L, a, b, X, Y). All of them have the same weight. In this context, EMD is equivalent to an assignment problem. Each pixel of the first thumbnail (top middle) is assigned to exactly one pixel of the second one (bottom middle). To illustrate the EMD operation, the thumbnail on the bottom right represents the optimal permutation of the pixels from the first thumbnail to match the pixels of the second one. Pixels of the red region are moved towards their location in the second image. The white column of paper is also correctly moved from right to left.

3. Automatic matching threshold using an *a contrario* approach.

In this section, we consider the task of adaptive image retrieval. Our framework is the classical *query by example*. Even though this framework is known to have some limitation (see for instance [21]), it is nevertheless usually one of the building blocks of a querying system and it enables us to address the problem we are concerned with : how to decide which answers are meaningful when querying a very large amount of data. More precisely, we tackle the following problem: which images from a database should be matched with a query image?

3.1. *A contrario method*

Our approach is similar to the one in [44,23] : we match two images as soon as their proximity is unlikely to be due to chance. This is an application of the general principle of *a contrario* methods: we do perceive events that are unlikely in a white noise situation. Such methods have been successfully applied to several tasks in computer vision following the work of [15]. A recent exposition can be found in the monograph [16]. In the field of shape matching, related ideas can be found in [38,25]. In short, the method works as follows. We first introduce a probabilistic model of generic images, called a *background model*. Given a query image, we then match it to images that are at a distance that is too small to be plausible under the background model. Such matches will be called *meaningful matches*. We now detail how this principle can be applied to the retrieval of color images.

3.2. *Setting*

We write Q for a color query image, and $\mathcal{B} = \{T_1, \ldots, T_{m_\mathcal{B}}\}$ for a database composed of $m_\mathcal{B}$ target color images. For two images Q and T, their EMD distance (as detailed in the previous section) is denoted as $d(Q,T)$. The goal here is to automatically compute a threshold for these distances that is able to distinguish between matches and non-matches. Using a classical statistical setting, we want, for each image T, to develop a test for the hypothesis H_1:

$$H_1 : \text{“ } T \text{ is similar to } Q \text{ ”},$$

relying on the quantity $d(Q,T)$.

3.3. *Meaningful matches*

We make use of an *a contrario* approach similar to the one of [44]. The test relies on a *background model*[2] $\mathcal{M}$ of generic images. The threshold on distances is fixed in such a way that it is very unlikely that an image generated by $\mathcal{M}$ would match the query. In this sense, matches are rare events with respect to $\mathcal{M}$. More precisely, we fix the rate of false alarms when testing H_1 against H_0:

$$H_0 : \text{“ } T \text{ follows the background model ”}.$$

The main idea is to fix the matching thresholds by bounding the probability that two images are matched by chance. Assume that we are able to compute the probability $Pr(d(Q,T) \leq c)$, for any $c \geq 0$, where T is an image distributed according to $\mathcal{M}$ (we write Pr_{H_0} for this probability). Recall that $m_\mathcal{B}$ is the number of images in the database $\mathcal{B}$. For $\epsilon > 0$, we say that

T is an ϵ-meaningful match of Q if $d(Q,T) \leq c_Q$,

where c_Q is defined as

$$c_Q = \sup \left\{ c : Pr_{H_0}\left(d(Q,T) \leq c\right) \leq \frac{\epsilon}{m_\mathcal{B}} \right\}. \quad (4)$$

[2] In this setting, the term *background* has no connection with background models used in object recognition and which refer to models of non-objects.

3.4. *Control of the number of false matches*

The upper bound $\epsilon / m_\mathcal{B}$ in the previous definition is chosen because it ensures that:

if all images T_i from the database $\mathcal{B}$ follow the model $\mathcal{M}$, then the expectation of the number of ϵ-meaningful matches is less than ϵ.

The proof of this fact is quite simple and relies on the linearity of the expectation. Indeed, if $T_1, \ldots, T_{m_\mathcal{B}}$ are distributed according to $\mathcal{M}$, then, by writing E for the mathematical expectation, we obtain,

$$\begin{aligned} &E(\text{there exists } i \text{ such that } d(Q,T_i) \leq c_Q) \\ &\qquad \leq \sum_{1 \leq i \leq m_\mathcal{B}} Pr_{H_0}(d(Q,T_i) \leq c_Q) \\ &\qquad \leq m_\mathcal{B} \epsilon / m_\mathcal{B} = \epsilon. \end{aligned}$$

In short, this means that ϵ-matches are defined so that, on the average, there are fewer than ϵ casual matches when querying a generic database having the same size as $\mathcal{B}$. This is possible because the matching threshold depends upon the size of the database. In practice, the value $\epsilon = 1$ is always considered in this paper, that is, the expected number of false matches is fixed at one. In what follows, the term meaningful match is used instead of 1-meaningful match.

3.5. *Background model*

Of course, the background model $\mathcal{M}$ now has to be defined. The main idea when applying *a contrario* methods is to use a background model in which the basic elements (in our case the spatial color content) are independently distributed. Drawing our inspiration from [17,44], we choose a model leading to images that have similar color marginal distributions as images from the database $\mathcal{B}$. Since the distance d only takes thumbnails into account, we seek a background model for thumbnails.

The first idea was to use a uniform white noise model, in which all pixels of a thumbnail are drawn independently and identically in the $CIELab$ space following the color distribution observed on $\mathcal{B}$. This was not satisfactory, since too many images were then meaningful matches of the query. Our interpretation of this fact is that, even when they are perceptually different, images share some common structures (e.g. the presence of homogeneous regions). In short, the high number of matchings follows from the fact that two random natural images are often closer one to another than they are to realizations of white noise.

A more structured model is needed, with geometric features similar to the ones encountered in real images, such as homogeneous regions and edges. The model we consider is the dead leaves model ([42,6], consisting in the superimposition of random objects, which provides a simple way to generate random images with uniformly colored regions. The random objects are assumed to be independently and identically distributed (i.i.d.). This model has been experimentally proved to be well adapted to the structure of natural images when using power law distributions for the size of objects [1,35]. For the sake of simplicity, we restrict ourselves to an isotropic model where objects are simply disks. The radius of these disks are i.i.d. random variables with density $f(r) \propto r^{-\gamma}$. The model is then characterized by three parameters : the scaling parameter γ, and r_0, r_1 the minimal and maximal sizes of objects [3]. We chose $\gamma = 3$, a typical value for natural images, $r_0 = 1$ and r_1 of the same magnitude as the dimensions of thumbnails. Each object is colored independently of its size and independently of other objects. Observe that this implies that the marginal color distribution of the model is the same as for objects. The distribution of the color of objects (and therefore also the color marginals of the model) is then learned from the color distribution of pixels of the thumbnails from the database. Samples of the model (using color marginals from the database of illuminated manuscripts used in the experimental section) are displayed in Figure 3.

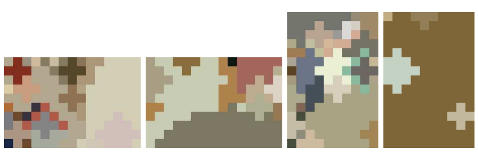


Fig. 3. Samples of the background model $\mathcal{M}$ using color marginals from the illuminated manuscript database used in the experimental section.

The last step, in order to be able to compute meaningful matches of a given query image, is to compute the matching threshold c_Q such that $Pr(d(Q,T) \leq c_Q) \leq \epsilon/m_{\mathcal{B}}$ where T follows $\mathcal{M}$. This quantity relies on geometrical properties of the dead leaves model for which no explicit formula are known. We therefore rely on numerical simulations. The probability is approximated by a frequency over a large number of samples. The choice of the number of samples will be addressed in Section 4.4.

3 One can get rid of r_0 by considering a limit model yielding the same small scale structure as in natural images [24], but this model is unnecessarily complex for the modeling of coarse representations of images.

3.6. *Discussion*

A more classical approach to the statistical testing of hypothesis H_1 (an image from the database is similar to the query) would be to use a probabilistic model for the images that are similar to the query and, for example, to perform a Bayesian test. This would imply the building up of specific models for each category of color images that we deal with, a task that is not compatible with a generic content-based querying system. An alternative is provided by *a contrario* methods, as explained above.

A major advantage of the proposed approach to the matching problem is that the threshold automatically adapts to both the query and the database. First, the threshold is expected to become more conservative as the database gets larger, which can be crucial when dealing with very large databases. Second, since the threshold is computed using color marginal learned on the database $\mathcal{B}$, it adapts to the scarceness or the commonness of the query image. This can be especially useful when dealing with specialized bases. More generally, we will observe in the experimental section that for a fixed ϵ the matching threshold value c_Q varies very much from one query image to another. In fact, this is one of the main advantages of the proposed method : the threshold is only driven by the parameter ϵ, which intuitively corresponds to a number of false matches when submitting a query and is therefore easier to choose than an upper bound on distances. In all experiments to be performed in this paper, we choose $\epsilon = 1$.

4. Experiments

Tests have been performed on two databases. The first one is an illuminated manuscript database provided by the *Institut de Recherche et d'Histoire des Textes* (IRHT), thereafter referred to as the IRHT database. This database contains 1500 high quality and color-calibrated illuminations. The second database, the CLIC [43] database, consists in 15000 non-specialized color photographs and has been gathered by researchers from the French *Com-*

missariat à l'Energie Atomique (CEA). The first database is especially adapted to the purpose of this paper since it contains a large variety of colorful arrangement of patterns. As stressed in the introduction, we believe our approach to be particularly adapted to the querying of cultural heritage databases since the spatial organization of colors summarizes the artistic composition. The CLIC database is less varied in terms of spatial organization despite its larger size. Yet it is quite useful for studying the behavior of the proposed matching threshold. Indeed, it contains very large series of images with similar compositions, such as landscapes in which the sky more or less covers the top half of the picture.

Besides section 4.1 where the choice of thumbnail size n is discussed, all experiments to be presented here use $n = 150$ pixels (10×15 or 15×10). Except where mentioned, tests are performed using parameters $\alpha = 0.5$, $\delta_c = 15$ and $\delta_s = 5$ for the distance d_e (see Equation (3)). The parameters used for the background model are $\gamma = 3$, $r_0 = 1$ and $r_1 = 30$, as justified in Section 3. The expected number of false detection is fixed at one, that is $\epsilon = 1$.

A point that is rarely discussed and that concerns all spatial color methods is how to manage different image ratios. Various choices are available to handle this problem. In the experiments, we chose to separate portrait and landscape orientations. However EMD has no constraints on signature length and total weight (see [47]), and therefore we could also have compared images with different ratios. Yet, this would transforms the assignment problem into a transportation problem leading to a bigger computational load since transportation problems are usually solved using the simplex method.

In Section 4.1, we illustrate the compromise between thumbnail size, result accuracy and computational time. In Section 4.2, we illustrate the effect of the parameter α (see Equation (3)) that balances color and spatial content, and in Section4.3, we compare our results with a region-based method relying on a segmentation followed by the use of EMD, as discussed in Section 2. Both experiments are performed on queries from the IRHT database. Next, we explain in Sectionsec:variability how to numerically compute the matching threshold and we study its variability according to both the query and the database. In Section4.6, we present several query examples from the IRHT and CLIC databases. Before proceeding, let us mention that our study does not include any quantitative evaluation of the results, for instance through the use of a pre-classified database. Yet the definition of the matching threshold would have enabled the comparison between predicted and obtained false matchings. Moreover, it would have been interesting to evaluate the efficiency of the threshold using ROC curves, for instance. However, to the best of our knowledge, none of the existing predefined classification of image databases relies on the spatial organization of colors. Therefore, the efficiency of the retrieval method and the precision of the matching threshold are essentially illustrated by examples, comparisons and a study of the variability of the matching threshold.

4.1. *Size of thumbnails*

In this section we illustrate the effect of the size of thumbnails on some query examples. As explained in Section 2, solving the assignment problem (2) for images having a standard resolution is not computationally feasible. Two query examples have been computed with the following thumbnails size $n = 150, 600, 2400$ (corresponding to 10×15, 20×30 and 40×60 for horizontal thumbnails). Results are displayed in Figure 4 and 5. One can see on these two examples that increasing the thumbnail size has a low impact on results. We observed similar results on other queries from the two databases used in this experimental section. This illustrates that an order of magnitude of several hundreds pixels is enough to achieve coarse description of the spatial organization of colors. Using a PC under Linux with Pentium IV running at 4.3gHz, and C programming code, computational time for querying a database of 1000 images is 18 seconds for $n = 150$, compared to about 10 minutes and 8 hours for $n = 600$ and 2400 respectively. This is coherent with the $O(n^3)$ complexity upper bound of the Hungarian method used to solve the assignment problem. In the rest of the paper, thumbnails of $n = 150$ pixels are always used.

4.2. *Color-spatial weigthing*

As detailed in Section 2, we use EMD to measure the distance between signatures of images. The weighting parameter $\alpha \in [0, 1]$ biases the distance d_e between two features (see Equation (3)) towards either the spatial or the color constraint. If the user wishes to be severe on spatial organization (pixels cannot move freely on the image), α must be set at a low value. The higher the value, the more the

Fig. 4. Three query results with different thumbnail sizes (Left column, from top to bottom: $n = 150, 600, 2400$). The query image is framed in red and is followed by the three closest results. Results and their order slightly change according to the thumbnail size.

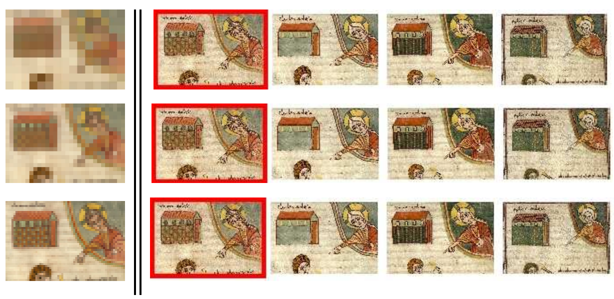


Fig. 5. Same layout as in Figure 4. Results and their order do not change in this case.

objects can move and be split apart. The extreme case where $\alpha = 1$ corresponds to a color histogram method, with EMD as a similarity measure. Two results are displayed in Figure 6, using different values for α.

Fig. 6. **Color spatial weighting using parameter α.** Two query images (framed in red on the left) are followed by their three nearest matches with $\alpha = 0.5$ in Figure a) and $\alpha = 1$ in Figure b). This second case corresponds to color histogram retrieval methods using EMD. We see in Figure b) the loss in the spatial organization.

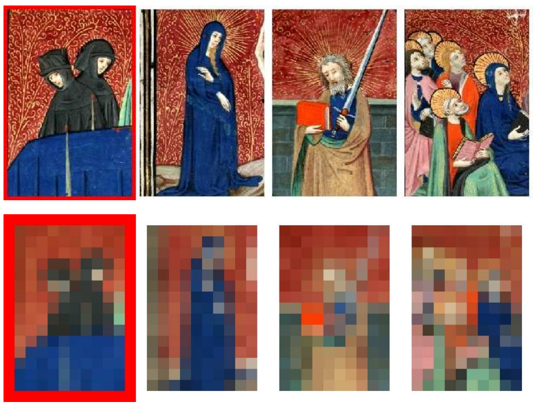
4.3. *Comparison with the use of a segmentation step before computing EMD*

In this section, we address the usefulness of a segmentation step before applying EMD, as proposed for instance in [47,19,32,41]. In our experiments, we choose the Blobworld segmentation [4] [10] that represents regions by using Gaussian mixtures. The original method uses texture and location to extract *blobs*. We only use color and location information for the results to be directly comparable with our method. We nevertheless observe similar retrieval results on the IRHT database when also using texture. The number of blobs per image is an important parameter. It is set by an MDL method (see [10] for details). For the sake of coherence in our comparison, we also use EMD to measure the distance between two blob signatures, the weight of each blob being equal to its size.

Two comparative results are displayed in Figures 7 and 8. In all our tests, the *blobs+EMD* method performs well on queries where simple objects stand on a homogeneous background. In this case both methods perform similarly. As explained before, since the weights of features (regions) are not equal, EMD breaks them apart when optimizing the cost of the mapping. Such a situation is displayed in Figure 7. On the other hand, our method performs better where segmentation methods fail. Indeed, wrong segmentations yield regions that are not well represented by a single feature, highly affecting the performances of the matching step. Figure 8 illustrates this drawback of region-based methods.

[4] Code available at: http://elib.cs.berkeley.edu/blobworld/

a) blobs+EMD

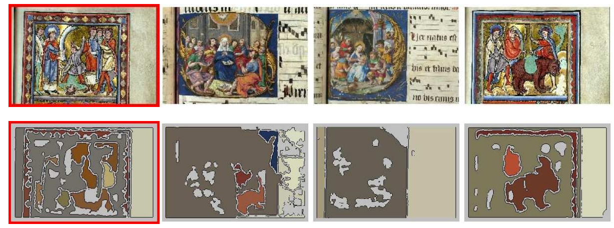


b) our method

Fig. 7. Comparison between a regions + EMD method and our method. A query image (framed in red) followed by its three nearest matches is shown on the first lines of Figure a) and b), according to each method. The blob representations and the thumbnails (10×15) of images are shown on the second line of Figure a) and b) respectively. Here the blob segmentation is relatively satisfactory and both methods perform similarly.

4.4. *Numerical computation of the matching threshold*

Recall that in Section 3.3, we introduced a method to compute a matching threshold when submitting a query image Q to a database $\mathcal{B}$. This threshold, c_Q, depends on both a positive parameter ϵ and the query Q. It enables to define the meaningful matches of Q as images T for which $d(Q,T) \leq c_Q$. Recall also that c_Q is defined by Equation (4). As explained at the end of Section 3.5, the practical computation of c_Q relies on numerical simulations. Namely, it is estimated as an empirical frequency. We first sample $n_{bg} \times m_{\mathcal{B}}$ realizations M_j of the background model, where $m_{\mathcal{B}}$ is the number of images in the database $\mathcal{B}$ and n_{bg} is an integer parameter. We then approximate c_Q by the empirical quantile of order $\epsilon/m_{\mathcal{B}}$ of distances from Q to M_j's. That is, denoting by B_Q the n_{bg}^{th} closest image from Q among $\{M_1, \ldots, M_{n_{bg} \times m_{\mathcal{B}}}\}$, c_Q is obtained as $c_Q = d(Q, B_Q)$.



a) blobs+EMD

b) our method

Fig. 8. Same layout as in Figure 7. Here the blob segmentation fails to give a satisfactory representation of the color spatial organization of each image. Images are clearly under-segmented, yielding large regions with average colors of different patterns, leading to unsatisfactory matching.

The next point is to investigate the number of samples that is needed in order to get correct approximations of c_Q. In order to do so, we perform the following experiment. Background databases of $n_{bg} \times m_{\mathcal{B}}$ realizations of the background model are synthesized, for $n_{bg} = 1, \ldots, 9$. Then, for a query image Q, we count how many images from $\mathcal{B}$ are meaningful matches. Let r_Q be this number of matches.

Figures 9 and 10 show the standard deviation of r_Q as a function of n_{bg}, when considering ten background databases for each n_{bg}. Standard deviations are divided by the average value of r_Q for $n_{bg} = 1$. Figure 9 deals with the IRHT database. This experiment suggests that, despite the high variability of realizations of the dead leaves model, $n_{bg} = 3$ is enough to obtain consistent results with different background databases, and is therefore the numerical value retained in this paper. Figure 10 deals with CLIC database and suggests that $n_{bg} = 3$ is not related to the specific IRHT content, and is also sufficient in the case of databases which are 10 times bigger.

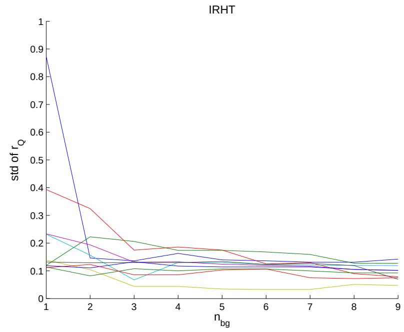


Fig. 9. Standard deviations of the number of retrieved images for 9 queries on the IRHT database. Each colored line corresponds to a different query image. The retrieval is made on a database of $m_{\mathcal{B}}$ images. The standard deviation of the rank of the last meaningful match (r_Q) is shown for background databases of size $m_{\mathcal{B}} \times n_{bg}$, for $n_{bg} = 1, \ldots, 9$. For each n_{bg}, ten background databases have been generated in order to estimate this standard deviation. Standard deviations are divided by the average value of r_Q for $n_{bg} = 1$.

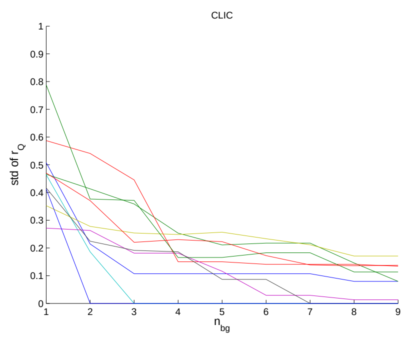


Fig. 10. Same layout as on Figure 9. Standard deviations of the number of retrieved images for 9 CLIC queries.

4.5. *Variability of the matching threshold*

In the following experiment, we computed all the 1500 possible IRHT queries Q with $n_{bg} = 3$. The bidimensional histogram of (c_Q, r_Q) is displayed in Figure 11 (recall that for a query Q, c_Q is the matching threshold and r_Q the number of matches). This experiment gives several insights into the proposed adaptive retrieval. The range of possible matching thresholds c_Q is large, as well as the range of the number of meaningful matches r_Q, which varies from 0 to 91 images. This shows that the matching threshold is quite different from a bound on distances or ranks and that it adapts to the query. This experiment also shows that r_Q and c_Q are not strongly correlated. A large r_Q indicates series of images in the database like the one appearing in Figure 12. We see on the histogram that this can happen with low and high values of c_Q. This example corresponds to the upper point in Figure 11, near $(0.3, 90)$.

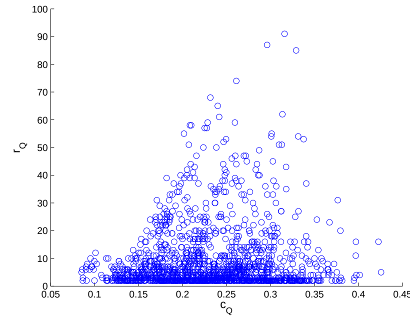


Fig. 11. Bidimensional histogram of (c_Q, r_Q) over all the 1500 *IRHT* queries. c_Q is the matching threshold and r_Q the number of images closer to the query than c_Q (meaningful matches).

Fig. 12. The query and the first 9 matches (among a total of 90 matches) for the upper point in Figure 11. The spatial organization of colors of the matches is very similar to the one of the query, denoting the presence of a large series of similar images in the database. The last thumbnail (framed in blue) is the background model thumbnail retained to compute the matching threshold.

In Figure 13, the same histogram is shown with the CLIC database, on 2000 different queries. In this case, there is a stronger correlation between r_Q and c_Q. Simultaneous high values for r_Q and low values for c_Q reveal the presence of very large series in this database which are mostly landscapes series.

4.6. *Examples of queries*

In the experiments, we display the query image, followed by meaningful matches from the database, in decreasing order up to the matching threshold

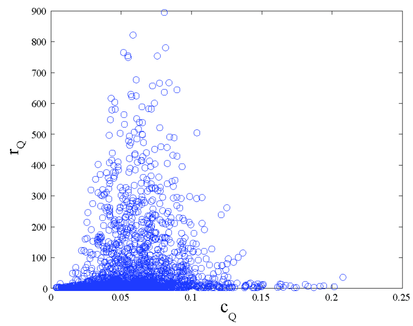


Fig. 13. Bidimensional histogram of (c_Q, r_Q) over 2000 *CLIC* queries. c_Q is the EMD distance matching threshold, and r_Q is the number of results closer to the query than c_Q (meaningful matches).

indicated by a grey thumbnail followed by the first five images that do not match the query.

Retrieval results with the IRHT illuminated manuscripts database are displayed in Figures 14 to 22. In these examples, the number of meaningful matches r_Q varies from 0 (second result of Figure 17) to 23 (Figure 16) according to the query and the the database. The matching threshold c_Q sometimes stops after or before where an observer would stop the list of results, but results are most of the time very relevant. The matching threshold is a good indication of where the results should stop and it also informs the user on the quantity of relevant images present in the database.

In Figure 22, a retrieval is performed with a sketch query made with a simple painting tool. If a user roughly remembers a color composition, querying can also be performed this way.

Retrieval results with the CLIC photo database are shown in Figures 23 to 28. The original CLIC database is divided into approximately a hundred categories such as *Landscape*, *City*, *Architecture*. Many images from different categories can have the same spatial organization of colors and this appears on many results. For instance, in Figure 26, the query belongs to the category *Architecture/Detail*. The first six results belong to the category *City/Building*.

An example which illustrates how EMD works on thumbnails is shown in Figure 27 with a synthetic image from the CLIC database. The blue region at the center of the query moves a little on the closest result. On the last result, the blue region breaks apart into four small regions. Even though its composition does not respect the query, its highly similar color distribution allows larger pixel movements. Increasing the α parameter would discard this image.

Even though the CLIC database is 10 times bigger than the IRHT database, the matching threshold still enables us to identify rare spatial organizations of colors. In Figure 28, from a query that is indeed quite unusual in the CLIC database, no match is returned. On the opposite, large series of similar composition can yield a high number of matches as in Figure 29, where a query has 72 meaningful matches. Only the first four matches and last five matches are visible here, as well as the first five images beyond the matching threshold.

5. Conclusion and future work

In this paper, we first introduced a method for the comparison of the spatial organization of colors between images, relying on a global optimization procedure through the use of the Earth Mover's Distance. The main claim here is that it is unnecessary to segment the image, a step which usually reduces the robustness of comparisons. Next, we apply the resulting distance to color image querying and address the decision stage of the system, that is the problem of deciding which images from a database match a query. This step is achieved by controlling the rate of false matchings using an *a contrario* approach. This work opens several perspectives. As mentioned earlier, we believe the method to be particularly adapted to the querying of cultural heritage databases. We are currently working with several artistic institutions to build large (typically over 100,000) specialized databases. We are also working on the installation of the system directly on web servers aiming at presenting the work of a collective artists studio [29]. For such applications, it will be interesting to investigate further the implications of the various parameters involved in the definition of distance (3) or to develop distances specially adapted to, for example, black and white photographs. Next, several accelerations of EMD computations have been recently proposed [31,3,39] and can be of interest for our approach. Another possible direction to speed up the query would be to impose constraints on the EMD between thumbnails, by forbidding large distances between pixels. This would yield sparse cost matrices and a faster computation of the assignment problem, see [8,9]. Alternatively, one could adopt a coarse to fine strategy. For such a strategy, the solution of the optimization problem at a coarse scale would constrain the computation at a finer scale. We also started to test the approach on thumbnails having more information than just one color at each pixel, for instance a color array or texture features. In another direction, it is of interest to investigate the use of EMD to compare the color content of small local image patches, as it has been done recently [39] for comparing the geometrical content of such local patches.

6. Acknowledgments

We thank Gilles Kagan from IRHT for providing the IRHT database and for his helpful comments. We also thank Farida Cheriet from LIV4D for her support. This work was partially supported by a CNRS ACI supervised by Anne-Marie Edde and Dominique Poirel from IRHT, a grant from the LIV4D laboratory, and a FQRNT grant from the Quebec government.

References

[1] L. Alvarez, Y. Gousseau, and J.-M. Morel. The size of objects in natural and artificial images. *Advances in Imaging and Electron Physics, Academic Press*, 111:167–242, 1999.

[2] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes. Proposal for a Standard Default Color Space for the Internet-sRGB. *Proceedings, IS&T/SID Fourth Color Imaging Conference: Color Science, Systems, and Applications*, pages 238–246, 1996.

[3] I. Assent, A. Wenning, and T. Seidl. Approximation techniques for indexing the earth mover's distance in multimedia databases. *22nd International Conference on Data Engineering (ICDE'06)*, 2006.

[4] S. Berretti, A. Del Bimbo, and E. Vicario. Weighted walkthroughs between extended entities for retrieval by spatial arrangement. *Multimedia, IEEE Transactions on*, 5(1):52–70, 2003.

[5] A. Del Bimbo. *Visual information retrieval*. Morgan Kufman Publishers, 1999.

[6] C. Bordenave, Y. Gousseau, and F. Roueff. The dead leaves model: a general tessellation modeling occlusion. *Adv. in Appl. Probab*, 38(1):31–46, 2006.

[7] B. Burford, P. Briggs, and J.P. Eakins. A Taxonomy of the Image: On the Classification of Content for Image Retrieval. *Visual Communication*, 2(2):123, 2003.

[8] R.E. Burkard and E. Çela. *Linear assignment problems and extensions*. Karl-Franzens-Univ. Graz & Techn. Univ. Graz, 1998.

[9] R.E. Burkard, M. Dell'Amico, and S. Martello. *Assignment Problems*. SIAM Monographs on Discrete Mathematics and Applications. To appear.

[10] C. Carson, S. Belongie, H. Greenspan, and J. Malik. Blobworld: image segmentation using EM and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 24(8):1026–1038, 2002.

[11] SK Chang, QY Shi, and CW Yan. Iconic indexing by 2-D strings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(3):413–428, 1987.

[12] H. Chen. An analysis of image retrieval tasks in the field of art history. *Information Processing and Management*, 37(5):701–720, 2001.

[13] G. Ciocca, R. Schettini, and L. Cinque. Image indexing and retrieval using spatial chromatic histograms and signatures. *CGIV*, pages 255–258, 2002.

[14] L.R. Conniss, A.J. Ashford, and M.E. Graham. *Information Seeking Behaviour in Image Retrieval VISOR I Final Report*. Institute for Image Data Research, University of Northumbria at Newcastle, 2000.

[15] A. Desolneux, L. Moisan, and J-M Morel. Meaningful alignments. *International Journal of Computer Vision*, 40:7–23, 2000.

[16] A. Desolneux, L. Moisan, and J.-M. Morel. From Gestalt Theory to Image Analysis: A Probabilistic Approach, volume 34 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 2008. in Press.

[17] A. Desolneux, L. Moisan, and J.M Morel. Edge detection by Helmohltz principle. *International Journal of Computer Vision*, 14:271–284, 2001.

[18] G. Dvir, H. Greenspan, and Y. Rubner. Context-based image modelling. *ICPR*, pages 162–165, 2002.

[19] G. Dvir, H. Greenspan, and Y. Rubner. Context-Based Image Modelling. *Proc. Int Conf. ICPR2002*, pages 162–165, 2002.

[20] J.P. Eakins, P. Briggs, and B. Burford. Image retrieval interfaces: A user perspective. *Proc. Third Int. Conf. on Image and Video Retrieval, Springer-Verlag Heidelberg, Dublin, Ireland*, pages 628–637, 2004.

[21] Y. Fang, D. Geman, and N. Boujemaa. An interactive system for mental face retrieval. In *7th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR'05)*, 2005.

[22] L. R. Ford and D. R. Fulkerson. Solving the transportation problem. *Management Science*, 3:24–32, 1956.

[23] Y. Gousseau. Comparaison de la composition de deux images, et application à la recherche automatique. In *proceedings of GRETSI 2003*, Paris, France, 2003.

[24] Y. Gousseau and F. Roueff. Modeling occlusion and scaling in natural images. *SIAM Multiscale Modeling and Simulations*, 6(1):105–134, 2007.

[25] W. E. L. Grimson and D. P. Huttenlocher. On the verification of hypothesized matches in model-basedrecognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(12):1201–1213, 1991.

[26] G. Heidemann. Combining spatial and colour information for content based image retrieval. *Computer Vision and Image Understanding*, 94:234–270, 2004.

[27] W. Hsu, T. S. Chua, and H. K. Phung. An integrated color-spatial approach to color image retrieval. In *3rd ACM Multimedia Conference*, pages 305–313, 1995.

[28] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih. Spatial color indexing and applications. *International Journal of Computer Vision*, 35(3):245–268, 1999.

[29] T. Hurtut, F. Cheriet, and M. Chronopoulos. A novel collaborative website and artworks database management system for artist-run centres. In *International Cultural Heritage Informatics (ICHIM)*, Oct. 2007.

[30] T. Hurtut, H. Dalazoana, Y. Gousseau, and F. Schmitt. Spatial color image retrieval without segmentation using thumbnails and the earth mover's distance. CGIV, Leeds, England, june 2006.

[31] P. Indyk and N. Thaper. Fast image retrieval via embeddings. *3rd Intl. Workshop on Statistical and Computational Theories of Vision*, 2003.

[32] F. Jing, M. Li, H. J. Zhang, and B. Zhang. An effective region-based image retrieval framework. *IEEE Trans. on Image Processing*, 13(5):699–709, 2004.

[33] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97, 1955.

[34] P. Lambert, N. Hervey, and H. Grecu. Image retrieval using spatial chromatic histograms. *CGIV*, pages 343–347, 2004.

[35] A. Lee, D. Mumford, and J. Huang. Occlusion models for natural images: A statistical study of a scale invariant dead leaves model. *International Journal of Computer Vision*, 41:35–59, 2001.

[36] E. Levina and P. Bickel. The earth movers distance is the Mallows distance: Some insights from statistics. *Proc. ICCV*, 2:251–256, 2001.

[37] M. S. Lew, editor. *Principles of Visual Information Retrieval*. Springer, 2001.

[38] M. Lindenbaum. An integrated model for evaluating the amount of data required for reliable recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 19(11):1251–1264, 1997.

[39] H. Ling and K. Okada. An efficient earth mover's distance algorithm for robust histogram comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 29(5):840–853, 2007.

[40] P. Lipson, E. Grimson, and P. Sinha. Configuration based scene classification and image indexing. In *IEEE Proc. Comp. Vis. Patt. Rec.*, pages 1007–1013, 1997.

[41] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma. Region-based image retrieval with high-level semantic color names. *IEEE Int. Multimedia Modelling Conference*, pages 180–187, 2005.

[42] G. Matheron. Modèle séquentiel de partition aléatoire. Technical report, CMM, 1968.

[43] P. A. Moellic, P. Hede, G. Grefenstette, and C. Millet. Evaluating content based image retrieval techniques with the one million images CLIC testbed. In *Proc. of PRCV*, 2005.

[44] P. Musé, F. Sur, F. Cao, Y. Gousseau, and J.-M. Morel. An *a contrario* decision method for shape element recognition. *International Journal of Computer Vision*, 69(3):295–315, 2006.

[45] G. Pass and R. Zabih. Comparing images using joint histograms. *Journal of Multimedia Systems*, 1999.

[46] B. G. Prasad, K. K. Biswas, and S. K. Gupta. Region-based image retrieval using intergrated color, shape, and location index. *Computer Vision and Image Understanding*, 94:193–233, 2004.

[47] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

[48] J. D. Rugna and H. Konik. Color coarse segmentation and regions selection for similar images retrieval. *CGIV*, pages 241–244, 2002.

[49] S. Siggelkow and H. Burkhardt. Improvement of histogram-based image retrieval and classification. In *IAPR International Conference on Pattern Recognition*, volume 3, pages 367–370, 2002.

[50] J. R. Smith and S. F. Chang. Tools and techniques for color image retrieval. In *Procedings of SPIE, Storage and Retrieval for Image and Video Databases IV*, volume 2670, pages 1630–1639, 1996.

[51] J. R. Smith and C.-S. Li. Image classification and querying using composite region templates. *Computer Vision and Image Understanding*, 75(1/2):165–174, 1999.

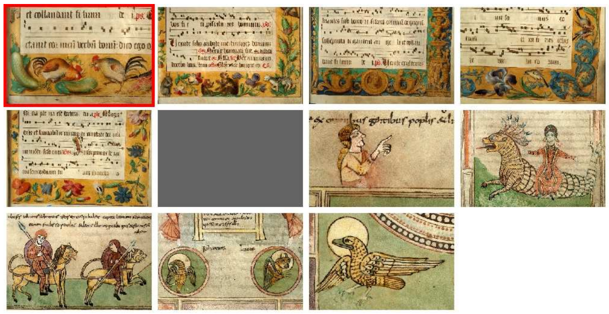
[52] M. Stricker and A. Dimai. Color indexing with weak spatial constraints. In *Procedings of SPIE, Storage and Retrieval for Image and Video Databases IV*, pages 29–40, 1996.

[53] Michael J. Swain and Dana H. Ballard. Color indexing. *Int. J. Comput. Vision*, 7(1):11–32, 1991.

[54] J. van de Weijer and C. Schmid. Coloring local feature extraction. In *European Conference on Computer Vision*, 2006.

[55] J. Z. Wang, J. Li, and G. Wiederhold. Simplicity: semantic-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 23(9), 2001.

[56] G. Wyszecki and WS Stiles. Color Science: Concepts and Methods, Quantitative Data and Formulae. *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition, by Gunther Wyszecki, WS Stiles,*, 2000.



a)

b)

Fig. 14. The query image is framed in red (top left) in figure a) and is followed (left from right and top to bottom) by its meaningful matches in decreasing order, up to the grey image which corresponds to the matching threshold. Beyond this limit, the next five nearest images are shown. The corresponding features (10×15 thumbnails) are shown in figure b). The last thumbnail (framed in blue) is the background model thumbnail retained to compute the matching threshold.

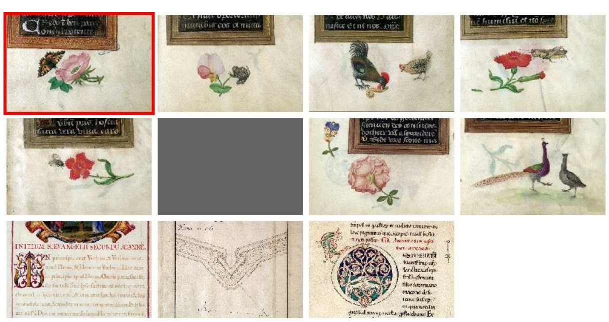


Fig. 15. The query image is framed in red (top left) in figure a) and is followed (left from right and top to bottom) by its meaningful matches in decreasing order, up to the grey image which corresponds to the matching threshold. Beyond this limit, the next five nearest images are shown.

Fig. 16. Same layout as in Figure 15. We see here that the matching threshold allows very variable number of matches according to the database content and the query. There are 23 meaningful matches. The database contains approximately a hundred white illuminations with a blue rectangle. On images that are not considered as matches, the blue rectangle has a different size and/or is not at the same location as in the query.

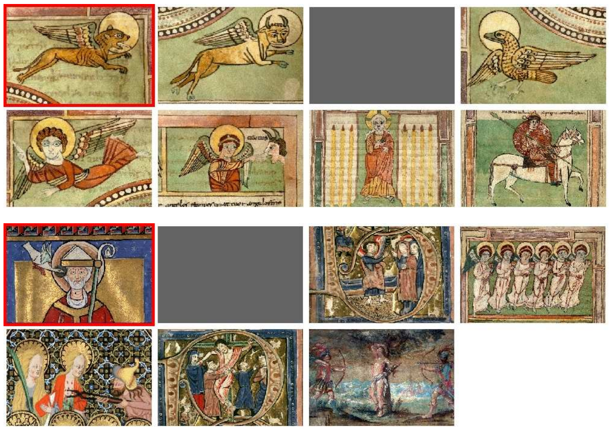


Fig. 17. Same layout as in Figure 15. Two examples where the matching threshold provides very few meaningful matches. On the second example, it gives no result and indeed no such spatial organization can be encountered in the database.

Fig. 18. Same layout as in Figure 15.

Fig. 19. Same layout as in Figure 15.

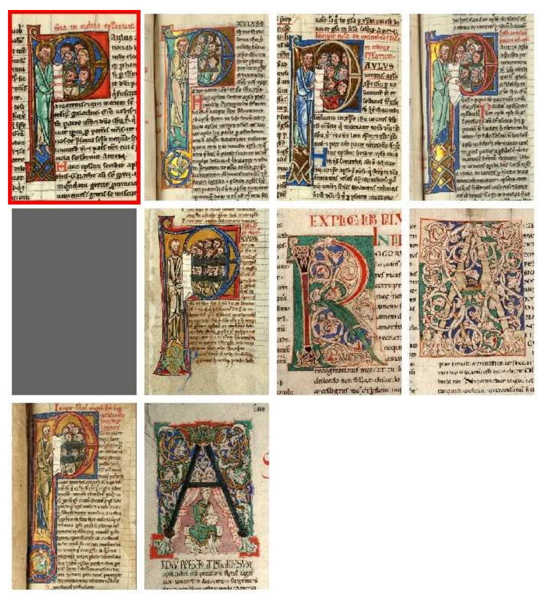


Fig. 20. Same layout as in Figure 15.

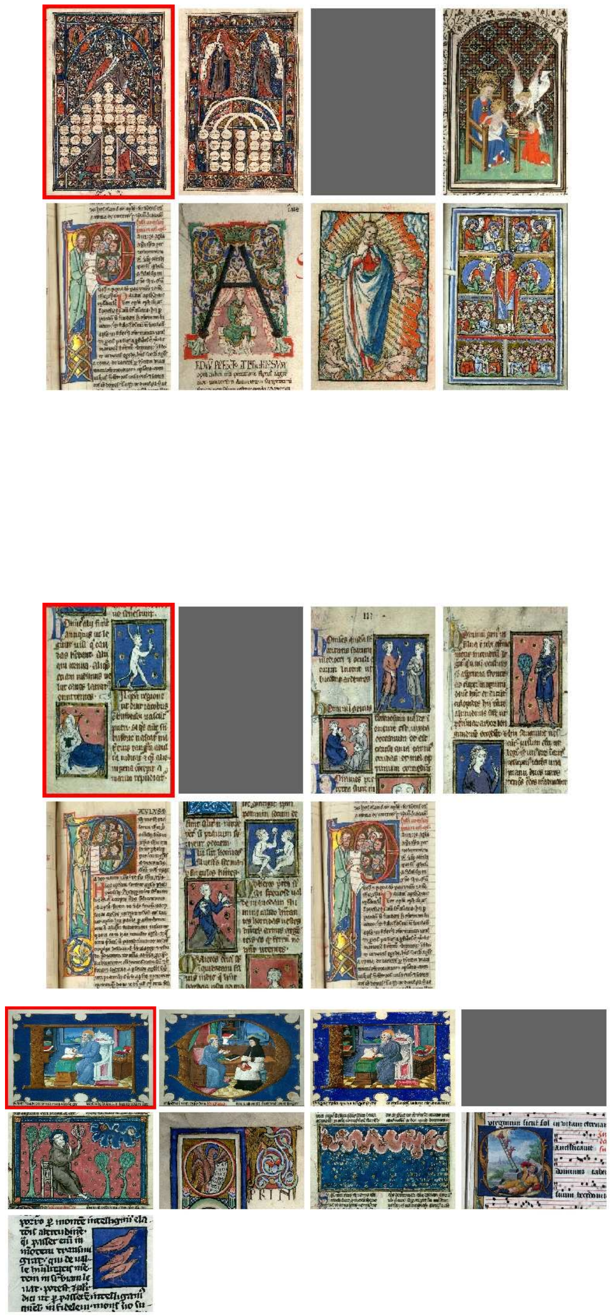


Fig. 21. Same layout as in Figure 15. Three examples with a small number of meaningful matches.

Fig. 22. Same layout as in Figure 15. Example of a retrieval from a sketch obtained with a simple painting software. In order for such a methodology to enable the retrieval of images we have a vague memory of, one can allow more flexibility on color through the parameter δ_C of the distance d_e (see (3)).

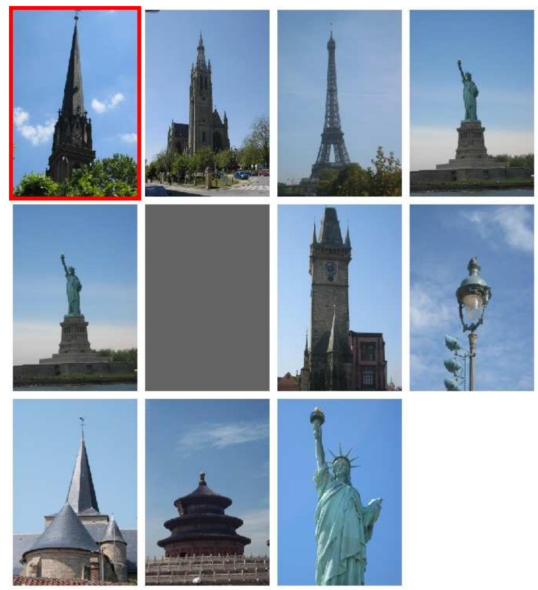


Fig. 23. Same layout as in Figure 15 with a query from the CLIC database.

Fig. 24. Same layout as in Figure 15 with a query from the CLIC database.

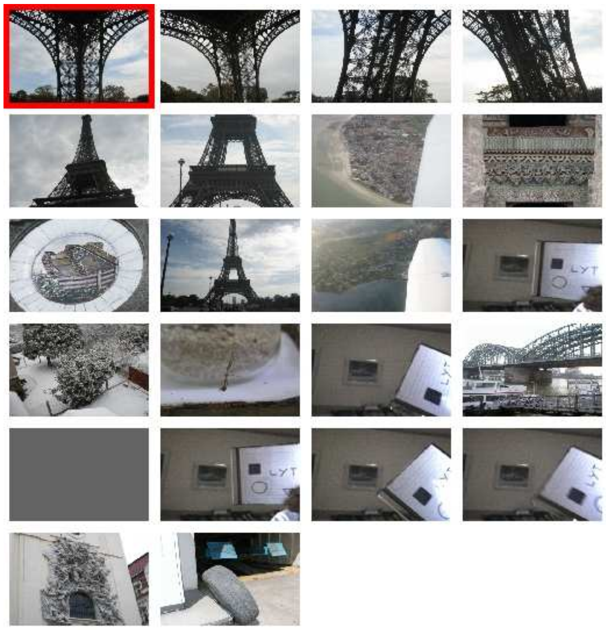


Fig. 25. Same layout as in Figure 15 with a query from the CLIC database.

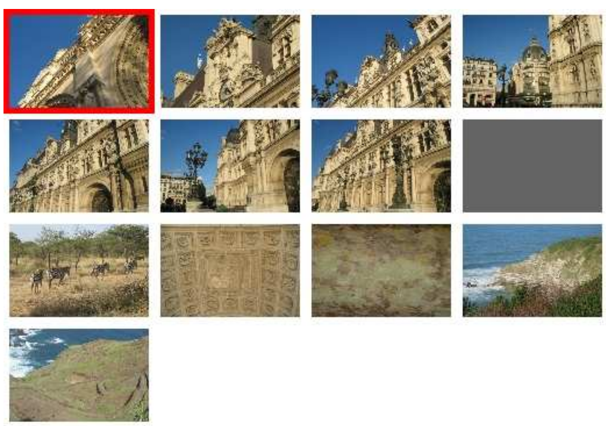


Fig. 26. Same layout as in Figure 15 with a query from the CLIC database.

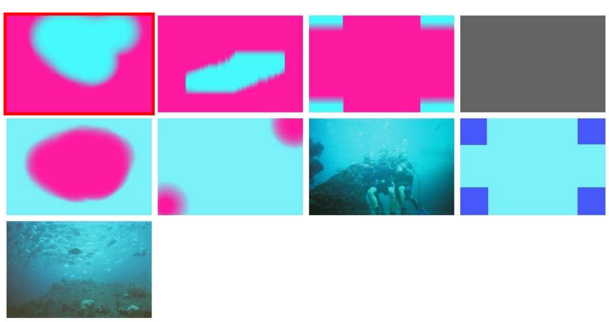


Fig. 27. Same layout as in Figure 15 with a query from the CLIC database.

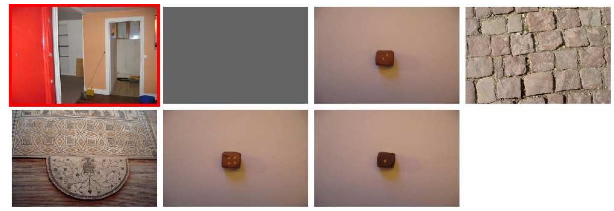


Fig. 28. Same layout as in Figure 15 with a query from the CLIC database. Even though this database is 10 times bigger than the IRHT database, the matching threshold still allows a query having a unique spatial organization of colors in the database to return no match.

Fig. 29. Same layout as in Figure 15 with a query from the CLIC database. Large series of similar spatial organization occur in this database. Here is shown a query which returns 72 results. Only the first four matches and last five meaningful matches are shown, as well as the next five images beyond the matching threshold.