Image Database Retrieval Using Wavelet Packets Compressed Data

Massimo Ancona[1], Walter Cazzola[1], and Paolo Raffo[1]

[1] DISI - Department of Informatics and Computer Science
Università degli Studi di Genova
Via Dodecaneso 35, 16146, Genova. Italy
{ancona|cazzola|raffo}@disi.unige.it

Marco Corvi[2]
[2] Elsag-Bailey Research Dept.
Via G. Puccini 2, 16154, Genova. Italy
corvi@elsag.it

Abstract. An efficient method for cross-correlating images, in querying operations over an image database is presented. The method relies on a multiresolution compression method, based on a variant of the wavelet packets best-basis algorithm of Coifman and Wickerhauser. It is shown that a searched image can be correlated with the compressed images of the database in a fraction of the time required by using the traditional cross-correlation function computed on the original bitmap image.

1 Introduction

Image database query is going to play a major role due to the growing importance of multimedia databases [1, 2]. In principle, one would like to retrieve images based on the *pictorial* comparison. This is however considered a hard task due to the large amount of the information enclosed in images; and for computational purposes, current image search engines makes use of simple features that are easily extracted from images: texture and color distribution, geometrical features such as shape, lines and corners distributions and/or meta-information associated with the images [3, 4].

From a theoretical point of view, a retrieval task can be considered as *establishing an ordering in the set of images in the database*, the topmost images in the ordering being the result of the query. The ordering criterion depends on the searched image (*the model*) and how two images should be compared (*the similarity measure*) [5, 6].

To cope with the large amount of information, researchers have used a multistage approach, in which a partial ordering is refined in successive steps only at the topmost positions, which are the most relevant. This approach allows to screen out, at each step, images that do not qualify as query result. The first steps should be the fastest and simplest as they analyze many images. These are usually based on meta-information or simple features. Later steps can be more elaborate and more accurately discriminating.

In this work, we propose an algorithm that relies on the *Best-Basis* multiresolution analysis and provides an accurate image pictorial comparison, that can be integrated in a image query engine [7, 8, 9, 10, 11]. The comparison feature is based on the *cross-correlation* between the image model and the images in the database, and works directly on the compressed data not requiring the reconstruction of the images. The *cross-correlation* is a general measure of the pictorial similarity of images and has been used for pattern matching problems in which the two images differ by a translation since the cross-correlation can be efficiently computed using the properties of the Fourier transform for all possible (relative) translations of the images [12]. It is still applicable when the images under comparison are related by more complex transformation, rotations, dilations, skews, but the computational cost becomes quickly unacceptable [13, 14]. The multiresoution analysis of images provides a faster method for a pictorial image comparison based on the cross-correlation [15] and a multiresolution approach to image querying offers many advantages over other techniques. First of all, wavelet decompositions give very good image approximations with just a few coefficients; these coefficients provide information that is independent of the original image resolution. Thus, the use of wavelets, specifically wavelet packets, allows the resolution of the query and the target to be effectively decoupled. Moreover, wavelet decompositions are fast and easy to compute, requiring linear time in the size of the image and very little code. Moreover, the running time and storage of the multiresolution method are independent of the resolution of the database images. In addition the information required by the query algorithm can be extracted from a wavelet packet compressed version of the image directly, allowing the creation of a convenient database of compressed images.

This pictorial comparison can be carried out directly on the compressed data. This is particularly important due to the necessity of storing the images in the database in compressed form.

Wavelet compressions are alternative schemes to the more conventional JPG, that are now getting attention thanks to their potential [16, 17] and the Best-Basis algorithm can be effectively used to implement a wavelet image compression [18, 19, 20].

We assume that the image in the database are stored compressed with a Best-Basis algorithm, with a common wavelet packet analysis. The model image used in the query-search is analyzed with the same wavelet packet algorithm. To compare it with each image in the database, only the coefficients most relevant to the image in the database, are considered. This amounts to decomposing the model image using the same basis used for the specific database image. The measure of the comparison is the cross-correlation of the two compressed images, i.e., of the images reconstructed using only the most relevant coefficients. By using the orthonormality of the best basis, this cross-correlation is evaluated directly from the coefficients.

Some details of the Best-Basis algorithm and of a possible compression scheme are described in section 2. In section 3 we discuss our algorithm of query search. We have tested the algorithm on a database of about 150 logo images; numerical results are reported in section 4.

2 Best-Basis

The Coifman and Wickerhauser's Best-Basis algorithm was originally developed for signal compression [7]. First, the method expands the signal ($\underline{x}$) into a b-tree (two-dimensional quad-tree) structured library of orthonormal bases, for example a redundant set of *wavelet packets*. Each node of the tree represents a subspace with different time-frequency localization characteristics. The computational complexity of this decomposition is $O(n \log n)$.

The best basis is defined as the set of orthogonal nodes in the tree that minimizes a particular cost function $\mathcal{M}$. An appropriate cost function should measure a signal's concentration of information in a given basis; a natural choice for this is the *Shannon's entropy* function.

Given a sequence $x = \{x_j\}$, one can define the Shannon-Weaver entropy of x by

$$\mathcal{H} = -\sum_j p_j \log p_j$$

where $p_j = \left|x_j\right|^2 / \|x\|^2$ and we define $p \log p = 0$ if $p = 0$. Unfortunately, the minimum cost is not rapidly computable, and, in general, the computational complexity for searching the minimum is not even low. However, when the cost function is additive, it is possible to reduce the computational complexity by employing a *divide & conquer* strategy [21]. The searching algorithm starts by marking all the bottom nodes; their total information cost is an initial value which we will try to reduce. Whenever a parent node is of lower information cost than the children, it is marked; but if the children have lower information cost, the parent is assigned the lower total information cost of the children.

In this inductive step any node is examined no more than twice, one as a child and one as a parent.

The algorithm terminates when the root is reached. The root node will return the minimum information cost of any basis subset below itself. This search has a complexity of $O(n \log n)$. Finally, after all the nodes have been examined, we take the topmost marked nodes, which constitutes the best basis. This is equivalent to a depth-first search and requires no more operations than the number of nodes in the tree, $O(2n)$.

A classical fact about entropy is that $\exp \mathcal{H}(x)$ is proportional to the number of coefficients needed to represent the signal to a fixed mean square error.

Unfortunately $\mathcal{H}$ is not an additive function, but if we consider

$$\mathcal{H}(x) = \|x\|^{-2} \lambda(x) + \log \|x\|^2,$$

where $\lambda(x) = -\sum |x_j|^2 \log |x_j|^2$, we may observe that minimizing $\mathcal{H}(x)$ is equivalent to minimizing $\lambda(x)$. Moreover, $\lambda(x)$ is an additive cost function, then it is possible to applicate the divide and conquer algorithm. The computational complexity of the wavelet packet transform is $O(n \log n)$ and the search for a global minimum for the function cost converges in $O(n \log n)$ operations, for n-sample signal.

The outcome of the decomposition of image consists of the best basis coefficients and their positions in the nodes in the b-tree. A compression scheme is envisionable by keeping the largest coefficients, motivated by the fact that the entropy of the decomposition is minimized by the choice of the best basis [21].

3 Same Basis

Now we are going to show our pattern matching algorithm, called *Same Basis*, used for accelerating database image query. It involves both on-line and off-line processing.

During off-line processing we gather images that will constitute our database. First of all we compute the average intensity of each image and normalize each of them so that the average intensity of the whole pixels is zero. This operation makes matching insensitive to the variations of background intensity.

After choosing a pair of quadrature mirror filters [22] and a proper cost function [21], we decompose all the images with the *Best-Basis* algorithm, selecting for each of them its own best orthonormal basis. Then we compress these bases deleting all the coefficients whose absolute value is above a threshold; this threshold may be different for one or other image because what is really needed is that the number of saved coefficients is the same for each image. Practically we are going to use only a subset of the basis to evaluate the similarity of images. This technique correspond to an orthogonal projection of the signal in a subspace of lower dimensions.

Now, the output of each decomposition includes the basis, (described by the ordinal numbers of the nodes of the tree, marked by the *Best-Basis* algorithm, in which there are some coefficients not deleted during compression), the basis coefficients and the positions in their own node of the tree.

During on-line processing we obtain the pattern to be searched in the database, and normalize it as we did before for other images. We decompose the pattern with the same wavelet packet library (tree) used for database models; then we select the nodes of the tree and the coefficients corresponding to the basis selected for the first image in the database: practically we choose the same orthonormal basis to represent the model and the pattern. Cross-correlation between the coefficients of the two bases is performed to evaluate the similarity degree of the two images. Pattern decomposition is performed once only, while selection and similarity evaluation are then repeated for each image of the database; the one having the higher result is the most similar to the pattern.

This strategy is suggested by the intuitive consideration that if test image is not much different from the model, then, using the same base to represent both of them, the respective coefficients are not so different.

This fact makes us think that little perturbations correspond to little variations in the coefficients only; even if the best basis of the signal is different from that of the model, the latter can represent well enough the signal to analyze.

In order to compare two $n \times n$ images, it would be enough to evaluate the cross-correlation on the images and this operations would take $O(n^2)$ operations; in our algorithm, by using orthonormality of the basis of model and pattern, the cross-correlation can be evaluated directly on the coefficients:

$$CC(T,I) = \frac{\sum_{i,j} c_{i,j} d_{i,j}}{\sqrt{\sum_{i,j} c_{i,j}^2}\sqrt{\sum_{i,j} d_{i,j}^2}}$$

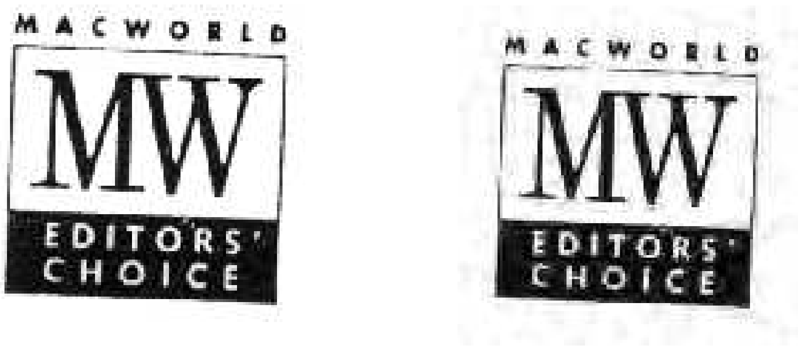
in which $c_{i,j}$ and $d_{i,j}$ are the coefficients of the two bases.

If a compression was performed the number of coefficients is smaller than the number of pixels of the original image. So cross-correlation requires only $O(kn^2)$ operations, where $k < 1$ is the compression rate of coefficients. Moreover, input/output time decreases, too because images stored in the database occupies less space than the original ones do.

4 Experimental results

We have conducted experiments to verify the accuracy and computational efficiency of the proposed algorithm, by an COMPAQ ALPHA Station 4000/4. In these experiments we created 3 databases: the first consists of 126 images, whose size is 128×128 pixels, representing gray scale logos; the others contain 126 gray scale images each gotten scanning the back of a £ 50000 banknote; the size of images is 64×64 pixels and 256×256 pixels, respectively.

In Fig. 1 a 128×128 reference image and the computed reconstruction is shown.



(a) Image n.69 (b) Reconstruction

Fig. 1. Example of database images and its reconstruction.

The compactly supported orthonormal wavelets used in our study have been chosen from [23]. The selection of filter coefficients depends both on the required accuracy and on the time constraints. Generally, larger is the order of filters than more accurate is the signal analysis and synthesis, but decomposition time may increase too much. So we chose second order of Daubechies' filters, whose length is twice their order and whose compression capability we found being good.

To decide how many coefficients should be saved, we repeatedly compressed and reconstructed some images, each time changing the compression threshold. Then, compression ratio was evaluated by a proper quality measure and comparing energy of deleted and saved coefficients. A commonly used measure is the *PSNR* (*Peak Signal to Noise Ratio*), by which the square of the differences of the corresponding pixels in the two images is evaluated. Given an image I and a compressed one E, whose sizes are $n \times n$, the formula is:

$$PSNR = 20 \cdot \log_{10} \left[\frac{\max_{i,j} I(i,j)}{\left(\frac{1}{n^2} SQD(i,j)\right)^{\frac{1}{2}}} \right]$$

Size	*Space*	*Coefficients*	*Nodes*	*Position*
64	4364	678	1052	1052
128	16446	1906	609	1906
256	65804	6824	3196	6824

Table 1. Occupied space by decomposed images

where

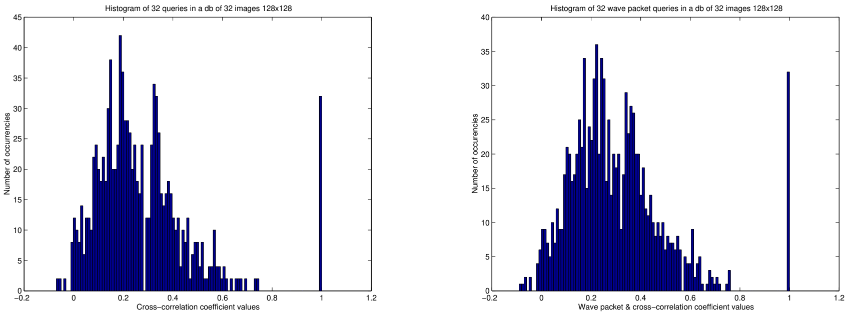
$$SQD(i,j) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (I(i,j) - E(i,j))^2$$

The PSNR value decreases when the differences between I and E grows.

Saving only 10% of coefficients of images whose energy ranges within 20000 and 29000, we obtained that $21.66 < PSNR < 35.89$ and the ratio between the energy of deleted and saved coefficients is within $1/40$ and $1/11$. Our first step is to decompose and to compress the reference images sets selecting only 10% largest coefficients; images are so stored in three files, as described in section 3. The basis coefficients in the first file should be stored with a precision of 32 bit, but this would require too much space; so we truncated them after the third decimal number; we normalized them as integers (by multiplying them for 1000) and appended another integer for each coefficient, to store the sign. The second file contains an integer in $[0,\ldots,\sum_{i=0}^{k} 2^i]$ where $k = \log_2 n$ and n is the linear size of images; the position of a coefficient is stored as an integer in $[0,\ldots,n^2]$; so the number of bits stored in the last two files is not the same for each image: if the integer is over 256, 16 bits are necessary, instead of 8.

In Table 1 the medium value of space occupied by the original image and by the three files, each of them considered as a binary file, is shown. As you can see from the table, storing an image compressed with *Best-Basis* method make us use half, or even a quarter, of the original space.

The evaluation of cross-correlation between two compressed images produces numerical results as good as those between two not compressed images. In fact in Table 2 we show that the range of mean and mean square differences between the respective values obtained in a query with compressed and not compressed images is small and boundaries values are quite little. As for the query time we can observe that as much larger the images grows, so much higher the time difference between our method and the classical one becomes. In Tables 3 and 4 we show the results of some queries in the three databases we described before; for each cell there are two values: the minimum and the maximum we found in our experiments. In particular, in the case of wavelet packet query, time of Input/Output, time of the decomposition of the test image, time of Same Ba-



(a) Correlation coefficient queries

(b) Wave packet & correlation coefficient queries

Fig. 2. Experimental results of 32 queries in a database of 126 images 128×128.

sis selection and Cross evaluation are shown; while in Cross-correlation query only Input/Output and evaluation time are shown.

In order to give a statistical description of similarity between patterns in the 3 databases we show in Fig. 2 two histograms; the first (a) regards cross-correlation values obtained during 32 queries in a database of 126 images of 128×128 size; while (b) refers to the same 32 queries performed with wavelet packet decomposition joint with cross-correlation. The diagrams obtained are similar: this means that even with compressed images cross-correlation results are still good.

5 Conclusions

We have presented a method for searching in an image database, based upon wavelet packets and cross-correlation as a similarity measure between patterns. We have used specifically feature sets based on the QMF wavelet packets decomposition because of the discrimination performance and the benefits the representation offers in a database application.

Size	*Mean (min-max)*	*Mean Square (min-max)*
64	0.0120 - 0.0372	0.0201 - 0.0381
128	0.0165 - 0.0261	0.0188 - 0.0268
256	0.0235 - 0.0291	0.0243 - 0.0302

Table 2. Differences between Cross correlation evaluated on entire or compressed data

Size	*I/O*	*Decomp.on*	*Same ba*	*Cross corr.*
64	10 - 11	0.27 - 0.29	1.57 - 1.73	0.02 - 0.04
128	22 - 24	1.16 - 1.22	6.02 - 6.59	0.12 - 0.14
256	123	5.43 - 5.57	26.0 - 27.3	0.50 - 0.55

Table 3. Maximum and minimum time value for Wavelet Packet queries

The algorithm we have described is fast, requires only a small amount of data to be stored for each target image and is remarkably effective. We have found that our method performs a query more quickly than cross-correlation only might do. Moreover, numerical values of matchings, obtained by our method, are quite similar to those given by cross-correlation query; being this latter a good similarity measure, this fact means that our method has a very good degree of accuracy in finding the correct match.

References

[1] Flickner, M., Sawhney, H.S., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by Image and Video Content: The QBIC System. IEEE Computer **28** (1995) 23–32

[2] Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., Shu, C.F.: Virage Image Search Engine: An Open Framework for Image Management. In Sethi, I.K., Jain, R.C., eds.: Storage and Retrieval for Image and Video Databases IV. Volume 2670 of SPIE Conference on Wavelet Applications., SPIE Press (1996) 76–87

[3] Jain, A.K., Vailaya, A.: Image Retrieval Using Color and Shape. Pattern Recognition **29** (1996) 1233–1244

[4] Manjunath, B., Ma, W.: Texture Features for Browsing and Retrieval of Image Data. IEEE Transactions on Pattern Analysis and Machine Intelligence **18** (1996) 837–842

[5] Santini, S., Jain, R.: Similarity Queries in Image Databases. In: Proceedings of CVPR'96, IEEE International Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA (1996)

[6] Santini, S., Jain, R.: Similarity Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence **21** (1999) 871–883

[7] Coifman, R.R., Wickerhauser, M.V.: Best-adapted Wave Packet Bases. Technical Report, Yale University (1990)

Image test	*Input/Output*	*Cross correl.*
64	19.89 - 20.88	0.295 - 0.358
128	79.707 - 81.48	1.361 - 1.436
256	315.86 - 317.87	6.662 - 6.855

Table 4. Maximum and minimum value of Cross correlation queries

[8] Chen, S.S., Donoho, D.: Basis Pursuit. Technical Report, Department of Statistics, Stanford University (1994)

[9] Chen, S.S., Donoho, D., Saunders, M.: Atomic Decomposition by Basis Pursuit. Technical Report, Department of Statistics, Stanford University (1995)

[10] Coifman, R.R., Wickerhauser, M.V.: Entropy Based Algorithms for Best Basis Selection. IEEE Transactions on Information Theory **32** (1992) 712–718

[11] Lu, J., Algazi, V.R., Estes, Jr., R.R.: Comparison of Wavelet Image Coders Using Picture Quality Scale (PQS). In Szu, H.H., ed.: Wavelet Applications II. Volume 2491. (1995) 1119–1130

[12] Wood, J.: Invariant Pattern Recognition: a Review. Pattern Recognition **29** (1996) 1–17

[13] sheng Chen, Q., Defrise, M., Deconinck, F.: Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994) 1156–1168

[14] Reddy, B.S., Chatterji, B.: An FFT-Based Technique For Translation, Rotation, And Scale-Invariant Image Registration. IEEE Transactions on Image Processing **5** (1996) 1266–1271

[15] Albanesi, M.G., Lombardi, L.: Wavelets for Multiresolution Shape Recognition. In Del Bimbo, A., ed.: Proceedings of ICIAP'97. LNCS 1311, Florence, Italy (1997) 276–283

[16] Nelson, L.J.: Wavelet Based Image Compression: Commercializing the Capabilities. Advanced Imaging (1996)

[17] Hilton, M.L., Jawerth, B.D., Sengupta, A.: Compressing Still and Moving Images with Wavelets. Multimedia Systems **2** (1994) 218–227

[18] Coifman, R.R., Meyer, Y., Quake, S.R., Wickerhauser, M.V.: Signal Processing and Compression with Wavelet Packets. In Byrnes, J.S., Byrnes, J.L., Hargreaves, K.A., Berry, K., eds.: Wavelets and Their Applications. Volume 442 of NATO ASI Series C: Mathematical and Physical Sciences. Kluwer Academic Publishers (1994) 363–379

[19] Wickerhauser, M.V.: Picture Compression by Best-Basis Sub-Band Coding. Preprint, Yale University (1990)

[20] Taswell, C.: Image Compression by Parameterized-Model Coding of Wavelet Packet Near-Best Bases. In Szu, H., ed.: SPIE Conference on Wavelet Applications. Volume 2491., SPIE Press (1995) 153–161

[21] Wickerhauser, M.V.: Lectures on Wavelet Packet Algorithms. (1991)

[22] Meyer, Y., Ryan, R.D.: Wavelets: Algorithms and Applications. SIAM (1993)

[23] Daubechies, I. CBMS-NSF Series in Applied Mathematics. In: Ten Lectures on Wavelets. SIAM (1992)