Image Indexing & Retrieval Using Intermediate Features

Mohamad Obeid'
Equipe MIIRE ENIC, Telecom Lille I
Cité Scientifique, Rue G. Marconi
59658 Villeneuve d'Ascq France
e-mail :obeid @ enic.fr

Bruno Jedynak
USTL de Lille
Laboratoire de Statistique et Probabilités
U.F.R. de Mathématiques Bât. M2,
59658 Villeneuve d'Ascq Cedex
e-mail : Bruno.Jedynak@ univ-lille1 fr

Mohamed Daoudi
Equipe MIIRE ENIC, Telecom Lille I
Cité Scientifique, Rue G. Marconi
59658 Villeneuve d'Ascq France
e-mail : daoudi@enic.fr

ABSTRACT
Visual information retrieval systems use low-level features such as color, texture and shape for image queries. Users usually have a more abstract notion of what will satisfy them. Using low-level features to correspond to high-level abstractions is one aspect of the semantic gap.

In this paper, we introduce intermediate features. These are low-level "semantic features" and "high level image" features. That is, in one hand, they can be arranged to produce high level concept and in another hand, they can be learned from a small annotated database. These features can then be used in an image retrieval system.

We report experiments where intermediate features are textures. These are learned from a small annotated database. The resulting indexing procedure is then demonstrated to be superior to a standard color histogram indexing.

Keywords
Intermediate feature, semantic, images retrieval

1. INTRODUCTION
Recent advances in computing and communication technology are taking the actual information processing tools to their limits. The last years have seen an overwhelming accumulation of digital data such as images, video, and audio. Internet is an excellent example of distributed databases containing several millions of images. Other cases of large image databases include satellite and medical imagery, where it is often difficult to describe or to annotate the image content.

MM'01, Sept. 30-Oct. 5, 2001, Ottawa, Canada.

Techniques dealing with traditional information systems have been adequate for many applications involving alphanumeric records. They can be ordered, indexed and searched for matching patterns in a straightforward manner. However, in many scientific database applications, the information content of images is not explicit, and it is not easily suitable for direct indexing, classification and retrieval. In particular, the large-scale image databases emerge as the most challenging problem in the field of scientific databases.

The Visual Information Retrieval (VIR) systems are concerned with efficient storage and record retrieval. In general, a VIR system is useful only if it can retrieve acceptable matches in real-time. In addition to human-assigned keywords, VIR systems can use the visual content of the images as indexes, e.g. color, texture and shape features. Recently, several systems combine heterogeneous attributes to improve discrimination and classification results: QBIC[1], Photobook[2], Virage[3]. While these systems use low-level features as color, texture and shape features for image queries, users usually have a more abstract notion of what will satisfy them using low-level feature to correspond to high-level abstractions is one aspect of the semantic gap [4].

An interesting technique to bridge the gap between textual and pictorial descriptions to exploit information at the level of documents is borrowed from information retrieval, called Latent Semantic Analysis (LSA) [S]. First, a corpus is formed of documents (in this case, images with a caption) from which features are computed. Then by singular value decomposition (SVD), the dictionary covering the captions is correlated with the features derived from the pictures.

The search is for hidden correlation of feature and caption. The image collection consists of ten semantic categories of five images each. While LSA seems to improve the results of content-based retrieval experiments, this improvement is not great perhaps du the small size of image collection (50 jpeg images).

In this paper, we introduce intermediate features. These are low-level "semantic features" and "high level image" features. That is, in one hand, they can be arranged to produce high level concept and in another hand, they can be learned from a small annotated database. These features can then be used in an image retrieval system.

We report an experiment where an intermediate features are textures. These are learned from a small annotated database. The resulting indexing procedure is then demonstrated to be superior to a standard color histogram indexing method.

The reminder of this paper is organized as follows. In Section 2, we present the images database, while Section 3 the learning method is presented. In section 4, we show how images are indexed. In section 5 retrieval image system is presented. Section 6, presents the evaluation of our retrieval system. Finally, Section 7 presents our conclusion.

2. IMAGES DATABASE

Our database is made of various tourism images:

Beach scenes, snowy mountain, landscape forest, archeological site, cities. Ideal image retrieval systems for such database, when presented with an archeological site image would certainly return images of archeological site as opposed to beach images or snowy-mountain images. In order to put some "semantic knowledge" into the system we have chosen manually six natural-textures corresponding to sky, snow, sea, rocks, vegetation and sand. These are our intermediate features.

In our experiments, we use a database of 500 jpeg tourism images of size ranging from 768*512 to 1536*1024.

3. LEARNING INTERMEDIATE FEATURES

Six intermediate features: sky, water, snow, rock, vegetation and sand are defined in our application (figure I). In order the build the basic model of each texture, we extract manually from the tourism image database six sets of images.

Each one contains about 50 training images.

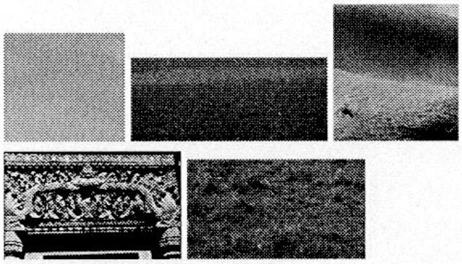

Figure 1: Example of training images.

For each natural texture, we build an empirical 3D-color histogram made of 32' bins. Let's note $p_i(r,g,b)$ the proportion of pixels with value in the bin (r, g, b) in any of the images of natural-texture t_i., $(1<=i<=6)$.

4. INDEXING IMAGES

An image is indexed by a vector $(w_1,w_2,w_3,w_4,w_5,w_6)$ representing the estimate proportion of texture t_i $(1<=i<=6)$ in this image.

The procedure for indexing on image is as follow: for each pixel s, compute f(s) that is the natural-texture t_i for which the quantity $p_i(r,g,b)$ is maximum. This is the estimated natural-texture at s. However, if the value $p_{t(s)}(r,g,b)$ is not significant then the pixel s is not labeled.

Finally, w_i is the proportion of pixels classified with texture t_i

a) b)

Figure 2. a) The query image b) Each pixel is labeled according to the estimated natural texture.

Table 1. Colors definition

Color	Red	Green	Blue	While	Violet	Yellow
Natural texture	Rock	Vegetation	Waler	Snow	Sky	Sand

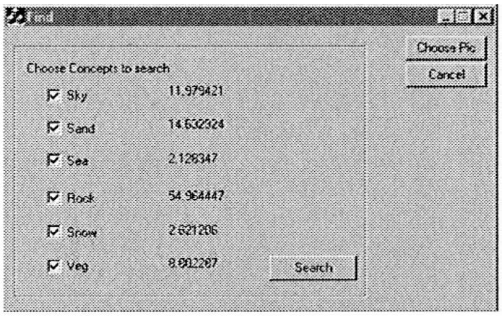

Figure 3. Snapshot of our System interface

For the example in figure 2, the result is:

Natural texture	Rock	Veg	Water	Snow	Sky	Sand
Percentage	154.96	8.8	2.12	12.62	II.97	14.63

The image is represented by the vector W=[54.96 8.8 2.12 2.62 II.97 14.631.

5. IMAGE RETRIEVAL SYSTEM

When two images I and I' are indexed respectively by the vectors $(w_1,w_2,w_3,w_4,w_5,w_6)$ and $(w'_1,w'_2,w'_3,w'_4,w'_5,w'_6)$, we define their distance to be $d(I,I')=\sum_{i=1}^{6}(w_i - w'_i)^2$

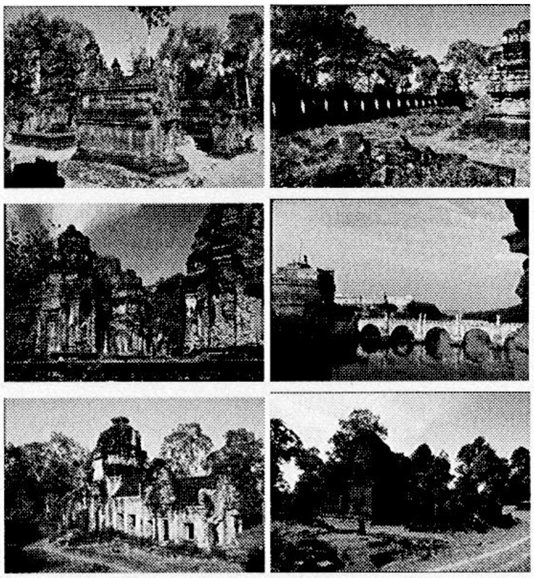

Figure 4. Ten top results obtained from the query image in figure 2.

In figure 4, we present the closest images from the one in figure 2a) using this distance.

As show" in figure 5, this system supports also the search by one or more intermediate features.

6. EVALUATION OF IMAGE RETRIEVAL SYSTEM

In order to evaluate the performance of the proposed technique, we have used the performance criterion in [6] that is defined as follows:

The precision P_r of a ranking method for some cutoff point r is the fraction of the top r ranked images that are relevant to the query

$$P_r = \frac{\text{number retrieved that are relevant}}{\text{Total number retrieved}}$$

In contrast, the recall R, of a method at some value r is the proportion of the total number of relevant images that were retrieved in the top r.

$$R_r = \frac{\text{number retrieved that are relevant}}{\text{Total number relevant}}$$

The interpolated precision is the maximum precision at this and all higher recall levels.

Figure 5 shows the precision-recall graphs for the image in figure 2. There are 25 retrieved images and 10 relevant. From this figure, it is obvious that our method is at least as performant as the histogram method using 256 bins [7]. The histogram with four bins require a" index size similar to ours but is clearly less efficient.

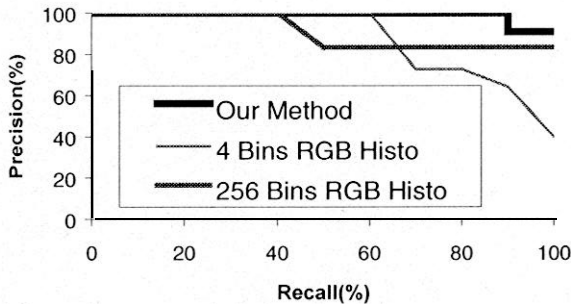

Figure 5. Recall-precision curve

7. CONCLUSION

The semantic gap is one of scientific problem for new VIR system. This paper contributes to this problem by some preliminary results.

In this paper, we report experiments using intermediate features. They are teamed from a small annotated database. The resulting indexing procedure is then demonstrated to be superior to a standard color histogram indexing method.

8. BIBLIOGRAPHIE

[1] C. Faloutous and al. Efficient and Effective Querying by Image Content. Journal of Intelligent Information Systems 3, pp. 231-262. 1994

[2] A. Pentland, R. W. Picard, and Scalaroff, Photobook: tools for content based manipulation of image database In: Proceeding of SPIE, Storage and Retrieval for Image and video Database II, No 2185, San Jose, CA, pp. 34-47. 1994

[3] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. C Jain and C. Shu, Virage image search engine : an open framework for image management" In: Proceeding. of SPIE, Storage and Retrieval for Image and video Database IV, San Jose, CA, pp. 76-87.1996

[4] Arnold W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain "Content-based Image Retrieval at the End of The Early Years" IEEE Trans. PAM1 Vol. No. 12, 2000.

[5] R. Zhao, W. Grosky " From features to semantic: some Preliminary Results" Proc. Intel Conf Multimedia Expo 2000.

[6] Ian H. Witten, Alistair Moffat, Timothy C.Bell "Managing Gigabyte?, Evaluation retrieval effectiveness, pp.188-191. Morgan Kaufman" Publishers, Inc, 1999.

[7] Swain and D. Ballard Color indexing. International Journal of Computer Vision, Vol. 7 No 1, 1991.