

Content Based Image Retrieval based on Wavelet Transform coefficients distribution

Mathieu Lamard, Guy Cazuguel, Gwénoél Quéllec, Lynda Bekri, Christian Roux, Béatrice Cochener

Abstract—In this paper we propose a content based image retrieval method for diagnosis aid in medical fields. We characterize images without extracting significant features by using distribution of coefficients obtained by building signatures from the distribution of wavelet transform. The research is carried out by computing signature distances between the query and database images. Several signatures are proposed; they use a model of wavelet coefficient distribution. To enhance results, a weighted distance between signatures is used and an adapted wavelet base is proposed. Retrieval efficiency is given for different databases including a diabetic retinopathy, a mammography and a face database. Results are promising: the retrieval efficiency is higher than 95% for some cases using an optimization process.

Index Terms—CBIR, wavelets, diabetic retinopathy, signature

I. INTRODUCTION

Images have always been used in medicine for teaching, diagnosis, and management purposes. Now medical imaging systems produce more and more digitized images in all medical fields: visible, ultrasound, X-ray tomography, MRI, nuclear imaging, etc... Thus for instance, Lund University hospital produces 15,000 new digital X-ray images per day [1]. These images are very interesting for diagnostic purposes: they are directly related to the patient pathology and medical history. However the amount of images we can access nowadays is so huge that database systems require efficient indexing to enable fast access to images in databases. Automatic image indexing using image digital content (Content-Based Image Retrieval) is one of the possible and promising solutions to effectively manage image databases [2], [3]. The purpose 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. Thus, in the CBIR framework, images are used as queries and a retrieval system returns the most similar images in the database. We propose in this work a signature derived from coefficients of the wavelet transform. Standard wavelets (used in JPEG-2000) are tested and we also introduce an adapted wavelet to improve results. We also demonstrate that outcomes can be enhanced by optimizing parameters like weights.

Mathieu Lamard and Béatrice Cochener are with Univ Bretagne Occidentale, Brest, F-29200 France Mathieu.Lamard@univ-brest.fr

Guy Cazuguel, Gwénoél Quéllec and Christian Roux are with ENST Bretagne, GET-ENST, Brest, F-29200 France

Mathieu Lamard, Guy Cazuguel, Gwénoél Quéllec, Béatrice Cochener and Christian Roux are with Inserm, U650, Brest, F-29200 France

Lynda Bekri and Béatrice Cochener are with CHU Brest, Service d’Ophtalmologie, Brest, F-29200 France

The setup of the article is as follows. Section II-A describes the databases we used for evaluation. Section II-B describes the signature design process with standard or adapted wavelet base. Section II-C propose a distance between two signatures. Section III presents results and Section IV the conclusion and perspectives.

II. MATERIAL AND METHODS

A. Databases

Several databases are used in order to test our algorithms.

1) *Diabetic retinopathy database*: The diabetic retinopathy (DR) database contains retinal images of diabetic patients, with associated anonymized information on the pathology. Diabetes is a metabolic disorder characterized by sustained inappropriate high blood sugar level. This progressively affects many blood vessels, which may lead to serious renal, cardiovascular, cerebral and also retinal complications. Different lesions appear on the damaged vessels, which may lead to blindness. The database is made up of 63 patient files containing 1045 photographs altogether. Images have a definition of 1280 pixels/line for 1008 lines/image. They are lossless compressed images. Patients were recruited in Brest University Hospital since June 2003 and images were acquired by experts using a Topcon Retinal Digital Camera (TRC-501A) connected to a computer. An example of image series is given in figure 1. The disease severity level [4] (ranging from 0 to 5) of each patient is also available.

2) *Mammography database*: The DDSM project [5] is a collaborative effort involving the Massachusetts General Hospital, the University of South Florida and Sandia National Laboratories. The primary purpose of the database is to facilitate research in the development of aid computer algorithms screening ([6] for example). The database contains approximately 2,500 studies. Each study includes two images of each breast, along with some associated patient information (age at time of study, ACR breast density rating) and image information (scanner, spatial resolution). An example of image series is given in figure 2. The expert diagnosis (‘normal’, ‘benign’ or ‘cancer’) is also available.

3) *FaceDataBase*: This database [7] is used as a benchmark database. It includes ten different images for each of the 40 distinct subjects [7]. For some subjects, the images were taken at different times, with different lightnings, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position. Fig. 3 shows the 10 face images of the same subject. Even though it is not our main purpose,

This material is presented to ensure timely dissemination of scholarly and technical work.

Copyright and all rights therein are retained by authors or by other copyright holders.

All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright.

In most cases, these works may not be reposted without the explicit permission of the copyright holder.

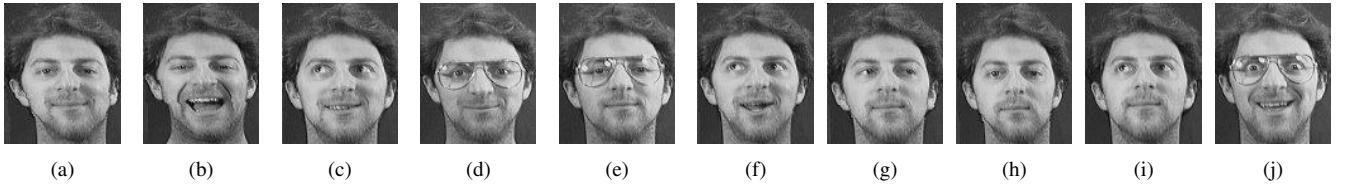


Fig. 3. Ten images sequence of the same person's face

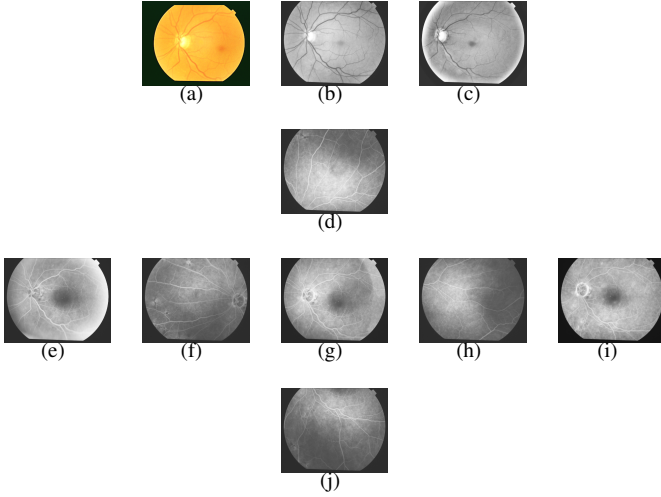


Fig. 1. Photograph series of a patient eye

Images (a), (b) and (c) are photographs obtained by applying different color filters. Images (d) to (j) form a temporal angiographic series: a contrast product is injected and photographs are taken at different stages (early (e), intermediate (d), (f), (g), (h), (j) and late (i)). For the intermediate stage, photographs from the periphery of the retina are available.

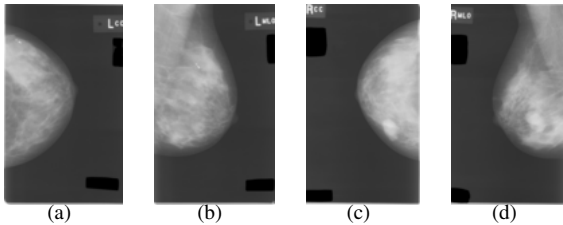


Fig. 2. Image series from a mammography study

(a) and (b) are images of the left breast, (c) and (d) are images of the right one.

this database can be easily classified. Images belong to the same class if and only if they represent the same subject. The 400 database images are then classified in 40 classes.

B. Signature

1) *Wavelets decomposition*: Wavelets are families of functions generated from one single prototype function (mother wavelet) ψ by dilation and translation operations: ψ is constructed from the so-called scaling function ϕ , satisfying

the two-scale difference equation

$$\phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} h(k)\phi(2t - k) \quad (1)$$

where $h(k)$ are the scaling coefficients. Then, the mother wavelet $\psi(t)$ is defined as

$$\psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} g(k)\phi(2t - k) \quad (2)$$

where the wavelet coefficients $g(k) = (-1)^k h(1 - k)$ [8]. Several different sets of coefficients $h(k)$ can be found, which are used to build a unique and orthonormal wavelet basis [8]. The wavelet transform represents the decomposition of a function into a family of wavelet functions $\psi_{m,n}(t)$ (where m is the scale/dilation index and n the time/space index). In other words, using the wavelet transform, any arbitrary function can be written as a superposition of wavelets. Many constructions of wavelets have been introduced in mathematical [9] and signal processing literature (in the context of quadrature mirror filters) [10]. In the mid eighties, the introduction of multiresolution analysis and the fast wavelet transform by Mallat and Meyer provided the connection between the two approaches [11]. The wavelet transform may be seen as a filter bank and illustrated as follows, on a one dimensional signal $x[n]$:

- $x[n]$ is high-pass and low-pass filtered, producing two signals $d[n]$ (detail) and $c[n]$ (coarse approximation)
- $d[n]$ and $c[n]$ may be subsampled (decimated by 2: $\downarrow 2$), otherwise the transform is called translation invariant wavelet transform
- the process is iterated on the low-pass signal $c[n]$

This process is illustrated in figure 4. We have extract information at several scales (subbands) plus an approximation of the signal (the last $c[n]$).

In the case of images, the filtering operations are both per-

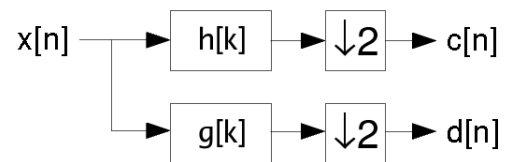


Fig. 4. Two-channel filterbank involving subsampling

formed on rows and columns, leading to the decomposition shown in figure 5.

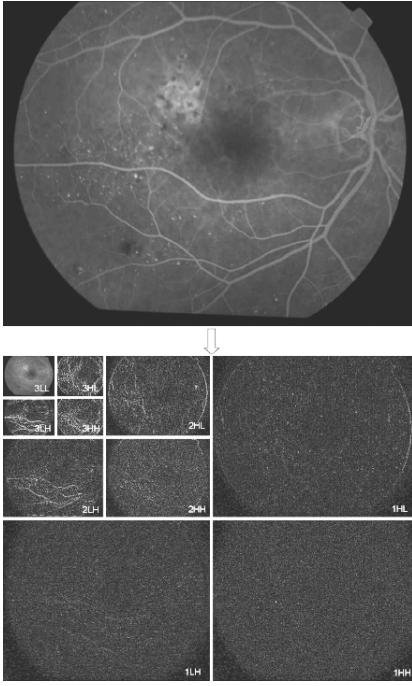


Fig. 5. Wavelet transform of an image

There are three subbands at each scale, depending on whether the rows/columns were high-passed or low-passed

2) *Wavelet Choice*: unlike other time frequency transformations, the function basis is not defined by the wavelet method. So we can choose the mother wavelet adapted to studied images. Classical wavelets have been tested (Haar, Daubechies 9/7, Le Gall 5/3, Daubechies 4-tap orthogonal...). We also propose an automatic way to find an adapted wavelet. In 1994, Sweldens introduced a convenient way to satisfy all the desired properties of wavelets by reducing the problem to a simple relation between the wavelet and scaling coefficients. This approach is called the lifting scheme [12]. It permits to generate any compactly supported biorthogonal wavelet. An interesting property of biorthogonal wavelet filters is that they allow a perfect reconstruction of decomposed images. All the more, the lifting scheme makes the wavelet transform faster, hence it is use in the Jpeg-2000 compression standard. The optimization process is based on a genetic algorithm followed by a Powell direction set descent.

3) *Subband Characterization*: Wouwer et al. [13] show that for textured images, coefficients of the wavelet transform are distributed with a generalized Gaussian law on each subband. The law density (see equation (3) and Fig. 6) has parameters:

- α : a scale factor, it corresponds to the standard deviation of the classical Gaussian law.
- β : a shape parameter (Fig. 6) (its value is 2 for a Gaussian distribution)

$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-\left(\frac{|x|}{\alpha}\right)^\beta}, \quad (3)$$

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, z > 0$$

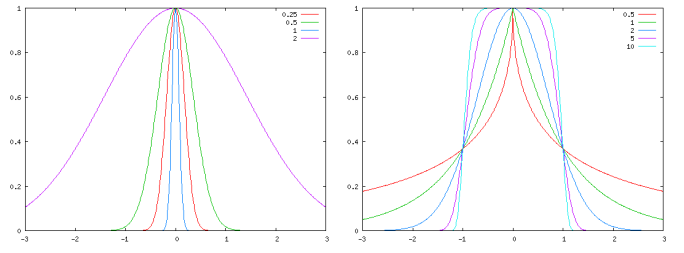


Fig. 6. Generalized Gaussian function

The first plot (on the left) illustrates the role of the parameter α , the second the role of β

We propose to characterize image texture by using a maximum likelihood estimator ($\hat{\alpha}, \hat{\beta}$) of the distribution law for coefficients of each subband of the wavelet decomposition.

These estimators are defined like this: given $\mathbf{x} = (x_1, \dots, x_L)$ the wavelet coefficients for one subband. We suppose that each x_i are independent. Varanasi and Aazhang [14] demonstrated that ($\hat{\alpha}, \hat{\beta}$) is the unique solution of equation (4).

$$\begin{cases} \hat{\alpha} = \left(\frac{\hat{\beta}}{L} \sum_{i=1}^L |x_i|^\beta \right)^{\frac{1}{\beta}} \\ \frac{\sum_{i=1}^L |x_i|^\beta \log |x_i|}{\sum_{i=1}^L |x_i|^\beta} - \frac{\log \left(\frac{\hat{\beta}}{L} \sum_{i=1}^L |x_i|^\beta \right)}{\hat{\beta}} - \frac{\Psi(\frac{1}{\hat{\beta}})}{\hat{\beta}} = 1 \end{cases} \quad (4)$$

where the function Ψ is defined by $\Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$. The $\hat{\beta}$ value should be find first by a Newton-Raphson algorithm. If the Newton-Raphson iterative algorithm is well initialized, it converges with few iterations. We define the image signature as the set of couple obtained for each ($\hat{\alpha}, \hat{\beta}$) in each subband.

C. Distance

To compute the distance between two signatures, we estimate the divergence of the coefficient distribution law for each subband. The distance is a weighted sum of the divergences between corresponding subbands of two images. The most classical divergence measure is the Kullback-Leibler distance (see equation (5)).

$$D(p(X; \theta_q) || p(X; \theta_i)) = \int p(X; \theta_q) \log \frac{p(X; \theta_q)}{p(X; \theta_i)} dx \quad (5)$$

By using the expression of the generalized Gaussian law density, we obtain equation (6).

$$D(p(X; \alpha_1, \beta_1) || p(X; \alpha_2, \beta_2)) = \log \left(\frac{\beta_1 \alpha_2 \Gamma(\frac{1}{\beta_2})}{\beta_2 \alpha_1 \Gamma(\frac{1}{\beta_1})} \right) + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma(\frac{\beta_2+1}{\beta_1})}{\Gamma(\frac{1}{\beta_1})} - \frac{1}{\beta_1} \quad (6)$$

Some subbands are more relevant for CBIR than others depending on the type of image. To take into account that, we introduce a weight for each divergence measure. If the data base is not classified we could use a relevance loop to adjust all the weights. If the database is classified (which is our case), these weights can be automatically learnt. In such a case we use genetic algorithm in order to define the best set of weights.

TABLE I
MEANS PRECISION FOR A 5 IMAGE WINDOW

base	decomposition level number	generalized gaussians classical wavelets (Daubechies 9/7)	generalized gaussians adapted wavelets	histogramms
faces	2	93.05%	95.5%	88.8%
retinas	3	43.7%	46.1%	41.0%
mammography	3	69.4%	-	70.0%

III. RESULT

The previous algorithms are independently tested on the three databases (see II-A). For evaluation and learning purpose, databases have to be classified. We choose the same criteria for the retinopathy and mammography database: the disease severity level. There are 6 levels for the diabetic retinopathy and 3 for the mammography. For the face database, images representing the same person belongs to the same class (40 classes).

The evaluation criteria is the precision.

- 1) Each image in the database is used as a query image.
- 2) The algorithm find the five images of the database closest to the query image.
- 3) Precision is computed for this query. It is the percentage of image belonging to the query image class.
- 4) After each query, we compute the mean precision. This criteria is communicated to the optimization process in order to find the best weights and wavelets (genetic algorithms).

In Table I, results are given for the best set of parameters (weight and decomposition number) of an 5 images window. This number (5) is the best compromise for physicians to perform a diagnosis. Too much images are difficult to interpret and too few are not representative of a pathology. Results are compared with the previous CBIR histograms algorithm detailed in [15]. These parameters are computed once during the training phase. The adapted wavelet results for mammography are not available due to a too long computing time. To find the adapted wavelet, the genetic algorithm needs to decompose the entire base with numerous wavelet candidates, our computing resources couldn't perform that within a rational time. The time needed to find the best wavelets and the best weight for the face database is few minutes and 1 days for the retina base with our 4 CPUs system. As soon as the wavelet and the weight are found a query has a quick answer (few seconds), because only the query image has to be decomposed.

IV. CONCLUSION AND DISCUSSION

Results show that this method can give good results. The adapted wavelet is a little better than the standard ones and histograms methods. Score for faces and mammography data are very good. Results for the retina database are encouraging but not very high. In fact, in diabetic retinopathy, lesions are

the only elements that determine the disease severity level. Some lesions like microaneurisms are very small (15x15 pixels): our global method is not suited to detect such lesions. And most of the time, physicians need more than one image to perform their diagnosis. Our goal is now to fuse semantic (clinical data) and numeric images to enhance results. Preliminary results show that the precision can reach 79.5% for the retinas database.

REFERENCES

- [1] http://www.storagetek.com/solutions/case-studies/casestudy_page2136.html.
- [2] C. Nastar, "Indexation d'images par le contenu: un etat de l'art," in *CORESA'97*, March 1997.
- [3] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, December 2000.
- [4] C. Wilkinson, F. Ferris, R. Klein, and al., "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.
- [5] M. Heath, K. Bowyer, and D. K. et al, "Current status of the digital database for screening mammography," *Digital Mammography, Kluwer Academic Publishers*, pp. 457–460, 1998.
- [6] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1186–1191, 1993.
- [7] F. Samaria and A. Harter, "Parameterisation of a stochastic model for human face identification," in *2nd IEEE Workshop on Applications of Computer Vision, Sarasota FL*, december 1994.
- [8] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- [9] J. Combes, A. Grossmann, and P. Tchamitchian, *Wavelets: Time-Frequency Methods and Phase Space*, 2nd ed. Springer-Verlag, 1989.
- [10] T. Nguyen and P. Vaidyanathan, "Two-channel perfect-reconstruction fir qmf structures which yield linear-phase analysis and synthesis filters," *IEEE transactions on acoustics speech and signal processing*, vol. 37, no. 5, pp. 676–690, may 1989.
- [11] S. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE transactions on acoustics speech and signal processing*, vol. 37, no. 12, pp. 2091–2110, 1989.
- [12] W. Sweldens, "The lifting scheme: a custom-design design construction of biorthogonal wavelets," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 2, pp. 186–200, 1996.
- [13] G. Wouwer, P. Scheunders, and D. Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Processing*, vol. 8, pp. 592–598, April 1999.
- [14] M. Varanasi and B. Aazhang, "Parametric generalized gaussian density estimation," *J. Acoust. Soc. Amer.*, vol. 86, pp. 1404–1415, 1989.
- [15] M. Lamard, W. Daccache, G. Cazuguel, C. Roux, and B. Cochener, "Use of jpeg-2000 wavelet compression scheme for content-based ophthalmologic retinal retrieval," in *Proceedings of the 27th annual international conference of IEEE engineering in medicine and biology society*, september 2005.