KNOWLEDGE AND PERCEPTION: A CONTRIBUTION TO DETECT THE SHAPE OF OBJECTS IN NOISY IMAGES.

Sandra DENASI, Giorgio QUAGLIA
and Salvatore RIZZO

Istituto Elettrotecnico Nazionale "Galileo Ferraris"
Strada delle Cacce 91, I-10135 Torino, Italy

RÉSUME

Cette comunication décrit une approche qui évalue la forme d'un objet dans une image avec bruit par un ensemble de procès perceptifs et cognitifs. Chaque point de l'image est marqué avec un ensemble d'attributs qui dérive des caractéristiques intrinsèques de l'image, de la connaissance a-priori sur l'objet et des raisonnements sur ces attributs. L'application d'un groupe de règles, dérivées des lois de la perception et de la connaissance, formule hypothèses sur la forme en modifiant ces attributs jusqu'á la réalisation d'une marque stable pour chaque point sur l'appartenance ou moins á l'objet.

SUMMARY

This paper describes an approach that evaluates the shape of objects in noisy images by means of a set of perceptive and cognitive processes. Each pixel is labeled with a set of attributes that come from the intrinsic properties of the image, from the a-priori knowledge about the object and from reasonings about these attributes. The application of a set of rules, derived from the perception laws and from the a-priori knowledge, builds hypotheses of shape updating these attributes until the achievement of a stable label for each pixel as belonging or not to the object.

1. Introduction

The accurate estimation of the boundaries of complicated structures in noisy images is still a challenging task, mainly if the classification and the evaluation of these structures depend on the boundary accuracy.

Medicine, astronomy and meteorology are typical fields of applications where the automatic interpretation of data is performed through the estimation of the boundaries. The conventional approach is local edge detection using edge templates, followed by edge linking to obtain curve segments, and finally gap filling to obtain boundaries. The effect of the noise is usually deleterious with loss of accuracy and reliability. Inefficient is the reduction of the noise effects by filtering, because the benefits are always combined with a greater inaccuracy in the boundary estimation, due to the edge blur and to the contrast reduction.

Several approaches have been suggested in order to overcome this drawback. They make use of edge-preserving algorithms for the smoothing of the noise [1], of heuristic search algorithms [2,3,4], and of relaxation techniques [5].

This paper proposes an innovative approach, which, like a human observer, classifies the pixels of the image as belonging to the object or to the background, according to the information available about the object and according to the laws of perception. Each pixel is labeled with a collection of attributes, which contribute to its classification. Each attribute depends on the local luminance properties, on the relationships with surrounding pixels and on the dynamic behavior in time.

2. Human behavior in the boundary detection

The human interpretation process of noisy images to detect the accurate boundary of an object or of a complex structure is a sophisticated task, that is performed in a sequence of stages that reach an accurate evaluation updating an initial primal perception. This updating process consists of mental processes operating at three different levels of detail, as it is shown in figure 1.

The goal of the first level is the localization of the most perceptible structure. This task is performed looking at the most evident area in the image (attraction area). *Most evident* means that this area, compared to its surroundings, is more noticeable because of its luminance, of the presence of textures, or of particular business of edges and regions.

The processes in the next level are addressed to single out an area of interest, if only a part of the structure must be analyzed. This task is essentially driven by the experience of the observer.

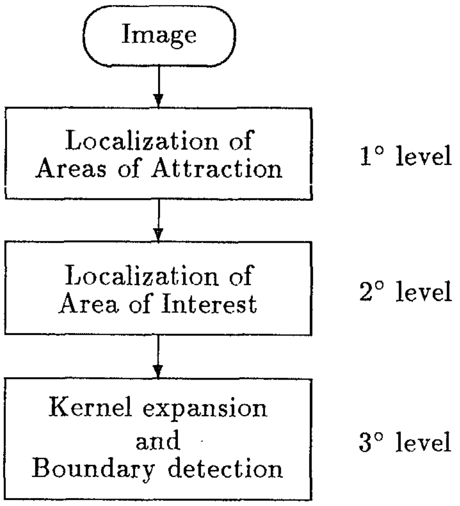

Figure 1: Levels of detail.

At the last level the observer analyzes in detail the approximated boundary detected at previous levels, in order to refine its position. Since the human perception places the edges where it finds the maximum luminance gradient and, if the object is in motion or is changing its shape, it can neglect the noise effects and can track the boundary in the confused areas [6], he localizes the edges firstly where they are more evident. Then, using sure and supposed boundary segments, he formulates initial hypotheses of pixel classification in the ambiguous areas. Each pixel is candidated either as *object* or as *non-object* according to the a-priori knowledge about the image formation, the expected shape and the dynamic behavior of the object, that is the knowledge about the domain. The hypotheses must be consistent with the context. The incongruities impose the reconsideration of the initial hypotheses and this task of cyclic formulation and verification continues until a definite status is reached. Once this status is reached, it is very difficult for the observer to introduce changes in its interpretation.

3. The proposed approach

The concept that directs the development of the proposed approach is based on labeling each pixel with a set of attributes that come from the intrinsic properties of the image, from the a-priori knowledge of the domain and from reasonings about subsets of these attributes. The application of a set of rules, derived from the perception laws and from the a-priori knowledge, can update these attributes until the achievement of a stable label for each pixel as belonging or not to the object.

The system architecture is structured in a common data base and in a set of knowledge sources, which interact each other under the supervision of a control strategy module (figure 2). The common data base is structured as a *blackboard* where are stored the image pixels, the pixel attributes and the primal classification labels. The knowledge sources are modules that provide knowledge to the control module and are activated by the current content of the blackboard [7].

The human behavior suggests that the control strategy should be performed with a set of tasks that interact each other and formulate a succession of hypotheses. These tasks can be arranged in three groups: each one is directed to the accomplishment of those goals, that reflect the human approach at the different levels, that are described in the previous section.

The attraction area is localized analyzing the image according to its global luminance distribution. The luminance histogram is examined for the occurrence of a bimodal distribution. This bimodality is an indication of the presence of an object that is distinguishable from the background and suggests a threshold value to segment the image to produce an approximated localization of the object. If the histogram is not clearly bimodal, more sophisticated criteria must be activated to localize the object. They will be based on the edges or homogeneous regions distribution. The horizontal and vertical business, that is the count of edges or homogeneous regions along these directions, will be evaluated in order to limit the search area. The a-priori knowledge about the object (i.e. expected position, expected shape, expected behavior) can also be exploited for this restriction.

Then, an area of interest is pointed out whenever a single part of a complex structure must be analyzed. The a-priori knowledge of this structure and particular information can be used at this purpose.

A last group of processes is devoted to the accurate evaluation of the contour pixels. They produce pixel labels taking into account:

- the luminance value, the position within the image and related to the surrounding structures, the amplitude and the direction of the gradient, the edge and region segmentation, and other features reflecting the intrinsic attributes of the image;
- the perceptive laws of similarity, of proximity, of good continuation, of closure, and of common

fate;

- the domain knowledge about the object shape, about its behavior and its peculiarities.

The reasoning on the labels of each pixel, compared to the surrounding labels, creates hypotheses of upgrading the area of interest, hypotheses of edges and edges aggregation, hypotheses of region clusters, and finally hypotheses of pixel classification as *object* or *non-object*. These hypotheses are included in the blackboard. The process stops when a single hypothesis, representing the result of the shape detection, is created.

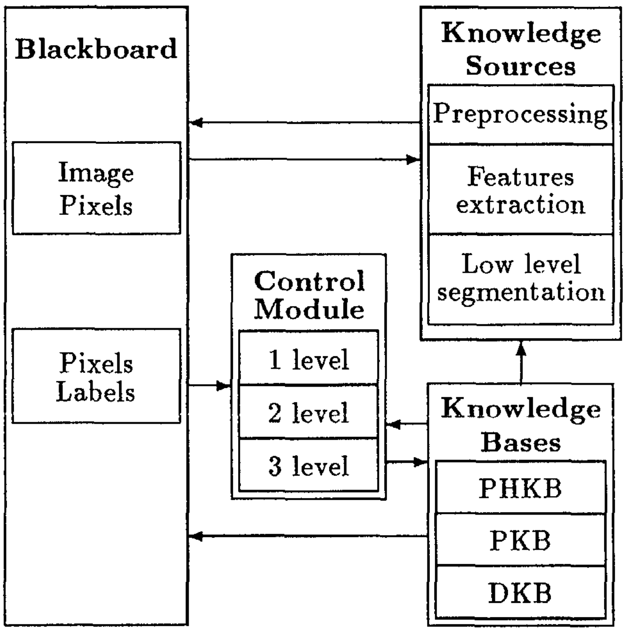

Figure 2: System architecture.

4. The application

The proposed strategy is under investigation on a sequence of sixteen scintigraphic images of the heart. This clinical test permits the visualization of the distribution of a radioactive tracer in the heart chambers. The resulting images present a low signal-to-noise ratio which is intrinsic to the physical process of the image acquisition.

The strategy developed for the localization of the left ventricle (LV) is organized in three successive stages:

- localization of the *attraction area* associated with the cardiac region;
- localization of the *area of interest*, that is the area corresponding to the LV, and identification of an initial set of pixels which have a great probability to belong to the ventricle; these pixels will be the kernel for the aggregation of all the ventricle points;
- refinement and increase of this kernel until the estimation of the complete shape.

The domain knowledge is included in the knowledge data base DKB, and consists of the knowledge about the morphology (anatomical structures, shapes, spatial relationships), about the physiology (dynamic behavior), and about the pathology (diseases, anomalous movements). While the information about the image formation, about the noise, and about the synchronization between the cardiac cycle and the image acquisition system are included in the physical knowledge data base PHKB.

First stage

The localization of the most noticeable area of the image is a task that is usually performed by perceptive faculties, which make possible the individuation of an approximated shape of an object on a background. In our application, the attraction area is identified with the cardiac shadow which can be easily perceived analyzing dynamically the sequence. An indication of the movement can be evaluated by the images of amplitude and phase of the first component of the Fourier transform of the corresponding points of each image of the sequence. Since the heart is the only moving structure, the corresponding amplitude image is adequate to locate its points.

Second stage

The restriction of the attraction area to the area of interest is a task that can be carried on exploiting the information available about the morphology of the structure. Since the movement of the ventricles is in phase opposition respect to the atria movement, the area corresponding to the left ventricle is easily determined as the right part of the ventricles area, as pointed out in figure 3.

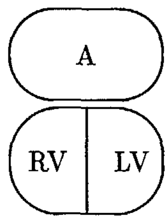

Figure 3: Localization of Area of Interest.

Pixels belonging to each one of the previous areas are labeled as A if lie in the atria area or as V if lie in the ventricular area. Furthermore, the spatial relationships collaborate to add other labels as R (right) or L (left) according to the relative position of each pixel within the ventricle area.

A small set of pixels is then selected in the center of the area of interest according to the rule of selecting those points which lie near the center of gravity and are characterized by maximum values of luminance.

Third stage

Starting from this kernel of points in the LV, a process of aggregation with the surrounding points is undertaken until the accurate and complete accretion of the LV. This aggregation is performed only if each pixel underlies to the constraints attributed to the perceptive laws or to the medical knowledges.

In order to ease this aggregation task, it is possible to exercise a control strategy, which pushes, on one hand, the implementation of those heuristic techniques which firstly grant global suggestions (i.e. Fourier transform, histograms, intra-pixel relationships in space and in time) and successively apply the local techniques (i.e. gradients, luminances, thresholds), and, on the other hand, rules the pixel labeling beginning from those segments of the boundary which have a higher evidence.

In the boundary structure of the LV three approximated segments can be singled out as sketched in figure 4: the segment which bounds the external side of LV (P), the segment which splits the two ventricles near the septum area (S), and the last one lying in the confused division between the LV and the area of atria and great vessels (B).

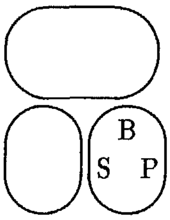

Figure 4: Parts of the left ventricle.

Each segment is distinguishable by peculiar features and must be analyzed selecting the most suitable heuristics. Since P is the segment that better separates the ventricle from the background, it should be easier to detect a reliable boundary. According to the perceptive laws, the boundary follows those points where the luminance gradient is maximum. Furthermore, the gradient orientation in these points must be consistent with the orientation of the neighboring points, since the LV is a convex, continuous and close structure. These consistency constraints, that can be forced to the neighboring points (in space within the image or in time between images), are a strong way to recover the effect of noise.

These primal hypotheses of aggregation in the less blurred part give rise to some hypotheses about the ventricle shape. These indications, together with the rules coming from the a-priori knowledge, help the boundary evaluation in S and B segments. Where the ventricle is confused with the surrounding structures, so that it is difficult to distinguish them, only the knowledge of an expert physician can supply the necessary suggestions for the most reliable boundary detection.

5. Conclusion

This paper proposes a new approach to the evaluation of the shape of objects or structures in noisy images. The main drawbacks introduced by the noise on the boundary evaluation process can be overcome using AI techniques in addition to the conventional ones. In this way, the capabilities of the perception laws and of the a-priori knowledge about the structure under investigation are exploited in order to direct the boundary detection in the confused areas. Finally, the proposed architecture allows a good generalization because of it re-usability, that can be performed changing the knowledge sources.

References

[1] J. Biemond and J.J. Gerbrands, "An edge-preserving recursive noise-smoothing algorithm for image data," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 9, pp. 622–627, October 1979.

[2] H. Elliot, D.B. Cooper, F.S. Cohen, and P.F. Symosek, "Implementation, interpretation, and analysis of a suboptimal boundary finding algorithm," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 4, pp. 167–182, March 1982.

[3] D.B. Cooper and F.P. Sung, "Multiple-window parallel adaptive boundary finding in computer vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 5, pp. 299–316, May 1983.

[4] S. Denasi and G. Quaglia, "A spatio-temporal segmentation technique for the evaluation of the movement in dynamic angioscintigraphic images," in *Proc. 11th GRETSI*, (Nice), June 1987.

[5] S.W. Zucker, R.A. Hummel, and A. Rosenfeld, "An application of relaxation labeling to line and curve enhancement," *IEEE Trans. on Computers*, vol. 26, pp. 394–403, April 1977.

[6] K. Koffka, *Principles of Gestalt Psychology.* New York: Harcourt, Brace & World, 1963.

[7] E. Rich, *Artificial intelligence.* McGraw-Hill, 1987.