Pattern Recognition and Document Processing

Combining Resolution and Granularity for Pattern Recognition

P. Bottoni, L. Cinque, S. Levialdi, L. Lombardi+, P. Mussio

Dipartimento di Scienze dell'Informazione, Universita' di Roma,
Via Salaria 113, 00198 Roma, Italy
+ Dipartimento di Informatica e Sistemistica, Universita' di Pavia,
Via Abbiategrasso 209, 27100 Pavia, Italy

Abstract. This paper combines two approaches for shape recognition to reduce the memory needed to store the salient features of the image and the time to describe images. A first approach uses pyramidal multiresolution to provide successively condensed representations of the information of input images. The second approach describes shapes by typical characteristic patterns at different levels of abstraction granularity, corresponding to different levels of detail at which an image can be studied. Since different patterns of interest are visible in a restricted range of resolution levels, the combination of the two allows the recognition of such patterns to be performed at the minimum sufficient level, thus saving costly description of irrelevant details.

1. Introduction

This paper addresses the problem of defining the efficient strategies for recognizing objects in digital images, based on the structural description of shape.
Shape description can be considered as a first step in pictorial object recognition. It has been shown that humans can often recognize object in line drawings as quickly and accurately as in color photographs [1]. This means that recognition can be essentially achieved from the contour shape and therefore it is often unnecessary to provide the gray-level intensity, color, texture, etc. of the objects in an image. This is advantageous for a computer vision system since processing contour alone reduces complexity and increases efficiency. If recognition is to be based on contours a method is required to represent the contour shape.
A multiresolution approach, to shape detection and coding, has been proposed, based on pyramids [2]. Pyramids provide successively condensed representations of information present in input images [3]. This approach has proved useful in supporting planning strategies for edge and contour detection: a coarse-to-fine search is performed at a coarse scale to quickly locate potentially interesting regions; further search is restricted to the detected regions at a finer resolution to verify the presence of specific features.
A shortcoming of this approach is that the contour code may be influenced by noise and object rotation.
On the other hand, an approach was proposed to describe shapes by detecting typical characteristic patterns at different levels of granularity in abstraction, corresponding to the different levels of detail at which an image can be studied [4]. Descriptions at each level are based on those at the previous level, providing more and more synthetic

descriptions of the same image at a given resolution level. A drawback of this approach is that several irrelevant details present in the image have however to be taken into consideration and described.
This paper combines the two techniques to derive descriptions at different levels of granularity for each level of resolution in the pyramid, so reducing the mentioned drawbacks. Since different patterns of interest are visible in a restricted range of resolution levels, the combination of the two techniques allows the search for such patterns to be performed at the minimum sufficient level, thus saving the costly description of irrelevant details. Pyramidal resolution is taken as analogous of optical resolution, granularity levels as analogous of the different levels of detail at which an image can be studied. Their combination allows a reduction of the amount of memory needed to store salient features of the image and a reduction of time in the retrieval of images with particular features. In fact, since different patterns of interest are visible in a certain range of resolution levels, the search for such patterns can be performed at the minimum sufficient level, thus saving the costly description of irrelevant details.
Applications can be envisaged in the field of retrieval of images presenting some patterns from image databases.

2. Shape characterization through resolution levels

An image may be subdivided into subimages in a number of different ways: in particular, if quadrants are considered and this operation is recursively performed until a single pixel is reached, then a set of hierarchical levels is created each one having a power of 2 number of elements: these levels correspond to versions of the image at different resolution (thus using multiresolution systems) [2]. A pyramidal structure is particularly suited to implement this approach, each plane of the pyramid contains an image at a different resolution: higher at the base and coarser as raising towards the apex.
In order to extract the different images corresponding to the pyramidal levels we apply a technique known as the Gaussian pyramid [5], and to characterize the object shape present in the image at the base of the pyramid we will use a labeling approach.
Our technique is based on the analogy with a heat diffusion process on the object by computing numerical values for every pixel of the object in order to obtain the object contour temperature. Initially, the contour of the object is heated at time t_0 at a given energy value g; at each iteration, new neighbouring pixels will be contaminated provided they lie in the interior of the object.
After a time t_f, the simulated diffusion process is halted, and the temperature on the contour is measured; the temperature value of the contour pixels can be used to characterise a shape-related code [6]. The contour elements that have local maxima correspond to local convexities while those having local minima correspond to local concavities: we may see that the values distributed along the contour depend on the shape local configuration.
As shown in Fig.1, we can obtain a contour shape characterisation, at each level of resolution, based on the results of the diffusion process: these characterizations are given in terms of labels attached to different contour segments.
We define an alphabet of contour labels CL = {w, c, s, x, y} where, the symbols w, c, s, x, y denote, respectively the geometrical primitives very concave, concave, straight, convex and very convex.
Labeling provides the working image for deriving descriptions of different granularity.

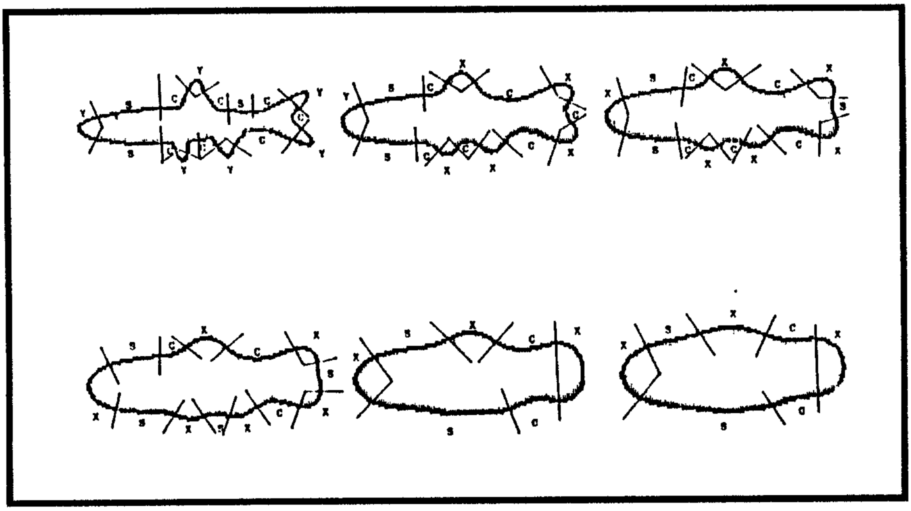

Fig.1 A picture resuming the results of contour shape characterisation on six pyramidal layers

3. Shape characterization through granularity levels

An *image description* is a vector (tuple) of values (the features of the image) assumed by a set of variables - the attributes considered useful to classify the image at hand [7]. A global description is a tuple consisting of a name denoting the image and of the set of features an observer exploits to describe the image as a whole. Some features are names of the image characteristic patterns, i.e. of those shapes that an observer identifies as relevant for the image interpretation. In turn, characteristic patterns are also described by tuples , comprising the name of the characteristic pattern and the names of its sub-patterns - whenever the characteristic pattern can be decomposed by the relation "part-of". This decomposition process is rooted in the descriptions of non decomposable, elementary patterns.

As a result, the image is described at different levels of *abstraction granularity* by a set of tuples named *description scheme* (ds). The most abstract description is the global one, the less abstract (also called the *primal description*) is the set of tuples describing the elementary patterns. Descriptions at two subsequent levels of granularity are related by the part-of relation.

This process is automated exploiting the concept of *attributed symbol* and *Conditional Attributed Rewriting System* (CARW). The finite set of names used in image description is assumed as an alphabet and the attributes/features as attributes/values of the symbols. Descriptions become attributed symbols and description schemes are organized as strings of attributed symbols. CARWs formalize the relation "part-of" by mapping a description d_{g-1} (at level of granularity $_{g-1}$) into d_g at the next level.

A knowledge manager interpreter KM exploits a suitable coding of CARWs to synthesize less abstract descriptions into more abstract ones so as to organize them into description schemes.

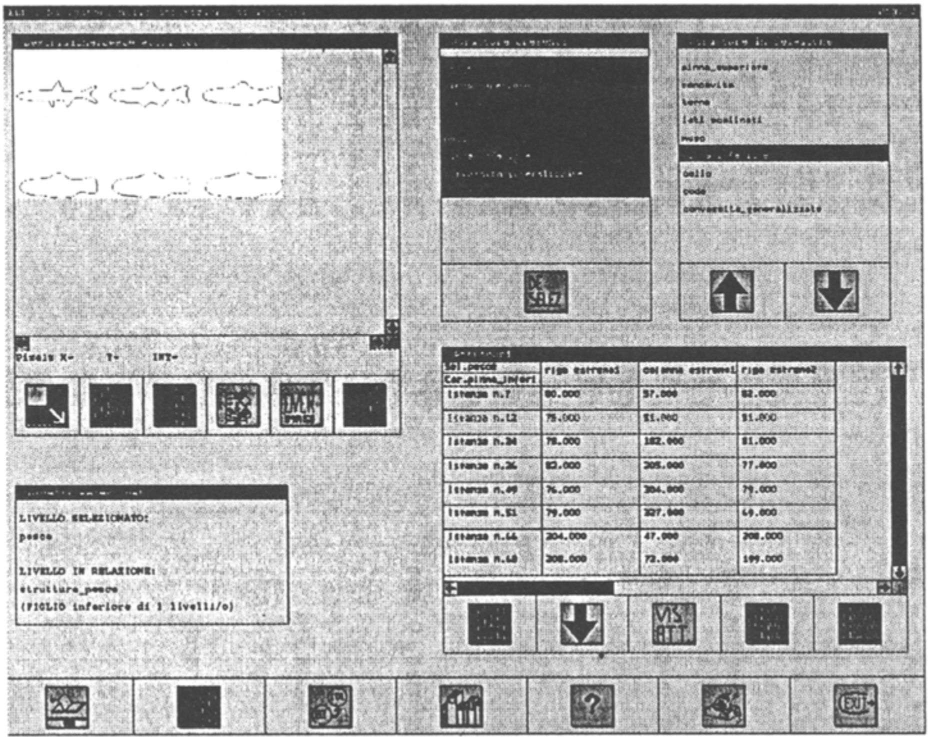

Fig. 2 A screen dump resuming the results of contour description

For instance, binary images can be described by this process assuming as elementary patterns those black pixels in which the contour changes its direction locally. Each such pixel is described by a code resuming the pixel 8-neighborhood information, and by its coordinates. In [8] it is demonstrated that this primal description allows the reconstruction of the original binary image; in the same paper an algorithm is also described which maps an arbitrary binary image into its primal description.

Figure 2 shows the result of applying the interpreter KM to the binary images obtained at the different resolution levels. Five different windows appear in the figure, displaying different views of the obtained description schemes. The graphical window in the upper left corner contains the silhouettes of the recognised features at different levels of resolution. In particular, the nose, upper and lower fins, and tails in the original image and at the first level of resolution appear.

4. Combining Resolution and Granularity

Fig.3 shows how the technique for the characterisation of a shape at different resolution and at different levels of granularity are combined .

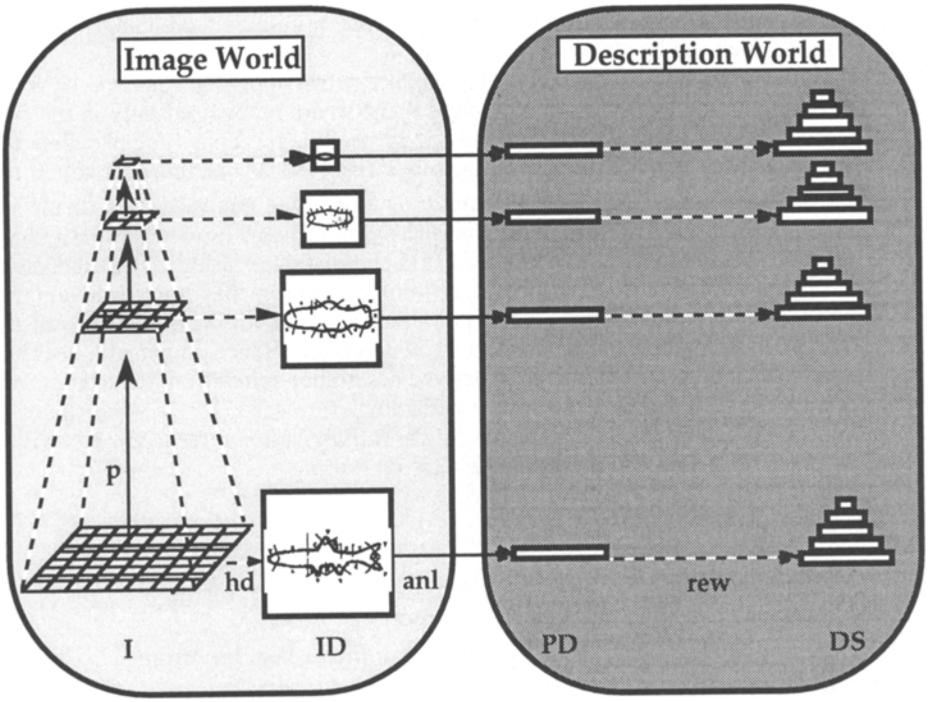

Fig.3 Combining the two techniques

On the left hand side the Image World schematizes the multiresolution approach. Each level of the pyramid contains an image at a different resolution (the base with the original image $i \in I$ and progressively other resolution levels until the apex by the repeated application of the function *p*), the heat diffusion functions (*hd*) map every image in *I* into a contour labeled shape $id \in ID$ as shown inside the squares. The right hand side represents the Description World where, for each resolution a primal description $pd \in PD$ is obtained by applying a function *anl* to the corresponding *id*.
The functions *p* and *hd* are computed by programs implementing the algorithms mentioned in section 2 and the function *rew* by the interpreter KM.
The function *anl* can be computed by extending the pre-existing tools for translating a binary image into a primal description. More precisely, each image *id* is translated into a binary image, memorizing in a table T the label and coordinates of each contour pixel in *id*. Then the algorithm to extract a primal description *pd'* is applied to the binary image. Finally for each element in *pd'* (i.e. a description of a single pixel belonging to the contour where it changes direction) the coordinates (x,y) are extracted. This couple is used as an index to find in T the label *l* that was associated with the described pixel in the image *id*. *l* is added as a fourth feature to the element at hand. The overall primal description *pd* is thus obtained.

5. Conclusion

We have presented a method which, by combining a multiresolution approach for contour shape labelling with a method for characterizing shapes at different granularities, allows one to establish which is the minimum convenient level of resolution at which can be classified.
For the particular example considered in this paper, when applying a same CARW for fish description and recognition it was found that fish are recognised only in the first four levels of resolution (the first four shown in Fig.1). This suggests that the detection of fish can be obtained on the image resolved at the fourth level in the pyramid, thus reducing the amount of data to be stored and analysed by a factor k^3, where k is the reduction factor. Note that the pyramid has a blurring effect, which reduces the details in the contour of shapes. This elimination of details is advantageous up to the point in which the blurring eliminates details that are irrelevant for recognition. The size of the primal description depends on both the dimensions of the image and on the roughness of the contour to be described. Hence, in general, even the sizes of the primal description and of the derived description scheme are reduced as well as the complexity of the strategies for their deduction.
This work has been partially supported by the Italian National Reasearch Council, Progetto Finalizzato "Sistemi Informatici e Calcolo Parallelo".

6. References

1. I.Biederman and G. Ju "Surface versus edge-based determinants of visual recognition", *Cognitive Psychol.* 20, 1988, pp. 38-64.
2. V.Cantoni, S.Levialdi, "Multiprocessor computing for images", *Proc. of the IEEE*, vol.76, n.8, pp.959-969, 1988.
3. A.Rosenfeld, "Some useful properties of pyramids", in *Multiresolution Image Processing and Analysis*, A. Rosenfeld (Ed.), Springer-Verlag, 1984, pp. 2-5
4. P. Bottoni, P.Mussio, M. Protti, "Metareasoning in the determination of image interpretation strategies",*PRL* , vol.15, n.2, 1993, pp.177,190.
5. P.J.Burt, "Smart sensing within a pyramid vision machine", *Proc. of the IEEE*, vol.76, n.8, pp.1006-1015, 1988.
6. L.Cinque, L.Lombardi, "Shape description and recognition by a multiresolution approach", to appear in *Image and Vision Computing*
7. R.M.Haralick, L.G.Shapiro, "Glossary of Computer Vision Terms", *Pattern Recognition*, vol.24, n.1, pp.69-93, 1991
8. P. Mussio, D. Merelli, M. Padula."An approach to the Definition, Description and Extraction of structures in Binary Digital Images"*CVGIP*, vol. 31, pp. 19-49, New York, Academic Press, 1985.