Dijkstra Algorithm for Accelerating Deformable Template Matching

Liu Zhuofu [1,2] **Liao Zhenpeng** [1,2] **Sang Enfang** [1]

[1] College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, 150001, P. R. China
[2] Institute of Engineering Mechanics, China Earthquake Administration, Harbin, 150001, P. R. China

Abstract

An improved method for deformable template-based object detection is presented. The energy function optimization problem is tackled using the Dijkstra Algorithm (DA) by adding a shape restriction and improving the noise-resistance ability so that robustness and high detection efficiency are acquired. This approach has been successfully tested on the underwater object detection. The experimental results show that when DA is used the speedup is significant while accuracy is maintained.

Keywords: Deformable template, Dijkstra Algorithm, underwater object detection

1. Introduction

Model-based shape matching is a well-known problem in image recognition. Early research in this area concentrated mainly on rigid shape matching, where the matched shapes were obtained by applying simple transformations such as translation, rotation, scaling, and affine transformation to the model template, which can be recovered using correlation-based matching or the Hough transform.

Because of the rigidness of the above approaches, their utility is limited. In most applications, an exact geometric model of the object is not available because of the variability in the imaging process and inherent within-class variabilities. Deformable template matching is more versatile and flexible in dealing with the deficiencies of rigid shape matching. Rubber template[1] and elastic template [2] are regarded as the original form of deformable template theories, which can deform itself to fit the obvious image features. But the common problem in deformable template detection is controlling the computational complexity of object detection and model fitting.

Snake model is one of the most widely used deformable template algorithms [3]. But it contains shortages such as sensitive to original position, poor noise resistance and occlusion resistance. Based on Snake model, the energy function of this paper is redefined to fully use the global information. So the global structures are gotten when dealing with local discontinuing and deformable images. The algorithm is insensitive to deformation, noise and contract variance. Efficiency, accuracy and stability are also its advantages.

2. Deformable template

Deformable templates can be partitioned into two classes. The free-form templates can represent any arbitrary shape as long as some general regularization constraints (continuity, smoothness, etc.) are satisfied. On the other hand, parametric deformable templates are capable of encoding a specific characteristic shape and its variation. The shape can be characterized by a parametric formula or using a deformation modes. Parametric deformable templates are used to detect underwater objects in this paper.

2.1. Parametric representation of deformable template

In parametric deformable templates, the deformation of templates can be represented by a set of parameters, T_0 is the original template, T_d is the deformable template, $d = (\xi, \theta, s, b)$ is deformable parameters [4-6]:

$$T_d(x, y) = T_0\left(S \cdot \left(\left((x, y) + (D^x(x, y), D^y(x, y))\right) \cdot R_\Theta(x, y)\right) + P\right) \quad (1)$$

where $(D^x(x, y), D^y(x, y))$ is the deformation factor $R_\Theta(x, y)$ is the rotation angle factor S is the scale factor $P = (P^x, P^y)$ is the displacement factor. Without any loss of generality, it is assumed that the template is drawn on a unit square $\Omega = [0,1]^2$. The points in the square are mapped by the function $(x, y) +$ $(D^x(x, y), D^y(x, y))$ where the deformation functions $D^x(x, y)$ and $D^y(x, y)$ are continuous and satisfy the Dirichlet and Neumann boundary conditions:

$$D^x(0, y) \equiv D^x(1, y) \equiv D^y(x, 0) \equiv D^y(x, 1) \equiv 0 \quad (2)$$

The space of deformation functions is spanned by the following orthogonal bases:

$$e_{nm}^{x}(x,y)=\left(2\sin(\pi nx)\cos(\pi my),0\right) \tag{3}$$

$$e_{nm}^{y}(x,y)=\left(0,2\cos(\pi mx)\sin(\pi ny)\right) \tag{4}$$

where *m*, *n*=1, 2,…. Specifically, the deformation function is chosen as follows:

$$D(x,y)=(D^{x}(x,y),D^{y}(x,y))=\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{\zeta_{mn}^{x}e_{mn}^{x}+\zeta_{mn}^{y}e_{mn}^{y}}{\lambda_{mn}} \tag{5}$$

where $\lambda_{mn}=\alpha\pi^{2}(n^{2}+m^{2}),\ m,n=1,2,\cdots$ are the normalizing constants. The parameters, $\zeta=\{(\zeta_{mn}^{x},\zeta_{mn}^{y}),m,n=1,2,\cdots\}$, are the projections of the deformation function on the orthogonal basis. The illustration of deformable templates is shown in Fig. 1[6].

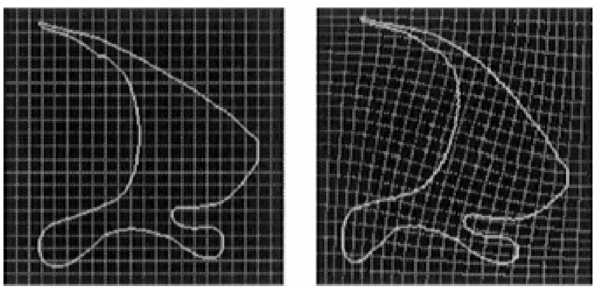


Fig.1 Illustration of deformable templates

2.2. Energy function

The basic idea of object detection based on deformable template is the energy function composed of internal energy and external energy: (i) the amount of deformation of the template is described by internal energy, and (ii) the gray gradient belongs to the content of external function. Thus the problem of object detection is changed to the optimization of energy function.

As the energy function of Snake model is used in the field of optical images and the definition is complex, it cannot be applied to acoustic image directly. According to the characteristic of acoustic images, the function is redefined:

(1) Internal energy function:

$$E_{\text{int}}=\sum_{m}\sum_{n}\frac{\sqrt{\zeta_{mn}^{x\,2}+\zeta_{mn}^{y\,2}}}{2\sigma} \tag{6}$$

where ζ_{mn}^{x} and ζ_{mn}^{y} are deformable parameters, σ is the variance of deformable parameters. The reason why the internal function is defined in this way is that θ,s,b is the rigid transformation, and the final shape is only determined by $\zeta=\{\zeta_{mn}^{x},\zeta_{mn}^{y},m=1,2,\cdots,M,n=1,2,\cdots,N\}$, the greater the values of ζ and M,N are, the bigger the deformation is. The physical meaning of internal function is the discrepancy between the deformable template and the original template.

(2) External energy function:

$$E_{ext}=\frac{1}{n}\sum_{x,y}|\,gradf(x,y)\,|\cos(gradf(x,y),e) \tag{7}$$

where $e=i\cos\varphi+j\sin\varphi$ is the normal of contour of deformable template, $gradf(x,y)$ is the gradient at (x,y), $(gradf(x,y),e)$ is the angle between $gradf(x,y)$ and e, n is the number of pixels on the template. The external function contains the direction as well as the value of the gradient. This makes the edge location more accurate. The physical meaning of external energy is the trend of deformation between original template and deformable template. The bigger the value is, the more dissimilar the two templates are.

(3) Total energy function:

$$E_{total}=\omega_{1}E_{\text{int}}+\omega_{2}E_{ext} \tag{8}$$

where ω_1 and ω_2 ($\omega_1+\omega_2=1$) are the weights assigned to the internal energy and external energy, respectively, which adjust the proportion of the internal energy and external energy.

3. Optimization algorithm

The energy function to be optimized in (8) is not unimodal. In fact, it is a very complex function with many local extrema[5] over the deformation parameter space. The minima for this function can, in principle, be obtained by using the Monte Carlo Relaxing Algorithm (MCRA), the Metropolis Algorithm (MA) or the Stochastic Diffusion Algorithm (SDA). In all such algorithms, the minima are obtained by constructing an ergodic Markov chain whose stationary distribution has support over only the modes of the a posteriori probability density function. However these approaches achieve the optimal solution at the cost of excessive computing. Dijkstra Algorithm (DA), one of the graph algorithms, is employed to quickly find the good solutions. At the coarse stage, as the dominative part is the deformation of the edge, ω_1 is greater than ω_2, that is internal energy is the major part of the total energy and the threshold is set to a big value; At the fine stage, finer energy fields are used because more accurate localization is desired. The impact of external energy is increased so that ω_2 is greater than ω_1 and the threshold is set to a small value.

The algorithm for deformable template-based object detection is summarized as follows:

Step I:

Initiate the energy threshold in the coarse level T_c and the energy threshold in the fine level T_f , the maximum number of iteration C , deformation parameters $d = (\xi, \theta, s, b)$ and assign the weights ω_1 and ω_2 ;

Step II:

Perform the coarse level detection, the coarse detection process is:

Loop: Calculate the total energy function E_{total} , and adjust deformation parameters $d = (\xi, \theta, s, b)$:

If $E_{total} \le T_c$

Go to step III;

Elseif iterating number $> C$

Report "detection failed";

Else

Use the Dijkstra algorithm to update the deformation parameters. Go to Loop;

Endif

Step III:

Perform the fine level detection, and the procedure is the same as described in step II

If $E_{total} \le T_f$

Output the detected object and stop the program

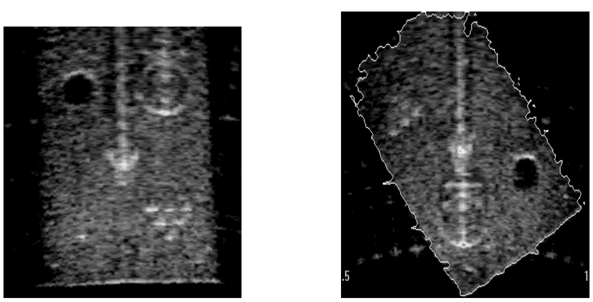
Else

Continue the detection procedure

Endif

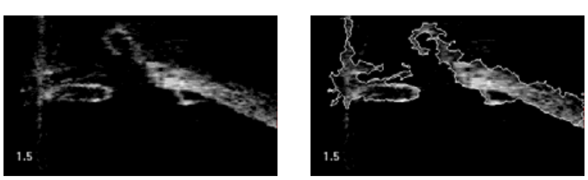
4. Experimental results

The proposed deformable template-based object detection scheme has been applied to different objects of underwater acoustic images: from Fig.2 to Fig.4.



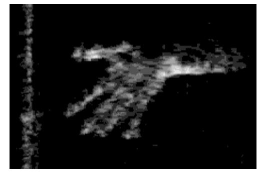
(a) Original image (b) Detection of the rotated image

Fig.2 Detection of the rigid object

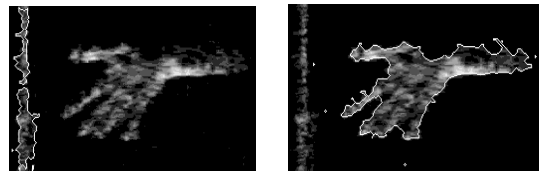


(a) Original image (b) Result of detection

Fig.3 Detection of complex objects (hook and ring)



(a) Original image



(b) Steel tube detection (c) Hand detection

Fig.4 Detection of the expected object

Fig.2 shows the detection of single object, where Fig.2 (a) is the acoustic image of steel plate, Fig.2(b) is the detection of the rotated image which gray level and contrast have been changed. From Fig.2 it can be seen that the proposed method is robust to rotation and contrast.

Fig.3 shows the detection of complex object, where Fig.3 (a) is the original image with two objects (hook and ring) and the aim is to detect the two objects from background. Fig.3 (b) is the detection that satisfied the demand.

Two objects (hand and steel tube) are shown in Fig.4 (a) and the goal is to detect them from background separately. The two objects in Fig.4 (a) are very near in spatial position, and one of the objects (hand) is irregular and obviously concave. The results of image detection are shown in Fig.4 (b) and Fig.4(c), which demonstrate the good performance of the proposed scheme in image detection.

Table 1 Comparison of efficiency of different optimization algorithms

Detected objects	MCRA (s)	MA (s)	SDA (s)	DA(s)
Rotated steel plate	15.8	21.3	10.4	7.5
Hook and ring	78.9	—	51.6	12.8
Steel tube	65.2	73.8	41.5	9.1
Hand	—	—	66.6	20.9

Deformable template-based detection is equal to the optimization of the energy function, which is a multi-extrema, nonlinear and multi-variance function. The optimization problem was tackled using Monte Carlo Relaxing Algorithm (MCRA), the Metropolis Algorithm (MA) or the Stochastic Diffusion

Algorithm (SDA) in past literatures. To compare the execution time, Dijkstra Algorithm (DA) is compared with these methods to detect interested objects shown in Fig.2 to Fig.4. Table 1 is the comparison result where "—" represents the failed detection or non-optimal result.

Table1 shows that MCRA and MA require a good initialization of the template and are sensitive to deformable parameters. Otherwise a suboptimal solution is obtained. SDA belongs to the techniques of relaxing stochastic algorithms and can obtain the optimal solutions if the successive iteration is enough. But this will take much execution time. The shortest path and parallel idea have been used in Dijkstra algorithm to improve executing efficiency.

5. Conclusion

A Dijkstra algorithm-based deformable template approach is presented. With definitely physical meaning, the modified energy function simplifies the original equation and accelerates the calculation. The shape of the deformable objects is defined by the length of the line detection and the angle at the vertices.

Dijkstra algorithm is used as the optimization algorithm to calculate the energy function. The results of real underwater object detection show that the method is robust to noise and deformation and fit for the detection of complex objects.

6. References

[1] B. Widrow, "The rubber mask technique," *Pattern Recognition*, pp175-211, 1973.

[2] M. Fischler, R. Elschanger, "The representation and matching of pictorial structure," *IEEE Trans. Computer*, pp 57-92, 1973.

[3] A. Bimbo, P. Pala, "Visual image retrieval by elastic matching of user sketches," *IEEE Trans. Pattern Anal. Mach. Intell*, pp 121-132, 1997.

[4] A K. Jain, D. Zongker, "Representation and recognition of handwritten digits using deformable templates," *IEEE Trans. Pattern Anal. Mach. Intell*, pp 1386- 1390, 1997.

[5] Y. Amit, U. Grenandar, M. Piccioni, "Structural image restoration through deformable template," *Journal American Statistical Assn*, pp 376-378, . 1991.

[6] M. Kass, A. Witkin, D. Ferzopoulos, "Snakes: active contour model," *Proc. First International Conference Computer Vision*, pp 259-268, 1987.

7. Acknowledge

This work was supported in part through a Post-doctorate Fund of Heilongjiang Province, Science and Technology plan project of Heilongjiang Province GZ04A126 and Harbin Engineering University Science Fund HEUFT0408.