Unlevel-Sets: Geometry and Prior-based Segmentation

Tammy Riklin-Raviv[1], Nahum Kiryati[1], and Nir Sochen[2]

[1] School of Electrical Engineering
[2] Dept. of Applied Mathematics
Tel Aviv University, Tel Aviv 69978, Israel

Abstract. We present a novel variational approach to top-down image segmentation, which accounts for significant projective transformations between a *single* prior image and the image to be segmented. The proposed segmentation process is coupled with reliable estimation of the transformation parameters, without using point correspondences. The prior shape is represented by a *generalized cone* that is based on the contour of the reference object. Its *unlevel* sections correspond to possible instances of the visible contour under perspective distortion and scaling. We extend the Chan-Vese energy functional by adding a shape term. This term measures the distance between the currently estimated section of the generalized cone and the region bounded by the zero-crossing of the evolving level set function. Promising segmentation results are obtained for images of rotated, translated, corrupted and partly occluded objects. The recovered transformation parameters are compatible with the ground truth.

1   Introduction

Classical methods for object segmentation and boundary determination rely on local image features such as gray level values or image gradients. However, when the image to segment is noisy or taken under poor illumination conditions, purely local algorithms are inadequate. Global features, such as contour length and piecewise smoothness [16], can be incorporated using a variational segmentation framework, see [1] and references therein. The handling of contours is facilitated by the level set approach [17]. In the presence of occlusion, shadows and low image contrast, prior knowledge on the shape of interest is necessary [20]. The recovered object boundary should then be compatible with the expected contour, in addition to being constrained by length, smoothness and fidelity to the observed image.

   The main difficulty in the integration of prior information into the variational segmentation process is the need to account for possible pose transformations between the known contour of the given object instance and the boundary in the image to be segmented. Many algorithms [4, 6, 5, 14, 19, 13] use a comprehensive training set to account for small deformations. These methods employ various statistical approaches to characterize the probability distribution of the shapes.

They then measure the similarity between the evolving object boundary (or level set function) and representatives of the training data. The performance of these methods depends on the size and coverage of the training set. Furthermore, none of the existing methods accommodates perspective transformations in measuring the distance between the known instance of the object and the currently segmented image.

We suggest a new method which employs a *single* prior image and accounts for significant *projective* transformations within a variational segmentation framework. This is made possible by two main novelties: the special form of the shape prior, and the integration of the projective transformations via *unleveled* sections. These allow concurrent segmentation and explicit recovery of projective transformation in a reliable way. Neither point correspondence nor direct methods [12] are used. The prior knowledge is represented by a *generalized cone*, which is constructed based on the known instance of the object contour. When the center of projection of a camera coincides with the vertex of the generalized cone, we are able to model the effects of the scene geometry.

We use an extension of the Chan-Vese functional [3] to integrate image data constraints with geometric shape knowledge. The level set function and the projective transformation parameters are estimated in alternation by minimization of the energy functional. The additional energy term that accounts for prior knowledge is a distance measure between a planar (not necessarily horizontal) section of the generalized cone and the zero-crossing of the evolving level set function. Correct segmentation of partly occluded and corrupted images is demonstrated based on a prior image taken with different perspective distortion. The transformation parameters are recovered as well and are in good agreement with the ground truth.

2   Unlevel-Sets

2.1   Previous framework

Mumford and Shah [16] proposed to segment an input image $f: \Omega \rightarrow \mathbb{R}$ by minimizing the functional

$$E(u, C) = \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \lambda \frac{1}{2} \int_{\Omega - C} |\nabla u|^2 dx dy + \nu |C| \ , \qquad (1)$$

simultaneously with respect to the segmenting boundary C and the piecewise smooth approximation u, of the input image f.

When the weight λ of the smoothness term tends to infinity, u becomes a piecewise constant approximation, $u = \{u_i\}$, of f. We proceed with

$$E(u, C) = \frac{1}{2} \sum_i \int_{\Omega_i} (f - u_i)^2 dx dy + \nu |C| \qquad \cup_i \Omega_i = \Omega, \quad \Omega_i \cap \Omega_j = \emptyset \quad (2)$$

In the two phase case, Chan and Vese [3] used a level-set function $\phi \in \mathbb{R}^3$ to embed the contour $C = \{x \in \Omega | \ \phi(x) = 0\}$, and introduced the Heaviside

function $H(\phi)$ into the energy functional:

$$E_{CV}(\phi, u_+, u_-) = \int_\Omega \left[(f - u_+)^2 H(\phi) + (f - u_-)^2 (1 - H(\phi)) + \nu |\nabla H(\phi)| \right] dxdy \tag{3}$$

where

$$H(\phi) = \begin{cases} 1 & \phi \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Using Euler-Lagrange equations for the functional (3), the following gradient descent equation for the evolution of ϕ is obtained:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \, \text{div} \, (\frac{\nabla \phi}{|\nabla \phi|}) - (f - u_+)^2 + (f - u_-)^2 \right] . \tag{5}$$

A smooth approximation of $H(\phi)$ (and $\delta(\phi)$) must be used in practice [3]. The scalars u_+ and u_- are updated in alternation with the level set evolution to take the mean value of the input image f in the regions $\phi \geq 0$ and $\phi < 0$, respectively:

$$u_+ = \frac{\int f(x, y) H(\phi) dxdy}{\int H(\phi) dxdy} \qquad u_- = \frac{\int f(x, y)(1 - H(\phi)) dxdy}{\int (1 - H(\phi)) dxdy} \tag{6}$$

2.2 Shape prior

The energetic formulation (3) can be extended by adding a prior shape term [7]:

$$E(\phi, u_+, u_-) = E_{CV}(\phi, u_+, u_-) + \mu E_{shape}(\phi), \quad \mu \geq 0. \tag{7}$$

We present two novel contributions to this framework. One is a reformulation of the distance measure between the prior and the evolving level-set function, outlined, in a preliminary form, in the rest of this subsection and finalized in subsection 2.5. The other is our unique way of embedding the prior contour within the energy functional, motivated in subsections 2.3-2.4, and formulated in subsection 2.5.

Initially, the shape-term we incorporate in the energy functional measures the non-overlapping areas between the prior shape and the evolving shape. Let $\tilde{\phi}$ be the level set function embedding a prior shape contour. Then

$$E_{shape}(\phi) = \int_\Omega \left(H(\phi(x, y)) - H(\tilde{\phi}(x, y)) \right)^2 dxdy \tag{8}$$

Note that we do not enforce the evolving level set function ϕ to resemble $\tilde{\phi}$, instead we demand similarity of the regions within the respective contours. Minimizing this functional with respect to ϕ leads to the following evolution equation:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \, \text{div} \, (\frac{\nabla \phi}{|\nabla \phi|}) - (f - u_+)^2 + (f - u_-)^2 - 2\mu \left(H(\phi) - H(\tilde{\phi}) \right) \right] \tag{9}$$

This shape-term is adequate when the prior and segmented shapes are not subject to different perspective distortions. Otherwise, the shape-term should incorporate the projective transformation, as detailed in subsections 2.5-2.6. However, a few key concepts should be introduced first.
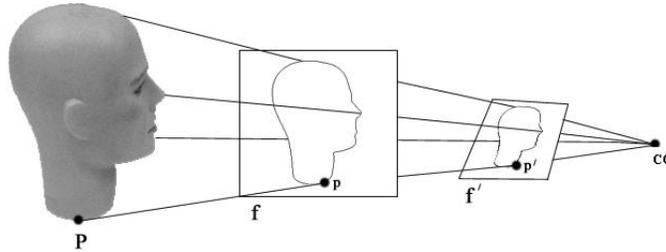
Fig. 1. The cone of rays with vertex at the camera center. An image is obtained by intersection of this cone with a plane. A ray between a $3D$ scene point P and the camera center CC pierces the plane in the image points $p \in f$ and $p' \in f$. All such image points are related by planar homography, $p' = H_p p$. See [11].

2.3   Plane to plane projectivity

An object in a $3D$ space and a camera center define a set of rays, and an image is obtained by intersecting these rays with a plane. Often this set is referred to as a *cone of rays*, even though it is not a cone in the classical sense. Now, suppose that this cone of rays is intersected by two planes, as shown in Fig. 1. Then, there exists a perspective transformation H mapping one image onto the other. This means that the images obtained by the same camera center may be mapped to one another by a plane projective transformation [8, 11, 9].

Let f and f' be the first and the second image planes, respectively. Let K denote a 3×3 internal calibration matrix. Consider two corresponding points, $p \in f$ and $p' \in f'$, expressed in homogeneous coordinates, which are two distinct images of the $3D$ object point $P = (X, Y, Z)$, taken with the same camera. Their relation can be described by $p' = KRK^{-1}p + \frac{1}{Z}K\mathbf{t}$. R is a 3×3 rotation matrix and $\mathbf{t} = [t_x, t_y, t_z]$ is a translation vector. Thus, for any given K, the homography matrix H_p, such that $p' = H_p p$, can be recovered simply by estimating the values of R and $\mathbf{t}$. Since only the plane transformation is important for the segmentation process, when the camera internal parameters are not known, K can be set to the identity matrix, implying that the optical axis is normal to the image plane f and the focal length is 1.

2.4   Generalized cone

A generalized cone[3] or a conical surface, is a ruled surface generated by a moving line (the generator) that passes through a fixed point (the vertex) and continually intersects a fixed planar curve (the directrix). Let $P_v = (X_v, Y_v, Z_{vertex})$ denote the cone vertex, and let $p_v = (x_v, y_v)$ be the projection of the vertex on the directrix plane. We set, without loss of generality, $X_v = x_v$ and $Y_v = y_v$.

---

[3] The concept of generalized cone (or cylinder) in computer vision has been introduced to model $3D$ objects [2, 15]. Its geometrical properties have been intensively investigated, see [10, 18] and references therein.

Now, consider a directrix, $C = p(s) = (x(s), y(s))$ which is a closed contour, parameterized by arc-length s, of an object shape in the plane $Z = Z_{plane} = 0$. The generalized cone surface is the ruled surface defined by:

$$\Phi(r, s) = \Phi((1 - r)p(s) + rp_v) = (1 - r)Z_{plane} + rZ_{vertex} \qquad (10)$$

where r varies smoothly from 1, that corresponds to the vertex, via 0, the directrix, to some convenient negative value.

When the vertex of the generalized cone is located at the camera center, the definition of the generalized cone coincides with that of the cone of rays, presented in subsection 2.3. It follows that by planar slicing of the generalized cone, one can generate new image views as though they had been taken with a camera under the perspective model. There is, however, one exception to this analogy. The intersection of a cone and a plane is either a closed curve, an open curve or a point. In projective geometry terminology, the latter two correspond to projection of finite points in the first image plane to infinity. We do not consider ideal points and planes at infinity. Phrasing it explicitly, our only concern is the mapping of a given closed curve to another closed curve.

2.5   Reformulation of the energy functional

The shape-term in the energy functional (7) is now extended to account for projective transformations. The evolution of the level-set function, given the prior contour and an estimate of the pose parameters, is considered in this subsection. The recovery of the pose parameters, given the prior contour and the curve generated by the zero-crossing of the estimated level-set function, is described in subsection 2.6.

Following subsection 2.2, $\tilde{\phi}$ embeds the prior contour. For reasons that will soon be explained, it is referred to as the *unlevel-set function* and will take the form of a generalized cone. Let $\tilde{C} = \{x, y | \tilde{\phi}(x, y) = 0\}$ be the prior contour in f, and let T_p be a pose transformation applied to the unlevel-set function $\tilde{\phi}$:
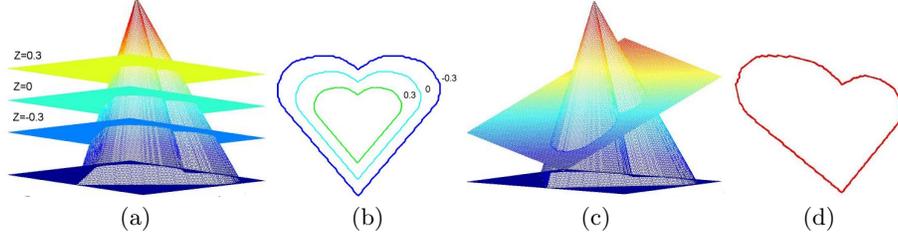
$$(x', y', T_p(\tilde{\phi}))^T = R(x, y, \tilde{\phi})^T + \mathbf{t} . \qquad (11)$$

The evolving contour in the image to be segmented f' is iteratively compared with $\tilde{C}' = \{x', y' | T_p(\tilde{\phi}) = 0\}$ which is the zero-crossing of the transformed unlevel-set function. Note, that instead of changing the pose of the intersecting plane and maintaining the generalized cone fixed, we rotate the generalized cone around its vertex and translate it, while keeping the intersecting plane fixed. Next, we apply the Heaviside function to the transformed unlevel-set function. Thus, the shape-term of the energy functional (7) becomes

$$E_{shape}(\phi) = \int_{\Omega} \left(H(\phi) - H(T_p(\tilde{\phi})) \right)^2 dxdy \qquad (12)$$

and the gradient descent equation, derived similarly to (9), is

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \, \mathrm{div} \, (\frac{\nabla \phi}{|\nabla \phi|}) - (f - u_+)^2 + (f - u_-)^2 - 2\mu \left(H(\phi) - H(T_p(\tilde{\phi})) \right) \right]$$
$$(13)$$

Fig. 2. (a) A generalized cone is sliced by three planes, at $Z = 0.3$, $Z = 0$ and $Z = -0.3$.(b) The resulting intersections. (c) A generalized cone is intersected by an inclined plane: $ax + by + cz + d = 0$. (d) The resulting contour.

2.6   Recovery of the transformation parameters

In order to solve (13), one has to evaluate ϕ simultaneously with the recovery of the transformation T_p of the unlevel-set function $\tilde{\phi}$. The transformation parameters are evaluated via the gradient descent equations obtained by minimizing the energy functional (12) with respect to each parameter. We demonstrate this for the special cases of pure translation and rotation.

Translation Translation of an image plane along the principal axis t_z results in scaling: As the planar section of the generalized cone is closer to the vertex, the cross-section shape is smaller, see Figs. 2a-b. Thus, a scale factor can be incorporated into the energy functional, in compatibility with the scene geometry, simply by translation. Equivalently, one can move the generalized cone along the principal axis, while the plane remains stationary at $Z = 0$. In the case of pure scaling, $T_p(\tilde{\phi})$ is reduced to $\tilde{\phi} + t_z$. Substituting this expression into the shape-term (12) of the energy functional, and minimizing with respect to t_z, gives the following equation:

$$\frac{\partial t_z}{\partial t} = 2\mu \int_\Omega \delta(\tilde{\phi} + t_z)(H(\phi) - H(\tilde{\phi} + t_z))dxdy \tag{14}$$

To account for general translation $\mathbf{t} = (t_x, t_y, t_z)^T$, we can substitute the expression for $T_p(\tilde{\phi})$ (11) in (12), with $R = I$, where I is the identity matrix. The shape term takes the form

$$E_{shape}(\phi) = \int_\Omega (H(\phi)(x, y) - H(\tilde{\phi}(x + t_x, y + t_y) + t_z))^2 dxdy$$

and the translation parameters t_x and t_y can be recovered similarly to t_z.

Rotation Consider a tilted planar cut of the generalized cone, as shown in Figs. 2c,d. The resulting contour is perspectively deformed, as a function of the inclination of the intersecting plane and its proximity to the vertex of the

cone. Equivalently, one may rotate the generalized cone around its vertex, and zero-cross to get the same perspective transformation.

Any rotation can be decomposed to rotations about the three axes (Euler's rotation theorem), and can be represented by a matrix $R = R_X(\alpha)R_Y(\beta)R_Z(\gamma)$ operating on a vector $(x, y, z)^T$:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\alpha & sin\alpha \\ 0 & -sin\alpha & cos\alpha \end{bmatrix} \begin{bmatrix} cos\beta & 0 & -sin\beta \\ 0 & 1 & 0 \\ sin\beta & 0 & cos\beta \end{bmatrix} \begin{bmatrix} cos\gamma & sin\gamma & 0 \\ -sin\gamma & cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Let η be some rotation angle corresponding to any of the angles α, β or γ. The general gradient descent equation for a rotation angle is of the form:

$$\frac{\partial \eta}{\partial t} = 2\mu \int_\Omega \delta(T_p(\tilde{\phi})) \left(H(\phi) - H(T_p(\tilde{\phi})) \right) \left[\frac{\partial z'}{\partial x'}\frac{\partial x'}{\partial \eta} + \frac{\partial z'}{\partial y'}\frac{\partial y'}{\partial \eta} + \frac{\partial z'}{\partial \eta} \right] dxdy \tag{15}$$

Note that $z = \tilde{\phi}(x, y)$ and $z' = T_p(\tilde{\phi})$. The partial derivatives for $\eta = \beta$, for example, are

$$\frac{\partial x'}{\partial \beta} = -x\ cos\beta\ sin\gamma - y\ sin\beta\ sin\gamma - z\ cos\beta$$

$$\frac{\partial y'}{\partial \beta} = x\ sin\alpha\ cos\beta\ cos\gamma + y\ sin\alpha\ cos\beta\ sin\gamma - z\ sin\alpha \tag{16}$$

$$\frac{\partial z'}{\partial \beta} = x\ cos\alpha\ cos\beta\ cos\gamma + y\ cos\alpha\ cos\beta\ sin\gamma - z\ cos\alpha$$

and similarly for $\eta = \alpha$ and $\eta = \gamma$. The values of $\partial z'/\partial x'$ and $\partial z'/\partial y'$ are derived numerically from the cone surface values.

2.7   The *unlevel-set* algorithm

We summarize the proposed algorithm, for concurrent image segmentation given a prior contour, and recovery of the projective transformation between the current and prior object instances.

1. The inputs are two images f and f' of the same object, taken with the same camera, but under different viewing conditions. The boundary $\tilde{C}$ of the object in f is known. The image f' has to be segmented. The image plane of the first image f is assumed to be perpendicular to the principal axis, at distance 1 from the camera center. The second image plane, of f', is tilted and shifted relative to the first one.
2. Given the contour $\tilde{C}$, construct a generalized cone, using the expression in (10) with $Z_{vertex} = 1$.
3. Choose some initial level-set function ϕ, for example a standard right cone.
4. Set initial values (e.g. zero) for α, β , γ , t_x, t_y and t_z.
5. Compute the average gray level values of the object and background pixels, u_+ and u_-, using equation (6).
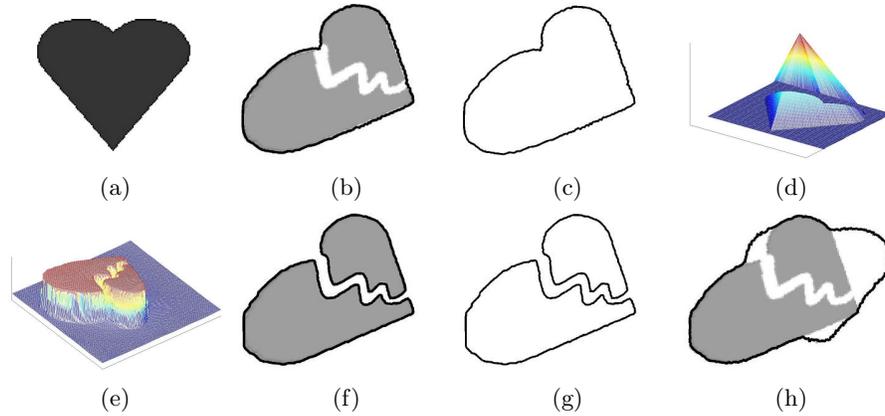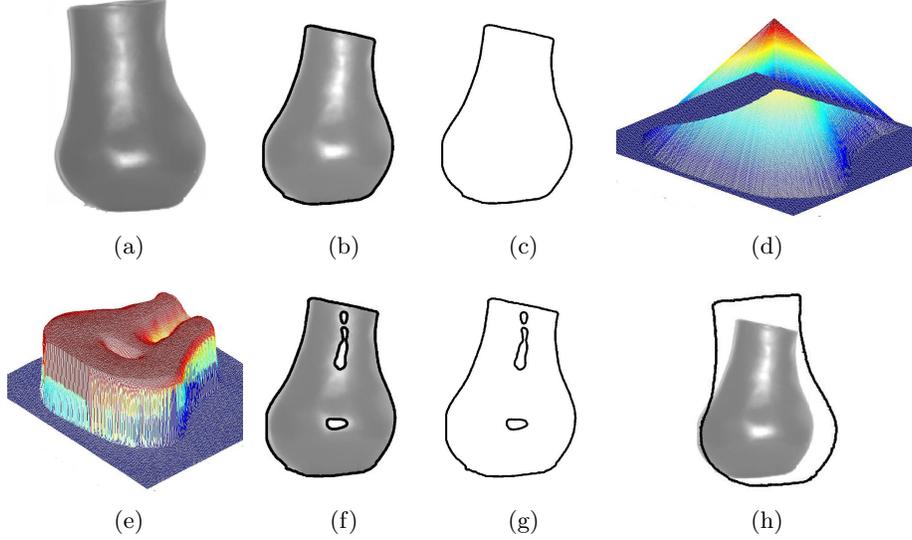
Fig. 3. Synthetic example. (a) Prior image. The contour is known (not shown). (b) Successful segmentation: the final contour is shown (black) on the transformed and corrupted image. (c) The final contour $\tilde{C}'$ obtained in (b). (d) Generalized cone $\tilde{\phi}$, based on the prior contour $\tilde{C}$. (e) Final level set function ϕ. (f) Wrong segmentation: prior knowledge was not used. (g) The final contour obtained in (f). (h) Wrong segmentation: the prior is used without incorporated the projective transformation.

6. Compute the values of $T_p(\tilde{\phi})$ according to equation (11), for the currently estimated transformation parameters.
7. Update ϕ according to the gradient descent equation (13).
8. Update $\mathbf{t}$, using (14) for t_z and similar equations for t_x and t_y, and (15) for α, β and γ, until convergence.
9. Repeat steps 5-8 until convergence.

3   Experimental Results

To demonstrate our model, we present segmentation results on various synthetic and real images. Relative scale and pose parameters between the image of the known contour and the image to be segmented have been estimated and compared to the ground-truth, where available. The strength of this algorithm is expressed by its weak sensitivity with respect to the parameters of the functional. We use $\nu = 50$, $\mu = 25$ unless otherwise stated. Exclusion of the shape prior knowledge from the functional means setting μ to zero.

Consider the synthetic images shown in Figs. 3a,b. Only the contour of the object in Fig. 3a (not drawn) was known in advance and used as prior. The object in Fig. 3b was generated from Fig. 3a by rotation and translation with the following parameters: $R_X(\alpha) = 0.3^0$, $R_Y(\beta) = -0.3^0$ and $R_Z(\gamma) = 60^0$ with scale factor of 0.9. It has also been broken and lightened. Note the significant perspective distortion despite the fairly small rotations around the X and Y axes. The black contour in Fig. 3b is the result of the segmentation process. For clarity, the final contour is presented by itself in Fig. 3c. The generalized cone

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

Fig. 4. Real image with synthetic transformation. (a) Prior image. The contour is known (not shown). (b) Successful segmentation: the final contour (black) on the transformed image. (c) The final contour $\tilde{C}'$ obtained in (b). (d) Generalized cone $\tilde{\phi}$, based on the prior contour $\tilde{C}$. (e) The final level set function ϕ. (f) Wrong segmentation: prior knowledge was not used. (g) The final contours obtained in (f). (h) Wrong segmentation: the prior is used without incorporating the projective transformation.

ϕ that was constructed, based on the known image contour, using Eq. (10), is shown in Fig. 3d. Fig. 3e shows the final evolving level-set function ϕ. It is worth emphasizing that ϕ and $T_p(\tilde{\phi})$ resemble in terms of their Heaviside functions - that is by their zero-crossings (the final contour), but not in their entire shapes. The estimated transformation parameters are: $\hat{R}_X(\alpha) = 0.38^0$, $\hat{R}_Y(\beta) = -0.4^0$, $\hat{R}_Z(\gamma) = 56.6^0$ and $\hat{t}_z = -0.107$ - which corresponds to scaling of 0.893. When no shape prior is used, each part of the broken heart is segmented separately (Figs. 3f-g). Segmentation fails when the prior is enforced without recovery of the transformation parameters, as shown in figure 3h.

We next consider real images, Figs. 4a-b, where the black contour around the object in figure 4b is again the segmentation result. The final contour itself is shown in Fig. 4c. The transformation between the images was synthetic, so that the calculated parameters could be compared with the ground-truth. The transformation parameters are: $R_X(\alpha) = -0.075^0$, $R_Y(\beta) = 0.075^0$ and $R_Z(\gamma) = 9^0$ with scaling factor of 0.8 . Compare with the recovered transformation parameters: $\hat{R}_X(\alpha) = -0.063^0$, $\hat{R}_Y(\beta) = 0.074^0$, $\hat{R}_Z(\gamma) = 7.9^0$ and scaling of 0.81. The generalized cone $\tilde{\phi}$, based on the given jar contour, and the final level set function ϕ are shown in Figs. 4d-e respectively. The jar shown is black with white background. Thus, without using the prior, the bright specular reflection spots spoil the segmentation, as shown in Figs. 4f-g. Again, when the prior is enforced,
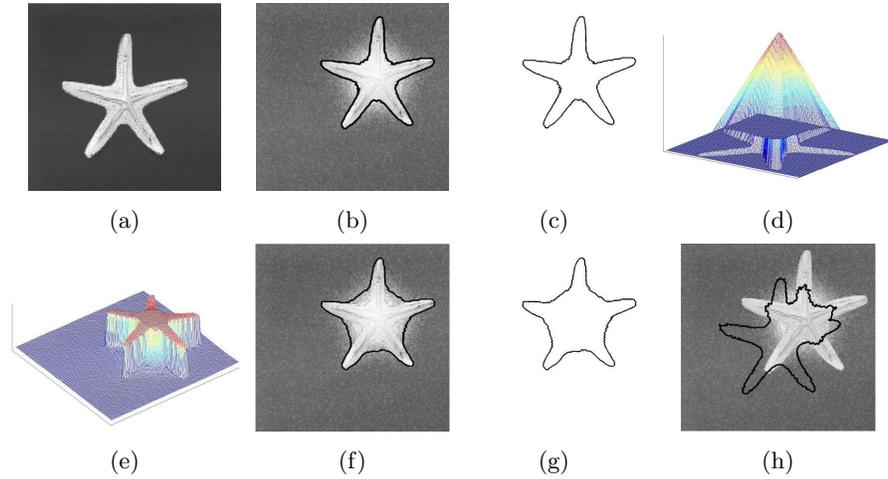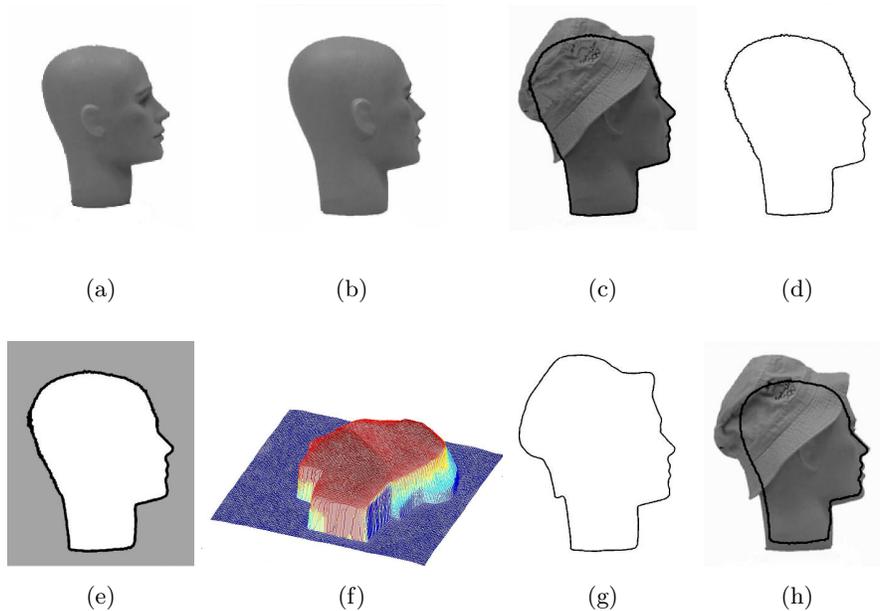
Fig. 5. Real image with synthetic noise. (a) Prior image. The contour is known (not shown). (b) Successful segmentation: the final contour (black) on the transformed image. (c) The final contour $\tilde{C}'$ obtained in (b). (d) Generalized cone $\tilde{\phi}$, based on the prior contour $\tilde{C}$. (e) Final level set function ϕ. (f) Wrong segmentation: prior knowledge was not used. (g) The final contours obtained in (f). (h) Wrong segmentation: the prior is used without incorporating the projective transformation.

but the transformation parameters are not recovered, segmentation fails as seen in Fig. 4h.

To check simultaneous translations along the X, Y and Z axes we applied our algorithm to the images shown in Figs. 5a-b. The noisy Fig. 5b is segmented correctly (black contour) in spite of the significant translation with respect to the prior. No preprocessing alignment has been performed. The functional parameters in this case were $\mu = 13$ and $\nu = 40$. The recovered transformation parameters are: $t_x = 19.54$, $t_y = -18.8$, $t_z = 0.08$.

Finally, we demonstrate the method using a real object (mannequin head), which has actually been rotated, moved and occluded, as seen in Figs. 6a-c. The algorithm is able to segment the head precisely, in spite of the covering hat which has color similar to that of the mannequin. The segmenting contour accurately traces the profile of the mannequin, despite the significant transformation. Since the actual transformation was not measured, then in order to confirm the recovered transformation parameters, Fig. 6e shows the zero-crossing of the transformed generalized cone together with the final segmenting contour (Fig. 6d).

Translation and rotation of non-planar objects may reveal previously hidden points and hide others. Therefore, the visible contour in a new instance of the object might be significantly different from the reference. However, as seen in the jar and mannequin examples, for moderate transformations of these non-planar objects, promising segmentation results are obtained.

Fig. 6. Real example. (a) Reference image (mannequin head). The contour is known (not shown). (b) New instance of the mannequin head, rotated and translated. (c) Successful segmentation: the final contour (black) on the transformed mannequin head. The segmentation is precise despite the covering hat. (d) The final contour $\tilde{C}'$ obtained in (b). (e) The final contour as in (d), drawn on the Heaviside function of the transformed generalized cone: $H(T_p(\tilde{\phi}))$. This shows the compatibility between the calculated and actual transformation parameters. (f) Final shape of the evolving level set function ϕ. (g) Final contour obtained without using a shape prior. (h) Final contour obtained using the prior but without recovery of the transformation parameters.

4   Discussion

Detection of an object in a corrupted image, based on a reference image taken with from a different view-point, is a classical challenge in computer vision. This paper presents a novel approach that makes substantial progress towards this goal. The key to this accomplishment is the unique integration of scene geometry with the variational approach to segmentation. The reference shape is the foundation of a generalized cone. In principle, the zero level set of an evolving function, related to the image features, is matched with *unlevel* sections of the generalized cone that correspond to projectively deformed views of the shape.

The suggested algorithm successfully accounts for scale and pose variations under the perspective model, including rotation outside the image plane, without using point correspondence. The algorithm converges empirically even for fairly large transformations and significantly corrupted images. Promising segmentation results and accurate numerical estimation of the transformation parameters, suggest this model as an efficient tool for segmentation and image alignment.

References

1. G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, 2002.
2. T.O. Binford. Visual perception by computer. In *Proc. IEEE Conf. Systems and Control*, December 1971.
3. T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, February 2001.
4. Y. Chen, S. Thiruvenkadam, H.D. Tagare, F. Huang, and D. Wilson. On the incorporation of shape priors into geometric active contours. In *VLSM01*, pages 145–152, 2001.
5. D. Cremers, T. Kohlberger, and C. Schnorr. Nonlinear shape statistics via kernel spaces. In *DAGM01*, pages 269–276, 2001.
6. D. Cremers, T. Kohlberger, and C. Schnorr. Nonlinear shape statistics in mumford-shah based segmentation. In *ECCV02*, volume II, pages 93–108, 2002.
7. D. Cremers, N. Sochen, and C. Schnorr. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In *Intl. Conf. on Scale-Space Theories in Computer Vision*, pages 388–400, June 2003.
8. O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.
9. O. Faugeras, Q.T. Luong, and T. Papadopoulo. *The Geometry of Multiple Images*. MIT Press, 2001.
10. D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
11. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
12. M. Irani and P. Anandan. All about direct methods. In W. Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory and Practice*. Springer-Verlag, 1999.
13. M. Leventon, O. Faugeraus, W. Grimson, and W. Wells III. Level set based segmentation with intensity and curvature priors. In *Workshop on Mathematical Methods in Biomedical Image Analysis Proceedings*, pages 4–11, June 2000.
14. M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR00*, volume I, pages 316–323, 2000.
15. D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, 1982.
16. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–684, 1989.
17. S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
18. K.G. Rao and G. Medioni. Generalized cones: Useful geometric properties. In *CVIP92*, pages 185–208, 1992.
19. A. Tsai, A. Yezzi, Jr., W.M. Wells, III, C. Tempany, D. Tucker, A. Fan, W.E.L. Grimson, and A.S. Willsky. Model-based curve evolution technique for image segmentation. In *CVPR01*, volume I, pages 463–468, 2001.
20. S. Ullman. *High-Level Vision: Object Recognition and Visual Cognition*. MIT Press, 1996.