A Robust Active Contour Model for Natural Scene Contour Extraction with Automatic Thresholding

Kian Peng, Ngoi[1] and Jiancheng, Jia[2]

[1] Defence Science Organisation, 20 Science Park Drive, Singapore 0511
[2] Nanyang Technological University, Nanyang Avenue, Singapore 2263

Abstract. An active contour model is proposed for contour extraction of objects in complex natural scenes. Our model is formulated in an analytical framework that consists of a colour contrast metric, an illumination parameter and a blurring scale estimated using hermite polynomials. A distinct advantage of this framework is that it allows for automatic selection of thresholds under conditions of uneven illumination and image blur. The active contour is also initialised by a single point of maximum colour contrast between the object and background. The model has been applied to synthetic images and natural scenes and shown to perform well.

1 Introduction

Shape analysis is very important in computer vision and has been widely used for recognition and matching purposes in many applications. Important information about the shape of the object can be provided by its contour. Active contour models have been successfully applied to the problem of contour extraction and shape analysis since its introduction by Kass et. al. [4] as snakes. These active contour models [1,11] are attracted to image features such as lines and edges, whereas internal forces impose a smoothness constraint. However, existing models are not well adapted for object segmentation in a complex image scene as they are easily distracted by uneven illumination, image blur, texture and noise. Several noteworthy variants were therefore proposed to allow active contour models to handle a complex scene. One class of active contour models [14] uses *a priori* knowledge about the expected shape of the features to guide the contour extraction process. When no knowledge of the expected object geometry is available, Lai et. al. [6] proposed a global contour model based on a regenerative shape matrix obtained by shape training. Region-based strategies to guide the active contours to the object boundaries were also proposed by Ronfard [10]. To distinguish the target object in a scene comprising of multiple objects, a local Gaussian law based on textural characteristics of objects [2] was introduced.

However, there are still several unresolved issues in applying active contour models to a complex natural scene. One issue is that current active contour models have not considered the effects of blurring. For example, Dubuisson et. al. [3] assumed that moving objects in an outdoor scene are not blurred. The reason is that blurring corrupt salient features like edges, lines and image contrast which is used

to guide the active contour to the object's boundary. Blurring also makes the location the exact edge's position difficult. Unfortunately image blur cannot be prevented in many applications. In a scene consisting of multiple objects, the limited depth of field of the camera's lens will cause some objects to be out of focus. Objects can also be blurred by motion when lighting conditions do not permit a fast shutter speed.

The performance of current active contour models also depends on the initial contour position and the proper selection of internal parameters and thresholds [13]. The common experience among users is that if the initial position and thresholds are not chosen properly, the active contour may be trapped in a local minima resulting in a sub-optimal solution. The selection of internal parameters and thresholds are done heuristically through several times of trial and error. Thus, for a complex natural scene, the active contour has to be initialised near the object's boundary and this requires human interaction and knowledge of the object's shape. To overcome the problem of initial position, Lai et. al. [6] had proposed the use of Hough Transform to initialise the active contour. However, computations and processing time are increased. Vieren [12] used a "border snake" to initialise and enclosed the object as it entered the scene by its periphery.

In this work, an active contour model robust to uneven illumination and image blur has been proposed to extract object contours in complex natural scenes. Our model is formulated in an analytical framework consisting of a colour contrast metric, a blurring parameter estimated using hermite polynomials and an illumination parameter. The distinct advantage of our model over current models is its robustness to an arbitrary complex natural scene achieved by the automatic threshold selection. Our active contour model is initialised by a single point in the object of interest with no *a priori* knowledge of the shape. However, the point of initialisation must lie in the region with maximum colour contrast between the object and background. It is argued that locating a point with maximum contrast is computationally more efficient than initialising an entire contour near the object's boundary.

2 Model Formulation

In this section, the proposed automatic threshold framework will be presented. Details of the other aspects of the model can be found in [8].

2.1 Colour Contrast

The logarithmic colour intensity has been chosen to determine the minimum threshold value. The purpose is to provide a high gain at the low and central portion of the camera's response and decreasing gain at the positive extremes. This is consistent with our human eye which is sensitive to low and normal lighting but saturates when the lighting is bright. The colour contrast between a point of colour intensity (r,g,b) and a reference point with colour intensity (r_o,g_o,b_o) is defined as

$$C = \ln^2[(1+r)/(1+r_o)] + \ln^2[(1+g)/(1+g_o)] + \ln^2[(1+b)/(1+b_o)] \quad (1)$$

2.2 Automatic Threshold Selection

The threshold of the active contour has to be adaptively determined because of uncertainties introduced by noise, uneven illumination and image blur. Setting the correct threshold will ensure that the active contour remain at the desired edges of the object. For real images, we characterise edges by a Gaussian model with a scale σ_b. This scale parameter is related to our human experience of blurness. A high value of σ_b will result in a very blurred edge while a sharp edge has a small value. The Gaussian edge model used is

$$I(x,y) = L_e + \frac{C}{2}\mathrm{erf}\,[(x\cos\theta + y\sin\theta - d)/\sigma_b] \quad (2)$$

where

L_e	: *the mean value of the edge's intensity*
C	: *the colour contrast of the edge in (1)*
θ	: *the orientation of the edge*
d	: *perpendicular distance from the centre of window function*
erf	: *error function*

The Gaussian blur which is one of the most commonly encountered description of blur has been assumed. It has been shown that the blur due to camera defocuses can well be approximated by a 2-D Gaussian kernel [9]. The case of Gaussian blurring is also easy to tackle mathematically. In this respect, the Hermite polynomials are chosen as a basis since it is orthogonal to the Gaussian window [7].

An input image *f(x,y)* can be decomposed into the sum of several images using a window function *w(x,y)* given as

$$f(x,y) = \frac{1}{h(x,y)} \sum_{(p,q)\in S} f(x,y) \bullet w(x-p, y-q) \quad (3)$$

where *h(x,y)* is the periodic weighting function

$$h(x,y) = \sum_{(p,q)\in S} w(x-p, y-q) \quad (4)$$

for all sampling position *(p,q)* in sampling lattice S. The windowed image is then approximated by a polynomial with orthonormal the basis functions φ given as

$$w(x-p, y-q)[f(x,y) - \Sigma\Sigma f_{m,n-m}(p,q)\varphi_{m,n-m}(x-p, y-q)] = 0 \quad (5)$$

If the window function is chosen to be Gaussian with a window spread of σ, the coefficients of the polynomial can be evaluated as

$$f_{m,n-m}(p,q) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)a_{m,n-m}(x-p, y-q)\,dxdy \quad (6)$$

where

$$a_{m,n-m}(x,y) = \frac{1}{\sqrt{2^n m!n!}} \frac{1}{\pi\sigma^2} H_m(\frac{x}{\sigma}) H_{n-m}(\frac{y}{\sigma}) \exp(\frac{-(x^2+y^2)}{\sigma^2})$$

$H_n(x)$ is the Hermite polynomial of degree n in x.

It has been shown [5,7] that the analysis function $a_{m,n-m}$ of order n is the n^{th} order derivative of the Gaussian i.e.

$$a_n(x) = \frac{1}{\sqrt{2^n n!}} d^n/d(\frac{x}{\sigma})^n \, [\frac{1}{\sigma\sqrt{\pi}} \exp(-x^2/\sigma^2)] \tag{7}$$

From (6) and (7), the hermite coefficients of a blurred edge with a scale of σ_b up to the third order can be written as

$$\begin{aligned}
f_0 &= L_e + (C/2)\mathrm{erf}[(d/\sigma)/\sqrt{1+(\sigma_b/\sigma)^2}\,] \\
f_1 &= [1/\sqrt{1+(\sigma_b/\sigma)^2}\,](C/\sqrt{2\pi}\,)\exp(-(d/\sigma)^2/(1+(\sigma_b/\sigma)^2)) \\
f_2 &= [1/\sqrt{1+(\sigma_b/\sigma)^2}\,]^2(C/\sqrt{2\pi}\,)[(d/\sigma)/\sqrt{1+(\sigma_b/\sigma)^2}\,]\exp(-(d/\sigma)^2/(1+(\sigma_b/\sigma)^2)) \\
f_3 &= [1/\sqrt{1+(\sigma_b/\sigma)^2}\,]^3(C/\sqrt{2\pi}\,)(\tfrac{1}{\sqrt{6}})[(2(d/\sigma)/1+(\sigma_b/\sigma)^2)-1]\exp(-(d/\sigma)^2/(1+(\sigma_b/\sigma)^2))
\end{aligned} \tag{8}$$

We derived the edge parameters from the hermite coefficients as

$$\sigma_b/\sigma = [(2e_2^2/e_1^2 \pm \sqrt{6}\, e_3/e_1) - 1]^{1/2} \tag{9}$$

$$d/\sigma = e_2/e_1(2e_2^2/e_1^2 \pm \sqrt{6}\, e_3/e_1)^{-1} \tag{10}$$

$$C = e_1\sqrt{2\pi}\,(2e_2^2/e_1^2 \pm \sqrt{6}\, e_3/e_1)^{-1/2}\exp(e_2^2/e_1^2(2e_2^2/e_1^2 \pm \sqrt{6}\, e_3/e_1)^{-1}) \tag{11}$$

where the n^{th} order local energy is $e_n^2 = \Sigma f_n^2$

From (9), there are two possible solutions. The first solution known as the near solution

$$\sigma_b/\sigma = [(2e_2^2/e_1^2 + \sqrt{6}\, e_3/e_1) - 1]^{1/2} \tag{12}$$

is valid when the centre of the window function is near the edge i.e.

$$(d/\sigma)^2/(1+(\sigma_b/\sigma)^2) \le 1/2 \tag{13}$$

and the far solution

$$\sigma_b/\sigma = [(2e_2^2/e_1^2 - \sqrt{6}\, e_3/e_1) - 1]^{1/2} \tag{14}$$

is valid for all other cases.

The threshold value for our Gaussian edge model occurs when the rate of change of gradient is maximum. If the window is sufficiently close to the edge point such that (12) is satisfied, an analytical expression to select the threshold automatically can be written as

$$\begin{aligned} thld &= |C| \exp(-d/\sigma_b)^2 \\ &= \sqrt{2\pi}\sqrt{1+(\sigma_b/\sigma)^2}\ \exp[-1/(2(\sigma_b^2/\sigma^2+1))]\, e_1 \end{aligned} \tag{15}$$

The parameter e_1 in (15) is the energy of the first order hermite polynomials and is obtained by convoluting the Gaussian mask with the image. Physically, e_1 can be interpreted as an illumination measure. In other words, a strong illumination increases the value of e_1 while a weak illumination reduces e_1 correspondingly. The value of e_1 therefore acts as a scaling agent of the threshold when illumination changes. For a constant illumination, the threshold of the active contour also increases with image blur. Setting the threshold under constant illumination is known as scale thresholding. It is observed in (15) that when the ratio of the image blur and the window function (σ_b/σ) is greater than 1, the threshold (*thld*) increases almost linearly with blurring.

3 Model Evaluation

In this section, the automatic threshold framework proposed in the previous section will be evaluated using synthetically generated images. The performance between our model and current active contour models using fixed thresholds is also compared.

3.1 Effects of Uncorrelated Errors on Image Blur estimates

The uncorrelated errors in the measured energies of e_1, e_2 and e_3 occur mainly due to image noise, quantisation of filter or polynomial coefficients . For small uncorrelated errors in energies, it is possible to express the error in the estimate of image blur as

$$\begin{aligned} \Delta\sigma_b &= (\partial\sigma_b/\partial e_1)\Delta e_1 + (\partial\sigma_b/\partial e_2)\Delta e_2 + (\partial\sigma_b/\partial e_3)\Delta e_3 \\ &= [\sigma^2/2\sigma_b][2(e_2^2/e_1^2) + \sqrt{6}\, e_3/e_1]^{-2}[(4e_2^2/e_1^2 + \sqrt{6}\, e_3/e_1^2)\Delta e_1 - (4e_2/e_1^2)\Delta e_2 - (\sqrt{6}\, e_3/e_1)\Delta e_3] \end{aligned} \tag{16}$$

Figure 1 shows the error in the estimation of the edge's scale for an error of 5% in e_1, e_2 *and* e_3 plotted by the analytical expression in (16). From Figure 1, the error is minimum when the window spread is approximately equal to the image blur ($\sigma_b/\sigma = 1$). After that, the error increases rapidly as the ratio of σ_b/σ becomes smaller. This is due to the large spread of the window function resulting in information redundancy. Estimation accuracy is poor because the edge's

information forms only a small portion of the overall data. However if the spread of the window function is smaller than the image blur ($\sigma_b/\sigma > 1$) estimation accuracy is still more reliable as compared to ($\sigma_b/\sigma < 1$). For a small spread of the window function ($\sigma_b/\sigma > 1$), all the information obtained still belongs to the edge. This shows that incomplete information has greater reliability than redundant information. From Figure 1, we set the following bound to ensure an accuracy of 86%

$$0.5 \leq (\sigma_b/\sigma) \leq 2.0 \tag{17}$$

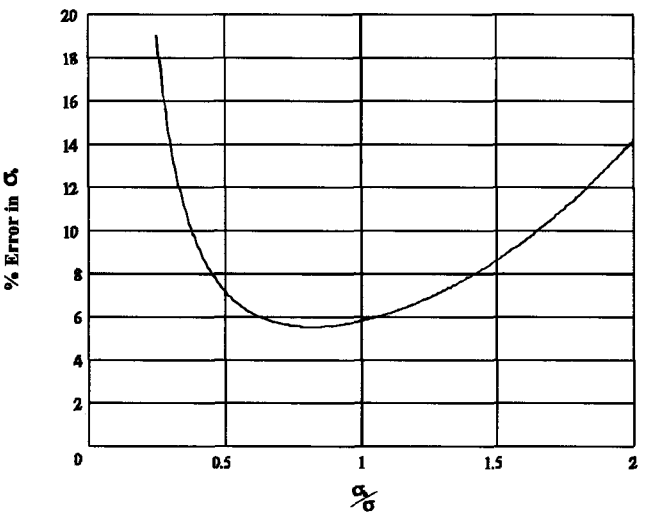

Fig. 1. : Error in the estimation of the image blur for $C = 5, d = 0.25, \sigma = 8$

3.2 Minimum Threshold

It can be seen from Figure 1 that when $\sigma_b/\sigma < 0.5$, the estimate of the image blur and the threshold of the active contour is unreliable. This happens when the object's boundary is not within the region of interest of the active contour. In this instance, a minimum threshold value ($thld_{min}$) is set to allow the active contour to move towards the object's boundary.

The contrast of a colour (r,g,b) from a reference colour (r_o,g_o,b_o) for a maximum colour deviation of $\{\Delta r, \Delta g, \Delta b\}$ and assuming that $r,g,b \gg 1$, can be expressed as

$$\text{thld}_{\min} = \ln^2(1 + \Delta r/r_o) + \ln^2(1 + \Delta g/g_o) + \ln^2(1 + \Delta b/b_o) \tag{18}$$

If $\max\{|\Delta r/r_o|, |\Delta g/g_o|, |\Delta b/b_o|\} < 1$, expanding the above equation gives

$$\begin{aligned} thld_{min} \approx (\Delta r/r_o)^2 - (\Delta r/r_o)^3 + 0.25(\Delta r/r_o)^4 + (\Delta g/g_o)^2 - (\Delta g/g_o)^3 \\ +0.25(\Delta g/g_o)^4 + (\Delta b/b_o)^2 - (\Delta b/b_o)^3 + 0.25(\Delta b/b_o)^4 \end{aligned} \quad (19)$$

If $|\Delta r/r_o|, |\Delta g/g_o|, |\Delta b/b_o|$ is sufficiently small, the minimum threshold ($thld_{min}$) required for the proper working of the active contour can be obtained from (19) as

$$thld_{min} = (\Delta r/r_o)^2 + (\Delta g/g_o)^2 + (\Delta b/b_o)^2 \quad (20)$$

The above equation serves as a guide to determine the minimum threshold value of the active contour in our model in terms of the maximum colour intensity fluctuations allowable within the object. When the ratio of σ_b/σ is less than 0.5, the minimum threshold criterion in (20) is used. Scale thresholding in (15) is used for the rest of the image edge's scale $(0.5 \le \sigma_b/\sigma \le 2.0)$.

3.3 Performance Evaluation on Image Blur

Simulations of synthetic images had been performed to evaluate the performance of the automatic thresholding framework on image blur. Firstly, a reference threshold must be obtained. This reference threshold is the value the active contour has to be set to remain at the object's boundary as the object blurs. Using the reference threshold, the performance of three thresholding methods (i.e. fixed thresholding, scale thresholding and the proposed automatic thresholding) under blurring is evaluated. The performances of current active contour models are similar to fixed thresholding because their thresholds do not change according to image blur.

The reference threshold is obtained using the following procedure. A step edge with a colour contrast of 4.84 is generated at a predetermined location. The image is subsequently blurred by a series of Gaussian kernels with a known scale σ_b. Each time the step edge is blurred by a Gaussian kernel, the contrast at the predetermined location is measured. For example, if the step edge at location 400 in the x-axis have an initial contrast of zero, a Gaussian blurring scale of 5 will give a contrast of approximately 0.5 at location 400 (see Figure 2). Setting the active contour threshold below 0.5 yields an edge location below 400 and a threshold higher than 0.5 results in an edge location greater than 400. Therefore, the active contour will contract if the threshold is less than 0.5 and expands outward if the threshold is greater than 0.5.

The performance of automatic thresholding, scale thresholding and fixed thresholding under image blurring is also shown on Figure 2. A window spread σ of 12 is used for the experiment. Scale thresholding is less reliable for small image blur $(0 < \sigma_b < 6)$ but gives better results for image blur from 6 to 24 in comparison to the reference threshold. The most accurate estimate occurs at an image blur of approximately 12 $(\sigma_b/\sigma = 1)$. This is consistent with the theoretical results in Figure 1. However, the performance of a using a fixed threshold for an active contour is better at smaller image blur and becomes progressively unreliable at higher image scale. The propose automatic thresholding framework combines the advantages of scale thresholding and fixed thresholding. In the framework, fixed thresholding is

used when the ratio of σ_b/σ is between 0.0 and 0.5 while scale thresholding is used when the ratio is between 0.5 and 2.0.

Figure 3 gives a comparison of the performance under blurring between the our model which uses automatic thresholding and current models with fixed thresholds under image blur. A window function with spread of 12 and fixed threshold of 0.5 is again used. It can be seen from Figure 3 that fixed thresholding gives an error of 6 pixels for an image scale of 24. This is compared to an error of only 2 pixels for our automatic thresholding scheme. Therefore our proposed model is more robust to the effects of blurring than current active contour models.

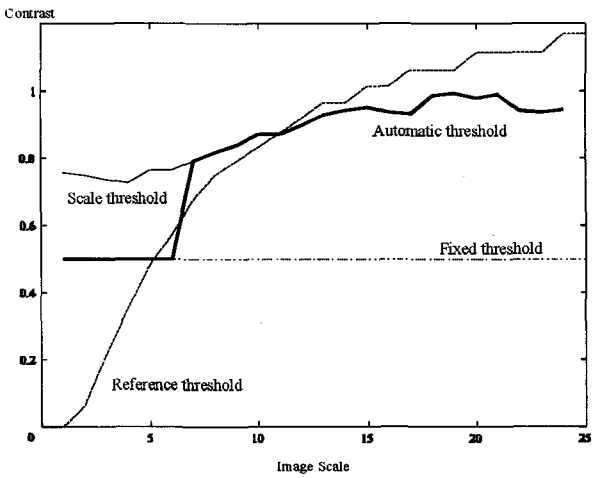

Fig. 2. Threshold of Active Contour Model with $\sigma = 12$

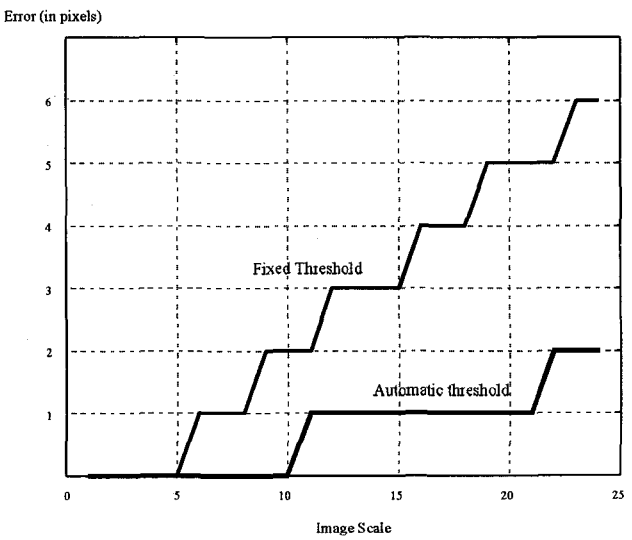

Fig. 3. Performance of automatic threshold against a fixed threshold

4. Experimental Results

Our active contour model has been applied to natural scenes. Figure 4 shows a complex image of a reptile on a relatively uniform background. The active contour is initialised at the body of the reptile as indicated by a cross. The high contrast between the reptile (brown hues) and the blue sky allowed the reptile's profile to be extracted accurately. Figure 5 shows a monochromatic image of three rocks with large intensity variations. The active contours are again initialised at the centre of the rocks and the intensity difference between the rocks and their background allowed reliable contours to be extracted. Figure 6 shows the contour of a vehicle extracted from a complex background with illumination and colour fluctuations.

For the contour extraction of a blurred vehicle in Figure 7, some rocks on the river bank have very similar colour and intensity as the vehicle. Our active contour cannot differentiate between the object and the background at these points without any *a priori* knowledge of the car's shape. Therefore, two volcanoes [4] denoted by crosses have been manually placed to push back the contours. It is observed that a reliable contour is still extracted despite the effects of image blur.

Figure 8 shows the performance of our model for a complex object (cottage) in a cluttered background. Noticed the different textures of the roof, windows, door and the background trees. Furthermore, shadows also reduce the colour contrast between some parts of the roof and the trees. The overall performance of our active contour model is still reliable despite some distortions at the roof and chimney. The active contour is initialised near the door of the cottage. It is noted that the point of initialisation is important for this image. For example, if the initialisation is at the roof of the cottage, the active contour will not be able to extract the lower half of the cottage. The reason is that the blue doors and windows have a greater colour contrast as compared to the background trees with respect to the colour of the roof. Therefore, under these circumstances, the active contour must be initialised at the point in the cottage where the colour contrast with the background is maximum.

5. Conclusion

In this paper, we present an active contour model for the contour extraction of objects in a natural image scene. Only local region computations are involved and no pre-processing, feature detection or knowledge of the object's geometry is required. Our model is formulated in an analytical framework that allows thresholds to be selected automatically. This makes our model more robust than current models which have heuristic thresholds. In the contour extraction of a natural scene, our active contour model is initialised by a single point with maximum colour contrast between the object and the background. Locating the point of maximum contrast is computationally more efficient than initialising the active contour near the object boundary as done by current models. Experimental results on synthetic and real images showed that the model is robust to blurred images as well as complex natural outdoor scenes.

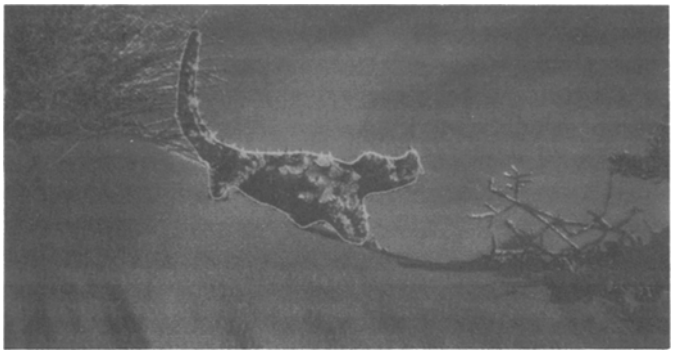

Fig. 4. Contour extracted of reptile

Fig. 5. Contour extracted of monochromatic rocks image

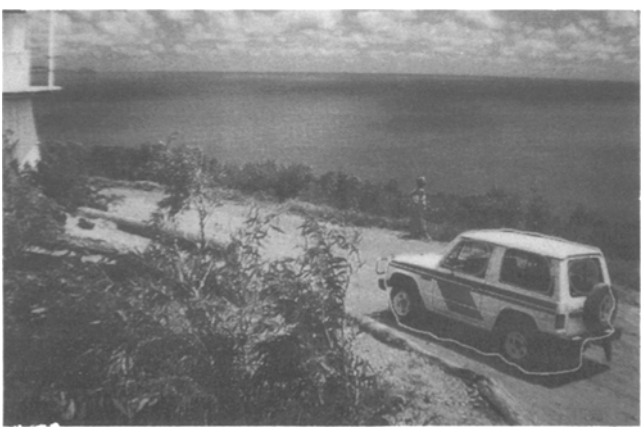

Fig. 6. Contour extracted of a vehicle

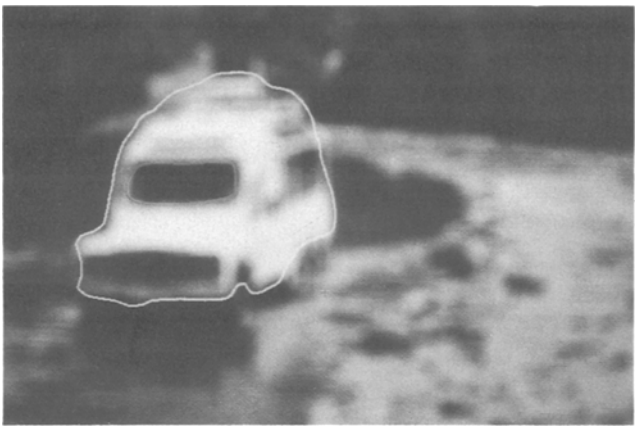

Fig. 7. Contour extracted of a blurred vehicle

Fig. 8. Contour extracted of a cottage

References

1. Cohen, L. D., Cohen, I.: Deformable Models for 3D Medical Images Using Finite Element and Balloons. Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1992) 592-598
2. Delagnes, P., Benois, J., Barba, D.: Active contours approach to object tracking in image sequences with complex background. Pattern Recognition Letters **16** (1995) 171-178
3. Dubuisson, M.P., Jain, A. K.: Contour Extraction of Moving Objects in Complex Outdoor Scenes. Int. Journal of Computer Vision **14** (1995) 83-105
4. Kass, M., Witkin, A., Terzopoulos, D.: Snakes : Active Contour Models. First Int. Conf. on Computer Vision (1987) 259-269

5. Kayargadde, V., Martens, J. B.: Estimation of Edge Parameters and Image Blur Using Polynomial Transforms. CVGIP:Graphical Models and Image Processing. **56** (1994) 442-461
6. Lai, K. F., Chin,R. T.: Deformable Contours: Modelling and Extraction. Int. Conf. on Computer Vision and Pattern Recognition (1994) 601-608
7. Martens, J. B.: The Hermite transform-theory. IEEE Trans. Accoust. Speech Signal Process **38** (1990) 1595-1606
8. Ngoi, K. P., Jia, J. C..: An Analytical Active Contour Model with Automatic Threshold Selection. Second Asian Conf. on Comp. Vision (1995) III72 - III76
9. Pentland, A.: A New sense for depth of field. IEEE Trans. Pattern Anal. Mach. Intelligence **9** (1987) 523-531
10. Ronfard, R.: Region-Based Strategies for Active Contour Models. Int. Journal of Computer Vision **13** (1994) 229-251
11. Terzopoulos, D., Szeliski, R.: Tracking with Kalman Snakes. Active Vision, MIT Press (1992) 3- 20
12. Vieren, C., Cabestaing, F., Postaire, J. G.: Catching moving objects with snakes for motion tracking. Pattern Recognition Letters **16** (1995) 679-685
13. Xu, G., Segawa, E., Tsuji, S.: A Robust Active Contour Model with Insensitive Parameters. Proc. of 4th Int. Conf. on Computer Vision (1993) 562-566
14. Yuille, A. L., Hallinan, P. W.,Cohen, D. S.: Feature Extraction from Faces Using Deformable Templates. Int. Journal of Computer Vision **8** (1992) 99-111