MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

Fast parallel multimedia database access based on wavelet multiresolution pyramidal decomposition

Jean Michel Marie-Julie - Hassane Essafi *
LETI (CEA – Technologies Avancees)

Abstract

In a multimedia database, information cannot been always accessed by numerical or textual attributes, we need often to analyze the content of images. Due to the high volume of data, a multimedia database is space and time consuming, the design of an efficient management system has to consider two major aspects: the storage space and the time retrieval by content. In order to reduce the data volume space and the time retrieval we have developed a method allowing, firstly, to store only compressed images in the database. Secondly, the retrieval process uses only compressed images.

This method has been evaluated either on monoprocessor workstation and on our SIMD parallel computer called SYMPHONIE (linear array processors, developed in our laboratory for infrared space applications). SYMPHONIE is an embedded multi- SIMD modular architecture based on MCM integration with 32 processors (32 bits) per chip so 128 processors can take place on a single VME board. SYMPHONIE is designed for real time vision and multimedia applications.

1 Introduction

The main principle for the access to image database, is to associate to the image, a semantic description of the image. With a such description, the system is able to process requests with a high semantic level. Several methods [4] has been proposed, as the use of textual descriptors or the indexation based on colors, shapes or geometric invariants. At the moment, textual description cannot be generated automatically by the system, the user has to describe himself the image content. So we can say this description is user dependent. The indexation based on color only allows queries based on color. With the indexation based on shape or geometric invariants, the problem is similar to the restriction of textual descriptor, primitives (edges, lines) must be identified at the storage step. At the moment, it is difficult to find a general purpose method, able to detect and discriminate major primitives of different objects in an image. The previous methods need a semantic analysis of image.

Another way to solve the problem of retrieval information from multimedia database, is to design an algorithm working directly with the image data, but it is time consuming because it is based on low-level process. A typically query, could be to find an user specified pattern in a large set of images. This query implies to evaluate each image with the pattern by using an appropriate algorithm. Due to storage needs, images are very often compressed, traditional methods used in pattern recognition which work with decompressed data, are not adapted; because the system has to decompress all images and then to match each of these with the pattern. This adds the decompression delay to the retrieval time. However, compression reduces the size of data, so working directly with compressed data will decrease the research time. There are two major parts in our work, first we have to choose a representation of the image directly available from the compressed data, and second we have to develop an adapted retrieval algorithm. The approach chosen is to analyze the images with the wavelet transform, this produces a multiresolution representation of the image. This representation is compressed by a vector quantization. Then at the research step, relevant levels from this representation can be extracted very fast. In order to match an user specified pattern and a candidate image, a specific multiresolution research scheme based on correlation has been developed. This reduces considerably the research time, compared to a monoresolution correlation. So this approach allows to work efficiently with wavelet compressed images for a fast retrieval based on content in image database. We present at first, the monoresolution approach to show its incompatibility with image database and content retrieval, then we present the retrieval based on the multiresolution representation generated by the wavelet transform. The third part presents the results obtained with the implementation on a scalar workstation and on a parallel system. We

*Address:DEIN – CE/S F91191 Gif-sur-Yvette Cedex. France. E-mail: {jmmjulie,hessafi}@gaap.saclay.cea.fr

finish with a conclusion and some results.

2 Monoresolution approach

Monoresolution approach requires matching pixel location in the candidate image and the pattern. This is accomplished by centering a correlation window on a location (X_L, Y_L) in the candidate image, the size of this window is equal to the pattern size. The correlation window is then systematically moved to all pixel locations within the candidate image, and the correlation between the pixels in the candidate image and those of the pattern is computed. So, an array of correlation coefficients is generated. Each correlation coefficient represents the degree of similarity of the pattern to a sub-area of the image at a specific location (X_L, Y_L) in the candidate image. When the maximum of the correlation coefficients is superior to a given threshold, this means that the pattern is present in candidate image at the position where the maximum correlation is achieved. In our process, the normalized correlation coefficient was chosen, mainly because it acts as a filter so that noise added to the images does not affect the correlation measure. Consequently, radiometric normalization of the images is not necessary.

The cross correlation algorithm [2][3] is very efficient, it is used in lots of industrial applications (robotics for instance), in our case it allows us to identify the image containing the pattern with high efficiency. However it is very time consuming, because we need to scan all the image. Its complexity is function of the image size and the pattern size. For instance, the matching of a 64×64 pixels pattern with a 512×512 pixels candidate image takes 150 seconds on a SPARCstation 5 cadenced at 110MHz. Furthermore the use of cross correlation for database adds an other dimension, in this case the time is multiplied by the number of candidate images in the database. If we consider a database containing 100 images, the processing for a 64×64 pixels pattern, takes 4 hours. So we can say that the monoresolution approach is not adapted to image database. In order to alleviate this drawback, we have proposed and evaluate a method based on image compression and multiresolution approach.

3 Image compression based on multiresolution approach

The use of a compressed image and a multiresolution approach reduces drastically the cost of cross correlation, by applying it to smaller representations of the image issued from a hierarchical spatial and spectral decomposition. This decomposition is obtained using wavelet transform. As described bellow this orthogonal transformation is a part of the compression algorithm.

3.1 Image Compression Algorithms

Image compression is classically achieved in three steps. As illustrated in figure 1, the image is first transformed in a set of decorrelated coefficients. Then, the transformed coefficients are quantized. Finally, the quantized values are entropy coded.

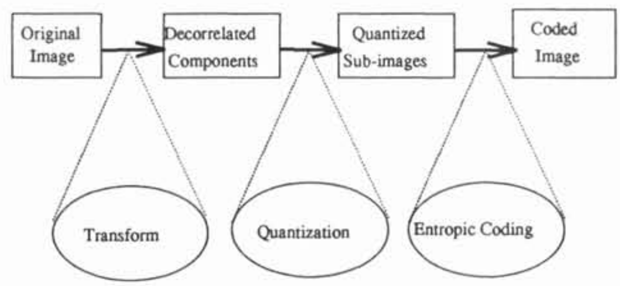

Figure 1: Image Compression Scheme

During the past few years several design algorithms have been developed for each step. The compression algorithms based on wavelets transform, vector quantization and Huffman coding provide one of the best trade-off between compression rates and quality.

3.2 Wavelet Transform

The use of the wavelet transform in image compression offers two essential advantages. First, the coefficients are well decorrelated thanks to the good localization of the wavelet functions both in space and frequency. And the multiresolution aspect of the wavelet transform is adapted to the vector quantization : each sub-image can be quantized using the appropriate size of the quantization vector.

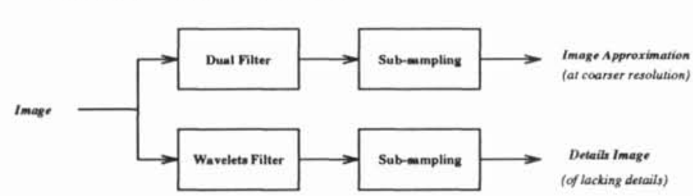

Figure 2: Wavelet filter convolution and decimation.

The most popular wavelet transform algorithm is based on a subband coding scheme (see figure 2), i.e it uses linear filter convolutions and image shrinking. It proceeds as follows: First, the signal is decomposed into a low-pass and a high-pass version; then, both versions are decimated with a factor 2 [1]. The low-pass version corresponds to the image texture (approximation) at the inferior resolution. And The high-pass version is equivalent to the missing details between the two resolutions.

3.3 Vector Quantization

Quantization is the most essential part of common image compression algorithms. In fact, it is the part which provides the major bits rate reduction, since each block of the decorrelated image is substituted by reference of the closest entry of a predefined codebook. The block size and codebook size depends on the desired compression rate.

This algorithm is achieved in two steps. First, the pixels, organized into k-dimensional vectors (blocks), are compared to every vector (named centroid) of a predefined catalog (named codebook). Then each vector of pixels is replaced by the reference of its closest centroid.

4 Retrieval and storage algorithm description

4.1 Storage

During the storage phase the image is first compressed using wavelet and vector quantization. Then only the compressed image is stored, this includes data from the compressed image and the codebook (one unique codebook can be used within a class of images). This compressed image can be viewed as a pyramidal representation thanks to wavelet transform. The retrieval algorithm is based on comparison between the pyramidal representations of the pattern and of the candidate image.

4.2 Retrieval

The first step of the algorithm is to decompose the pattern using the wavelet transform in order to obtain a pyramidal decomposition (see figure 3, the figure illustrates this process using only two levels). Each sub-image of the nth level of the pattern is correlated with its corresponding sub-image of the candidate image. The correlation associated with a defined threshold allows even to discard the candidate image, even to confirm the matching by considering the level number n-1 of the pyramid and reducing the search space in the image (see figure 3). Each level of the pyramid contains 4 spectral images. Only reduced sub-images presenting the same spectral properties are cross-correlated. The depth of the pyramid, the sub-images selected and the bounds of the threshold for the correlation are determined by the pattern geometrical and spectral properties. The number of decompositions applied to the pattern depends on its size (sub-images at the top of the pyramid, are constrained not to be smaller than 5×5 pixels). The correlation process is only applied to compatible part of the pyramid. This means that for a given image, only a portion of the pyramid is composed, this portion contains the same levels (in term of size, spectral properties) as the pattern pyramid. In order to avoid a useless correlation, the amount of information contained by each sub-images of the pattern is evaluated.

5 Implementation

The compression allows to reduce the database space by a factor of 10 thanks to the wavelet transform and the retrieval time by a factor of 200 and more thanks to the multiresolution research. Furthermore this algorithm is well suited for data-parallel architectures. We have compared result between, the monoresolution and the multiresolution approach on a scalar workstation (SPARCstation 5, cadenced at 110 Mhz) and then on a SYMPHONIE SIMD parallel system. For instance, we consider a database containing 100 512×512 pixels images, and a 64×64 pixels pattern. Results are given in the table bellow. We can observe an important reduction of the research time (table 1) (10 minutes against 4 hours obtained with the monoresolution approach), using the multiresolution approach, because correlation is applied on reduced representations of the image. A first stage consists in evaluating the correlation on high reductions of both image and pattern (8×8 pixels, 90 % images are eliminated). Then the system compares bigger representations of the pattern and the image (16×16 pixels), another part of candidate images is eliminated (8 %). The last stage is just a confirmation concerning 2 % of the database (table 2).

	100 images	1000 images
SPARC	654 s	6540 s
SYMPHONIE	2,08 s	20,8 s

Table 1: Total retrieval times of a 64×64 pattern, using a SPARC and an 128 PE SYMPHONIE system.

6 Conclusion

The use of multiresolution is a very interesting way for the retrieval by content in image databases. Furthermore, this approach is fully parallel and allows to retrieve images in real time using parallel architecture like SYMPHONIE. For instance, the retrieval in a 1000 images database, can be performed in 20 seconds with a SYMPHONIE system, including 8 VME boards.

References

[1] P. Moravie, H. Essafi, C. Lambert-Nebout, J-L. Basille, "Real-Time Image Compression using SIMD architectures", CAMP'95, Como Italy 18-20 September, 1995.

[2] Jean Michel Marie-Julie, Pascal Adam, Didier Juvin, "Real Time Stereovision Using Correlation on a Parallel SIMD Computer, Sympati 2", ICA3PP, Brisbane Australia 19-21 April, 1995.

[3] Christophe Mazzoni, Hassane Essafi, Patrick Julien, Olivier Jamet, "Stereocorrelation on Parallel Openvision System", ICA3PP, Brisbane Australia 19-21 April, 1995.

[4] A. Pentland, R.W. Picard, S. Sclaroff, "Content-Based Manipulation of Image Databases", International Journal of Computer Vision, 232-254 (1996).

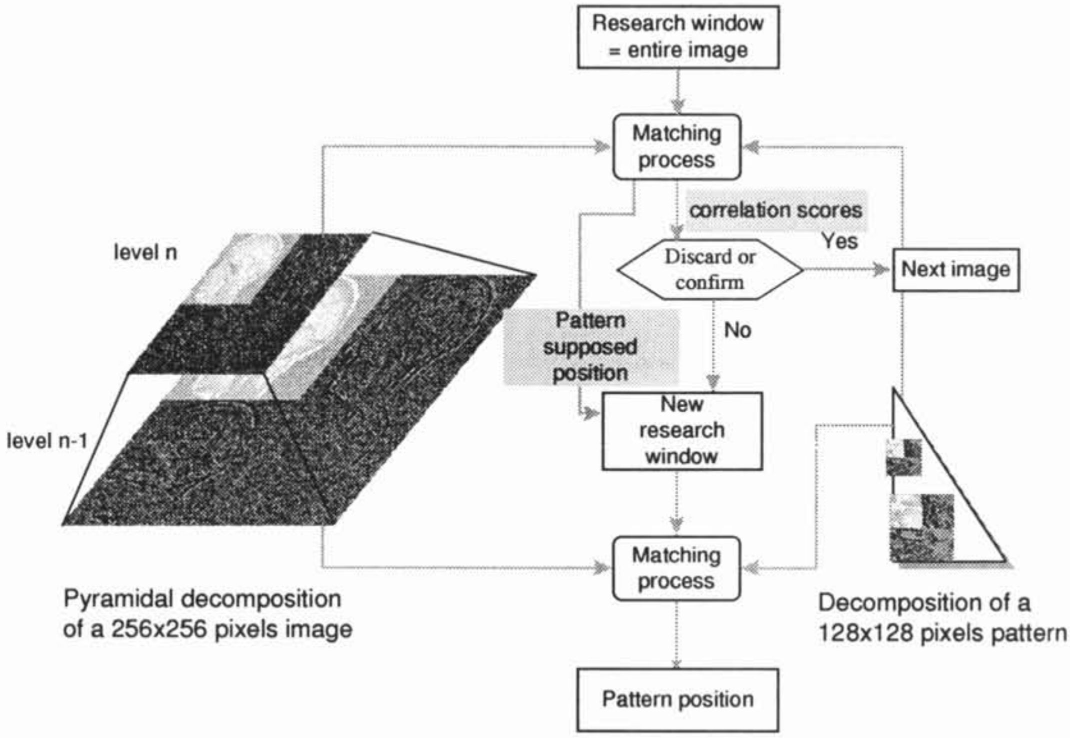

Figure 3: Multiresolution research approach

Stage	Pattern size	Image size	Candidate image number (nc)	SPARC		SYMPHONIE	
				for 1 image	for nc images	for 1 image	for nc images
Stage 1	8×8	64×64	90	4 s	360 s	4 ms	360 ms
Stage 2	16×16	128×128	8	8 s	64 s	32 ms	256 ms
Stage 3	32×32	256×256	2	40 s	80 s	272 ms	544 ms
Stage 4	64×64	512×512	1	150 s	150 s	920 ms	920 ms

Table 2: Partial times for the retrieval of a 64×64 pixels pattern in a 100 images (512×512 pixels) database