



**HAL**  
open science

## Spatial-textural medical image indexing based on vector quantization

John Richard Ordonez, Guy Cazuguel, John Puentes, Basel Solaiman,  
Christian Roux

► **To cite this version:**

John Richard Ordonez, Guy Cazuguel, John Puentes, Basel Solaiman, Christian Roux. Spatial-textural medical image indexing based on vector quantization. IEEE EMBS 2003: 25th annual international conference of the IEEE engineering in medicine and biology society, Cancun, Mexico, September 19-21, Sep 2003, Cancun, Mexique. hal-02128304

**HAL Id: hal-02128304**

**<https://hal.science/hal-02128304>**

Submitted on 8 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Spatial-textural Medical Image Indexing based on Vector Quantization

J. R. Ordonez<sup>1,2</sup>, G. Cazuguel<sup>1,2</sup>, J. Puentes<sup>1,2</sup>, B. Solaiman<sup>1,2</sup>, C. Roux<sup>1,2</sup>

<sup>1</sup>LATIM ERM 0102 - INSERM, Bat 2 bis, CHU Morvan, Brest, France

<sup>2</sup>Département I.T.I., ENST de Bretagne, B.P. 832, 29285 Brest-Cedex, France

*Abstract* — This paper addresses the problem of content base image retrieval in the compression domain. We characterize the images by a representation (signature) based on low level parameters. We propose the use of spatial moments to add spatial information to the signatures based on vector quantization. The developed signatures are a combination of the histogram of occurrences of the codewords, the center of gravity and the moment of inertia associated with the histogram. The signature of each image in the base is determined, and the research is carried out by calculating, using a given metric, the distance between the signatures in the base and the signature of the query image. The methods are applied to medical databases: ultra-sound esophagus and IRM tumor brain images, and improvements of up to 18 % in retrieval efficiency are found, with regard to known usual histograms techniques.

*Keywords*— Content based image retrieval, spatial-textural image indexing, vector quantization.

## I. INTRODUCTION

Image databases and image retrieval applications are playing a growing role in medical information systems. This becomes apparent when we consider the continued technical progress, bringing with it advances in affordable digital devices as well as a seemingly endless expansion of the internet. Consequently, the amount of images we are confronted with nowadays is so huge that, when facing the task of finding something useful in them, we are rapidly overtaxed. Therefore, visual database systems require efficient indexing to enable fast access to information (images and video sequences) in databases. Automatic or semi-automatic image indexing directly using image digital content (Content-Based Image Retrieval) is one of the possible and promising solutions to effectively managing image databases [1], [2]. This is particularly the case in the health field, where automatic indexing can also become a diagnosis assistance tool, by using images as queries in a knowledge base, complementary to the textual indexing carried out by an expert. The objective is to give an expert the possibility of carrying out a research "like a blind man" in the base, without formulating a semantic description of the image he is examining. This enhances the access to diagnostic possibilities with no *a priori* description. To do this, we propose to build signatures of the images, by carrying out an analysis of images by Vector Quantization (VQ), which may be used, in addition, for image compression. We are interested in VQ because it is naturally an indexing technique, as it maps image blocks to indices.

Furthermore, VQ allows introducing specific knowledge – mainly associated with textures – to the indexation signatures, through the process of codebooks creation.

Several indexing techniques associated with VQ [3], [4], have been proposed in the literature. Generally, they are based on histograms of occurrences and/or usage maps calculated over the entire image or specified regions. We did a study of the utilization of different similitude metrics to compare the histograms, as well as the effect of the size of blocks in the retrieval efficiency, with interesting results [5]. However, none of these techniques introduces any spatial information into the signatures that characterize the images. In a first approach to solving this problem, we proposed to use a multi-resolution decomposition based on Quadtree principles. The retrieval efficiency was improved, but the signature size increased, and the approach became sensitive to noise, translation and rotation operations. In this paper, we continue investigating this subject and develop global signatures based on spatial moments associated with the histograms. Image retrieval experiments on gray-level medical databases show that the proposed signatures outperform conventional histogram techniques. The paper is organized as follows: Section II briefly describes the principles of VQ. Section III introduces our novel signatures and comparison metrics in the VQ compressed domain. Section IV presents experimental results obtained by applying our algorithm, and we conclude in Section V.

## II. VECTOR QUANTIZATION METHODOLOGY

VQ is a technique proposed for signal compression and coding [6]. It is based on Shannon's theory [7] that coding systems can perform better if they operate on vectors or groups of symbols rather than on individual symbols. In VQ-based image compression, an image is decomposed into fixed-size non-overlapping blocks of  $n \times n$  pixels (the source blocks). Each block is considered as an  $N$ -dimensional vector (with  $N=n^2$ ) in the  $\Sigma^N$  vector space. VQ is a mapping from  $\Sigma^N$  into a finite subset  $\Omega$  of  $\Sigma^N$ , where  $\Omega = \{W_1, \dots, W_i, \dots, W_M\}$  is a set of predefined vectors, or blocks. These vectors  $W_i$  are called **codewords** and the set  $\Omega$  is the vector quantization **codebook**. The distances between each source block  $X$  and the codewords is measured in order to select a representative codeword  $W_i$  for  $X$ . Then, only the label (or the index) of this codeword in the codebook is kept to create a compressed image. Thus, the compressed image is a matrix of indices. To reconstruct the image, each index is replaced by the codeword in the codebook which is in fact a

lookup table. Usually the reconstructed image is only an approximation of the original one, since VQ is a lossy compression. The main problem in this approach is the design of the codebook. Several algorithms have been proposed to create the codebooks [6]: Lloyd, LBG, k-means, fuzzy clustering, the Kohonen learning algorithm, etc. We have chosen the Kohonen Self Organizing Feature Map (SOFM), based on neural networks, for this purpose. The main interesting properties of the SOFM are: (i) self-organizing algorithm: it does not need to classify the training image blocks (unsupervised learning), (ii) ability to form ordered topological feature maps: neighboring neurons on the map have close weight vectors in the gray-level pixel space, i.e. they are visually similar. Therefore, it is possible to perform a classification of the image pixels by segmenting the SOFM directly. Another interest of these topological properties is their robustness against transmission errors [8], (iii): quantization is performed in the gray-level pixel space, and the visual aspect of images is preserved, which is very important for heavily textured images.

### III. APPLICATION OF VQ TO IMAGE INDEXING

In [5], we introduced an image descriptor, based on the SOFM referred to as the "Signature Map" (SM). This image descriptor is a matrix  $[m_{ij}]$ , which has the same topology as the SOFM codebook. The value of each element  $m_{ij}$  of the SM is the number of occurrences, in the compressed image, of the codeword located at  $i, j$  in the SOFM (normalized 2D histogram of the codewords used). This signature gives information about the contents of the textures of the image. Intuitively, if two similar images use the same codewords in an equivalent proportion, then they have close signatures. Each image in the base is analyzed by using the relevant dictionary for the type of images considered, and an SM is associated with them. A query image is analyzed according to the same dictionary and its SM is compared with the SMs of the base. With the purpose of adding space information to the signature and in order to decrease the sensitivity of the methods based on the Quadtree decomposition, we propose a new approach, which takes into account the localization of the blocks associated to the codewords occurrences histogram. Spatial moments of the blocks are computed, coupled with each component of the SM. More precisely, we calculate the center of gravity and the inertia moment of the blocks represented by each class (bin) of the occurrence histogram (see Fig. 1 and equations (1)). The final signature comprises the SM and two vectors containing the associated moments. Normally, the SOFM (Kohonen's codebook) is composed of 256 codewords ( $16 \times 16$  neurons), so the SM will be a histogram of 256 bins. To reduce signature size, we propose to group blocks in macro-blocks of 4 by 4 neighbor blocks in the SOFM, reducing the signature size from a histogram of 256 to a 64 bin size; due to the utilization of macro-blocks, the spatial component is reduced as well,

from a vector size of 256 to 64. In addition, the spatial component can be reduced by half, taking only the X component or the Y component of spatial moments. To compare the signatures, we tested at least twenty combinations of distance measures (L1, L2, correlation, intersection, quadratics form, etc.), combining textural (histograms) and spatial components (center of gravity and inertia moment). In this paper we only present the results obtained according to the distances expressed in equations (2).

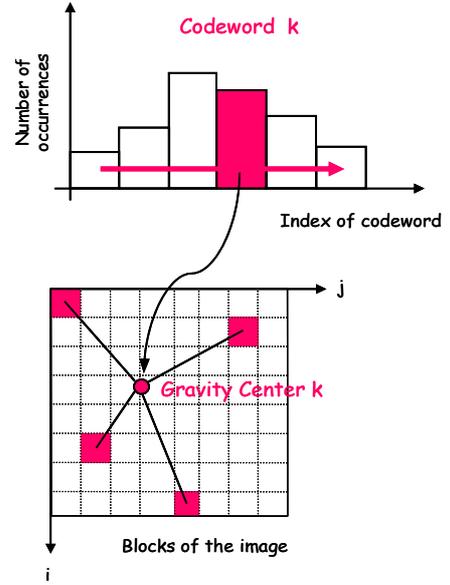


Fig. 1. Creation of signatures.

*Spatial moments:*

$$G_x = \frac{1}{M} \sum_{i=1}^M x_i \quad G_y = \frac{1}{N} \sum_{j=1}^N y_j \quad (1a)$$

$$I_x = \frac{1}{M} \sum_{i=1}^M x_i^2 \quad I_y = \frac{1}{N} \sum_{j=1}^N y_j^2 \quad (1b)$$

*Distances:*

$$d_a(P, Q) = \sum_{i=1}^{256} \|H_i^P - H_i^Q\| \quad (2a)$$

$$d_b(P, Q) = \sum_{i=1}^{256} \|H_i^P - H_i^Q\| \times \|\vec{G}_i^P - \vec{G}_i^Q\| \quad (2b)$$

$$d_c(P, Q) = \sum_{i=1}^{64} \|R_i^P - R_i^Q\| \quad (2c)$$

$$\text{where } R_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \|\vec{G}_i - \vec{U}_j\|$$

$$d_d(P, Q) = \sum_{i=1}^{256} \|H_i^P \times \vec{G}_i^P - H_i^Q \times \vec{G}_i^Q\| \quad (2d)$$

$$d_e(P, Q) = \sum_{i=1}^{64} \|H_i^P \times \vec{G}_i^P - H_i^Q \times \vec{G}_i^Q\| \quad (2e)$$

In (1),  $M$  and  $N$  are the number of blocks by row and column in the original image respectively. In (2),  $H$  represents the signature maps,  $G$  the centers of gravity,  $U$  the positions of the blocks associated with the components of the SM (see Fig. 1),  $I$  the inertia moments and  $\|X\|$  the Euclidean distance. In (2a), the Euclidean distance between SMs – probability of occurrence of a given codeword – is computed. Equation (2b) gives the Euclidean distance between SMs weighted by the Euclidean distance between centers of gravity. The idea behind this equation is that two images will be similar when the SMs “and” the gravity centers are similar. Equation (2c) is based on the difference between the mean radiuses, i.e., the dispersion of codewords around the centers of gravity, associated with each bin of the SMs. Finally, (2d) and (2e) give the differences between the product – component by component – of the SMs and the gravity centers, that is to say, the difference between centers of gravity will be significant when there is a significant number of blocks associated with the centers of gravity; in (2e), signature size has been reduced to a quarter, by reducing the size of the codebook via the utilization of macro-blocks of  $4 \times 4$  neighbors. This was the case too in (2c).

#### IV. RESULTS

For performance evaluation purposes, we tested our algorithms over two medical image databases: (i) a base of esophagus ultra sound images composed of 8 sequences of 30 images in each, and (ii) a subset of the brain tumor image database from the Harvard Medical School at the Brigham and Women's Hospital [9], composed of 10 sequences of 10 images. For both databases, each sequence is regarded as a class of similar images. This classification was necessary to be able to apply the classical measurements of retrieval efficiency. The performance evaluation of retrieval is based on the “average precision vs. the average recall” [10] and the “retrieval efficiency” graphics defined in [11], which for this classification corresponds to a value of average precision evaluated with a threshold “ $T = \text{number of images per class}$ ”. Figures 2 and 3 show retrieval efficiency as a function of the codewords' block size. There are 5 curves representing the various types of signatures and metrics proposed in (2). In [5], we showed that (2a) gives good retrieval results for many kinds of medical databases. For this reason, we take it as a reference for the new proposed signatures. We can observe that (2b) always outperforms the other signatures, except for large block size codewords, where the spatial moments do not provide significant information to characterize images; so the efficiency decreases to values similar to those produced by (2a). Another interesting result is the behavior of the reduced size signatures like (2c) and (2d) which give good retrieval efficiency with only a quarter the size of previous signatures. These results are confirmed in figures 4 and 5

where the average Precision-Recall graphs are shown, the curves (2b), (2c) are over (2a) in fig. 4 (outperforming), and the curves (2c) and (2e) are very near to (2a) in fig. 5 (same performing with a quarter size signature). The signatures based on the inertia moment give similar results to those obtained using the centers of gravity.

#### V. CONCLUSION

Using spatial moments for content-based image retrieval based on vector quantization allows introducing spatial information into image characterization by signatures. It improves retrieval efficiency up to 18 % in the tested databases, compared to typical VQ indexing methods based on histogram of occurrences and Euclidean metrics like (2a). Signature sizes can be reduced to a quarter of the usual sizes based on the histograms of codeword occurrences, without reducing retrieval performance. In addition, the proposed signatures seem to be less sensitive than Quadtree-decomposition methods to noise problems in images, and translation and rotation operations, improving retrieval capacities. We have developed similar approaches for the compression standards JPEG [12] based on DCT, and JPEG-2000 [13] based on wavelets, with similar results. We have found similar retrieval behavior in other medical bases (retinal angiographies, venous thromboses, video-endoscopies, faces, etc). So, the use of this kind of spatial information is of great interest for content-based image retrieval.

#### REFERENCES

- [1] F. Idris, S. Panchanathan, “Review of image and video indexing techniques”, *Journal of Vision Comm. and Image Representation*, 8, 2, pp. 146-166, 1997.
- [2] C. Nastar, “Indexation d'Images par le Contenu: un Etat de l'Art”, In CORESA'97, Paris, March, 1997.
- [3] F. Idris, S. Panchanathan, “Storage and retrieval of compressed images”, *IEEE Trans. Consumer Electronics*, 41(3):937-941, 1995.
- [4] L. Zhu and A. Zhang, “Keyblock: an approach for content-based image retrieval”, 2000.  
Available: <http://www.acm.org/sigs/sigmm/MM2000/ep/Zhu/>
- [5] J. R. Ordóñez, G. Cazuguel, J. Puentes, B. Solaiman and C. Roux, “Medical image indexing and compression based on VQ: image retrieval efficiency evaluation,” in *Proceedings of 23<sup>rd</sup> Annual International Conf. of the IEEE EMBS*. Istanbul - Turkey, October 25-28, 2001
- [6] A. Gersho and R. M. Gray, *Vector Quantization and Signal Processing*. Kluwer Academic Publishers Boston, 1992.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., New York, 1991.
- [8] O. Aitsab, R. Pyndiah and B. Solaiman, “Joint optimization of multidimensional SOFM,” *3<sup>rd</sup> International Workshop on Image/Signal Processing*, pp. 3-6, Manchester, UK, 1996.
- [9] Available: <http://splweb.bwh.harvard.edu:8000/pages/ppl/warfield/tumorbase/>
- [10] H. Müller, W. Müller, D. Squire, S. Marchand-Maillet, T. Pun, “Performance evaluation in content-based image retrieval: Overview and proposals,” *Elsevier Preprint*, 15 August 2000.

- [11] B. M. Mehre, M. S. Kankanhalli, A. D. Narasimhalu, G. C. Man. "Color matching for image retrieval," *Pattern Recognition Letters*, 16, 325-331, 1995
- [12] J. R. Ordóñez, G. Cazuguel, J. Puentes, D. Poulouin, B. Solaiman and C. Roux, "Joint spatial-spectral indexing of medical images exploiting the JPEG-DCT compression scheme," in *Proceedings*

*of the IEEE: Second Joint EMBS-BMES Conference*. Houston, Texas – USA, October 23-26, 2002.

- [13] J. R. Ordóñez, G. Cazuguel, J. Puentes, B. Solaiman and C. Roux, "Indexation d'images médicales basée sur les informations spatiale et spectrale extraites de JPEG-2000," in *Proceedings of CORESA'03*. Lyon, France, January 16-17, 2003.

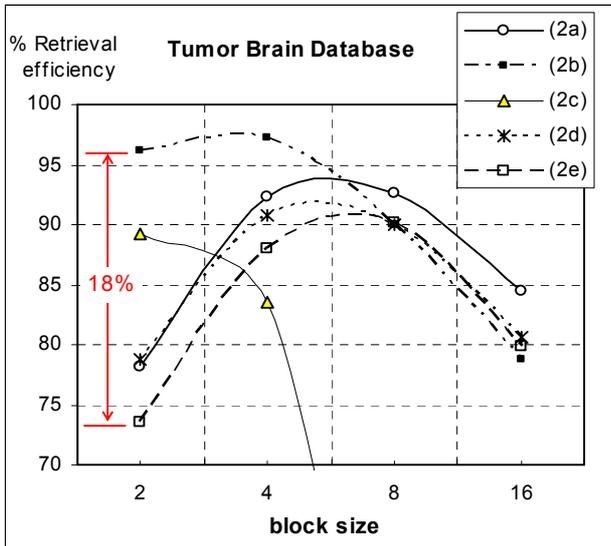


Fig. 2. Retrieval efficiency vs. size of codeword for the brain tumor database. Curves A, B, C, D and E correspond to the results based on (2a): Euclidean distance between the SMs, (2b): product of Euclidean distance between the components of the SMs and the centers of gravity, (2c): Euclidean distance between the mean radius of centers of gravity, (2d): Euclidean distance between the centers of gravity weighted by the components of the SMs with 256 bins and (2e) Euclidean distance between the centers of gravity weighted by the components of the SMs with 64 bins, respectively.

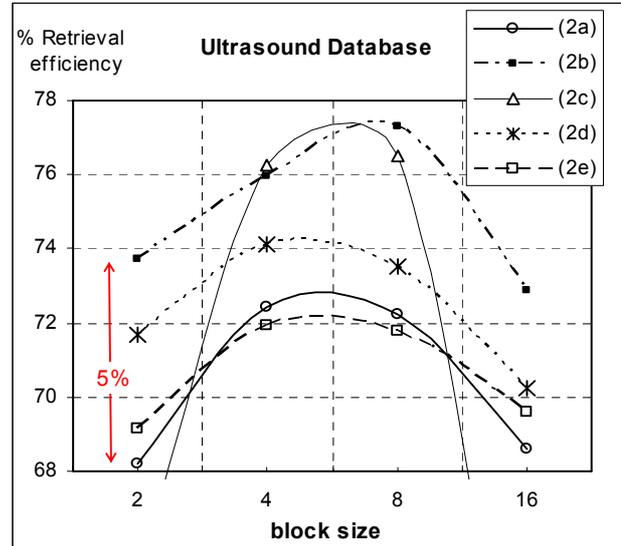


Fig. 3. Retrieval efficiency vs. size of codeword for the esophagus ultrasound database. Curves A, B, C, D and E correspond to the results based on (2a): Euclidean distance between the SMs, (2b): product of Euclidean distance between the components of the SMs and the centers of gravity, (2c): Euclidean distance between the mean radius of centers of gravity, (2d): Euclidean distance between the centers of gravity weighted by the components of the SMs with 256 bins and (2e) Euclidean distance between the centers of gravity weighted by the components of the SMs with 64 bins, respectively.

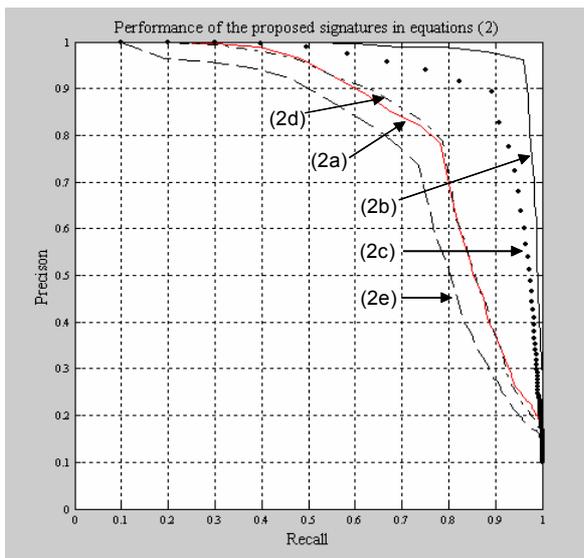


Fig. 4. Average Precision-Recall on brain tumor database using codewords of 2x2 pixels.

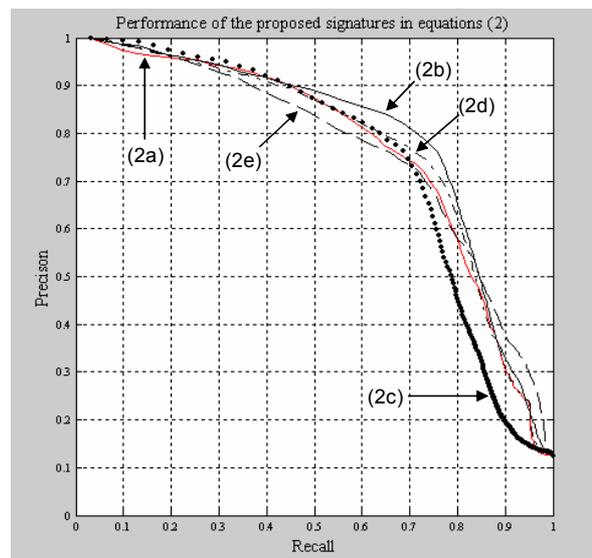


Fig. 5. Average Precision-Recall on esophagus ultrasound database using codewords of 4x4 pixels.