



# Wavelet optimization for content-based image retrieval in medical databases

Gwénolé Quéléec, Mathieu Lamard, Guy Cazuguel, Béatrice Cochener,  
Christian Roux

► **To cite this version:**

Gwénolé Quéléec, Mathieu Lamard, Guy Cazuguel, Béatrice Cochener, Christian Roux. Wavelet optimization for content-based image retrieval in medical databases. *Medical Image Analysis*, Elsevier, 2010, 14 (2), pp.227-241. <10.1016/j.media.2009.11.004>. <hal-00515295>

**HAL Id: hal-00515295**

**<https://hal.archives-ouvertes.fr/hal-00515295>**

Submitted on 6 Sep 2010

**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.

# Wavelet Optimization for Content-Based Image Retrieval in Medical Databases

G. Quéllec<sup>a,c</sup> M. Lamard<sup>b,c,\*</sup> G. Cazuguel<sup>a,c</sup> B. Cochener<sup>b,c,d</sup>  
C. Roux<sup>a,c</sup>

<sup>a</sup>*INSTITUT TELECOM; TELECOM Bretagne; UEB; Dpt ITI, Brest, F-29200 France;*

<sup>b</sup>*Univ Bretagne Occidentale, Brest, F-29200 France;*

<sup>c</sup>*Inserm, U650, IFR 148 ScInBioS, Brest, F-29200 France;*

<sup>d</sup>*CHU Brest, Service d'Ophthalmologie, Brest, F-29200 France;*

---

## Abstract

We propose in this article a Content-Based Image Retrieval method for diagnosis aid in medical fields. In the proposed system, images are indexed in a generic fashion, without extracting domain-specific features: a signature is built for each image from its wavelet transform. These image signatures characterize the distribution of wavelet coefficients in each subband of the decomposition. A distance measure is then defined to compare two image signatures and thus retrieve the most similar images in a database when a query image is submitted by a physician. To retrieve relevant images from a medical database, the signatures and the distance measure must be related to the medical interpretation of images. As a consequence, we introduce several degrees of freedom in the system so that it can be tuned to any pathology and image modality. In particular, we propose to adapt the wavelet basis, within the lifting scheme framework, and to use a custom decomposition scheme. Weights are also introduced between subbands. All these parameters are tuned by an optimization procedure, using the medical grading of each image in the database to define a performance measure. The system is assessed on two medical image databases: one for diabetic retinopathy follow up and one for screening mammography, as well as a general purpose database. Results are promising: a mean precision of 56.50%, 70.91% and 96.10% is achieved for these three databases, when five images are returned by the system.

*Key words:* Content-Based Image Retrieval, custom wavelet, lifting scheme, diabetic retinopathy, mammography

---

\* Tel.: +33 298018110, Fax.: +33 298018124

*Email addresses:* gwenole.quellec@telecom-bretagne.eu (G. Quéllec),

# 1 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. These images are very interesting for diagnostic purposes: they are directly related to the patient pathology and medical history, and may be used for diagnosis aid. 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 - CBIR) is one of the possible and promising solutions to effectively manage image databases (Nastar, 1997). In recent years, CBIR has been applied to several applications such as industrial (Tao et al., 2008), forensic (AlGarni and Hamiane, 2008), remote sensing (Hafiane et al., 2006; Xie, 2004) and medical applications (Müller et al., 2004). In a medical context, the potential applications of CBIR are categorization (i.e. retrieving images of the same anatomical regions) (Rahman et al., 2007; Horsthemke et al., 2007; Greenspan and Pinhas, 2007), data mining (Siadat et al., 2005), retrieving visually similar lesions (Kim et al., 2006; El-Naqa et al., 2004; Doyle et al., 2007; Balmashnova et al., 2007), typically for educational purposes, and finally diagnosis aid (Dy et al., 2003; Pourghassem and Ghassemian, 2008; Antani et al., 2004), which is the application we will focus on in this article. The purpose is to let an expert carry out a search in the database, without formulating a semantic description of the image he is examining: the expert simply places the image as a query to a system, which retrieves the most similar images in the database. Finally, the expert chooses the most similar images, from his point of view, in the set of retrieved images, and accesses the associated information.

To compare the query image to another image in the database, feature vectors characterizing their numerical content, called signatures, are extracted. Then, comparing these two images simply comes down to comparing their signatures. In recent CBIR systems, images are mostly characterized visually by their shape (Pourghassem and Ghassemian, 2008; Dai and Zhang, 2005; Zhang and Lu, 2005; Antani et al., 2004), their edges (Rahmani et al., 2008; Rahman et al., 2007; Banerjee and Kundu, 2003; Han and Guo, 2003), textural features (Rahman et al., 2007; Kokare et al., 2006; Chun et al., 2003; Rallabandi and Rallabandi, 2008; Pourghassem and Ghassemian, 2008; Oliveira et al., 2007; Muneeswaran et al., 2006; Yu et al., 2005) or color histograms (Rahman et al., 2007; Yap and Paramesran, 2006; Khanh et al., 2003; Konstantinidis et al.,

---

mathieu.lamard@univ-brest.fr (M. Lamard),  
guy.cazuguel@telecom-bretagne.eu (G. Cazuguel),  
beatrice.cochener@chu-brest.fr (B. Cochener),  
christian.roux@telecom-bretagne.eu (C. Roux).

2005). Several features proposed in the literature are invariant to translation (Bishnu et al., 2005), rotation (Kokare et al., 2006; Bishnu et al., 2005; Rallabandi and Rallabandi, 2008; AlGarni and Hamiane, 2008; Sastry et al., 2004; Xie, 2004), size (Bishnu et al., 2005; AlGarni and Hamiane, 2008; Lo et al., 2007) or illumination (Greenspan and Pinhas, 2007). In medical applications, rotation and size invariance are mostly necessary if we want to retrieve images similar in shape: typically, if we want to retrieve anatomical regions (Rahman et al., 2007; Horsthemke et al., 2007) or if we want to characterize lesions or patterns within a user-defined ROI (Kim et al., 2006) or within an automatically segmented region (Yu et al., 2005; Siadat et al., 2005; Antani et al., 2004). In this article, we propose a diagnosis aid system intended to retrieve images with the same pathology and at the same disease severity level, not simply representing the same anatomical region. For that purpose, we have to characterize images by the lesions they contain. And we do not want the user to define ROIs: medical images sometimes contain hundreds of lesions, it would be too time consuming and physicians would not use such a system. Also, we do not want to rely on segmentation to characterize the lesions, since it is hard to design a robust segmentation algorithm for each type of lesions. So we propose to characterize images by their textural content, at different scales: it allows us to characterize lesions of different sizes, in a generic fashion. Our approach is invariant to translation. For many medical imaging modalities, such as eye fundus imaging or mammography, the orientation and the scale of images are approximately normalized, due to the acquisition protocol. If for a particular imaging modality they are not, then they should be normalized prior to image retrieval; this possible issue is not addressed in the article.

The outstanding challenges of CBIR in medical applications are the following. First, how to bridge the gap between the extracted low-level features and the high-level perception of similarity of the physicians (Rahman et al., 2007)? We addressed this issue by automatically adapting the system parameters to a high-level criterion, using a supervised learning procedure. Then, how to integrate the proposed retrieval systems with the compression and archiving standards (Greenspan and Pinhas, 2007)? We addressed this issue by defining a wavelet-based image signature that may be extracted as images are compressed by a wavelet-based compression standard, such as JPEG 2000 (Taubman and Marcellin, 2001). Then, how to combine visual features with medical metadata (Greenspan and Pinhas, 2007; Antani et al., 2002; le Bozec et al., 2000; Shao et al., 2004)? And finally, how to efficiently retrieve images in large databases (Oliveira et al., 2007; Dy et al., 2003; Iakovidis et al., 2008)? We addressed the last two issues in a previously published article (Quellec et al., 2008): we used knowledge discovery trees to retrieve complete patient files consisting of possibly several images with metadata, thanks to wavelet-based image signatures, which enables a fast retrieval and an increased retrieval performance.

Wavelets have been widely used in image processing since the publication of the JPEG 2000 standard, including in the CBIR field (Xiong and Huang, 2002;

Do and Vetterli, 2002). Do and Vetterli, in particular, proposed a texture retrieval method based on the global distribution of wavelet coefficients in a texture image. We explore in this article two approaches to define image signatures and their associated distance measures from the global distribution of wavelet coefficients in images. Both approaches are based on a specific model for wavelet coefficient distributions: the first one is based on histograms and, like Do’s approach, the second one is based on generalized Gaussian functions. In fact, this article generalizes Do’s image representation, designed for texture images, to make it general, in particular to make it applicable to real-world medical images, and automatically tunable, so that it can be adapted to the studied images and pathologies. To adapt this image representation, we propose to use a custom wavelet basis and a custom decomposition scheme. The two approaches are evaluated on two medical image databases: one for diabetic retinopathy follow up and one for screening mammography. To show that the method is robust to illumination and slight orientation variations, it was also evaluated on a face database.

The setup of the article is as follows. The signatures and distance measures are described in section 2. Section 3 introduces our approach to search for an optimal wavelet basis. We discuss the use of an optimal wavelet transform for both image indexing and compression in section 4. Sections 5 and 6 explain how our method is evaluated and calibrated. And the method is applied to the three databases in section 7. We end with a discussion and conclusion in section 8.

## 2 Proposed image signatures and distance measures

We propose to build image signatures from their wavelet transform. The wavelet transform is indeed convenient since it is highly tunable: in particular we can use a custom basis of functions, called wavelets, to decompose images. This wavelet basis will be tuned in order to improve the retrieval performance of the system. To implement the wavelet transform, we use the lifting scheme Sweldens (1996), which is particularly suited to design a custom wavelet basis (see section 2.2). It is also computationally efficient, hence its use in compression, for instance. Two different image signatures are then extracted from these decompositions in section 2.4. The associated distance measures are described in section 2.5. But first, let us set out the basics of multiresolution analysis.

## 2.1 Multiresolution analysis

A multiresolution approximation is a nested sequence of linear spaces  $(V_j)_{j \in \mathbb{Z}}$  for approximating functions  $f \in L_2(\mathbb{R})$  such that:

- $V_j \subset V_{j+1}, \forall j \in \mathbb{Z}$
- $\bigcup_{j=-\infty}^{+\infty} V_j$  is dense in  $L_2(\mathbb{R})$  and  $\bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$
- $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, \forall j \in \mathbb{Z}$
- $f(x) \in V_j \Rightarrow f(x - 2^{-k}) \in V_j, \forall j, k \in \mathbb{Z}$

The approximation of a function  $f \in L_2(\mathbb{R})$  at a resolution  $2^j$  is defined as the projection of  $f$  on  $V_j$ . There is a function  $\Phi \in L_2(\mathbb{R})$ , called scaling function, such that for any  $j \in \mathbb{Z}$ ,  $(x \mapsto \Phi_{jk}(x) = \sqrt{2^j} \Phi(2^j x - k))_{k \in \mathbb{Z}}$  is a basis of  $V_j$ . The scaling function  $\Phi$  satisfies the refinement relation of equation 1.

$$\Phi(x) = 2 \sum_{k=-\infty}^{\infty} h_k \Phi(2x - k) \quad (1)$$

The additional information available in the approximation of  $f$  at the resolution  $2^{j+1}$  as compared with the resolution  $2^j$ , is given by the projection of  $f$  on the orthogonal complement of  $V_j$  in  $V_{j+1}$ , noted  $W_j$ . There is a function  $\Psi \in L_2(\mathbb{R})$ , called wavelet function, such that for any  $j \in \mathbb{Z}$ ,  $(x \mapsto \Psi_{jk}(x) = \sqrt{2^j} \Psi(2^j x - k))_{k \in \mathbb{Z}}$  is a basis of  $W_j$ . The wavelet function  $\Psi$  satisfies the refinement relation of equation 2.

$$\Psi(x) = 2 \sum_{k=-\infty}^{\infty} g_k \Phi(2x - k) \quad (2)$$

$h_k$  and  $g_k$  are the coefficients defining the family of the wavelet functions. The wavelet and the scaling filters are noted  $G = (g_k)_{k \in \mathbb{Z}}$  and  $H = (h_k)_{k \in \mathbb{Z}}$ , respectively.  $\{V_j\}_{j \in \mathbb{Z}}$  and  $\{W_j\}_{j \in \mathbb{Z}}$  constitute what we call a multiresolution analysis framework.

## 2.2 Biorthogonal wavelets and the lifting scheme

To reconstruct a signal decomposed by the procedure above, we define a dual multiresolution analysis framework  $(\{\tilde{V}_j\}_{j \in \mathbb{Z}}, \{\tilde{W}_j\}_{j \in \mathbb{Z}})$  associated with a dual scaling function  $\tilde{\Phi}$  and a dual wavelet function  $\tilde{\Psi}$ , that are biorthogonal to  $\Phi$

and  $\Psi$  (see equation 3).

$$\begin{cases} \langle \tilde{\Phi}_{jk}, \Phi_{jk'} \rangle = \delta_{kk'} \\ \langle \tilde{\Psi}_{jk}, \Psi_{jk'} \rangle = \delta_{kk'} \\ \langle \tilde{\Psi}_{jk}, \Phi_{jk'} \rangle = 0 \\ \langle \tilde{\Phi}_{jk}, \Psi_{jk'} \rangle = 0 \end{cases} \quad (3)$$

$\tilde{\Phi}$  and  $\tilde{\Psi}$  satisfy refinement relations such as equations 1 and 2, with coefficients  $\tilde{H} = (\tilde{h}_k)_{k \in \mathbb{Z}}$  and  $\tilde{G} = (\tilde{g}_k)_{k \in \mathbb{Z}}$  respectively. A function  $f \in L_2(\mathbb{R})$  can be written as follows:

$$\begin{aligned} x \mapsto f(x) &= \sum_{j,k \in \mathbb{Z}} \langle f, \tilde{\Psi}_{j,k} \rangle \Psi_{j,k}(x) \\ &= \sum_{k \in \mathbb{Z}} \langle f, \tilde{\Phi}_{J,k} \rangle \Phi_{J,k}(x) \\ &+ \sum_{j \leq J, k \in \mathbb{Z}} \langle f, \tilde{\Psi}_{j,k} \rangle \Psi_{j,k}(x), J \in \mathbb{Z} \end{aligned} \quad (4)$$

The lifting scheme allows the construction of compactly supported biorthogonal wavelets with compactly supported duals, i.e. such that the refinement filters  $H$ ,  $G$ ,  $\tilde{H}$  and  $\tilde{G}$  are finite filters (a finite number of coefficients are non-zero). It relies on a simple relationship between all multiresolution analysis frameworks that share the same scaling function: let  $\{\Phi^0, \tilde{\Phi}^0, \Psi^0, \tilde{\Psi}^0\}$  be an initial set of biorthogonal scaling and wavelet functions, then a new set  $\{\Phi, \tilde{\Phi}, \Psi, \tilde{\Psi}\}$  can be found according to equation 5 (Sweldens, 1996).

$$\begin{cases} \Phi(x) = \Phi^0(x) \\ \Psi(x) = \Psi^0(x) - \sum_k s_k \Phi^0(x - k) \\ \tilde{\Phi}(x) = 2 \sum_k \tilde{h}_k^0 \tilde{\Phi}(2x - k) + \sum_k s_{-k} \tilde{\Psi}(x - k) \\ \tilde{\Psi}(x) = 2 \sum_k \tilde{g}_k \tilde{\Phi}(2x - k) \end{cases} \quad (5)$$

Coefficients  $s_k$  of equation 5 can be freely chosen. Hence we can start with an initial set of biorthogonal filters and generate new biorthogonal filters by the previous relation. As a consequence, the lifting scheme is particularly suited to design a custom wavelet basis: we can generate several biorthogonal filters, using an optimization procedure, in order to find the wavelet basis maximizing a criterion (the retrieval performance of the system, in our application).

Claypoole et al. (1998) proposed a filter bank adapted from equation 5. It uses the lazy wavelet (Sweldens, 1996) as initial biorthogonal filter and two filters  $P = (p_i)_{i \in \mathbb{N}}$  and  $U = (u_i)_{i \in \mathbb{N}}$  are defined from the  $s_k$  coefficients. One step in the lazy wavelet transform consists in splitting the signal into its odd and even indexed samples. One step of the wavelet decomposition is given as follows:

- split: the signal  $x[n]$  is split into its odd  $x_o[n]$  and even  $x_e[n]$  coefficients.
- predict: we generate the wavelet coefficients  $d[n]$  as the error in predicting  $x_o[n]$  from  $x_e[n]$  using predictor operator  $P$ :  $d[n] = x_o[n] - P(x_e[n])$ .
- update: combine  $x_e[n]$  and  $d[n]$  to obtain scaling coefficients representing a coarse approximation to the original signal  $x[n]$ . This is done by using an update operator  $U$ :  $c[n] = x_e[n] + U(d[n])$ .

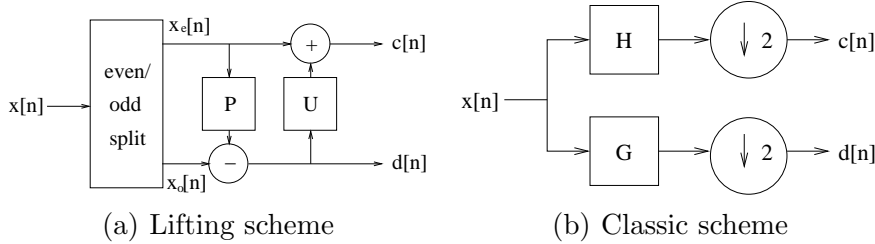


Fig. 1. The two-channel filter bank for the lifting scheme, compared with the classic scheme. Both filters are equivalent.

The filter bank is depicted in figure 1. Let  $N_p$  and  $N_u$  be the length of the linear filters  $P$  and  $U$ . (Claypoole et al., 1998) demonstrated that, to form a biorthogonal wavelet, filters  $P$  and  $U$  only have to satisfy the following conditions:

$$\begin{cases} \sum_{i=1}^{N_p} p_i = 1 \\ \sum_{i=1}^{N_u} u_i = \frac{1}{2} \end{cases} \quad (6)$$

There are then  $N_p + N_u - 2$  undetermined coefficients. Wavelets designed by the lifting scheme have a support length equal to  $s/t$  given in equation 7, where  $s$  is the support of the low-pass filter and  $t$  is the support of the high-pass filter.

$$\begin{cases} s = 2(N_p + N_u) - 3 \\ t = 2N_p - 1 \end{cases} \quad (7)$$

For instance, the following wavelets have been evaluated in the proposed system:

- the Le Gall 5/3 biorthogonal wavelet, which is used in the JPEG 2000 standard part I (support=5/3) (le Gall and Tabatabai, 1988)
- the Daubechies 9/7 biorthogonal wavelet, also in the JPEG 2000 standard (support=9/7)
- the orthogonal Haar wavelet (support=2/2)
- Cubic B-spline wavelet (support=4/2) (Cohen et al., 1992)
- the Daubechies 4-tap orthogonal wavelet (support=4/4) (Daubechies, 1988)
- the Daubechies 6-tap orthogonal wavelet (support=6/6) (Daubechies, 1988)
- Adapted wavelets (see section 3)



### 2.3 Decomposition of images

In the case of images, a separable wavelet decomposition is commonly used; it means that the one-dimensional filtering operations described above are performed on both rows and columns separately (the same filters are used along both directions). Three subimages of coefficients are obtained at each level  $l$ :  $lHL$ ,  $lLH$  and  $lHH$ , and a unique approximation subimage at the last level  $N_l$ :  $N_lLL$ . These coefficients are obtained from the decomposition of  $(l - 1)LL$  ( $0LL$  is the raw image):

- $lHL$  contains high frequency coefficients along columns and low frequency coefficients along lines,
- $lLH$  contains low frequency coefficients along columns and high frequency coefficients along lines,
- $lHH$  contains high frequency coefficients along columns and along lines,
- $lLL$  contains low frequency coefficients along columns and along lines, it is decomposed if  $l < N_l$ .

Subimages of coefficients are called subbands. An image decomposed on  $N_l$  levels thus consists of  $3 \times N_l + 1$  subbands. An example of image and its wavelet transform is shown in figure 2.

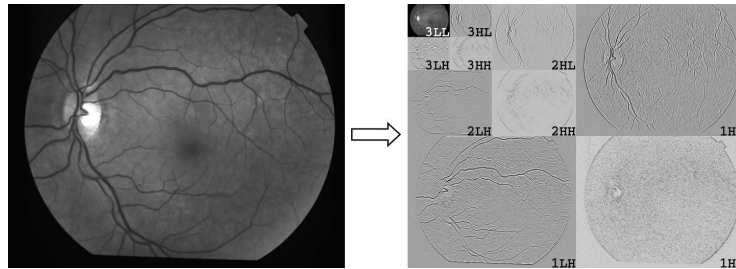


Fig. 2. Wavelet transform of an image - Pyramidal decomposition scheme

Such a decomposition scheme is said to be pyramidal. Other decomposition schemes are conceivable, indeed it is possible to decompose other subbands of level  $(l - 1)$  in addition to  $(l - 1)LL$ , and the choice of the subbands to decompose can be different at each scale. JPEG 2000 standard part II (JPEG, 2001), for instance, lets us specify any decomposition scheme. In particular, the following four schemes are specified in the standard, they are illustrated in figure 3:

- pyramidal decomposition
- SPACL decomposition (Signal Processing And Coding Lab, University of Arizona)
- packet decomposition
- FBI decomposition (Federal Bureau of Investigation, for fingerprint image compression)

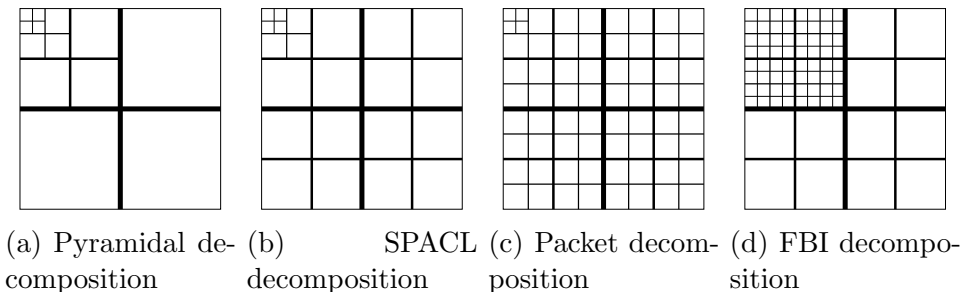


Fig. 3. JPEG 2000 - part II decomposition schemes. The line width indicates the decomposition level (the wider the lower).

Typically, the interest of using a different decomposition scheme is to find a basis relative to which the given image or collection of images has the lowest information cost, for compression purposes. Also, some subbands may emerge in which it is easier to discriminate between images of different classes, for classification or image retrieval purposes.

## 2.4 Image signatures

In order to characterize the image texture at different scales, we characterize the distribution of the wavelet coefficients in each subband of such a decomposition.

### 2.4.1 Histogram Representation

The first approach we consider to define an image signature is to build a histogram of the transformed image coefficients in each subband. If images are decomposed with a pyramidal scheme on  $N_l$  levels, they consist of  $3 \times N_l + 1$  subbands: there are 3 subbands of details at each scale  $l \leq N_l$  ( $l$ HH,  $l$ HL and  $l$ LH) plus an approximation ( $N_l$ LL),  $3 \times N_l + 1$  histograms are thus built. The signature is a vector formed by the concatenation of these histograms. The size of the resulting signature is relatively large. Indeed, let  $NB$  be the number of bins in each histogram, if  $N_l=4$  and  $NB = 32$ , the signature consists of  $NB \times (3 \times N_l + 1) = 416$  coefficients. For this reason, we propose an alternative method, in the next paragraph, that alleviates the problem.

### 2.4.2 Generalized Gaussian Representation

To characterize textures, [Wouwer et al. \(1999\)](#) proposed to model the distribution of transformed coefficients in each subband with zero-mean generalized Gaussian functions. A generalized Gaussian function may be expressed as in equation 8:

$$\begin{cases} p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} \\ \Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, z > 0 \end{cases} \quad (8)$$

We propose to discover to what extent this representation holds for any image dataset. The goodness of fit of generalized Gaussians to the actual distribution of wavelet coefficients is assessed by Kolmogoroff-Smirnov (K-S) tests ([von Mises, 1964](#)). For a reference wavelet basis (the Daubechies 9/7 wavelet basis), a K-S test is performed for each subband of the wavelet decomposition of each image in the dataset. The K-S tests measure the maximal absolute difference  $D$  between the cumulative distribution of the data, i.e. the wavelet coefficients in a given subband of a given image, and the cumulative distribution function (CDF) of the model, i.e. the generalized Gaussian function that best fits the distribution of the data. The CDF of a zero-mean generalized Gaussian is given below:

$$\begin{cases} cdf(x; \alpha, \beta) = \frac{1}{2} + \text{sign}(x) \frac{\gamma\left(\frac{1}{\beta}, \left(\frac{|x|}{\alpha}\right)^\beta\right)}{2\Gamma\left(\frac{1}{\beta}\right)} \\ \gamma(s, z) = \int_0^z e^{-t} t^{s-1} dt, z > 0 \end{cases} \quad (9)$$

We propose to build signatures consisting of the maximum likelihood estimator ( $\hat{\alpha}, \hat{\beta}$ ) of a generalized Gaussian distribution in each subband where the generalized Gaussian representation holds, and an histogram of the transformed image coefficients in the others. Intuitively, the last subband has no reason to have a specific distribution, since it is mostly an approximation of the image. For the scenario presented in the last paragraph, assuming the representation holds for each high frequency subband but not for the last subband, the signature consists of  $2 \times (3 \times N_t) + NB = 56$  coefficients. To compute the maximum likelihood estimator of a distribution, we applied the method proposed by [Do and Vetterli \(2002\)](#), extending the work of [Varanasi and Aazhang \(1989\)](#). This method relies on the Newton-Raphson root finding algorithm ([Kay, 1993](#)), that sometimes diverges on our images. This phenomenon is more or less important depending on the wavelet basis used to decompose images. As a consequence, we applied a robust root finding method proposed in ([Press et al., 1992c](#)): a hybrid method combining Newton-Raphson and the bisection algorithm.

## 2.5 Distance Measures

Each of the signatures presented above are associated with a particular distance measure, as described in the following paragraphs. The proposed distance measures are designed in order to profit from the characteristic of their associated signature.

### 2.5.1 Distance Between Histogram Based Signatures

The distance used to compare two images  $Im_1$  and  $Im_2$  by the first approach is base on the  $l1$ -distance between histograms. [Varela \(2004\)](#) compared other distance measures between histograms, on several medical databases, and this distance measure turned out to give the highest precision. The distance measure is given in equations [10](#) and [11](#):

$$d(Im_1, Im_2) = \sum_{i=1}^{3N_l+1} \lambda_i (H_i^1 - H_i^2) \quad (10)$$

$$H_i^1 - H_i^2 = \sum_{j=1}^{NB} H_i^1(j) - H_i^2(j) \quad (11)$$

where  $H_i^n(j)$  is the value of the  $j^{th}$  bin of the  $i^{th}$  normalized histogram of image  $n$  and  $(\lambda_i)_{i=1..3N_l+1}$  is a set of tunable weights.

### 2.5.2 Distance Between Generalized Gaussian Based Signatures

[Do and Vetterli \(2002\)](#) used the Kullback-Leibler divergence to compare two generalized Gaussian distributions. It measures the difference from a query probability distribution  $q$  to an arbitrary probability distribution  $i$ , defined by parameters  $\theta_q$  and  $\theta_i$ , respectively. Its general form is given in equation [12](#):

$$D(p(X; \theta_q) || p(X; \theta_i)) = \int p(X; \theta_q) \log \frac{p(X; \theta_q)}{p(X; \theta_i)} dx \quad (12)$$

Injecting equation [8](#) in [12](#), we obtain equation [13](#):

$$\begin{aligned} & D(p(X; \alpha_1, \beta_1) || p(X; \alpha_2, \beta_2)) \\ &= \log \left( \frac{\beta_1 \alpha_2 \Gamma(\frac{1}{\beta_2})}{\beta_2 \alpha_1 \Gamma(\frac{1}{\beta_1})} \right) + \left( \frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma(\frac{\beta_2+1}{\beta_1})}{\Gamma(\frac{1}{\beta_1})} - \frac{1}{\beta_1} \end{aligned} \quad (13)$$

The proposed *distance measure* between two images  $Im_1$  and  $Im_2$  is a combination of equations 13 and 11 (it is not actually a distance measure since it is not symmetric):

$$d(Im_1, Im_2) = \sum_{i=1}^{3N_l} \lambda_i D(p(X; \hat{\alpha}_{i,1}, \hat{\beta}_{i,1}) || p(X; \hat{\alpha}_{i,2}, \hat{\beta}_{i,2})) + (H_{3N_l+1}^1 - H_{3N_l+1}^2) \quad (14)$$

where  $(\hat{\alpha}_{i,n}, \hat{\beta}_{i,n})$  is the maximum likelihood estimator for the  $i^{\text{th}}$  subband of image  $n$ .

The two couples (signature, distance measure) presented in this section will be compared experimentally in section 7.4. In order to understand whether the potential performance difference between the two couples come from the signature or the distance, the couple (generalized Gaussian representation,  $l_1$ -distance) will also be assessed: the area between two probability distribution functions, with parameters  $(\hat{\alpha}_{i,1}, \hat{\beta}_{i,1})$  and  $(\hat{\alpha}_{i,2}, \hat{\beta}_{i,2})$ , is estimated according to the trapezoidal rule on an interval adapted to both distributions  $([-20\hat{\alpha}_{i,1}; 20\hat{\alpha}_{i,1}] \cup [-20\hat{\alpha}_{i,2}; 20\hat{\alpha}_{i,2}])$ . But first, we will see in the following section how the wavelet transform can be tuned in order to adapt these generic image representations to a given database.

### 3 Wavelet Adaptation

A major interest of wavelet decompositions is that we can choose the wavelet basis used to decompose images. Thus it can be adapted to the images and pathologies studied. However, the design of matched wavelet bases is usually a difficult task, as several constraints have to be satisfied simultaneously, in particular perfect reconstruction and vanishing moments. Moreover, these properties might not be sufficient to process specific signals or images. Several studies have thus been carried out to adapt a wavelet to a signal of reference, with different constraints on the wavelet properties and on the adaptation criteria. Usually, the goal is to minimize the  $L^2$  norm between the signal and an approximation derived from its wavelet decomposition (Tewfik et al., 1992; Gupta et al., 2005), for compression purposes, for instance. These methods have been applied to medical classification problems, in order to match waveform signals (Maitrot et al., 2005; de Sobral Cintra et al., 2004). However, in our case, we do not intend to match a signal of reference: we want to find the wavelet basis that maximizes the retrieval performance of our system.

### 3.1 *Adapting the wavelet basis with a high level criterion*

We propose in this article a method to find, within the lifting scheme framework, the wavelet basis that maximizes a given high level criterion, such as the retrieval performance of our system. The lifting scheme is convenient since it generates biorthogonal wavelet bases, and it is known that biorthogonality implies perfect reconstruction of decomposed images (Mallat, 1999). Indeed, any couple of filters  $P/U$  (of size  $N_p$  and  $N_u$ , respectively) satisfying the simple relations of equation 6 defines a biorthogonal wavelet filter bank.

### 3.2 *First solution: adapting biorthogonal wavelets*

A first solution to adapt the wavelet filter, within the lifting scheme framework, is thus to browse the space of biorthogonal filter banks, by an optimization procedure, and to select the filter bank that maximizes the adaptation criterion. The first  $N_p - 1$  (resp.  $N_u - 1$ ) coefficients of  $P$  (resp.  $U$ ) can be freely chosen, and the last coefficient of both vectors are computed in order to satisfy equation 6; as a consequence, we have  $N_p + N_u - 2$  degrees of freedom. The optimization procedure we used is described in appendix A: it is based on a genetic algorithm and Powell's direction set method.

### 3.3 *Second solution: adapting biorthogonal wavelets with vanishing moments*

In an image compression application, the  $N_p + N_u - 2$  degrees of freedom may be tuned in order to minimize the prediction error. To achieve this goal, Claypoole et al. (1998) proposed a solution to eliminate the first  $N_p$  moments of the primal wavelet  $\Psi$  and the first  $N_u$  moments of the dual wavelet  $\tilde{\Psi}$ . In our application, we wish to find a filter bank that would imply a high retrieval performance. However, it can be useful to design wavelets with several vanishing moments: it makes it possible to reduce the dimension of our optimization problem. Another interest of designing wavelets with a desired number of vanishing moments is discussed in section 4: it allows the design of wavelets suitable for both indexing and compression. So, we propose a solution to find a trade-off between these two goals: maximizing the retrieval performance and eliminating the lowest order moments of the wavelet. First, we eliminate the first  $M_p$  (resp.  $M_u$ ) moments of the primal (resp. dual) wavelet, where  $M_p \leq N_p$  and  $M_u \leq N_u$ : it sets  $M_p$  and  $M_u$  degrees of freedom to design filters  $P$  and  $U$ , respectively. And the remaining  $N_p - M_p$  and  $N_u - M_u$  degrees of freedom, respectively, are set in order to maximize our high level-criterion. Note that the first solution described above is a particular case of this second solution, where  $M_p = M_u = 1$ ; the same optimization procedure (described

in appendix A) is used to find the  $(N_p - M_p) + (N_u - M_u)$  free coefficients left after eliminating the first wavelet moments. The way we adapt the predict filter ( $P$ ) and the update filter ( $U$ ) is detailed in the next sections.

### 3.3.1 Adapting the predict filter ( $P$ )

In order to control the regularity of the primal wavelet, we will eliminate its first  $M_p \leq N_p$  moments:

$$\int t^n \Psi(t) dt = 0, n = 0..M_p - 1 \quad (15)$$

For that purpose, we define the following matrix  $V$ , of size  $M_p \times N_g$ , where  $N_g = 2N_p + 1$  is the size of the wavelet filter  $G$  (see equation 7):

$$V(k, l) = (l + 1 - N_p)^k, k = 0..M_p - 1, l = 0..N_g - 1 \quad (16)$$

$$V.G = 0 \quad (17)$$

If the wavelet filter  $G$  satisfies equation 17, then the first  $M_p$  moments of the primal wavelet vanish (each line in  $V$  eliminates one moment) (Claypoole et al., 1998).

We would like to transform this system in order to determine filter  $P$ . Comparing the two equivalent filter banks depicted in figure 1, we see that filter  $G$  can be expressed as a function of  $P$ . Indeed,  $d[n]$ , the high-pass filtered version of the input signal  $x[n]$ , satisfies the relations  $d[n] = G(x[n]) = x_o[n] - P(x_e[n])$ : as a consequence, if for instance  $P = (p_0, p_1, p_2, p_3)$ , then  $G = (-p_0, 0, -p_1, 1, -p_2, 0, p_2)$  (see figure 4 (a)).

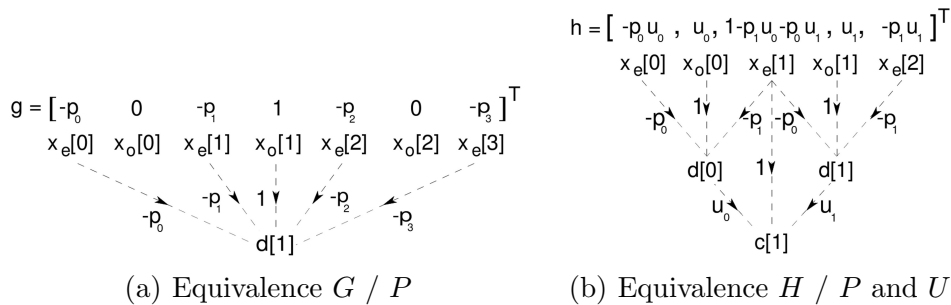


Fig. 4. Equivalence between the classic and the lifting scheme based filter bank.

Let the central coefficient in  $G$  be indexed by 0; we notice that the even coefficients in  $G$  are equal to 0, except the central coefficient which is equal to

1. As a consequence, we can reduce the dimension of the system above. Let  $C$  be the matrix containing the even coefficients in  $V$ , equation 17 becomes:

$$C \cdot \begin{pmatrix} p_0 & \dots & p_{N_p-1} \end{pmatrix}^t = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}^t \quad (18)$$

The right-hand side of this equation comes from the product of the central column of  $V$ , i.e.  $(0^0 = 1, 0^1 = 0, 0^2 = 0, \dots)^t$ , by the 1 in the central coefficient of  $G$  when solving equation 17 (the convention  $0^0 = 1$  was proposed by Claypoole et al. (1998)). Notice that the constraint associated with the first line in  $C$  is equivalent to the first biorthogonality relation (see equation 6). So we have  $M_p \leq N_p$  equations and  $N_p$  unknowns. We propose to generate coefficients  $(p_i)_{i=0..N_p-M_p-1}$  by the optimization module, in order to maximize the retrieval performance of the system (see appendix A) and to determine coefficients  $(p_i)_{i=N_p-M_p..N_p-1}$  thanks to equation 18. As a consequence, we have to solve a linear system  $A \cdot x = b$  of dimension  $M_p$ , where:

- $A$  is a square matrix of size  $M_p$ , which consists of the last  $M_p$  columns in  $C$ .
- Let  $B$  be the matrix containing the first  $N_p - M_p$  columns in  $C$ . The right-hand side  $b$  is equal to  $\begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}^t - B \cdot \begin{pmatrix} p_0 & \dots & p_{N_p-M_p-1} \end{pmatrix}^t$ .

We propose to solve this system by the LU method (Press et al., 1992a).

### 3.3.2 Adapting the update filter ( $U$ )

In order to control the regularity of the dual wavelet, we will eliminate its first  $M_u \leq N_u$  moments:

$$\int t^n \tilde{\Psi}(t) dt = 0, n = 0..M_u - 1 \quad (19)$$

As in the previous section, we design a matrix  $V'$  of size  $M_u \times N_{\tilde{g}}$  to eliminate the  $M_u$  moments of  $\tilde{G}$ . Since, we work with biorthogonal wavelets, filters  $\tilde{G}$  and  $H$  are related by equation 20 (Mallat, 1999) and  $N_{\tilde{g}} = N_h = 2(N_p + N_u) - 3$  (see equation 7).

$$\tilde{g}_k = (-1)^k h_k, \forall k \quad (20)$$

$V'$  is thus defined by equation 21 and  $\tilde{G}$  must satisfy equation 22.

$$V'(k, l) = (l + 2 - N_p - N_u)^k, k = 0..M_u - 1, l = 0..N_h - 1 \quad (21)$$



$$V'.\tilde{G} = 0 \quad (22)$$

We can express this system as a function of filters  $P$  and  $U$ . Indeed,  $c[n]$ , the low-pass filtered version of the input signal  $x[n]$ , satisfies the relations  $c[n] = H(x[n]) = x_e[n] + U(d[n]) = x_e[n] + U(x_o[n] - P(x_e[n]))$ : as a consequence, if for instance  $P = (p_0, p_1)$  and  $U = (u_0, u_1)$ , then  $H = (-p_0u_0, u_0, 1 - p_1u_0 - p_0u_1, u_1, -p_1u_1)$  (see figure 4 (b)) and according to equation 20,  $\tilde{G} = (-p_0u_0, -u_0, 1 - p_1u_0 - p_0u_1, -u_1, -p_1u_1)$ . Suppose that filter  $P$  has already been determined by the method proposed in section 3.3.1, then we can express equation 22 as a function of  $U$ . We propose a convenient solution below. Let  $C'$  be the matrix of size  $M_u \times N_u$  defined in equation 23, in which  $\tilde{g}_i^j$  is the factor of  $u_j$  in the expression of  $\tilde{g}_i$  (for instance, if  $\tilde{g}_i = 1 - p_1u_0 - p_0u_1$ , then  $\tilde{g}_i^0 = -p_1$ ,  $\tilde{g}_i^1 = -p_0$ ,  $\tilde{g}_i^2 = 0$ , etc.).

$$C'(k, l) = \sum_u V'(k, u)\tilde{g}_u^l \quad (23)$$

Then, equation 22 becomes:

$$C'.u = \left( -1 \ 0 \ \dots \ 0 \right)^t \quad (24)$$

The right-hand side of this equation comes from the product of the central column of  $V'$ , i.e.  $(0^0 = 1, 0^1 = 0, 0^2 = 0, \dots)^t$ , by the 1 in the central coefficient of  $\tilde{G}$ .

So we have  $M_u \leq N_u$  equations and  $N_u$  unknowns. We propose to generate coefficients  $(u_i)_{i=0..N_u-M_u-1}$  by the optimization module, in order to maximize the retrieval performance of the system (see appendix A) and to determine coefficients  $(u_i)_{i=N_u-M_u..N_u-1}$  thanks to equation 24. As a consequence, we have to solve a linear system  $A'.x = b'$  of dimension  $M_u$ , where:

- $A'$  is a square matrix of size  $M_u$ , which consists of the last  $M_u$  columns in  $C'$ .
- Let  $B'$  be the matrix containing the first  $N_u - M_u$  columns in  $C'$ . The right-hand side  $b'$  is equal to  $\left( -1 \ 0 \ \dots \ 0 \right)^t - B'. \left( u_0 \ \dots \ u_{N_u-M_u-1} \right)^t$ .

We also propose to solve this system by the LU method.

This concludes the description of the method we propose to adapt the wavelet basis with some high level criterion. We will now see how to adapt it in order to maximize the performance of a retrieval system.

## 4 Image indexing and compression

We envisage to extract the image signatures described above as images are compressed to be stored in a database: the same decomposition would be used for both indexing and compressing images. With this in mind, the use of the lifting scheme is particularly convenient, since it is used in compression standards such as JPEG 2000. In our application, we wish to find a filter bank that would imply a high retrieval performance and that may be used to compress images efficiently before they are stored in the database. That is why the second solution we proposed to adapt the wavelet (see section 3.3) is particularly suited: we build a wavelet basis with a desired number of vanishing moments, which ensures a good compression rate (Claypoole et al., 1998). In other words, the custom wavelet we build can be used to store images in a database efficiently.

## 5 Retrieval Performance

To tune the proposed retrieval system for a given database, we search for the system settings that maximize the retrieval performance. To evaluate the performance of a medical retrieval system, it is not possible to ask physicians to examine every image in the dataset and decide if it is relevant for each other image in this set. So, the evaluation of such systems simply relies on a classification of the images in the database, defined *a priori* by the physicians. However, images belonging to different predefined classes may be considered relevant by the physicians: for instance, two images from patients with different diseases, and consequently assigned to different classes, may contain lesions of the same type. So the retrieval performance we measure is a lower bound on the actual performance.

Let  $N$  be the number of images in the database and  $V_n$  the relevance of an image  $n$  for a query:  $V_n = 1$  if  $n$  and the query image belong to the same class,  $V_n = 0$  otherwise. The usual way to evaluate a retrieval system performance is the following (Müller et al., 2001): 1) each image in the database is presented to the system as a query, 2) the system ranks the  $N - 1$  remaining images by increasing order of the distance measure, 3) for each integer  $k = 1..N-1$ , the number of relevant images selected is computed:  $A_k = \sum_{n=0}^{k-1} V_n$ . Then, the following standard performance values are derived:

- the recall: the fraction of relevant images that are selected  $R_k = \frac{A_k}{\sum_{n=0}^{N-2} V_n}$
- the precision: the fraction of selected images that are relevant  $P_k = \frac{A_k}{k}$

To evaluate the system, precision-recall curves are usually plotted: they consist of linked points  $(P_k, R_k)$ ,  $k = 1..N-1$ . However, to optimize the system settings, a numerical value is needed: we measured the retrieval performance of the system by the mean precision evaluated at a cut-off rank  $k$  of interest, which is called mean precision at  $k$ .

Note that, to retrieve images in a large database, it may not be relevant to compare the query image to each image in the database: to speed up the retrieval process, some efficient search algorithms may be applied (Quellec et al., 2008; Oliveira et al., 2007; Dy et al., 2003; Iakovidis et al., 2008).

## 6 Retrieval system Calibration

In order to design an optimal wavelet family within the lifting scheme framework, for our retrieval system, we first tune the important parameters jointly with usual wavelet families, and the optimal value of these parameters are then used in the wavelet optimization step. The important parameters to be tuned are:

- $P_1$ : the model used to represent the distribution of coefficients in each sub-band (either histograms or generalized Gaussians) and the distance between models (either the  $l_1$ -distance or the Kullback-Leibler divergence),
- $P_2$ : the number of decomposition levels  $N_l$ ,
- $P_3$ : the decomposition scheme (either pyramidal, packet or SPACL decomposition).

Let  $P_4$  be the wavelet family. *a priori*, the weights in the distance measure have to be tuned for each set of important parameters (as described in appendix A). Each parameter is evaluated independently, the other three are set to a default value; the following default values are used:

- for  $P_1$ : generalized Gaussian model and Kullback-Leibler divergence,
- for  $P_2$ :  $N_l = 3$ ,
- for  $P_3$ : pyramidal decomposition,
- for  $P_4$ : Daubechies 9/7 wavelet.

For the purpose of training, the database is divided into a training and an evaluating set. The weights are set in order to maximize the mean precision at  $k$  on the training set. Then, we compute the mean precision at  $k$  on the evaluation dataset, using the weights found on the training set.

Finally, to design an optimal wavelet basis within the lifting scheme framework, we use the optimal values found for parameters  $P_1$ ,  $P_2$  and  $P_3$ , i.e. the values for these parameters that maximize the mean precision at  $k$  of the system, while the other parameters are set to their default values. Note that, to

find the optimal wavelet basis, we have to conduct a two-stage optimization: each time a wavelet basis is evaluated, we search for the optimal weight vector, as illustrated in figure 5. The wavelet basis is also designed in order to maximize the mean precision at  $k$  on the training dataset.

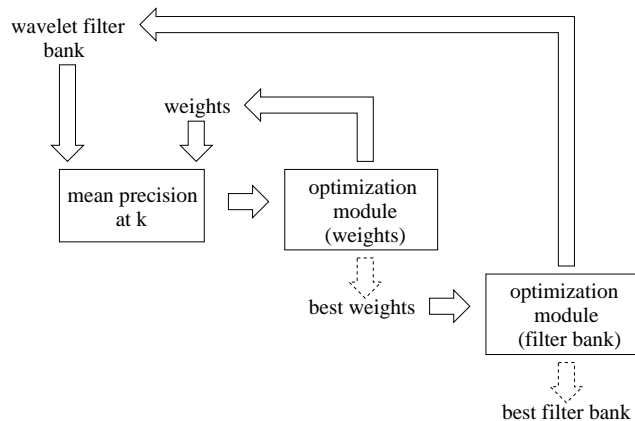


Fig. 5. Optimization procedure

## 7 Application

The retrieval method proposed below is evaluated on two medical databases. It is also evaluated on a face database, to show the robustness of the method to illumination and slight orientation variations. The databases are first described and the results we obtained are given in section 7.4.

### 7.1 Databases

#### 7.1.1 Diabetic Retinopathy Database (DRD)

This database has been built at the LaTIM laboratory<sup>1</sup> (Inserm U650) for research on diabetic retinopathy follow up. Diabetes is a metabolic disorder characterized by sustained inappropriate high blood sugar levels. This progressively affects blood vessels in many organs, which may lead to serious renal, cardiovascular, cerebral and also retinal complications. In the latter case, the pathology, namely diabetic retinopathy, can cause blindness. The database consists of 78 patient files: one for each diabetic patient seen at Brest University Hospital from June 2003 to April 2007 for an eye test. It contains 1261 photographs altogether, with associated anonymous information on the pathology. Images have a definition of 1280 pixels/line for 1008 lines/image. There are two possible orientations (right eye or left eye). Images have the

<sup>1</sup> <http://latim.univ-brest.fr>

same resolution. There are intra-image illumination variations but few inter-image illumination variations. They were acquired by experts using a Topcon Retinal Digital Camera (TRC-50IA) connected to a computer and are lossless compressed. Four types of images are obtained: color, red free, blue-light images and the angiographic sequences. An image series is given in figure 6 as an illustration. The disease severity level, ranging from 0 to 5, according to ICDRS classification (Wilkinson et al., 2003), was determined by experts for each patient.

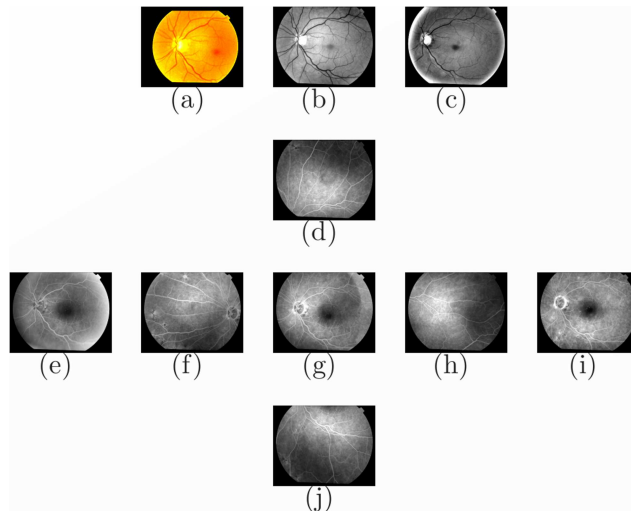


Fig. 6. Photograph series of a patient eye. Images (a), (b) and (c) are photographs obtained by applying different color filters on the camera lens. Images (d) to (j) form a temporal angiographic series: a contrast agent is injected and photographs are taken at different stages (early (e), intermediate (d), (f), (g), (h), (j) and late (i)). For the intermediate stage, photographs from the periphery of the retina are available.

### 7.1.2 Digital Database for Screening Mammography (DDSM)

The DDSM project (Heath et al., 1998), involving the Massachusetts General Hospital, the University of South Florida and the Sandia National laboratories, has built a mammographic image database for research on breast cancer screening. It consists of 2500 patient files. Each one includes two images of each breast, along with some associated patient information. The database consists of 10000 images. These images have a varying definition, of about 2000 pixels/line for 5000 lines/image. There are four possible orientations (see figure 7). Images have the same resolution for a given acquisition device. There are intra-image illumination variations but few inter-image illumination variations. An example of image sequence is given in figure 7. Each patient file has been graded by a physician. Patients are then assigned one of these labels: normal, benign and cancer.

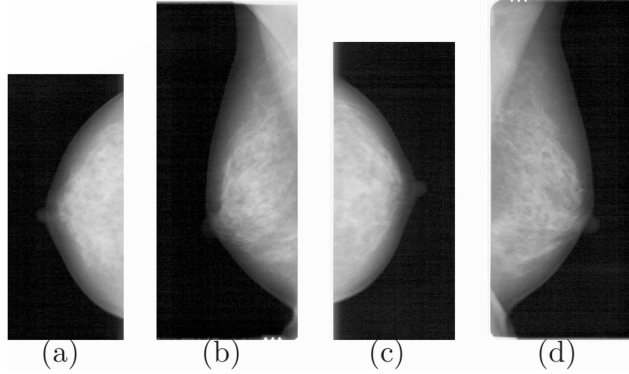


Fig. 7. Mammographic image sequence of the same patient. (a) and (b) are images of the left breast, (c) and (d) are images of the right one.

### 7.1.3 Face Database (FD)

The database consists of 400 images: ten photographs of 40 distinct subjects. For some of them, the images were taken at different times, with different lightings, 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 approximately frontal position. Images have a definition of 92 pixels/line for 112 lines/image. The database can be viewed on the website <sup>2</sup>. Figure 8 shows the 10 face images of a subject. Even though it is not our main purpose, this database can be classified easily. Images belong to the same class if and only if they represent the same subject.



Fig. 8. Image sequence of the same person's face

## 7.2 Class Definition

Concerning the two medical databases, we are mainly interested in the medical grade of each image. As a consequence, we associate each possible grade with a class: six classes are then defined in DRD and three in DDSM. In the face database, we associate each person with a class, so we have 40 classes of ten

<sup>2</sup> [www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html](http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html)

elements. The goal of the system is thus to retrieve images belonging to the same class.

### 7.3 Objectives of the retrieval system

In a medical application, the retrieval system is intended to be used to assist the diagnosis of physicians. Such a system would be used as follows: if a physician has a doubt on his diagnosis, he could send the acquired images to the retrieval system, that will display the  $k$  most similar images along with their medical interpretations. The physician would then analyze the retrieved images and, reasoning by analogy, these images may help him confirm or invalidate his diagnosis for the new case. In order to limit the time required by physician to analyze the retrieved cases, ophthalmologists in Brest University Hospital want the system to return only the five most similar images for a query. For comparison purposes, the same number of images is displayed for the other two databases. As a consequence, we want the five retrieved cases to be as relevant to the query as possible: we will tune the method parameters in order to maximize the mean precision when  $k=5$  images are returned by the system, i.e. the mean precision at  $k=5$ .

### 7.4 Results

The goodness of fit of the generalized Gaussian representation to the wavelet coefficient distribution, in each decomposition subband, is first reported for each of these databases: table 1 reports the average and standard deviation of the K-S statistic ( $D$ ) measured for each subband across all the images in a dataset. Figure 9 illustrates the goodness of fit of the representation for one image. The relevance of the generalized Gaussian representation is high for high frequencies and decreases as the frequency decreases.

To assess the performance of the method, each database is divided into a training and an evaluation dataset. The sizes of these datasets are chosen according to the number of images available for each database. For DRD and FD, the training dataset consists of 50% of the whole database; on DDSM, it consists of 5% of the database. For both medical databases, the training and the evaluation datasets have the same proportion of images from each class. For these two databases, images from the same person are always assigned to the same dataset, either the training or the evaluation dataset. For FD, all images in a given sequence are assigned to the same set. The images are assigned to the two datasets at random, with respect to the constraints above. The same training and evaluation datasets are used for each experiment. For wavelet adaptation, the size of filters  $P$  and  $U$ , respectively  $N_p$  and  $N_u$ ,

Table 1

Goodness of fit of the generalized Gaussian representation to the Daubechies 9/7 wavelet coefficient distribution in each decomposition subband. The distribution of the K-S statistics ( $D$ ), measured for each subband across all the images of a dataset, is reported: a small  $D$  indicates a good fit. We emphasized the subbands where an average  $D$  greater than 0.1 was measured.

database	DRD		DDSM		FD	
	mean	std	mean	std	mean	std
1HL	0.017	0.0027	0.014	0.0072	0.024	0.0066
1LH	0.009	0.0034	0.017	0.0095	0.032	0.0098
1HH	0.023	0.0070	0.027	0.0063	0.030	0.0102
2HL	0.007	0.0032	0.012	0.0062	0.038	0.0161
2LH	0.018	0.0043	0.017	0.0110	0.043	0.0258
2HH	0.038	0.0108	0.026	0.0062	0.047	0.0233
3HL	0.024	0.0066	0.009	0.0063	0.061	0.0239
3LH	0.026	0.1004	0.023	0.0156	0.074	0.0328
3HH	0.038	0.0227	0.019	0.0073	0.073	0.0329
3LL	<i>0.141</i>	<i>0.0358</i>	<i>0.225</i>	<i>0.0456</i>	<i>0.112</i>	<i>0.0323</i>

are both set to 4. At first, these filters only verify the biorthogonality relations (see section 3.2). The mean precision at  $k=5$  is given in table 2. To study the influence of the training dataset cardinality, we repeat the experiment with a training dataset consisting of 25% of the whole database for DRD and FD, and of 2.5% for DDSM. The scores obtained for the freely adapted wavelets (i.e. without forcing the moments to vanish) are reported in table 3. It emerges that the performance of the method increases with the size of the training dataset, we can thus expect an improvement as the databases will grow. We show in table 4 that we can also successfully adapt the wavelet if we eliminate some moments (see section 3.3). The experiment has been carried out on FD. The optimal weight vectors and  $P/U$  filters are reported in tables 5 and 6. The wavelet and scaling filters, respectively  $G$  and  $H$ , corresponding to filters  $P$  and  $U$ , are plotted in figure 10. The mean computation time to retrieve the 5 closest images is given in table 7 for the two types of signatures: histogram based (H) or generalized Gaussian based (GG). All experiments were conducted using an AMD Athlon 64-bit based computer running at 2 GHz, with the default values for  $P_2$ ,  $P_3$  and  $P_4$ , for comparison purposes.



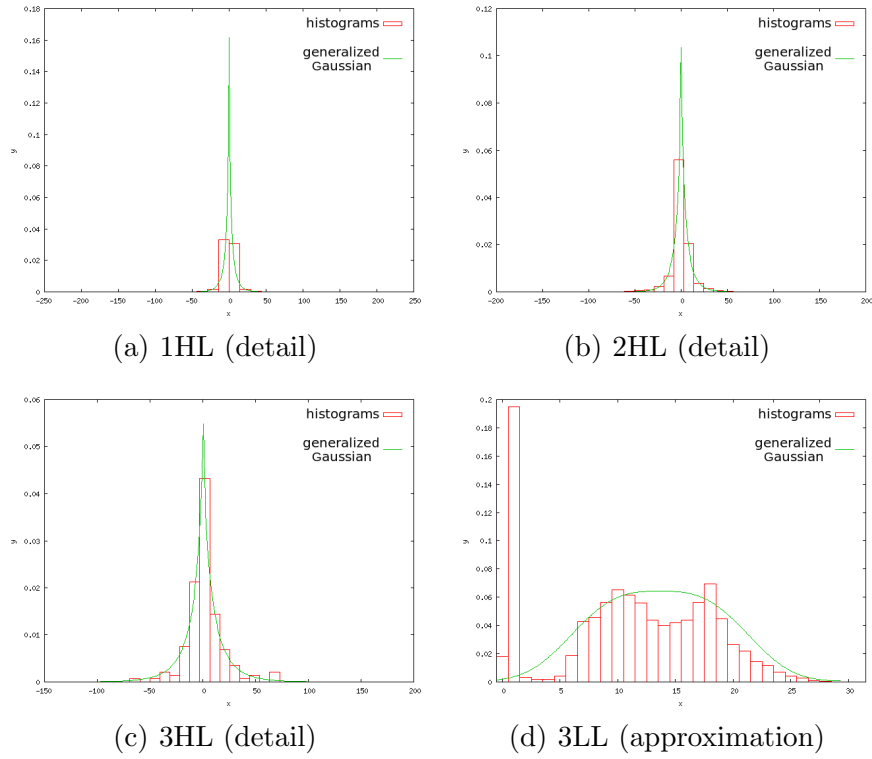


Fig. 9. Comparison of the generalized Gaussian model with the histogram model. The comparison was performed on a typical eye fundus red-free image (see figure 6 (b)).

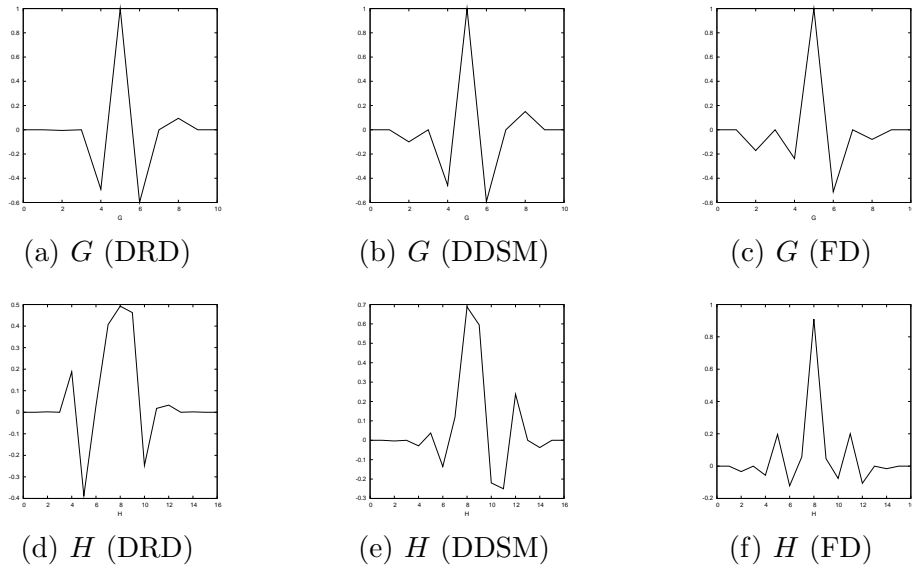


Fig. 10. Optimal wavelet and scaling filters obtained when the wavelet moments are not forced to vanish.

Table 2

Mean and standard deviation (std) of the precision at five (in %) obtained when varying the value of one important parameter of the method while setting the others to their default value. The default value for each parameter is underlined. GG and KL stand respectively for generalized Gaussians and Kullback-Leibler divergence.

database		DRD		DDSM		FD	
		mean	std	mean	std	mean	std
number of decom- position levels	0	<i>42.88</i>	19.73	<i>63.97</i>	24.96	94.20	9.07
	1	51.58	18.96	67.43	24.23	<b>95.10</b>	3.24
	2	52.42	18.93	67.72	25.77	95.50	1.99
	3	<b>54.17</b>	18.78	68.23	26.24	93.10	6.87
	4	54.12	18.89	<b>68.37</b>	25.82	86.40	9.54
	5	53.22	19.74	67.99	25.82	<i>78.70</i>	12.23
wavelet basis	Le Gall 5/3	52.90	19.37	68.17	25.53	89.20	8.01
	<u>Daubechies 9/7</u>	54.17	18.78	<b>68.23</b>	26.24	<b>93.10</b>	6.87
	Haar	<i>49.95</i>	19.20	<i>64.15</i>	24.21	<i>72.00</i>	15.67
	Cubic B-spline	53.30	18.77	66.34	24.71	85.70	10.84
	Daubechies 4	<b>56.10</b>	19.54	66.10	24.96	78.50	15.56
	Daubechies 6	56.14	19.95	64.95	25.76	80.00	12.35
signature / distance	histograms / <i>l1</i>	51.77	18.73	<b>69.75</b>	25.53	86.40	15.49
	GG / <i>l1</i>	<i>50.85</i>	18.93	<i>68.01</i>	25.82	<i>85.70</i>	14.31
	<u>GG / KL</u>	<b>54.17</b>	18.78	68.23	26.24	<b>93.10</b>	6.87
decom- position scheme	<u>pyramidal</u>	<b>54.17</b>	18.78	<i>68.23</i>	26.24	<b>93.10</b>	6.87
	wavelet packet	<i>51.36</i>	19.97	<b>68.63</b>	26.32	<i>85.70</i>	14.20
	SPACL	53.81	19.22	68.55	25.57	91.20	12.93
freely adapted wavelets (i.e. without forcing the moments to vanish)		<b>56.50</b>	19.26	<b>70.91</b>	25.10	<b>95.50</b>	4.66

Table 3

Influence of the cardinality of the training dataset on the precision at 5 of the retrieval method based on freely adapted wavelets. The large (resp. small) training dataset consists of 50% (resp. 25%) of the entire dataset for DRD and FD and of 5% (resp. 2.5%) for DDSM.

size	DRD		DDSM		FD	
	mean	std	mean	std	mean	std
small	52.42	21.89	70.75	26.24	94.40	10.47
large	56.50	19.26	70.91	25.10	95.50	4.66

Table 4

Mean precision at 5 of the retrieval method on the face database, as a function of the number adapted wavelet moments forced to vanish. Note that for  $M_p = N_p = M_u = N_u = 4$ , there is no wavelet adaptation.

	$M_u=1$	$M_u=2$	$M_u=3$	$M_u=4$
$M_p=1$	95.50	95.60	95.20	95.00
$M_p=2$	96.00	<b>96.10</b>	95.20	95.00
$M_p=3$	95.40	95.10	95.10	95.10
$M_p=4$	95.20	95.20	95.10	<i>94.20</i>

Table 5

Optimal weight vectors found for the retrieval method based on freely adapted wavelets (no vanishing moments enforced), using the optimal number of decomposition levels reported in table 2. The weights are gathered level by level:  $\{HH,HL,LH\}$ .

subbands	DRD	DDSM	FD
1st level	{2.030,0.3072,0.6189}	{17.73,43.32,13.78}	{0.4416,1.672,2.861}
2nd level	{22.62,3.553,2.705}	{26.31,8.748,3.524}	{ $\emptyset,\emptyset,\emptyset$ }
3rd level	{3.091,6.177,1.201}	{4.150,32.12,4.574}	{ $\emptyset,\emptyset,\emptyset$ }
4th level	{ $\emptyset,\emptyset,\emptyset$ }	{17.62,28.15,3.892}	{ $\emptyset,\emptyset,\emptyset$ }
<i>LL</i>	1	1	1

Table 6

Optimal filters  $P$  and  $U$  obtained when the wavelet moments are not forced to vanish.

DRD		DDSM		FD	
P	U	P	U	P	U
0.005319	-0.3874	0.09967	0.03717	0.1713	0.1964
0.4908	0.4068	0.4585	0.1180	0.2372	0.05720
0.5987	0.4627	0.5924	0.5949	0.5118	0.04635
-0.09482	0.01792	-0.1506	-0.2500	0.07974	0.2000

Table 7

Computation Time

database		DRD	DDSM	FD
wavelet transform		0.22s	1.99s	0.005s
H	computing histograms	0.03s	0.22s	0.0002s
	distance with each image in the database	0.31s	2.26s	0.109s
	overall time	0.56s	4.47s	0.114 s
	estimating $(\hat{\alpha}, \hat{\beta})$	4.35s	33.90s	0.03s
GG	distance with each image in the database	0.16s	1.16s	0.056s
	overall time	4.73s	37.05s	0.091s

As an illustration, figure 11 depicts the distribution of images from DDSM in the feature space.

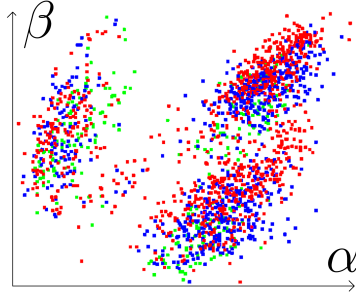


Fig. 11. Distribution of images from DDSM in the feature space, for a particular subband ( $HH2$ ), using the generalized Gaussian representation (image labels: green='normal', blue='benign', red='cancer'). Each cluster is associated with a particular acquisition device. Within these clusters, similar images are clustered (only approximately, because this figure depicts a projection of the feature space).

The proposed method is compared to image retrieval based on color (or gray level) histograms and to [Do and Vetterli \(2002\)](#)'s method, which are both particular cases of the proposed method. The color histogram based method is obtained by the proposed method without decomposition ( $N_l = 0$ ): the score obtained with this method is thus the first line of table 2. Do's method uses Daubechies 4-tap wavelets and a pyramidal transform; the same weight is assigned to each subband and the low-frequency subband is discarded. The proposed method is also compared to a retrieval method based on Zernike moments ([Khotanzad and Hong, 1990](#)), in which the influence of each moment in the distance measure is weighted using an SVM ([Mladenić et al., 2004](#)). The results are reported on figure 12. For information, the precision-recall curves obtained for each evaluated value of the system parameters ( $P_1$ ,  $P_2$ ,  $P_3$  and  $P_4$  - see section 6) are given in figure 13. On medical databases, we observe that the different parameters only affect the precision-recall curves for small recall values. The reason is that the learning procedure focuses on small recall values: it is supervised by the mean precision at five. Note that we obtain low precisions when the recall is high, because there are actually several clusters within each class (see figure 11 for instance), so two images from the same class can be far from each other in the feature space. Finally, we study the distribution of the images selected for a query, among the predefined classes. The repartition of the retrieved images among the different classes, as a function of the class of the query, is reported in table 8 for DRD.

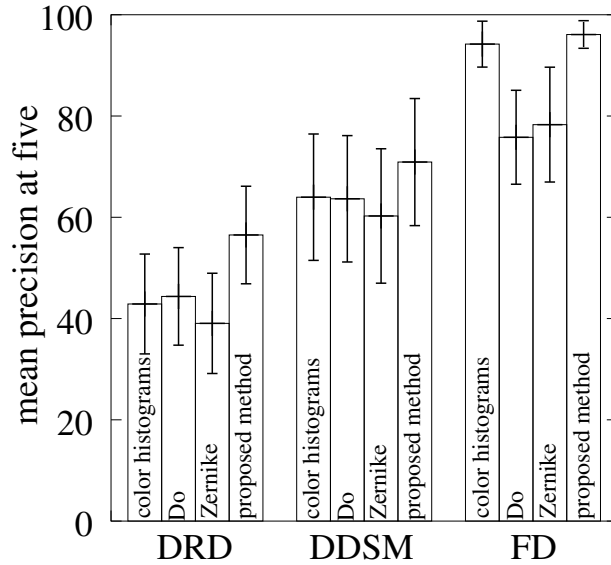


Fig. 12. comparison with other methods

Table 8

Confusion matrix - repartition of the retrieved images among the different classes, as a function of the class of the query

		class of the query					
		0	1	2	3	4	5
class of the results	0	53.03%	5.41%	5.60%	8.33%	6.59%	5.63%
	1	8.48%	57.88%	12.08%	10.71%	8.94%	7.57%
	2	10.91%	11.53%	55.56%	10.48%	8.00%	9.51%
	3	10.91%	8.94%	12.08%	54.05%	7.06%	7.18%
	4	8.79%	9.41%	6.76%	7.62%	58.82%	10.49%
	5	7.88%	6.82%	7.92%	8.81%	10.59%	59.61%

## 8 DISCUSSION AND CONCLUSION

In this article, two kinds of generic wavelet-based image signatures, with associated distance measures, have been evaluated in a CBIR system. They take advantage of the flexibility inherent in the wavelet transform framework to adapt the system to any specialized database. In particular, a way to adapt the wavelet transform to a high level criterion, within the lifting scheme framework, is proposed in this paper. It makes it possible to generate any wavelet transform, respecting the biorthogonality relations, and with a desired number of vanishing moments. A controlled random search, based on a genetic algorithm, is then performed in the predict and update filter space in order to find the optimal wavelet transform, and a similar search is performed in the

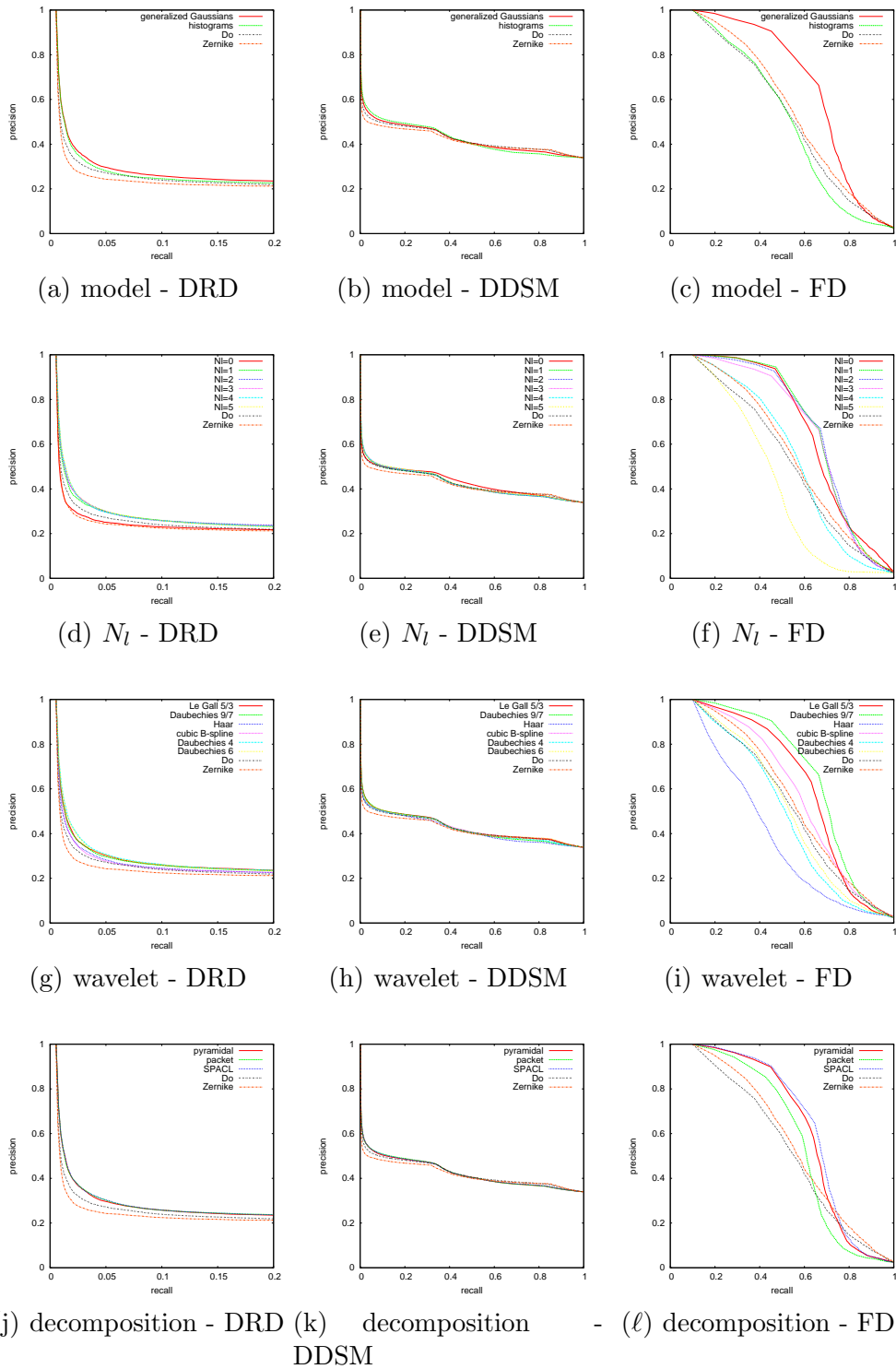


Fig. 13. Precision-recall Curves

distance weight vector space in order to maximize the precision of the system. The adapted wavelet transforms could also be used to store images in a compressed format, such as JPEG 2000, as their signature is computed. Indeed,

as table 4 shows, a high mean precision score can still be achieved while eliminating a desired number of moments, which ensures a good compression rate. It is particularly suited to JPEG 2000 standard part II (JPEG, 2001), which allows the user to define a custom wavelet basis and a custom decomposition scheme.

The system efficiency, measured with the mean precision when five images are returned by the system, i.e. the mean precision at five, is satisfactory: it reaches 56.50% on DRD, 70.91% on DDSM and 96.10% on FD. On DRD for instance, it means that on average two or three images out of the five retrieved by the system belong to the class of the query; and most images that do not belong to the right class are from nearby classes (see table 8). It is important to keep in mind that the scores of mean precision at five reported in this paper for medical datasets (56.50% on DRD and 70.91% on DDSM) are lower bounds on the actual performance of the system; since the system is evaluated through the use of predefined classes, it does not take into account the fact that images from different classes might be considered relevant by physicians (images at different severity scales can contain similar lesions for instance). Compared with the reference methods, the improvement is notable (see figure 12): for instance, performance increases by approximately a third on DRD. In particular, the proposed image representation is more efficient than Zernike moments, which are known to be rotation-, scale-, and translation-invariant; however, these invariances are not really important, since we are mainly interested in textures as well as lesions, which are details of these images.

It emerges from the precision-recall curves and the mean precision at five that the most influential parameters are the wavelet basis and the number of decomposition levels, which justifies our interest in finding an optimal basis within the lifting scheme framework. The optimization procedure always leads to a “better” wavelet basis than the standard bases evaluated. As for the decomposition scheme, it does not significantly influence the precision; as a consequence, we did not explore this trail any further. However, one could use the best basis algorithm (Coifman and Meyer, 1991) to find an optimal decomposition scheme.

A higher performance was obtained on the three databases when wavelet distributions were compared using the Kullback-Leibler (KL) divergence rather than using the  $l1$ -distance (see table 2: GG / KL *versus* GG /  $l1$ ). Using a generalized Gaussian (GG) representation is required to compute the Kullback-Leibler divergence: although this divergence was also defined for histograms, it is not suited for wavelet coefficient histograms that usually contain empty bins. However, we observe that representing the wavelet coefficient distribution by histograms does not lead to a high decrease of performance (see table 2: histograms /  $l1$  *versus* GG /  $l1$ ). From a computational point of view, both representations have their advantage. On one hand, computing the signatures is almost 20 times faster on average with the histogram based approach, although parameter estimation for generalized Gaussians could be approximated quickly from an adaptive histogram of each subband. But on the other hand,



it is twice as fast to compute the distance measure using generalized Gaussians. As a consequence, the larger the database size is, the more adequate the generalized Gaussian based approach is; and the larger the query image definition is, the more adequate the histogram based approach is.

Figure 11 illustrates another advantage of the generalized Gaussian representation: because the dimension of the feature space is much smaller than that of the histogram-based feature space, it is easier to visualize it and extract new knowledge. In particular, in the feature space of the mammography dataset, depicted in this figure, we can see that images from three different devices are clustered, and within these clusters, semantically similar images are close to each other. It suggests that if a dataset is large enough, we may be able to find relevant images for query images of different resolutions and from different modalities. On the other hand, we can see that normal images from the top-right cluster are close to the abnormal images from the bottom-right cluster. It suggests that if the cardinality of the dataset was smaller, then only images from the same modality should be looked for, in order to avoid retrieving semantically irrelevant images.

Although the face database is not a medical database, it is very interesting. Because we get a very high retrieval performance on this dataset, consisting of photographs of faces acquired with different illumination conditions and with slight orientation variations, we show that our method is robust to both these factors.

The adaptive framework proposed in this method turned out to be useful for the problems it was applied to. For instance, in DRD, more emphasis was given to the scales where the lesions are most visible. Moreover, in DDSM, the presence of lesions affects the distribution along specific directions: more emphasis was given to these discriminant directions (see table 5).

As mentioned in the introduction, the problem that remains to be solved, for such a CBIR system to be applied to any medical imaging modalities, is the normalization of the orientation and of the scale of images, when the variations are very large. However, the orientation normalization is mainly necessary in the presence of anisotropic textures or lesions, and the resolution of images is usually available in the metadata. Fortunately, if necessary, robust image registration techniques are available.

As a conclusion, because it returns at least two or three relevant images out of five in the worst case (DRD), we believe this system is a useful tool that can help physicians reducing the search space when looking for images in a database clinically similar to an image acquired during an examination.

## Appendix A - Optimization procedure

When we search for a weight vector between subbands or the most adapted wavelet, we have no preconceived ideas about either the fitness function or the free coefficients. Then we start the search by a controlled random process. Genetic algorithms ([Goldberg, 1989](#)) are amongst the quickest and most popular methods for finding rough estimations of the best local maxima within a given region of space. The general principle of genetic algorithms is to:

- generate a random population of solutions
- combine previous good solutions (crossover) and slightly modify some of them (mutation) to make the population evolve
- keep the fittest solutions at each generation

In particular, we used the steady state algorithm (at each generation, new individuals are created and added to the former population, the worst individuals of the resulting population being removed), with the following parameters:

- population size = 50
- maximum number of generations = 30
- selection methods: tournament selector (2 individuals are selected for crossover with a probability proportional to their score, and only the fittest is actually kept)
- crossover probability (CR) = 70% (CR: probability for two selected individuals to combine their genes and form two new individuals, otherwise they are cloned)
- mutation probability (MU) = 60% (MU: probability for a gene to be swapped with another after a crossover)

This configuration was adopted for its fast convergence.

Once approximations of the fitness function optima are found by a genetic algorithm, we use the best approximations as initial points for descents to reach the actual local minima. We used the Powell's Direction Set method ([Press et al., 1992b](#)). The interest of this method is that it does not require computing the function gradient, which is unknown. The descent typically leads to an improvement of 2 or 3% in precision over genetic algorithms alone, so this step is optional.

## References

- AlGarni, G., Hamiane, M., October 2008. A novel technique for automatic shoeprint image retrieval. *Forensic Science International* 181 (1).
- Antani, S., Lee, D. J., Long, L. R., Thoma, G. R., September 2004. Evaluation

- of shape similarity measurement methods for spine X-ray images. *Journal of Visual Communication and Image Representation* 15 (3), 285–302.
- Antani, S., Long, L. R., Thoma, G. R., 2002. A biomedical information system for combined content-based retrieval of spine x-ray images and associated text information. In: *Proceedings of the Indian Conference on Computer Vision, Graphics, and Image Processing*. pp. 242–247.
- Balmashnova, E., Platel, B., Florack, L. M. J., ter Haar Romeny, B. M., October 2007. Content-based image retrieval by means of scale-space top-points and differential invariants. In: *MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*. pp. 83–92.
- Banerjee, M., Kundu, M. K., November 2003. Edge based features for content based image retrieval. *Pattern Recognition* 36 (11), 2649–2661.
- Bishnu, A., Bhattacharya, B. B., Kundu, M. K., Murthy, C. A., Acharya, T., August 2005. Euler vector for search and retrieval of gray-tone images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 35 (4), 801–812.
- Chun, Y. D., Seo, S. Y., Kim, N. C., September 2003. Image retrieval using BDIP and BVLC moments. *IEEE Transactions on Circuits and Systems for Video Technology* 13 (9), 951–957.
- Claypoole, R., Baraniuk, R., Nowak, R., May 1998. Adaptive wavelet transforms via lifting. In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*. Vol. 3. pp. 1513–1516.
- Cohen, A., Daubechies, I., Feauveau, J., 1992. Bi-orthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* 45, 485–560.
- Coifman, R. R., Meyer, Y., 1991. Remarques sur l’analyse de Fourier à fenêtre (French. English summary) [Remarks on windowed Fourier analysis]. *Comptes Rendus de l’Académie des Sciences* 312, 259–261.
- Dai, S.-Y., Zhang, Y.-J., April 2005. Unbalanced region matching based on two-level description for image retrieval. *Pattern Recognition Letters* 26 (5), 565–580.
- Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* 41.
- de Sobral Cintra, R. J., Tchervensky, I. V., Dimitrov, V. S., Mintchev, M. P., September 2004. Optimal wavelets for electrogastrography. In: *Proceedings of the 29th IEEE EMBS Conference, San Francisco, USA*. Vol. 1. pp. 329–332.
- Do, M. N., Vetterli, M., February 2002. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image Processing* 11 (2), 146–158.
- Doyle, S., Hwang, M., Naik, S., Feldman, M., Tomaszewski, J., Madabhushi, A., October 2007. Using manifold learning for content-based image retrieval of prostate histopathology. In: *MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*. pp. 53–62.
- Dy, J. G., Brodley, C. E., Kak, A., Broderick, L. S., Aisen, A. M., March 2003.

- Unsupervised feature selection applied to content-based retrieval of lung images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (3), 373–378.
- El-Naqa, I., Yang, Y., Galatsanos, N. P., Nishikawa, R. M., Wernick, M. N., October 2004. A similarity learning approach to content-based image retrieval: application to digital mammography. *IEEE Transactions on Medical Imaging* 23 (10), 1233–1244.
- Goldberg, D. E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, Boston, MA.
- Greenspan, H., Pinhas, A. T., March 2007. Medical image categorization and retrieval for PACS using the GMM-KL framework. *IEEE Transactions on Information Technology in Biomedicine* 11 (2), 190–202.
- Gupta, A., Joshi, S. D., Prasad, S., 2005. A new method of estimating wavelet with desired features from a given signal. *Signal Processing* 85, 147–161.
- Hafiane, A., Chaudhuri, S., Seetharaman, G., Zavidovique, B., March 2006. Region-based CBIR in GIS with local space filling curves to spatial representation. *Pattern Recognition Letters* 27 (4), 259–267.
- Han, J. W., Guo, L., February 2003. A shape-based image retrieval method using salient edges. *Signal Processing: Image Communication* 18 (2), 141–156.
- Heath, M., Bowyer, K. W., Kopans, D., Kegelmeyer, W. P., Moore, R., Chang, K., MunishKumaran, S., 1998. Current status of the digital database for screening mammography. *Digital Mammography*, Kluwer Academic Publishers, 457–460.
- Horsthemke, W., Raicu, D., Furst, J., October 2007. Task-oriented medical image retrieval. In: *MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*. pp. 31–44.
- Iakovidis, D. D. K., Pelekis, N. N., Kotsifakos, E. E. E., Kopanakis, I. I., Karanikas, H. H., Theodoridis, Y. Y., 2008. A pattern similarity scheme for medical image retrieval. *IEEE Transactions on Information Technology in Biomedicine*In press.
- JPEG, 2001. Coding of still pictures - JPEG 2000 part2 ISO/IEC 15444-2. URL <http://www.jpeg.org/jpeg2000/j2kpart2.html>
- Kay, S. M., 1993. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Khanh, V., Hua, K. A., Tavanapong, W., July-August 2003. Image retrieval based on regions of interest. *IEEE Transactions on Knowledge and Data Engineering* 15 (4), 1045–1049.
- Khotanzad, A., Hong, Y. H., May 1990. Invariant image recognition by zernike moments. *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (5), 489–497.
- Kim, J., Cai, W., Feng, D., Wu, H., July 2006. A new way for multidimensional medical data management: Volume of interest (VOI)-based retrieval of medical images with visual and functional features. *IEEE Transactions on Information Technology in Biomedicine* 10 (3), 598–607.

- Kokare, M., Biswas, P. K., Chatterji, B. N., December 2006. Rotation-invariant texture image retrieval using rotated complex wavelet filters. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 36 (6), 1273–1282.
- Konstantinidis, K., Gasteratos, A., Andreadis, I., April 2005. Image retrieval based on fuzzy color histogram processing. *Optics Communications* 248 (4–6), 375–386.
- le Bozec, C., Zapletal, E., Jaulent, M. C., Heudes, D., Degoulet, P., 2000. Towards content-based image retrieval in a HIS-integrated PACS. In: *Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA)*. pp. 477–481.
- le Gall, D., Tabatabai, A., 1988. Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques. In: *Proc. of the International Conference on Acoustics Speech and Signal Processing (ICASSP)*. pp. 761–765.
- Lo, T.-W. R., Siebert, J. P., Ayoub, A. F., October 2007. An implementation of the scale invariant feature transform in the 2.5d domain. In: *MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*. pp. 73–82.
- Maitrot, A., Lucas, M.-F., Doncarli, C., March 2005. Design of wavelets adapted to signals and application. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. Vol. 4. pp. iv/617–iv/620.
- Mallat, S., 1999. *A Wavelet Tour of Signal Processing*. Academic Press.
- Mladeníć, D., Brank, J., Grobelnik, M., Milic-Frayling, N., January 2004. Feature selection using linear classifier weights: interaction with classification models. In: *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 234–241.
- Müller, H., Michoux, N., Bandon, D., Geissbuhler, A., 2004. A review of content-based image retrieval systems in medical applications - clinical benefits and future directions. *International journal of Medical Informatics* 73, 1–23.
- Müller, H., Müller, W., Squire, D. M., Marchand-Maillet, S., Pun, T., April 2001. Performance evaluation in content-based image retrieval: Overview and proposals. *Pattern Recognition Letters* 22, 593–601.
- Muneeswaran, K., Ganesan, L., Arumugam, S., Soundar, K. R., May 2006. Texture image segmentation using combined features from spatial and spectral distribution. *Pattern Recognition Letters* 27 (7), 755–764.
- Nastar, C., March 1997. Indexation d'images par le contenu: un etat de l'art. In: *CORESA'97*.
- Oliveira, M. C., Cirne, W., de Azevedo Marques, P. M., March 2007. Towards applying content-based image retrieval in the clinical routine. *Future Generation Computer Systems* 23 (3), 466–474.
- Pourghassem, H., Ghassemian, H., December 2008. Content-based medical image classification using a new hierarchical merging scheme. *Computerized Medical Imaging and Graphics* 32 (8), 651–661.

- Press, W., Teukolsky, S., Vetterling, W., Flannery, B., 1992a. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press.  
URL <http://www.library.cornell.edu/nr/bookcpdf.html>
- Press, W. H., Teukolsky, S., Vetterling, W., Flannery, B., 1992b. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, chapter 10.  
URL <http://www.library.cornell.edu/nr/bookcpdf.html>
- Press, W. H., Teukolsky, S., Vetterling, W., Flannery, B., 1992c. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, chapter 9.4.  
URL <http://www.library.cornell.edu/nr/bookcpdf.html>
- Quellec, G., Lamard, M., Bekri, L., Cazuguel, G., Cochener, B., Roux, C., March 2008. Recherche de cas médicaux multimodaux à l'aide d'arbres de décision. *Ingénierie et Recherche BioMédicale (IRBM)* 29 (1), 35–43.
- Rahman, M. M., Bhattacharya, P., Desai, B. C., January 2007. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. *IEEE Transactions on Information Technology in Biomedicine* 11 (1), 58–69.
- Rahmani, R., Goldman, S. A., Zhang, H., Cholleti, S. R., Fritts, J. E., November 2008. Localized content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (11), 1902–1912.
- Rallabandi, V. R., Rallabandi, V. P. S., October 2008. Rotation-invariant texture retrieval using wavelet-based hidden markov trees. *Signal Processing* 88 (10), 2593–2598.
- Sastry, C. S., Pujari, A. K., Deekshatulu, B. L., Bhagvati, C., December 2004. A wavelet based multiresolution algorithm for rotation invariant feature extraction. *Pattern Recognition Letters* 25 (16), 1845–1855.
- Shao, H., Cui, W.-C., Zhao, H., October 2004. Medical image retrieval based on visual contents and text information. In: *IEEE International Conference on Systems, Man and Cybernetics*. Vol. 1. pp. 1098–1103.
- Siadat, M.-R., Soltanian-Zadeh, H., Fotouhi, F., Elisevich, K., September 2005. Content-based image database system for epilepsy. *Computer Methods and Programs in Biomedicine* 79 (3), 209–226.
- Sweldens, W., 1996. The lifting scheme: a custom-design design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.* 3 (2), 186–200.
- Tao, D., Tang, X., Li, X., January 2008. Which components are important for interactive image searching? *IEEE Transactions on Circuits and Systems for Video Technology* 18 (1), 3–11.
- Taubman, D., Marcellin, M., 2001. *JPEG2000: Image Compression Fundamentals, Standards and Practice* (The International Series in Engineering and Computer Science). Kluwer Academic Publishers.
- Tewfik, A. H., Sinha, D., Jorgensen, P., March 1992. On the optimal choice of a wavelet for signal representation. *IEEE Trans. Inform. Theory* 38, 747–765.
- Varanasi, M. K., Aazhang, B., 1989. Parametric generalized Gaussian density estimation. *J. Acoust. Soc. Amer.* 86, 1404–1415.

- Varela, J. O., 2004. Indexation et recherche d'images par le contenu, utilisant des informations de compression d'images : application aux images médicales. Ph.D. thesis, ENST Bretagne - traitement du signal et télécommunication.
- von Mises, R., 1964. *Mathematical Theory of Probability and Statistics*. Academic Press, New York.
- Wilkinson, C., Ferris, F., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills, D., Kampik, A., Pararajasegaram, R., Verdaguer, J. T., 2003. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology* 110 (9), 1677–1682.
- Wouwer, G. V., Scheunders, P., Dyck, D. V., April 1999. Statistical texture characterization from discrete wavelet representations. *IEEE Trans. Image Processing* 8, 592–598.
- Xie, Z., November-December 2004. A rotation- and flip-invariant algorithm for representing spatial continuity information of geographic images in content-based image retrieval. *Computers and Geosciences* 30 (9–10), 1093–1104.
- Xiong, Z., Huang, T. S., September 2002. Wavelet-based texture features can be extracted efficiently from compressed-domain for JPEG2000 coded images. In: *Proceedings of the 2002 International Conference on Image Processing*. Vol. 1. pp. 481–484.
- Yap, P. T., Paramesran, R., February 2006. Content-based image retrieval using legendre chromaticity distribution moments. In: *IEE Proceedings - Vision, Image and Signal Processing*. Vol. 153. pp. 17–24.
- Yu, S.-N., Chiang, C.-T., Hsieh, C.-C., December 2005. A three-object model for the similarity searches of chest CT images. *Computerized Medical Imaging and Graphics* 29 (8), 617–630.
- Zhang, D., Lu, G., January 2005. Study and evaluation of different fourier methods for image retrieval. *Image and Vision Computing* 23 (1), 33–49.