International Journal of Pure and Applied Mathematics
Volume 61 No. 1 2010, 11-20

DETECTING AMOURPHOUSLY SHAPED OBJECTS IN A NOISY IMAGING ENVIRONMENT

Sándor Molnár[1] §, Péter Klinkó[2]
[1,2]Department of Informatics
Faculty of Mechanical Engineering
St. István University
Gödöllö, 2103, HUNGARY
[1]e-mail: molnar.sandor@gek.szie.hu
[2]e-mail: klinko.peter@gek.szie.hu

Abstract: In the present paper, we propose a method of detecting amorphously shaped objects in real-time ultrasonic images when high level of noise is present. In general, noisy imaging environment can lead to faulty object segmentation. Shape recognition algorithms can give ambiguous results even for detecting objects with permanent shapes. From prior knowledge, indexed object templates are introduced that are aimed to incorporate the most relevant object.

AMS Subject Classification: 68T10, 68U10
Key Words: shape detection, pattern recognition, segmentation

1. Introduction

Object recognition in image analysis is the most challenging task and still heavily studied especially in the computer vision discipline. The key to the successful object recognition is to detect certain features pertinent to the given object such as shape contour, texture or color that uniquely determine the presence and the location of the object. If no additional information is available about the detectable object other than its pure shape contour, edge detection analysis is widely used as the first step in extracting valuable shape information from the

Received: December 14, 2009
§Correspondence author

raw image. At first look, shape information is embodied in those regions where sudden changes can be observed in the image's intensity function. Wherever a sudden change is observed in the intensity profile, an edge is detected. The most common approach in edge detection algorithms is to convolve the raw image with some small-sized approximation of the derivative operator kernels followed by a threshold filtering. Depending on the approach the algorithms fall into two main groups. The most commonly used methods are based on either inspecting the first derivatives of the image intensity function (e.g. Sobel) or seeking the zero crossings in the second derivatives (Laplacian of Gaussian), see [6]. Although both approaches are equivalent, but methods based on inspecting the first derivatives are less sensitive to noise. One of the most preferred version of the first derivative approaches is developed by Canny [3], where some level of noise reduction and connectivity of neighboring pixels by threshold hysteresis are also taken into account. Another preferred convolution type of approach is the recently developed *SUSAN* algorithm [8] which does not require any type of inclusion of a derivative kernel, therefore the algorithm's sensitivity to noise is greatly reduced. In the presence of high background noise, only regions with high contrast can be successfully recovered while some parts are occluded or suppressed in the background. In noisy imaging environment, only by the means of edge filtering operators, we cannot successfully reconstruct the complete shape contour. The recovered shape has missing segments, and unnecessary artifacts can appear. In this case some interpolation is necessary for the missing line segments from the prior knowledge of the full shape in order to reconstruct the complete contour. Another disturbing issue is when the object lies in a more complex pattern structure, and only a fraction of the recovered edge pattern carries useful information about the single shape's contour to be extracted. Some of these issues were already addressed in early studies of pattern recognition by Hough [4] and others [5], [7], and a Hough transformation method was developed for detecting simple analytic curves (e.g. lines, circle, ellipse). In a later work by Ballard [2], the Hough transform was generalized for any non-analytic curves and composite shapes. Based on the object's contour, an R-table is constructed that contains information about the object's central position with respect to the image gradient direction at the boundary points. Each raw of the table corresponds to one gradient angle and consist of the object's center positions relative to the boundary point coordinates where the local gradient directions lie in the same interval. In contrast to the original approach of the Hough transform for simple shapes like circle or ellipse, the generalized R-table algorithm does not hold one to one correspondence between the object's center and the local gradient angle relative to the boundary

points. In the detection process for each pixel, the possible object positions are estimated and the corresponding accumulator array at the object position is incremented by the magnitude of the local gradient. The accumulator array can be further extended by considering that the object can be scaled by s. The resulted space is termed as Hough space in the literature [2], [1]. The actual object search is taking place in the Hough space by incrementing the accumulator array for all rotated or scaled object instances represented by the set of parameters. As the result of the accumulator approach, the object can be detected with good confidence even when some parts of the border are occulted [2] or the object is partially covered by overlapping [9]. If there is no prior knowledge about the shape to be detected, segmentation methods are widely used to extract the object's area from the actual image. The most often used segmentation methods in medical imaging are the so called level-set methods where the segmentation is based on a dynamical model that an initial curve (initial level-set) evolves in a force field description of the gradient image. To achieve segmentation, the algorithm and its variants require numerical integration of a first order pde $\partial_t\varphi + \mathbf{F}\nabla\varphi + v_N\left|\nabla\varphi\right| - v_K\left|\nabla\varphi\right|$ that pulls the initial curve towards the edges. $\mathbf{F}$ denotes the external force field, and v_N, v_K are the normal and curvature based speed functions. Although the method allows the segmentation of irregular shapes, but difficulties arise when the object lies in a complex pattern that cannot be simply excluded from the segmentation process. The object's internal pattern will also be segmented while the main object's outline is only partially recovered. The above mentioned issues clearly show that detecting an object with irregular shapes with no priori knowledge is rather cumbersome task especially when the object is considerably occluded. For processing large set of noisy images, we propose employing object templates as the first approximation to the irregular shape to be recovered. In the following sections, we first discuss the employed shape model followed by the shape detection algorithm, and then we present our results and future improvements of the method.

2. Formulation of the Model

In the previous section, we covered the main objectives of the shape recognition process. In the present study the shapes are irregular and cannot be simply reproduced by a similarity transformation of the single shape. In Figure 1, we illustrate the underlying concept of our approach for modeling amorphous shapes. Each instance of the shape is recovered from manual segmentation.

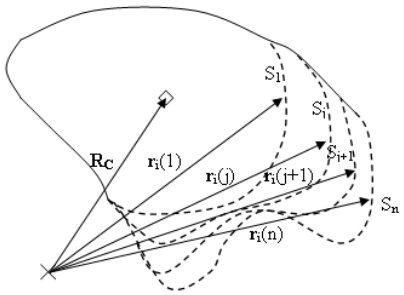


Figure 1: Formulation of the object templates used for modeling amorphous shapes. The set of single instances S_j are indexed allowing some parts of the shape to be continuously deformed throughout the reference points $\mathbf{r}_i(j)$ of the single shapes. $\mathbf{R}_C$ denotes the shape's central position.

The results of the subsequent segmentations are then collected into one group to form a set of object templates $S_j = \{\mathbf{r}_i(j)\}$. The templates and the reference points $\mathbf{r}_i(j)$ which compose the outline of the given shape S_j are both indexed. If necessary, the uniquely shaped regions of the shape sequence can be further deformed via the reference points or new shapes can be added to the existing template set later. The image pattern (e.g. outline of the shape) is represented by a collection of image points where the image intensity changes most rapidly (edge image). Therefore in our model instead of using the raw intensity function of the image as the descriptor of the pattern we use the gradient image. The smaller the gradient in magnitude at some edge points of the shape the less visually recoverable that part of the shape outline. When large amount of noise is present, some parts of the shape are barely visible and cannot be recovered by edge detection.

The object templates $S_j = \{\mathbf{r}_i(j)\}$ corresponding to the border of the shape form a closed domain Ω_j in the image domain I. We define the shape's binary masks $B_j(\mathbf{r})$ as follows

$$B_j(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \in \mathbf{\Omega}_j \subset I, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Let us denote the original and preprocessed image intensity functions by $\phi(\mathbf{r})$, $\varphi(\mathbf{r})$ and edge image by $\varphi_E(\mathbf{r})$. One choice of defining the edge image is to

threshold the magnitude of the gradient of $\varphi(\mathbf{r})$ by a control parameter α

$$\varphi_E(\mathbf{r}) = \begin{cases} 1 & |\nabla\varphi(\mathbf{r})| \geq \alpha \max\limits_I |\nabla\varphi(\mathbf{r}_k)|,\ 0 < \alpha < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

If high level of noise is present, the image gradient is taken after the noise reduction of the original image by convolving $\phi(\mathbf{r})$ with a Gaussian kernel $g(\mathbf{r})$:

$$\varphi(\mathbf{r}) = \phi(\mathbf{r}) * g(\mathbf{r}), \tag{2.3}$$
$$g(\mathbf{r}) = \frac{1}{2\pi\sigma^2} e^{-r^2/(2\sigma^2)}.$$

Note that the original defining space is now extended by setting periodic boundary conditions for $\mathbf{r} \notin I$. The gradient image spectrum is then obtained in the frequency space

$$F_D[\nabla\varphi(\mathbf{r})](\mathbf{k}) = F_D[\phi(\mathbf{r}) * \nabla\varphi(\mathbf{r})](\mathbf{k}),$$

or using complex notation $\varphi_G(\mathbf{r}) = \partial_x\varphi(\mathbf{r}) + i\partial_y\varphi(\mathbf{r})$

$$F_D[\varphi_G(\mathbf{r})](\mathbf{k}) = -\frac{1}{\sigma^2} F_D[\phi(\mathbf{r})](\mathbf{k})\, F_D[(x+iy)\, g(\mathbf{r})](\mathbf{k}). \tag{2.4}$$

The final result of the spectrum for the gradient image $\varphi_G(\mathbf{r})$

$$F_D[\varphi_G(\mathbf{r})](\mathbf{k}) = 2\pi \left(i\frac{k_x}{N_x} - \frac{k_y}{N_y} \right) F_D[\phi(\mathbf{r})](\mathbf{k})\, F_D[g(\mathbf{r})](\mathbf{k}), \tag{2.5}$$

where the subscript D corresponds to the DFT, and N_x and N_y are the image dimensions to directions x and y. The main advantage of the approach is that the convolution and the evaluation of the gradient image all together can be carried out in only one step in the frequency space. The complex gradient of the binary mask $b_j(\mathbf{r})$ can be also defined according to equation (2.5), but the smoothing function $g(\mathbf{r})$ only plays the role of making $B_j(\mathbf{r})$ differentiable.

The object's statistically best fit position is determined by the correlation function of the gradient edge images

$$\begin{aligned} C(\mathbf{r}) &= \langle b_j / \|b_j\|, \varphi_G \rangle \\ &= \frac{1}{\|b_j\|} \sum_k \left(b_j(\mathbf{u}_k) - \overline{b}_j \right)^* \left(\varphi_G(\mathbf{r} + \mathbf{u}_k) - \overline{\varphi}_G(\mathbf{r}) \right). \end{aligned} \tag{2.6}$$

Since only modifications to the original image template B_j make any change in the subsequent evaluation of $C(\mathbf{r})$, it is practical to normalize the correlation function $C(\mathbf{r})$ only for b_j. The image domain I is finite, and the periodic boundary conditions are required to ensure that $\int_0^{N_x} \int_0^{N_y} \partial_y\varphi(x,y)\, dydx = 0$ and

$\int_0^{N_y}\int_0^{N_x} \partial_x\varphi(x,y)\,dxdy = 0$. So we can further simplify equation (2.6) if the mean values $\overline{b}_j$ and $\overline{\varphi}_G$ are both set equal to zero

$$C(\mathbf{r}) = \frac{1}{\|b_j\|}\sum_k b_j^*(\mathbf{u}_k)\,\varphi_G(\mathbf{r}+\mathbf{u}_k). \tag{2.7}$$

The real part $\mathrm{Re}(C(\mathbf{r}))$ corresponds to the sum of the scalar products of the gradient vectors while the imaginary part $\mathrm{Im}(C(\mathbf{r}))$ comes from the sum of the vectorial products. From the correlation theorem, the spectrum of $C(\mathbf{r})$ is

$$F_D[C(\mathbf{r})](k) = \frac{1}{\|b_j\|}F_D^*[b_j(\mathbf{r})](\mathbf{k})\,F_D[\varphi_G(\mathbf{r})](\mathbf{k}). \tag{2.8}$$

The similarity between S_j and the image pattern is the measure of $c(\mathbf{r}) = |\mathrm{Re}[C(\mathbf{r})]|$ by our definition, and same shapes with opposite contrast are considered to be the same. The position $r_m = \max[c(\mathbf{r})]$ is the position where the shape template statistically best fits the image pattern. We must point out that a single shape template S_j can also be stretched or shrunk in both dimensions and rotated by θ, and therefore the correlation $C(\mathbf{r},j,\mathbf{a},\theta)$ must be recalculated for the transformed object template $S_j(\mathbf{a},\theta)$. As a result of scaling, the components of the complex gradients will be inversely scaled and the gradient phase will be shifted by θ. The complete object search is taking place in a closed domain of defining space for $C(\mathbf{r},j,\mathbf{a},\theta)$ by allocating a finite parameter set $\{\mathbf{a},\theta\}$ of interest. From the elements of $\{\mathrm{Re}[C(\mathbf{r},j,\mathbf{a},\theta)]\}$, we can arbitrarily construct a finite series c_k which always contains its maximum and minimum. Therefore $\max|\mathrm{Re}[C(\mathbf{r},j,\mathbf{a},\theta)]|$ always exists, and the convergency of the algorithm is guaranteed.

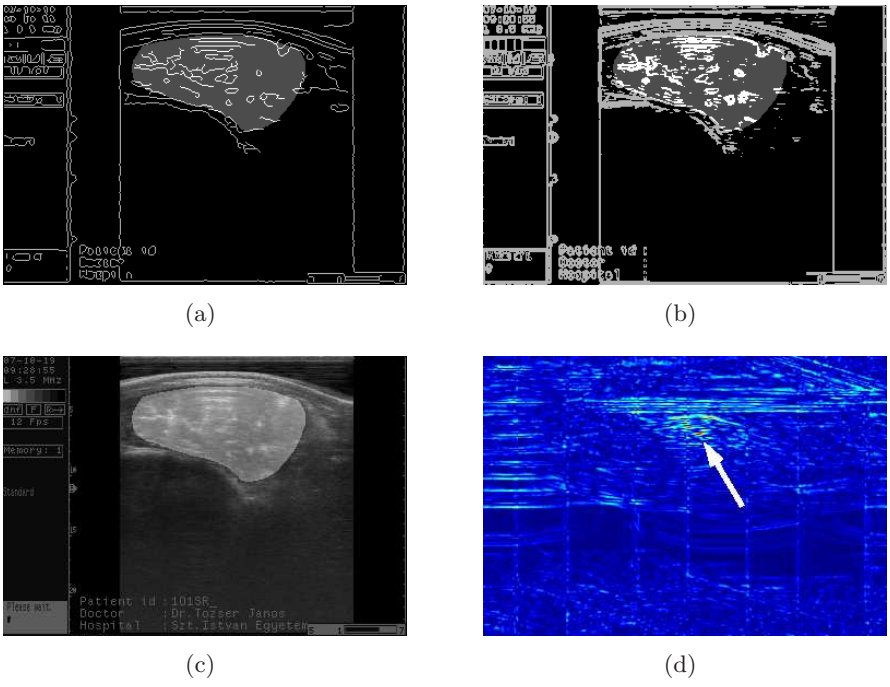
3. Results

Pattern or shapes are the indication of sudden changes in the gradient image compared to their surroundings. After calculating the gradient images φ_G and b_j with equation (2.5) for ϕ and B_j, then we proceed calculating the correlation function in equation (2.7) for a finite set of parameters of interest $\{\mathbf{a},\theta\}$. The parameter set is allocated by setting up a meshgrid in the domain formed by the maximum and minimum values of scaling and rotation parameters.

After sweeping through the set of parameters and shape templates, the position of the best fit template is found in the image space at $\mathbf{r}_m = \mathbf{R}_C$ where $c(\mathbf{r}_m,j_m,\mathbf{a}_m,\theta_m) = \max[c(\mathbf{r},j,\mathbf{a},\theta)]$. All calculations for equation (2.5)

and equation (2.8) were carried out in the frequency space with FFT that has significant speed advance over evaluating the above mentioned equations directly. The required computation time can be further reduced if the shape orientations are almost identical.

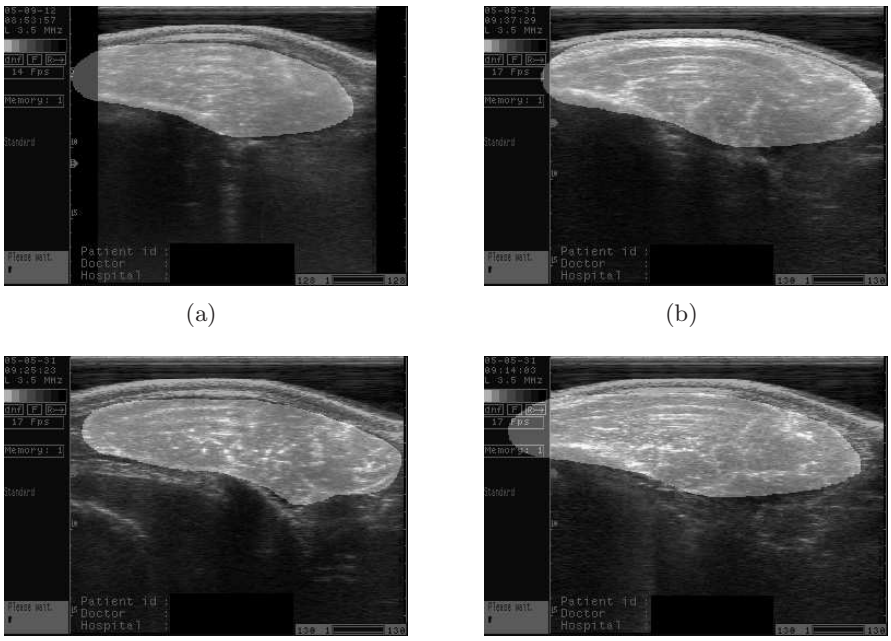
In Figure 2, we present our result for the object search of musculus longissimus dorsi in a real-time ultrasonic image. The object search was performed for the binary edge image $\varphi_G = \varphi_E$ defined in equation (2.2) with threshold



(a) (b)

(c) (d)

Figure 2: Shape detection of *musculus longissimus dorsi* from a real-time ultrasonic image (256×370, 8bit). The upper left and upper right images are edge images. On the upper left (a), the binary edge image is extracted with using the Canny edge filter. On the upper right (b), the edge image is simply obtained by thresholding the gradient image. The lower right (d) image shows the correlation $c(\mathbf{r}) = |\mathrm{Re}\,[C\,(\mathbf{r})]|$ in the image space. The solution for $c(\mathbf{R}_c) = \max |\mathrm{Re}\,[C\,(\mathbf{r})]|$ yields the position of the statistically best fit shape template (white arrow). The detected shape template is superimposed to the original image (c) and the edge images (a) and (b) at the best fit position.



(a) (b)

(c) (d)

Figure 3: More examples of detecting *musculus longissimus dorsi* from ultrasonic images (256×370, 8bit). In images (a), (b), and (d) the shape is successfully recovered. In image (d), the shape is also detected, but the result is not as good as in the other examples due to the fact that the shape outline is not well separable.

parameter $\alpha = 0.18$ (Figure 2(b)) and for a Canny edge image $\alpha = 0.15$ (Figure 2(a)). The original image (c) and the edge images (a) and (b) show that there are two huge gaps of the shape outline that cannot be recovered by the edge filtering on the left and on right side of the shape. The outline corresponding to these gaps are very faint and unidentified by edge filtering. These gaps with low contrast to their surrounding are partially comprised in the background noise and barely noticeable. It is also noticeable that inside the shape there is a complex pattern of edges that has to be ignored by the object detection. The image also contains letters, notes, and frames that further complicate the situation. The plot of the correlation $c(\mathbf{r})$ is presented in Figure 2(d) for the best fit template. As a result of having periodic boundary conditions allowed, $c(\mathbf{r})$ is plotted in the full image space. The position where $c(\mathbf{r})$ has its maximum value

is marked with a white arrow. In all images (a), (b), and (c) the detected shape is also indicated by overlaying the shape on the image at position $\mathbf{r}_m$. Almost identical results were found for both images (a) and (b), and even the central coordinates for the best fit shape position $\mathbf{R}_C = \mathbf{r}_m$ were nearly identical. The detected shape remarkably fits to the appropriate image area of interest in both cases. In Figure 3, more examples are presented for detecting the same object in several ultrasonic images. In Figure 3(d) gives an example when the shape outline cannot be recovered unambiguously, but the best fit shape still fits well.

4. Conclusion

In the present paper we discussed our method for detecting amorphously shaped objects in noisy images. The shapes are irregular and their outline is only partly recoverable from the background noise. For these situations, we proposed employing object templates that are uniquely shaped and can be further deformed via some reference points. As in common with the literature on the subject, we referred to the shape outline as the set of points in the image intensity function having the highest gradient in magnitude (edges). Our further goal is to recover the shape even when its domain is embedded in a complex pattern structure. Therefore the proposed method takes advantage of correlating the gradient edge image with the gradient edge mask obtained from the shape templates. Additional linear transformations are also taken into account (scaling and rotation) for dealing with perspective view. The method works well even with a few templates and were able to recover the shape of interest with good confidence regardless how the preprocessed gradient edge image was obtained. The method can be improved by applying a model of stretching tensor for modeling the applied deformations to the shape templates.

Acknowledgments

The authors would like to thank the support received from OTKA grant 68187.

References

[1] A.S. Aguado, E.Montiel, M.S. Nixon, *Pattern Recognition*, **35**, No. 5 (2002), 1083.

[2] D.H. Ballard, *Pattern Recognition*, **13**, No. 2 (1981), 111.

[3] J. Canny, *IEEE Trans. Pattern Anal. Mach. Intell.*, **8**, No. 6 (1986), 679.

[4] P.V.C. Hough, A method and means for recognizing complex patterns, *US Patent No. 3 069 654* (1962).

[5] C. Kimme, D.H. Ballard, J. Sklansky, *Communications of the ACM*, **18**, No. 2 (1975), 120.

[6] D. Marr, E. Hildnth, *Proc. Royal Soc. London Bulletin*, **B207** (1980), 187.

[7] S.D. Shapiro, *Computer Graphics Image Processing*, **8**, No. 2 (1978), 219.

[8] S.M. Smith, J.M. Brady, *International Journal of Computer Vision*, **23**, No. 1 (1997), 45.

[9] D.M. Tsai, *Image and Vision Computing*, **15**, No. 12 (1997), 877.