Randomize color signature employing locality-sensitive hashing (LSH)

Bahri abdelkhalak [1] and hamid zouaki[2]

[1]Faculty of Sciences, University Chouaib Doukkali,
Equipe: Modélisation mathématique et informatique décisionnelle
El Jadida, Morocco

[2]Faculty of Sciences, University Chouaib Doukkali,
Equipe: Modélisation mathématique et informatique décisionnelle
El Jadida, Morocco

Abstract

Visual signatures construction of images is an essential step for the CBIR system. When the database size become larger. Most existing algorithms (e.g. k-means, Kd-tree, Mean-shift) to build signature become unfavorable due to the prohibitive time and space requirements. In this paper we propose the randomize color signature based on the LSH technique. The proposed descriptor benefited of the effectiveness of LSH in terms of time and accuracy of clustering.
Experiments show the effectiveness of our approach.
Keywords: *CBIR, Color, signature, KD-Tree, Randomize, LSH.*

1. Introduction

The dramatic increase in the volume of image databases requires going beyond a manual annotation. To meet this obligation, the content-based approach, to directly extract the relevant information from the image, appears as an alternative textual approach. This new modality has opened opportunities for users.

Content based image retrieval (CBIR), instead of using text annotations as the basis for indexing and searching, uses visual features extracted from images, such as color, texture, shape and spatial relations of pixels. Unlike text annotations which are subject to human perception, these features make objective representations of images. Since early 1990's, many CBIR systems have been developed [1,2].

Color plays an important role in image search by content. This visual information is the most used search systems by content [3,4,5,6]. He is considered a powerful cue for content-based image retrieval (CBIR) [16][17] and is also an effective feature in image analysis because color is robust with noise, image orientation, and resolution [15,17,18,19]. These dimensional values make its discriminatory potential is greater than the value of gray scale images. Before selecting the appropriate color descriptor, the color must be determined first. In most image retrieval system based on visual features, *a histogram [7]* (or *a fixed-binning histogram*) is widely used as a visual feature descriptor due to its simple implementation and insensitivity to similarity transformation [14]. Several improvement of it has been developed [8,9,10,11]. However, in some cases, the histogram based indexing methods fail to match perceptual dissimilarity [12]. The performance of retrieval system employing a histogram as a descriptor severely depends on the quantization process in feature space because a histogram is inflexible under various feature distribution representations. To overcome these drawbacks, a clustering based representation, *signature* (or *adaptive-binning histogram*) has been proposed [12, 20, 21]. A signature compactly represents a set of clusters in feature space and the distribution of visual features. Color signatures that are developed by Rubner et al [12] that are obtained by classifying their pixels by the KD-Tree [13] method have present theirs efficiency in terms of accuracy compared to color histograms. However, this method becomes ineffective because of the time of pixel classification. Hurtut and Gousseau introduce in [22] the color thumbnail. He had be configured as follows. If it is possible to do, in the first step, the RGB channels of the Gamma factor are corrected for the tristimulus values. According to the additive laws Grassman [23], the tristimulus values can then be averaged to obtain the thumbnails. Then we convert these thumbnails to psycho-metric CIELab space. Felzenszswalb and Huttenlocher [24] proposed an efficient approach to grouping pixels in an image by making use of a spanning tree and showed that locally greedy grouping decisions can yield plausible results. This approach also revolves around the computation of multiple pair wise distance values.

Although signatures that are proposed in the literature present their effectiveness in terms of accuracy, these methods suffer the problem of the construction time that is important.

In order to improve the construction time of the signatures, a fast algorithm for construction color signature (Randomize color signature) is proposed. This algorithm use the locality- sensitive hashing (LSH[25,26]) data structure. We only give a brief review on it in the next section.

The method proposed in this paper avoids the computational costs associated with distance computation in favor of a randomized hashing approach which relies upon the locality preserving properties of the hashing function.

The rest of the paper is organized as follows. Section 2 Describe the Locality-sensitive hashing structure. Section 3 Present our contribution. Section 4 evaluates the construction time and the result search quality of our proposed signature experimentally. Finally, Section 8 concludes the paper.

2. Locality-sensitive hashing

The basic idea of LSH is to hash the input data so that similar items are mapped to the same buckets with high probability. The algorithm has two main parameters: the width parameter k and the number of hash tables L. Firstly, a family of hash functions H is defined; and then L group hash functions g is obtained by concatenating k randomly chosen hash functions L times from set H. In other words, the algorithm constructs L hash tables, each corresponding to a different randomly chosen hash function g and g is consisted with k randomly chosen hash functions in H.

Definition 1 (LSH). Given a distance r, approximation ratio c, probability values p_1 and p_2 such that $p_1 > p_2$, a hash function h(.) is (r, cr, p_1, p_2) locality sensitive if it satisfies both conditions below:

1. $If\ \|O_1, O_2\| \le r, \quad then\ Pr[h(O_1) = h(O_2)] \ge p_1$ (1)
2. $If\ \|O_1, O_2\| > cr, \quad then\ Pr[h(O_1) = h(O_2)] \le p_2$ (2)

LSH functions are known for many distance metrics. For l_p norm, a popular LSH function is defined as follows [28]:

$$h(o) = \left\lfloor \frac{a.o + b}{w} \right\rfloor \quad (3)$$

Where, ***o*** represents the d-dimensional vector; ***a*** is another d-dimensional vector so-called p-stable distribution [28]; ***a.o*** denotes the dot product of these two vectors. ***w*** is a sufficiently large constant, and finally, ***b*** is uniformly drawn from [0,w). To increase the gap between right and wrong detection, authors [10] have proposed to construct a family of hash function obtained by concatenating *M* functions h_i as g(p) = [h_1(p), h_2(p), . . . , h_M(p)], with *M* a fixed integer (> 1). The data structure is constructed by placing each point p from the input set into a bucket g_i(p), for j = 1, . . . , L. These hash functions do not allow to bring direct addressing. Classical method to handle this problem is to use a universal hash. Hash function $t_1^L : U^M$ {0, . . . , tableSize} is associated with each table *l*, and thus with each function g_l, *l=1, . . . ,L*. this function will be used to store the LSH buckets in an array of a fixed size (denoted tablesize).). All L hash tables use the same primary hash function V_1 (used to determine the index in the hash table) and the same secondary hash function V_2. These two hash functions have the form

$$V_1(a_1, a_2, \dots, a_k) = \left(\left(\sum_{i=1}^{k} r_i^{'} a_i\right) mod\ P\right) mod\ tableSize \quad (4)$$

$$V_2(a_1, a_2, \dots, a_k) = \left(\left(\sum_{i=1}^{k} r_i^{''} a_i\right) mod\ P\right) \quad (5)$$

Where r_i' and r_i'' are random integers, *tableSize* is the size of the hash tables, and *P* is a prime. In the current implementation, ai are represented by 32-bit integers, and the prime P is equal to $2^{32} - 5$. This value of the prime allows fast hash function computation without using modulo operations. If there are *l* hash tables, as *l* increases, more buckets are examined. Recall is improved but precision may become worse. As *tablesize* increases, the bucket size becomes smaller and more false positives are removed. Precision increases but recall degrades. Similarly, as *tablesize* decrease the bucket size becomes bigger and more true positives are retrieved, but search time is increases.

3. Randomize color signature

To compute the colour signature of a image, we first smooth each band of the image's RGB representation slightly to reduce possible colour quantization and dithering artifacts. We then transform the image into S-CIELab. At this point each image can be conceived as a distribution of points in CIELab, where a point corresponds to a pixel in the image.

Algorithm 1 randomize color signature via LSH

1. We smooth each band of the image's RGB representation slightly to reduce possible color quantization and dithering artifacts.
2. We transform the image into CIELab space.
3. We calculate the first and the second key to each image pixel using the (4) and (5) equation

respectively and we stock it in its bucket correspondent.

4. Each bucket contributes a pair ***(p, w)*** *to the signature representation of the image, where* ***p*** *is the mean of the cluster, and* ***w*** *is the fraction of image pixels in that bucket.*

The set of (pi, wi) for each bucket form the image signature.

The image signatures obtained by the approach presented in [27] and our approach can be of different sizes, and do not retain the same color. For this, the metrics that are based on the comparison component to component can't be applied. The distances of the cross-components are the only ones that can be applied to this kind of signatures. The Earth Mover's Distance(EMD) [27] cited the best metric of comparison for this type of signatures. We give a brief detail on it in the next paragraph.

In [27], the Earth Mover's Distance (EMD) is introduced as the smallest amount of work needed to match a set of weighted features $(f_i, w_i)_{i=1,\dots,n}$ 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 $P=(p_i, w_i)_{i=1,\dots,n}$ and $Q=(q_i, u_i)_{i=1,\dots,m}$ having same total weight is defined in [27] as:

$$EMD(Q,P) = \min_{f_{ij}} \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} f_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}}$$

With the following constraints:

$$f_{ij} \geq 0, 1 \leq i \leq m, 1 \leq j \leq n$$

$$\sum_{j=1}^{n} f_{ij} \leq w_i, 1 \leq i \leq m$$

$$\sum_{i=1}^{m} f_{ij} \leq u_j, 1 \leq j \leq n$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = \min\left(\sum_{i=1}^{m} w_i, \sum_{j=1}^{n} u_j\right)$$

The first constraint only allows movement of components from P to Q. The two first constraints limit the amount of displaced components of P, and quantity of components received by Q at their respective weights. The last constraint implies the maximum possible displacement of components.

4. Experimental Results

We have two tasks in the experiments. We compare the speed and accuracy obtained by our method over the signatures obtained by KD-tree. We adapt the hash parameters to be a compromise between construction time and signature qualities.

In order to characterize the performance of the proposed signature were carried out using three image databases, which are described in the following paragraphs.

First database: Wang database [28] which contains 1000 images extracted from the COREL database. The images are with the size of either 256×384 or 384 × 256. One image per category is presented on Fig1.

Fig.1: Examples extracted from the Wang collection.

Second database: Coil-100 [29], which contains 100 class, each class contains 72 images. One image per class is presented on Fig 2.

Fig.2: One image per class extracted from the Coil-100 database.

Third database: Oxford Flowers database is available online [30]. This dataset consists of 1360 labeled images of 17 categories, with 80 images per category. This dataset is very challenging because there are large variation between same category while small variation between different ones considering viewpoint, scale, and illumination.

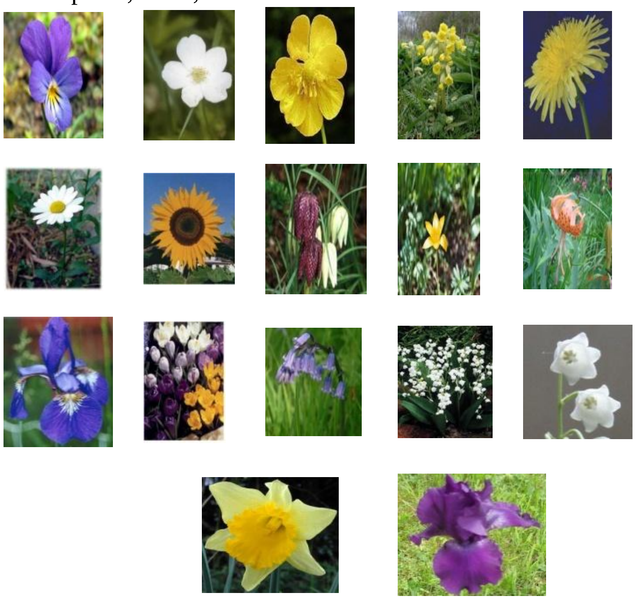

Fig.3: Sample images (one image per category) from Oxford Flowers database.

In this part, some experiments have been conducted to illustrate the effectiveness of the proposed method.

4.1 Real Time Implementation

A significant advantage of the proposed signature construction is the computational effort required scales linearly in the number of pixels and the operations required are simple and regular. In order to demonstrate this a real time of the building is calculated a quad core of an Intel processor running at 1.6 GHz PC with 4 GB memory and the Windows 7 operating system. The image retrieval system is built in Java. This rate includes the time taken for all phases of the algorithm, for example for an image of 384 × 256 pixels the construction time of signatures is presented in table-1.

Table-1 construction time result

Descriptor	**Construction time(s)**
Signature by KD-Tree	0.240
Our method	0.105

For the big size image database, the running time to build the signature of images becomes remarkable. The Fig 4 present an example for different size of databases.

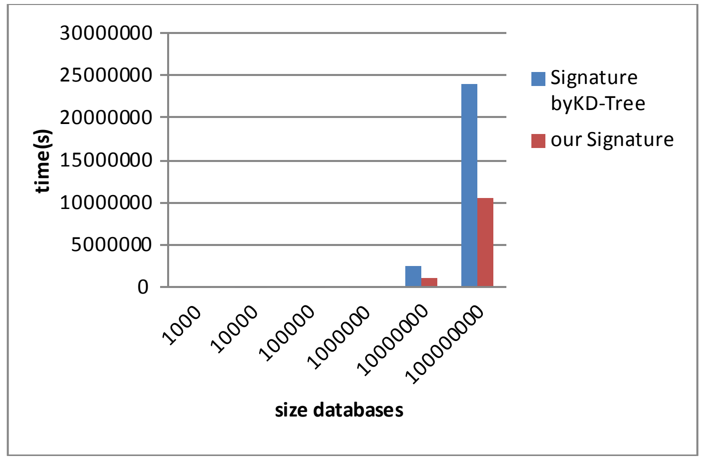

Fig.4: Running time to building signature of image databases for different size

From this figure we see that to build the signatures for a big database using Kd-Tree is more expensive in term of time. By using our method we have improve the construction time for the image signature more than 50% without degradation of its quality.

4.2 pertinence studies

The goal of this experiment is to show that, the effectiveness of our descriptor does not degrade in terms of accuracy. This is to evaluate and compare the accuracy of the results obtained with the descriptor proposed by Rubner et al [12], since it gives better results. The evaluation is carried out through the return of results on three image databases described above.

We use precision vs. recall figures, a standard evaluation technique for retrieval systems [31], for comparing the effectiveness of our algorithms. Precision (P) is the fraction of the retrieved objects which are relevant to a given query, and recall (R) is the fraction of the relevant objects which have been retrieved from the database. If R is the set of relevant objects to the query, A is the set of objects retrieved, and R_A is the set of relevant objects in the result set, then $P = |R_A|/|A|$ and $R = |R_A|/|R.|$

In the research process, we enter a query image by comparing it signature with theirs of image databases, we obtained a images set ranked and sorted using the EMD metric as a similarity measure. The figure 5 shows the curves precision / recall using signatures that we proposed and those proposed by Rubner et al. [12] for three image databases. We note that the experimental results are close in terms of accuracy. And consequently our descriptor is efficient and faster to build.

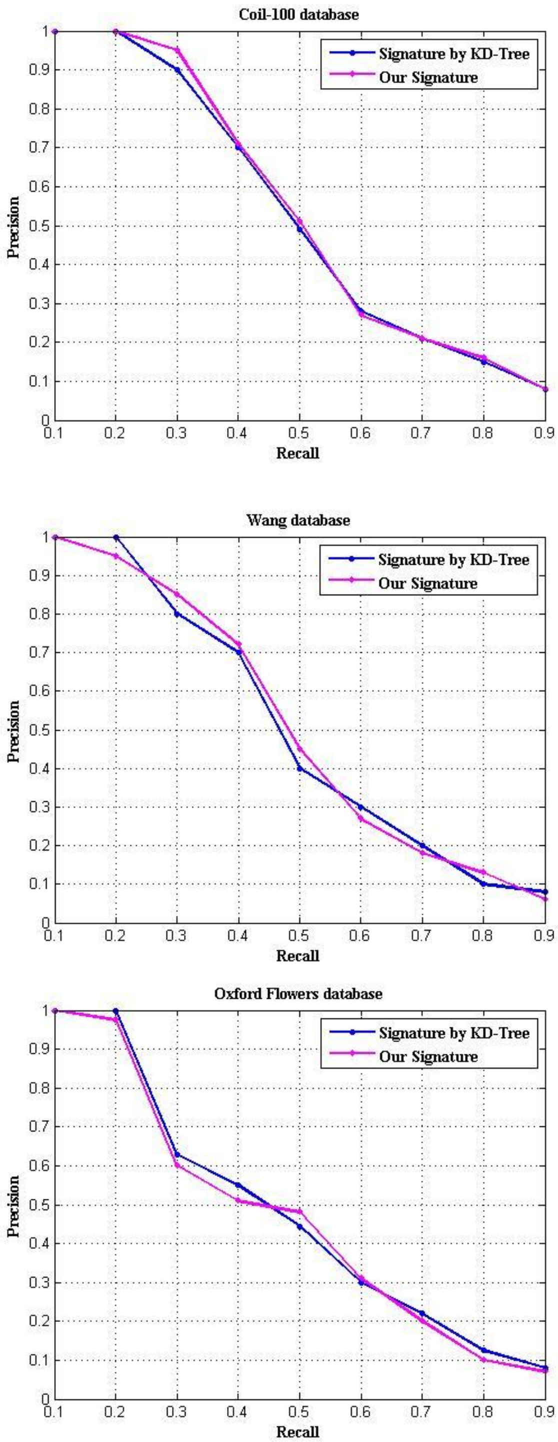

Fig.5: The precision vs. recall curves by our signature and the one proposed by Rubner et al. for three databases.

The Precision/recall curves presented on Fig 5 shows the results comparison between our signatures and theirs proposed by Rubner et al. Those results are almost the same for three databases. And then the randomize color signature that we have proposed present an efficiency of its construction time and the result quality of CBIR system.

5. Conclusion

This paper describes a new approach to building the image signatures which leverages the idea of randomized hashing. The method bypasses the computational effort associated with computing distances between feature vectors and the spatial subdivision vectors which comprises a significant fraction of the effort in other techniques. Advantage more, the search quality result using our signature not be degraded.

References

[1] J. R. Smith and S.-F. Chang, “VisualSEEk: A Fully Automated Content-Based Image Query System,” in ACM Multimedia, Nov. 1996, pp. 87–98.

[2] J. Laaksonen, M. Koskela, and E. Oja, “PicSOM - Self-Organizing Image Retrieval With MPEG-7 Content Descriptors,” IEEE Transactions on Neural Networks, vol. 13, no. 4, pp. 841–853, Jul 2002.

[3] M. Flickner, H. Niblack, W. Ashley, J. Dom, B. Gorkani, M. Hafner, D. Lee, J.and Petkovic, D. Steele, D. Yanker, and al. Query by image and video content : the qbic system. Computer, pages 23–32, 1995.

[4] Y. Gong. Image indexing and retrieval using color histogram. Multimedia Tools and Applications, Vol 2 :133–156, 1996.

[5] M. Stricker and A. Dimai. Spectral covariance and fuzzy regions for image indexing. MVA, pages 66–73, 1997.

[6] P. Lipson, E. Grimson, and P. Sinha. Configuration based scene classification and image indexing. Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR’97), 1997.

[7] Hum-Woo Yoo and al, "Visual information retrieval via content-based approach" Pattern Recognition 35 (2002) 749-769.

[8] Gong Y, "Image Indexing and retrieval using color histograms," Multimedia Tools and Applications, Vol 2, pp 133-156, 1996.

[9] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu, R. Zabih, “Image indexing using color correlograms,” CVPR, p. 762, 1997.

[10] P. Lambert, N. Hervey, H. Grecu, “Image retrieval using spatial chromatic histograms,” CGIV, pp. 343–347, 2004.

[11] G. Ciocca, R. Schettini, L. Cinque, “Image indexing and retrieval using spatial chromatic histograms and signatures,” CGIV, pp. 255–258, 2002.

[12]Y. Rubner and C. Tomasi. Perceptual metrics for image database navigation. PhD thesis,Stanford University, 2001.

[13] J.L. Bently. Multidimensional binary search tree used for associative searching. M.Communications of ACM, pages 509–517, 1975.

[14] A.W.M.Smeulders et al., “Content-based image retrieval at the end of the early years,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.22, no.12, pp.1349-1380, December 2000.

[15] W. Premchaiswadi and A. Tungkasthan, "A Compact Auto Color Correlation using Binary Coding Stream for Image

Retrieval," Proceedings of the 15th WSEAS international conference on Computers, 2011, pp. 430-436.

[16] S. Kiranyaz, M. Birinci, and M. Gabbouj, "Perceptual color descriptor based on spatial distribution: A top-down approach," Journal of Image and Vision Computing vol. 28 (8), 2010, pp. 1309–1326.

[17] S. Kiranyaz, M. Birinci, and M. Gabbouj, Perceptual Color Descriptors. Foveon, Inc. / Sigma Corp., San Jose, California, USA: Boca Raton, FL, CRC Press, 2012.

[18] Y. H. Lee, K. H. Lee, and H. Y. Ha, "Spatial Color Descriptor for Image Retrieval and Video Segmentation," IEEE Trans. Multimedia, vol. 5(3), 2003, pp. 358–367.

[19] R. Schettini, G. Ciocca, and S. Zuffi, "A Survey Of Methods For Colour Image indexing And Retrieval In Image Databases," Proc. Schettini A survey, 2001, pp. 312-322.

[20] A.Dorado and E.izquierdo, "Fuzzy color signature," IEEE Int'l Conference on Image Processing, vol.1, pp.433-436, 2002.

[21] W.K.Leow and R.Li,"The analysis and applications of adaptive-binning color histograms," Computer Vision and Image Understanding, vol.94, pp.67-91, 2004.

[22] Hurtut T., Gousseau Y. & Schmitt F. (2008)Adaptive image retrieval based on the spatial organization of colors, Computer Vision and Image Understanding (CVIU).

[23] Wyszecki G., Stiles W.S., (2000) Color science: concepts and methods, quantitative data and formulae.

[24]Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. Int. J. Comput. Vision, 59(2):167–181, 2004. ISSN 0920-5691. doi: http://dx.doi.org/10.1023/B:VISI.0000022288.19776.77.

[25]Gionis G., P. Indyk, and R. Motwani. Similarity search in high dimensions via hashing. In VLDB '99: Proceedings of the 25th International Conference on Very Large Data Bases, pages 518-529, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers Inc.

[26]Indyk P. and R. Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In STOC '98: Proceedings of the thirtieth annual ACM symposium on Theory of computing, pages 604-613, New York, NY, USA, 1998.

[27] Y. Rubner, C. Tomasi, and L.J. Guibas. The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision, 40(2) :99–121, November 2000.

[28] Wang J.Z., Li J. & Wiederhold G. (2001)SIMPLIcity: Semantics sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pages 947-963.

[29] Nene S.A., Nayar S.K. & Murase H. (1996)Columbia object image library (coil-100). Technical Report CUCS-006-96.

[30] M. Nilsback and A. Zisserman. 17 Category Flower Dataset. http://www.robots.ox.ac.uk/_vgg/data/flowers/17/, 2006.

[31] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval. Addison-Wesley, 1999.