

# A NOVEL INDEXING APPROACH FOR MULTIMEDIA IMAGE DATABASES

Mathieu Vissac<sup>†</sup>, Jean-Luc Dugelay<sup>†</sup> and Kenneth Rose<sup>††</sup>

<sup>†</sup>Institut Eurécom, Dept. of Multimedia Communication

06904 Sophia-Antipolis Cedex, France

Tel : +33 (0)4.93.00.26.{41.71}, Fax : +33 (0)4.93.00.26.27

Email : {vissac, dugelay}@eurecom.fr

<sup>††</sup>University of California, Dept. of Electrical and Computer Engineering

Santa Barbara, CA 93106, USA

Email : rose@ece.ucsb.edu

**Abstract** - This paper proposes an algorithm for *Content-based Image Indexing* whose formulation of similarity is borrowed from methods for fractal compression. Unlike most traditional block-based image indexing algorithms, the proposed method employs dynamic programming to exploit inter-block dependencies. A regularization constraint is globally imposed, and the overall optimal transformation relating two images is efficiently determined by application of the Viterbi algorithm. Preliminary experimental results on a sample of eighty binary images from the MPEG-7 database are presented.

## INTRODUCTION

Huge volumes of image data accumulate and are stored for diverse applications. These include medical imagery, satellite imagery, entertainment image data, etc. As manual search and retrieval of images in large databases is impractical, efficient automatic access techniques are needed. The problem is exacerbated by the need for content-based retrieval rather than the alphanumeric search typical of traditional database management. This is a major challenge that has been recognized by many researchers and developers [1, 2]. Several products are currently available on the internet, and offer good performance when queries are well captured by color indexing via histograms, etc. Nevertheless, the main difficulty is due to the wide diversity of users and queries whose needs are not satisfied by such simple search techniques. The human notion of similarity is subjective and hard to define precisely. It depends on the type of database, the context, the application and the user.

In this paper we propose to measure the similarity of two pictures, by applying local similarity principles from the fractal coding theory [3]. This approach to formalating similarity offers much flexibility, and the premise of this work is that it is a useful tool for approximating the subjective aspects of an image database query. In most fractal-based image indexing algorithms,

such as [4, 5], images are divided into blocks, and decisions are made independently for each block. In this work, however, the fractal local similarity measure is complemented by a regularization constraint that enforces coherency of consecutive matching decisions. The regularization constraint is globally and efficiently imposed by a dynamic programming procedure which is commonly known as the Viterbi algorithm [6, 7].

We assume the basic problem of query by example. The user presents a “query image”, also called the “example image”, and the algorithm searches the database of “test images” for images that are most similar to the query image.

## LOCAL SIMILARITY

As is commonly done in fractal image compression, the example image is partitioned into blocks that are referred to as range blocks. Block matching is then performed to associate (range) blocks of the query image with (domain) blocks of a candidate image in the database, by adjusting the transformation parameters. We consider the geometric parameters  $(a_i, b_i, c_i, d_i, e_i, f_i)$ :

$$\tau_i \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a_i & b_i & 0 \\ c_i & d_i & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \\ 0 \end{pmatrix}.$$

Thus, for each block of the example image and for each possible geometric transformation we search for the most similar block in the test image. The results are recorded as a list of scores, one per transformation.

$$\Delta(n, i) = \min_j \left( \sqrt{\frac{\sum_{\text{pixels}} (S_n - \tau_i \cdot D_j)^2}{\# \text{ pixels}}} \right)$$

where  $S_n$  is the  $n$ th block of the example image  
and  $D_j$  is the  $j$ th block of the test image

These scores are converted into a probabilistic representation by a function  $f$ , and for simplicity we use the linear function:  $f(x) = 1 - \frac{x}{255}$ . Specifically, block  $n$  is in state  $i$  with probability

$$O_n(i) = f(\Delta(n, i))$$

It should be emphasized that while the above formulation of similarity offers substantial flexibility which may capture non-trivial notions of similarity, it offers too many degrees of freedom. It is conceivable that non-similar images might “match” by mixing and transforming unrelated blocks. In order to eliminate this problem we introduce a requirement of “global coherence”. In other words, we wish to impose a degree of continuity in the block matching results. We propose to optimize the trade off between local matching scores and the global coherence constraint by application of a dynamic programming technique known as the Viterbi algorithm [6, 7].

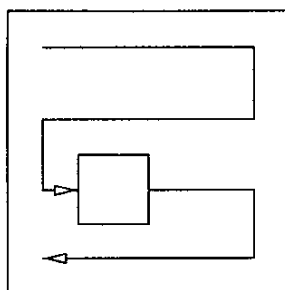
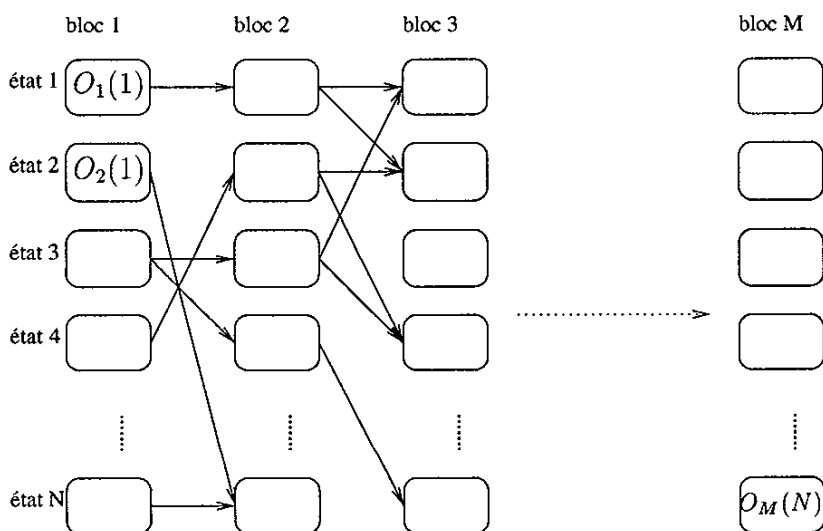


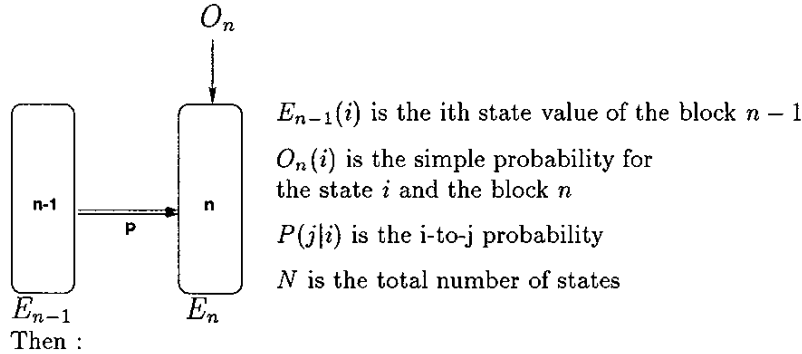
Figure 1: "Serpent scan"

## GLOBAL COHERENCY

The Viterbi algorithm (VA) finds the optimal path in the trellis of a Markov chain (random state machine). In our case VA is used to find the optimal sequence of block matching pairs for the given pair of example and test images. Since the optimality and simplicity of VA requires a one-dimensional Markov chain, we use the "serpent" scan of the image (figure 1). This path ensures that consecutively scanned blocks are neighbors on the image plane.



The state of the machine is determined by the transformation parameters used for matching the current block. Additional cost is incurred when consecutive blocks select “conflicting” states. VA finds the best (minimum overall cost) path in the trellis and thereby optimizes the trade-off between matching and continuity.



$$E_n(j) = \max_i (E_{n-1}(i) \times P(j|i) \times O_n(j))$$

### The Viterbi States

The states enumerated below were selected for the case of binary images on which we focus in the results section, but can be extended to the case of gray level, color or multi-spectral images. We define eight sub-states corresponding to the eight geometric transforms :

|   |                       |   |                           |
|---|-----------------------|---|---------------------------|
| 0 | Identity              | 4 | 2nd diag. reflection      |
| 1 | Horizontal reflection | 5 | $\frac{\pi}{2}$ rotation  |
| 2 | Vertical reflection   | 6 | $\pi$ rotation            |
| 3 | 1st diag. reflection  | 7 | $\frac{3\pi}{2}$ rotation |

### The transition probability

Transitions between these substates are strongly penalized to underline the importance of geometric continuity for “true” similarity between two images. Here we adopted a simple three-level cost (no transition, soft transition, and abrupt transition), but a more sophisticated training procedure to determine the transition probabilities is under current investigation.

## EXPERIMENTAL RESULTS

Preliminary experimentation has been performed on a subset of the MPEG-7 database: (*Trademark images captured by a scanner*, CD-Rom num. 10).

Retrieval efficiency is quantified as:

$$\text{efficiency} = \frac{N_{sir}}{\min(N_{si}, N_{ri})} \times 100\%$$

$N_{sir}$  is the number of similar images retrieved until the first dissimilar

$N_{si}$  is the total number of similar images in the base

$N_{ri}$  is the number of returned images for the query

Our preliminary results on this database produced efficiency varying from 89% to 100%. An illustrative example is given in Figure 2

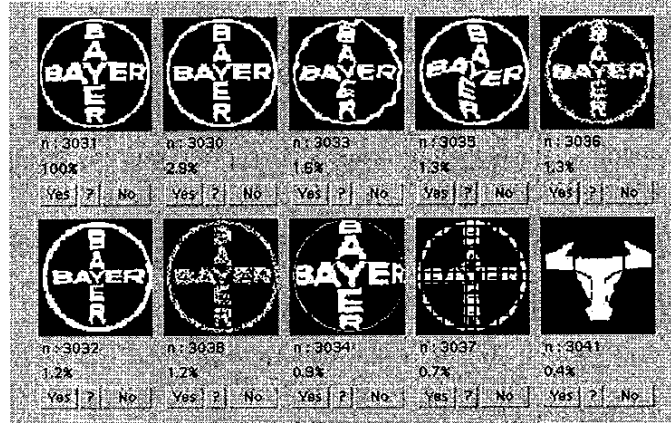


Figure 2: Exemple image : 3031 (efficiency = 100%)

## CONCLUDING REMARKS

The proposed fractal-based image retrieval algorithm shows considerable promise as a means to capture the subjective notion of image similarity. Preliminary simulation results support this promise.

The proposed algorithm can be extended and adapted to new types of queries and applications. The use of the Viterbi algorithm, allows easy extensions via a richer set of states including new transformations (photometric adjustments, zoom, etc.). In particular, in the case of gray-level images, it is natural to enrich the set of states and account for photometric transformations including "scale" and "offset".

## ACKNOWLEDGMENTS

This work was supported in part by a regional grant, and by Eurécom's industrial partners: Ascom, Cegetel, France Telecom, Hitachi, IBM France, Motorola, Swisscom, Texas Instruments, and Thomson CSF. K. Rose was also supported by the University of California MICRO program, Cisco Systems, Inc., Conexant Systems, Inc., Dialogic Corp., Fujitsu Laboratories of America, Inc., General Electric Co., Hughes Network Systems, Intel Corp., Lernout & Hauspie Speech Products, Lockheed Martin, Lucent Technologies, Inc., Panasonic Technologies, Inc., Qualcomm, Inc., and Texas Instruments, Inc.

## References

- [1] T. Randen and J.H. Husoy. Image content search by color and texture properties. In *ICIP-97*, volume 1, pages 580–583, 1997.
- [2] H. Yu and W. Wolf. A hierarchical, multi-resolution method for dictionary-driven content-based image retrieval. In *ICIP-97*, volume 2, pages 823–826, 1997.
- [3] A. E. Jacquin. Image coding based on a fractal theory of iterated contractive image transformation. *IEEE Transactions on Image Processing*, 1(1):18–30, January 1991.
- [4] A. Zhang, B. Cheng, and R. Acharya. A fractal-based clustering approach in large visual database systems. *Multimedia Tools and Applications*, (3):225–244, 1996.
- [5] J.M. Marie-Julie and H. Essafi. Image database indexing and retrieval using the fractal transform. In *ECMAST'97*, pages 169–182, 1997.
- [6] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [7] A. J. Viterbi. Errors bounds for convolutional codes and asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, IT-13:260–269, 1967.