Prior-based Segmentation by Projective Registration and Level Sets

Tammy Riklin-Raviv
Faculty of Engineering
Tel-Aviv University, Israel
tammy@eng.tau.ac.il

Nahum Kiryati
Faculty of Engineering
Tel-Aviv University, Israel
nk@eng.tau.ac.il

Nir Sochen
School of Mathematics
Tel-Aviv University, Israel
sochen@post.tau.ac.il

Abstract

Object detection and segmentation can be facilitated by the availability of a reference object. However, accounting for possible transformations between the different object views, as part of the segmentation process, remains a challenge. Recent works address this problem by using comprehensive training data. Other approaches are applicable only to limited object classes or can only accommodate similarity transformations.

We suggest a novel variational approach to prior-based segmentation, which accounts for planar projective transformation, using a single reference object. The prior shape is registered concurrently with the segmentation process, without point correspondence. The algorithm detects the object of interest and correctly extracts its boundaries. The homography between the two object views is accurately recovered as well. Extending the Chan-Vese level set framework, we propose a region-based segmentation functional that includes explicit representation of the projective homography between the prior shape and the shape to segment. The formulation is derived from two-view geometry. Segmentation of a variety of objects is demonstrated and the recovered transformation is verified.

1. Introduction

The coupled tasks of segmentation and object detection are essential for the extraction of semantic content from images. Prior knowledge on the shape of interest can significantly facilitate these processes, particularly when the object boundaries are not well defined.

Variational methods solve segmentation problems by means of energy minimization, integrating top-down and bottom-up information, see [1] and references therein. Specifically, the level-set framework [14] for curve evolution, has been successfully applied to extract complex object boundaries allowing an automatic change in the topology. Recent approaches incorporate a representation of a reference shape within the energy functional. Thus, the recovered object boundary should resemble the expected contour, in addition to being constrained by length, smoothness and compatibility with the image features.

The main difficulties in the integration of shape information into the segmentation process is the determination of a suitable shape representation and the adjustment of similarity measure between shape representations that can accommodate the possible group of transformations. The statistical approach [3, 4, 9, 11, 16, 18] addresses this challenge by using a comprehensive training set to account for shape deformations. It characterizes the probability distribution of the shapes and then measures 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. Several studies [5, 6, 16, 18] accommodate similarity transformations, however, besides [15] that handles perspectivity, none of the existing methods explicitly accounts for projective transformations between the prior shape and the shape of interest.

We suggest a novel variational approach to prior-based segmentation, that accounts for homographies, using a *single* reference object. Segmentation process is carried out concurrently with the registration of the prior shape to the shape of interest. Neither point correspondence nor direct methods [10] are used. The outcome of the algorithm includes the detection of the object of interest and a correct extraction of its boundaries. The planar projective transformation between the two object views is accurately recovered as well. This is accomplished by a novel extension of the Chan-Vese [2] level-set framework. The proposed region-based functional for segmentation includes an explicit expression of the projective homography between the prior shape and the shape to segment, which is derived from two-view geometry. It is supported by the introduction of a shape-similarity measure that admits general transformations, beyond similarity, into the functional. Segmentation of a variety of objects is demonstrated and the recovered transformation is verified.

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**

2. Level Set Framework

2.1. The Chan-Vese two-phase model

In the level-set formulation for curve evolution [14], a segmenting boundary C in an image plane $\Omega \subset \mathbb{R}^2$, is the zero level-set of a 3D function ϕ, $C = \{\mathbf{x} \in \Omega |\ \phi(\mathbf{x}) = 0\}$. Chan and Vese [2], inspired by the segmentation model of Mumford and Shah [13], suggested to segment an input image $f: \Omega \rightarrow \mathbb{R}$ by minimizing an energy functional based on a piecewise constant approximation of the image. They used the Heaviside function of the evolving level-set function $H(\phi)$ as an indicator for object and background regions in the image. Thus, they have turned the segmentation framework to be a region-based, establishing a cost functional with well defined integration limits:

$$\begin{aligned} E_{CV} &= \int_{\Omega} \left[(f-u_+)^2 H(\phi) + (f-u_-)^2 \left(1 - H(\phi)\right) \right. \\ &+ \left. \nu |\nabla H(\phi)| \right] d\mathbf{x}, \end{aligned} \tag{1}$$

where ν is a real and positive weight parameter. The scalars u_+ and u_- denote the average gray level values of the input image in the regions indicated by $\phi \geq 0$ and by $\phi < 0$ respectively. They are alternately updated with the evolution of the level-set function ϕ and take the form:

$$u_+ = \frac{\int f(\mathbf{x}) H(\phi) d\mathbf{x}}{\int H(\phi) d\mathbf{x}} \qquad u_- = \frac{\int f(\mathbf{x}) \left(1 - H(\phi)\right) d\mathbf{x}}{\int (1 - H(\phi)) d\mathbf{x}} \tag{2}$$

The gradient descent equation for the evolution of ϕ is derived using the Euler-Lagrange equations for the functional (1):

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \ \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (f-u_+)^2 + (f-u_-)^2 \right]. \tag{3}$$

Practically, a smooth approximation of the Heaviside function H_ϵ, rather than a step function, is used [2]:

$$H_\epsilon(\phi) = \frac{1}{2}(1 + \frac{2}{\pi} \arctan(\frac{\phi}{\epsilon})) \tag{4}$$

$$\delta_\epsilon(\phi) = \frac{dH_\epsilon(\phi)}{d\phi} = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + \phi^2}. \tag{5}$$

2.2. Prior shape model

The energy functional (1) can be extended by adding a prior shape term, e.g. [5, 15]:

$$E(\phi, u_+, u_-) = E_{CV}(\phi, u_+, u_-) + \mu E_{shape}(\phi) \quad \mu \geq 0. \tag{6}$$

One of the contributions of this work is the similarity measure between the prior and the evolving shape. The shape-term incorporated in the energy functional measures the non-overlapping areas between the prior shape and the zero level-set of ϕ. Let $\tilde{\phi} : \mathbb{R}^2 \rightarrow \{0, 1\}$ be a labeling function of the shape in the prior image. Note that $H_\epsilon(\tilde{\phi}) \rightarrow \tilde{\phi}$ as $\epsilon \rightarrow 0$. Thus, the shape term takes the form:

$$E_{shape}(\phi) = \int_{\Omega} \left(H_\epsilon(\phi(\mathbf{x})) - H_\epsilon(\tilde{\phi}(\mathbf{x})) \right)^2 d\mathbf{x} \tag{7}$$

Note, that we do not force the evolving level-set function ϕ to resemble $\tilde{\phi}$. Instead, we demand similarity of the regions within the respective contours. Since ϕ is not constrained to a distance function, the proposed similarity measure can accommodate *planar projective transformations* between the prior contour $\tilde{C}$ and the evolving segmenting contour C. The following evolution equation is obtained by minimizing (7) with respect to ϕ:

$$-\frac{\delta E_{shape}}{\delta \phi} = \frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) \left(H_\epsilon(\phi) - H_\epsilon(\tilde{\phi}) \right) \tag{8}$$

Hence, at each time step, ϕ will be changed in regions where there are inconsistencies between the object-background areas indicated by $H_\epsilon(\phi)$ and $H_\epsilon(\tilde{\phi})$. The spatial evolution in ϕ is controlled by ϵ, according to (4) and (5).

The shape-term (7) is adequate when the prior and segmented shapes are aligned. Otherwise, the representation of the prior shape-term should incorporate the projective transformation, as is detailed in section 4. However, a few geometric concepts should be reviewed first.

3. Projectivity

3.1. Planar projective homography

To draw the geometric relation between two corresponding shape contours, we review the concept of *planar projective homography*. The identification of geometric projectivity with algebraic homography is supported by the theorems of Desargues [17]. Planar projective homography (projectivity) is a mapping $\mathcal{H} : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ such that the points p_i are collinear if and only if $\mathcal{H}(p_i)$ are collinear (projectivity preserves lines) [8, 17].

The relation between corresponding views of points on a world plane Π in a 3D space, as is illustrated in Fig. 1, can be modeled by a planar homography induced by the plane. An explicit expression for the induced homorgraphy can be derived as follows. Consider two views p and p' of a world point $P \in \Pi$, in two camera frames f and f' respectively. We will denote their corresponding homogeneous coordinates by $\mathbf{x}$, $\mathbf{x}'$ and $\mathbf{X}$. Let $M = K[\ I \mid 0\ ]$ and $M' = K'[\ R \mid \mathbf{t}\ ]$ be the first and the second camera projection matrices (respectively), where R and $\mathbf{t}$ are the relative rotation and translation between the cameras and K and K' are the respective internal calibration matrices. Thus, $\mathbf{x} = K[\ I \mid 0\ ]\mathbf{X}$ and $\mathbf{x}' = K'[\ R \mid t\ ]\mathbf{X}$.

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**

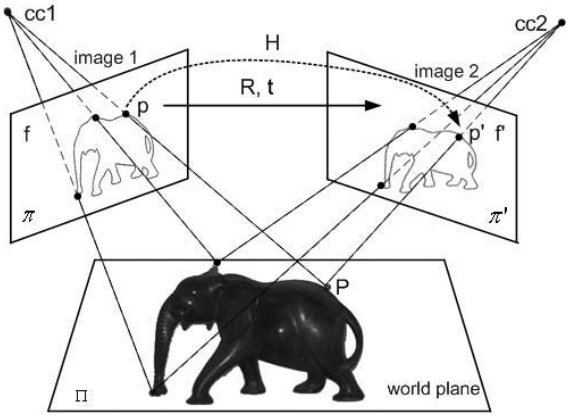


Figure 1. The homography induced by a plane. The ray that intersects the first image plane at a point p meets the world plane Π at a point P. The point P is projected to a point p' in the second image plane. The mapping from p to p' is the homography, denoted by $\mathcal{H}$, induced by the plane Π. Illustrated after [8]. In this research, the corresponding points p and p' are on corresponding planar contours.

Let $\mathbf{n}$ be the unit normal vector to the plane Π, and let $d > 0$ denote the distance of Π from the optical center of the first camera. The linear transformation from $\mathbf{x}$ to $\mathbf{x}'$ can be expressed as

$$\mathbf{x}' = K' \left(R + \frac{1}{d} \mathbf{t} \mathbf{n}^T \right) K^{-1} \mathbf{x} = \mathcal{H} \mathbf{x} \, . \tag{9}$$

Assuming calibrated cameras, we set $K = K' = I$. Thus, the planar homography matrix takes the form [7, 8, 12]:

$$\mathcal{H} = R + \frac{1}{d} \mathbf{t} \mathbf{n}^T . \tag{10}$$

The matrix $\mathcal{H}$ is determined by the translation and rotation between the two cameras $\{R, \mathbf{t}\}$, and by the structure parameters $\{\mathbf{n}, d\}$ of the world plane Π. Note that only the ratio $\mathbf{t}/d$ can be recovered from $\mathcal{H}$. We now show how the structure of the homography can be used to recover the relation between the camera frames and thus between the respective shape contours.

3.2. Implicit recovery of the homography

A nonsingular homography between two planes π,π' is determined (up to a scale factor) by four pairs of corresponding points $\mathbf{x} \in \pi$, $\mathbf{x}' \in \pi'$ with no collinear triplet of points in either plane. Consider the homography matrix

$$\mathcal{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3} \tag{11}$$

and let $\mathbf{x} = (x, y, 1)$ and $\mathbf{x}' = (x', y', 1)$ be the homogeneous representations of the points $\mathbf{x} \in \pi$, $\mathbf{x}' \in \pi'$ such that $\mathbf{x}' = \mathcal{H}\mathbf{x}$. The eight unknowns of $\mathcal{H}$ (the ratios of its nine entries) can be recovered by solving at least four pairs of equations of the form:

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \tag{12}$$

Classic approaches recover $\mathcal{H}$ by solving an overdetermined set of equations as in (12). The rela tive position of the image planes $(R, \mathbf{t})$, and the world plane orientation and pose $(\mathbf{n}, d)$ are recovered, by decomposition methods from the known homography matrix (see [7, 8, 12] and references therein).

Our novel approach calculates the homography directly in its *explicit* representation (10). No point correspondence is needed. Instead, we use the calculus of variations to recover the relative motion between the two camera frames, given two corresponding views of the shape of interest. Practically, since the recovery of the homography is done simultaneously with the process of object detection and segmentation, only the reference view is known in advance. The shape to segment is extracted by the segmentation process concurrently with the registration of the prior to it.

3.3. Explicit recovery of the homography

We now use the explicit formulation of the homography (10) to reformulate equation (12). We denote by γ, β and α the Euler rotation angles around the Z,Y and X axes respectively, in this order of rotation. They determine the relative orientation between the first and the second camera coordinate systems. The rotation matrix $R \in \mathbb{R}^{3\times 3}$ operating on a vector $(x, y, z)^T$ takes the form:

$$R = \begin{bmatrix} c_\beta c_\gamma & c_\beta s_\gamma & -s_\beta \\ s_\alpha s_\beta c_\gamma - c_\alpha s_\gamma & s_\alpha s_\beta s_\gamma + c_\alpha c_\gamma & s_\alpha c_\beta \\ c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma & c_\alpha s_\beta s_\gamma - s_\alpha s_\gamma & c_\alpha c_\beta \end{bmatrix} \tag{13}$$

where s_α is $sin(\alpha)$ and c_α is $cos(\alpha)$. The relative position between the two camera frames is determined by $\mathbf{t} = (t_x, t_y, t_z)$. Consider first the special case in which the world plane and the first image plane coincide, and d is set to 1. In this case, the normal to the world plane $\mathbf{n}$ is $(0, 0, 1)$. Substituting d and $\mathbf{n}$ in (10), we obtain the entries of the homography matrix $\mathcal{H}$: $h_{13} = R_{13} + t_x$, $h_{23} = R_{23} + t_y$, $h_{33} = R_{33} + t_z$ and $h_{ij} = R_{ij}$ otherwise.

Generally, the world plane is not perpendicular to the optical axis of the first camera: $\mathbf{n} \neq (0, 0, 1)$. As before, we

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**

represent $\mathbf{n}$ using the coordinate system of the first camera. The unit vector $\mathbf{n}$ is obtained by rotating the vector $(0,0,1)$ by an angle ξ around the y-axis, then by an angle ψ around the x-axis. Hence, $\mathbf{n} = (-\sin\xi, \sin\psi\cos\xi, \cos\psi\cos\xi)$. Substituting R, $\mathbf{t}$ and $\mathbf{n}$ in (10), we obtain the components of the homography matrix $\mathcal{H}$:

$$\begin{aligned}
h_{11} &= \cos\beta\cos\gamma - \tfrac{t_x}{d}\sin\xi \\
h_{12} &= \cos\beta\sin\gamma + \tfrac{t_x}{d}\sin\psi\cos\xi \\
h_{13} &= -\sin\beta + \tfrac{t_x}{d}\cos\psi\cos\xi \\
h_{21} &= \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - \tfrac{t_y}{d}\sin\xi \\
h_{22} &= \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma + \tfrac{t_y}{d}\sin\psi\cos\xi \\
h_{23} &= \sin\alpha\cos\beta + \tfrac{t_y}{d}\cos\psi\cos\xi \\
h_{31} &= \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - \tfrac{t_z}{d}\sin\xi \\
h_{32} &= \cos\alpha\sin\beta\sin\gamma - \sin\alpha\sin\gamma + \tfrac{t_z}{d}\sin\psi\cos\xi \\
h_{33} &= \cos\alpha\cos\beta + \tfrac{t_z}{d}\cos\psi\cos\xi
\end{aligned} \tag{14}$$

Hereafter, ξ and ψ are referred to as structure parameters.

4. The cost functional

4.1. Representation of the prior shape

The shape-term in the energy functional (6) is now extended to account for projective transformations. The evolution of the level-set function, given the prior contour and an estimate of the transformation parameters, is considered in this section. Consider a prior image $f : \Omega \to \mathbb{R}^+$ with labeled object and background regions Ω_+ and Ω_- respectively. The prior shape is represented by a function $\tilde{\phi} \in \mathbb{R}^3$:

$$\tilde{\phi}(x,y) = \begin{cases} 1 & f(x,y) \in \Omega_+ \\ 0 & f(x,y) \in \Omega_- \end{cases} \tag{15}$$

Let $\mathcal{T}_p : (\tilde{\phi}(x,y), x, y) \to (\tilde{\phi}'(x',y'), x', y')$ be a transformation of a point (x,y) and its $\tilde{\phi}$ value to a (projectively) equivalent point (x',y'), with $\tilde{\phi}'(x',y') = \tilde{\phi}(x,y)$. Expressions for x' and y' are obtained by substituting the explicit entries of the homography matrix (14) in (12). For example, when $\xi = \psi = 0$ and $d = 1$, the expressions are:

$$x' = \frac{c_\beta c_\gamma x + c_\beta s_\gamma y - s_\beta + t_x}{(c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma)\, x + (c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma)\, y + c_\alpha c_\beta + t_z} \tag{16}$$

$$y' = \frac{(s_\alpha s_\beta c_\gamma - c_\alpha s_\gamma)\, x + (s_\alpha s_\beta s_\gamma + c_\alpha c_\gamma)\, y + s_\alpha c_\beta + t_y}{(c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma)\, x + (c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma)\, y + c_\alpha c_\beta + t_z} \tag{17}$$

The evolution of the level-set function ϕ is determined by an energy shape term that measures the dissimilarity between its zero level-set and the zero level-set of the projectively transformed $\tilde{\phi}$. Formally,

$$\begin{aligned}
E(\phi) &= \int_\Omega (f-u_+)^2 H_\epsilon(\phi) + (f-u_-)^2 (1 - H_\epsilon(\phi)) \\
&+ \nu|\nabla H_\epsilon(\phi)| + \mu \left[H_\epsilon(\phi) - H_\epsilon(\mathcal{T}_p(\tilde{\phi})) \right]^2 d\mathbf{x}
\end{aligned} \tag{18}$$

and the gradient descent equation for the evaluation of ϕ is

$$\begin{aligned}
\frac{\partial\phi}{\partial t} &= \delta_\epsilon(\phi) \Big[(f-u_-)^2 - (f-u_+)^2 + \nu \,\mathrm{div}\Big(\frac{\nabla\phi_\epsilon}{|\nabla\phi_\epsilon|}\Big) \\
&+ \mu \Big(H_\epsilon(\phi) - H_\epsilon(\mathcal{T}_p(\tilde{\phi})) \Big) \Big].
\end{aligned} \tag{19}$$

4.2. Recovery of the transformation

In order to solve (19), one has to evaluate ϕ simultaneously with the recovery of the transformation $\mathcal{T}_p$ of the prior level-set function $\tilde{\phi}$. The transformation parameters (α, β, γ, t_x/d, t_y/d, t_z/d, ψ and ξ) are evaluated via the gradient equations obtained by minimizing the energy functional (18) with respect to each of them. The general gradient descent equation for each of the transformation parameters (denoted here by η) is of the form:

$$\frac{\partial\eta}{\partial t} = 2\mu \int_\Omega \delta(\mathcal{T}_p(\tilde{\phi})) \left(H_\epsilon(\phi) - H_\epsilon(\mathcal{T}_p(\tilde{\phi})) \right) \frac{\partial \mathcal{T}_p(\tilde{\phi})}{\partial\eta} d\mathbf{x} \tag{20}$$

where,

$$\begin{aligned}
\frac{\partial \mathcal{T}_p(\tilde{\phi})}{\partial\eta} &= \frac{\partial \mathcal{T}_p(\tilde{\phi})}{\partial x} \left(\frac{\partial x}{\partial x'}\frac{\partial x'}{\partial\eta} + \frac{\partial x}{\partial y'}\frac{\partial y'}{\partial\eta} \right) \\
&+ \frac{\partial \mathcal{T}_p(\tilde{\phi})}{\partial y} \left(\frac{\partial y}{\partial x'}\frac{\partial x'}{\partial\eta} + \frac{\partial y}{\partial y'}\frac{\partial y'}{\partial\eta} \right)
\end{aligned} \tag{21}$$

The values of $\partial\mathcal{T}_p(\tilde{\phi})/\partial x$ and $\partial\mathcal{T}_p(\tilde{\phi})/\partial y$ are derived numerically using the finite difference method. The expressions for $\partial x'/\partial\eta$ and $\partial y'/\partial\eta$ for each of the transformation parameters η are evaluated by differentiating the explicit expressions for x' and y' (generated by substituting equations (14) in (12))with respect to η. We use the implicit function theorem to derive the equations for $\partial x/\partial x'$, $\partial x/\partial y'$, $\partial y/\partial x'$, $\partial y/\partial y'$.

4.3. Algorithm

We summarize the proposed algorithm assuming the following setup. The input is two images f and f' of the same object, taken with the same camera, but under different viewing conditions. Contour points of the object to segment should be approximately co-planar. The segmentation of the reference object in f is known and used to construct the prior shape representation $\tilde{\phi}$ as in (15)

1. Choose an initial level-set function ϕ. Its zero level-set should form an initial contour within the image.
2. Set initial values (e.g. zero) for the transformation parameters α, β, γ, t_x, t_y, t_z, ξ and ψ. Set $d = 1$.
3. Compute the average gray levels of the estimated object and background pixels, u_+ and u_-, using (2).

4. Transform the prior shape representation $\tilde{\phi}$, by applying a coordinate transformation (substituting (14) in (12)) with the currently estimated parameters.

5. Update ϕ using the gradient descent equation (19).

6. Update the transformation parameters α, β, γ, t_x, t_y and t_z, ψ and ξ using the derivatives of (18) with respect to each, according to (20).

7. Repeat steps 3-6 until convergence.

5. Experimental Results

To demonstrate our algorithm, we present segmentation results on various images. Relative transformation parameters between the known shape and the shape to segment have been evaluated and verified. In all experiments we set: $dt = 0.1$, $\epsilon = 1$, $\mu = 1.5$ and $d = 1$. The contributions of the first three terms in the gradient descent equation of ϕ (19) are normalized to $[-1, 1]$, thus ν is set to 1.

Consider the synthetic reference image in Fig. 2a, that contains several components of different sizes and gray levels. The prior shape has been obtained by thresholding. The image to segment is shown in Fig. 2b. The initial segmenting contour is drawn on the image. Note, that this image is a noisy[1], transformed and corrupted version of the image in Fig. 2a. Successful segmentation result is shown in Fig. 2c (red contour). The recovered transformation parameters are validated by comparison to the actual transformation parameters in Table 1. The unregistered prior contour is drawn on the image to segment in Fig. 2d, to demonstrate the misalignment between the two images, Fig. 2a-b. Fig. 2e shows segmentation without a prior shape.

We next consider two views, Fig. 3a-b, of a real object, a toy elephant. The initial segmenting contour is drawn on the image to segment in Fig. 3a. The prior shape has been constructed from the reference image, Fig. 3b. The desired segmentation result is shown (in red) in Fig. 3c. The contour precisely tracks the elephant profile, recovering the missing part of its trunk. The segmenting contour in Fig. 3c is the zero level-set of the final level-set function (ϕ) shown in Fig. 3d. Fig. 3e demonstrates the significant transformation between the two elephant views. It shows the prior contour of the reference object drawn on the image to segment. Since the ground truth is unavailable, we compared the homography matrix generated from the recovered transformation parameters, with that obtained with point correspondence method, using 13 pairs of corresponding points on the object contour, that were manually selected (Table 2). Visual verification can be seen in Fig. 3f in which the projectively transformed prior shape is shown with the final segmenting contour (same as in Fig. 3c).

[1]Zero-mean Gaussian noise with STD=0.14

In the next example, Fig. 4a, the bottom-right part of the bottle opener assimilates the background. Nevertheless, the opener is extracted precisely, as shown in Fig. 4c, using the prior shape generated from the reference image shown in Fig. 4b. In Fig. 4d, we visually verify the recovered transformation by showing the projectively transformed prior shape with the final segmenting contour. Fig. 4e demonstrates the fairly significant projective transformation between the image to segment and the prior contour shape (red). When no prior is used, as seen in Fig. 4f, the segmenting contour traces (undesirably) the gradients within the object.

Finally, Fig. 5a-b are two views of Europe from a satellite, at night and day time respectively. The images were adapted from NASA public domain: *earthobservatory.nasa.gov*. Successful segmentation is demonstrated in Fig. 5c. The coast line of the continent and the islands nearby are precisely extracted, despite the neither constant nor smooth object gray levels. Note the significant transformation, Fig. 5e, between the prior and the image to segment. The accuracy of the recovered transformation is visually verified in Fig. 5d. Segmentation when no prior is used[2], is shown for comparison in Fig. 5f. The dark areas of the eastern Adriatic shore and Scandinavia are not well segmented.

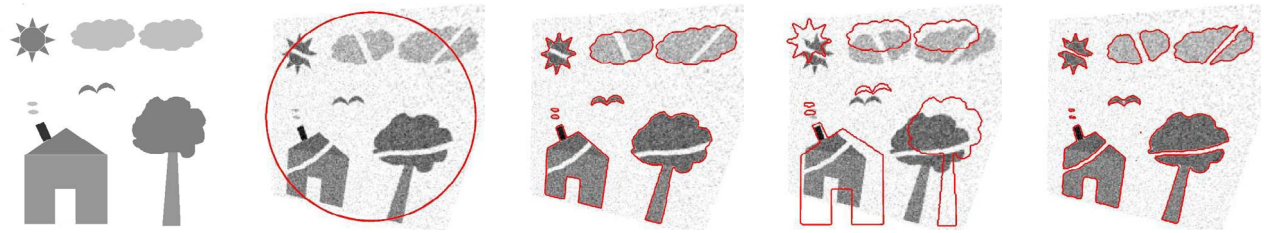
6. Summary

This paper presents a novel variational approach for object segmentation in the presence of noise, specular reflections and occlusions, given a reference image from a different view. Segmentation process is carried out concurrently with a registration of the prior instance of the object to the image to segment. The outcome of the algorithm includes the detection of the object of interest and a correct extraction of its boundaries. The planar projective transformation between the two object views is accurately recovered as well. The successful segmentation results and the reliable estimation of the transformation parameters, suggest this method as a promising tool for a wide range of applications. These include perspective-invariant search in image databases, registration and structure recovery in stereo imaging, and object tracking in video sequences.

Acknowledgments

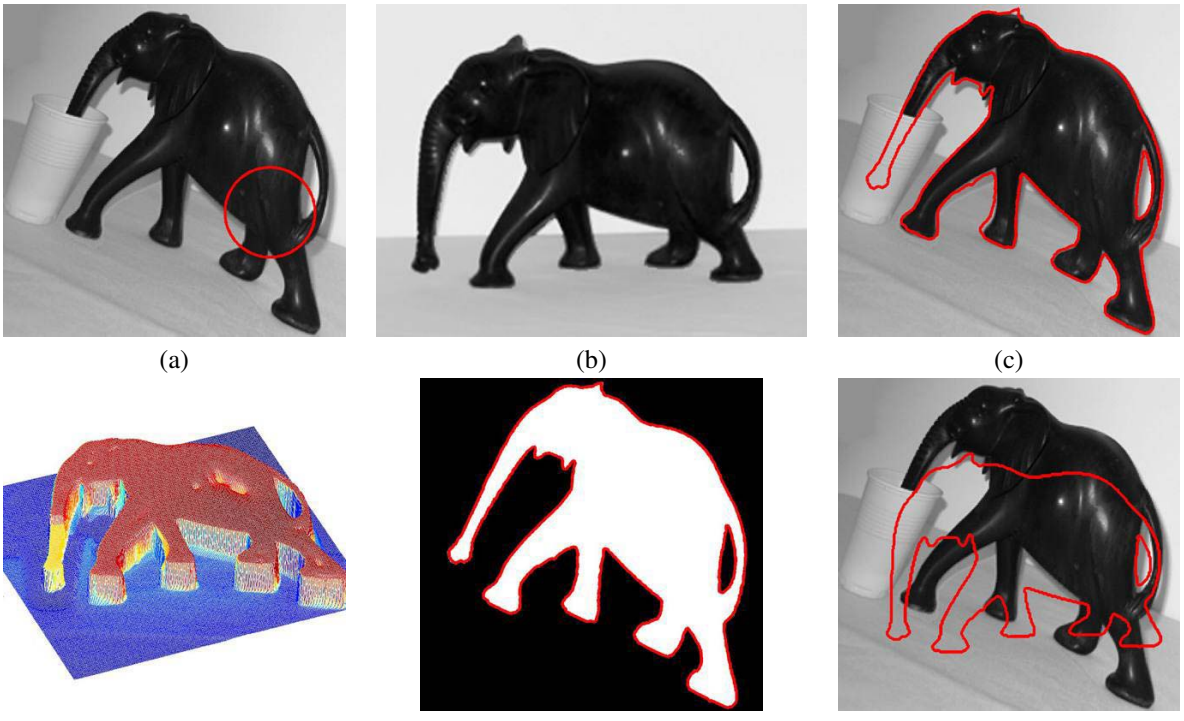
This research was supported by MUSCLE: Multimedia Understanding through Semantics, Computation and Learning, a European Network of Excellence funded by the EC 6th Framework IST Programme. It was also supported by the Israel Science Foundation.

[2]Similar segmentation result, without prior, appears in [2], Fig.11



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

Figure 2. (a) Reference image: the object is composed of several components in different sizes and gray levels. Prior shape was obtained by thresholding. (b) Noisy, transformed and corrupted version of the image in (a). The initial contour (red) is drawn on the image. (c) Correct segmentation. The final contour (red) is drawn on the corrupted, transformed and noisy image. (d) The unregistered prior shape contour (red) drawn on the input image. This demonstrates the misalignment between the reference (prior) image and the image to segment. (e) Segmentation result (red) when no prior shape is used.



(a) (b) (c)

(d) (e) (f)

Figure 3. (a) Input image (a toy elephant).The initial contour (red) used for the segmentation process is drawn on the image. (b) Reference (prior) image. (c) Successful segmentation: the final contour (red) on the projectively transformed elephant. (d) Final step of the evolving level-set function ϕ for a correct segmentation. (e) The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. This shows the compatibility between the calculated and actual transformation parameters. (f) The unregistered prior shape contour (red) drawn on the input image. This demonstrates the misalignment between the reference (prior) image and the image to segment.

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**

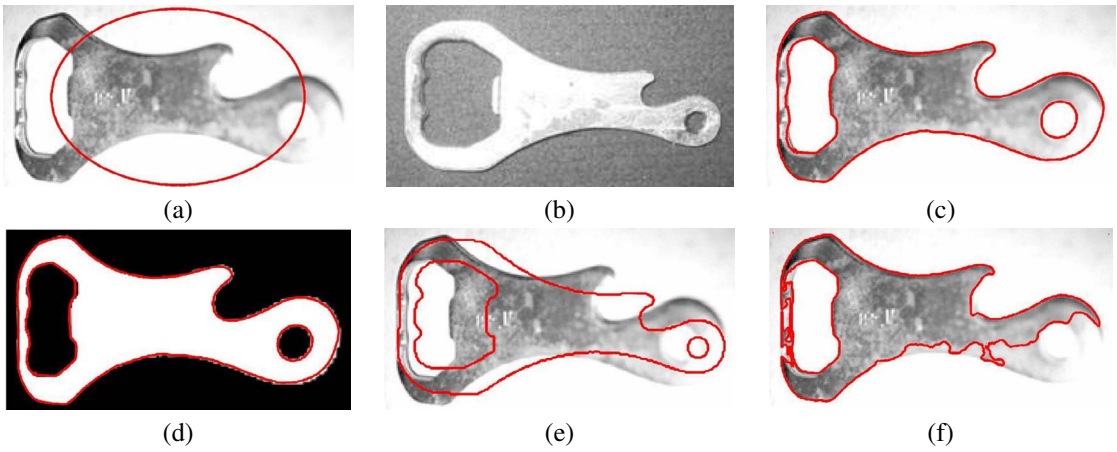


Figure 4. (a) Input image (bottle opener). The initial contour (red) used for the segmentation process is drawn on the image. (b) Reference (prior) image. (c) Successful segmentation: the final contour (red) on the projectively transformed opener. The segmentation accurately completes the missing contour part despite the specular reflection. (d) The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. This shows the compatibility between the calculated and actual transformation parameters. (e) The unregistered prior shape contour (red) drawn on the input image. (f) Segmentation result without a prior.

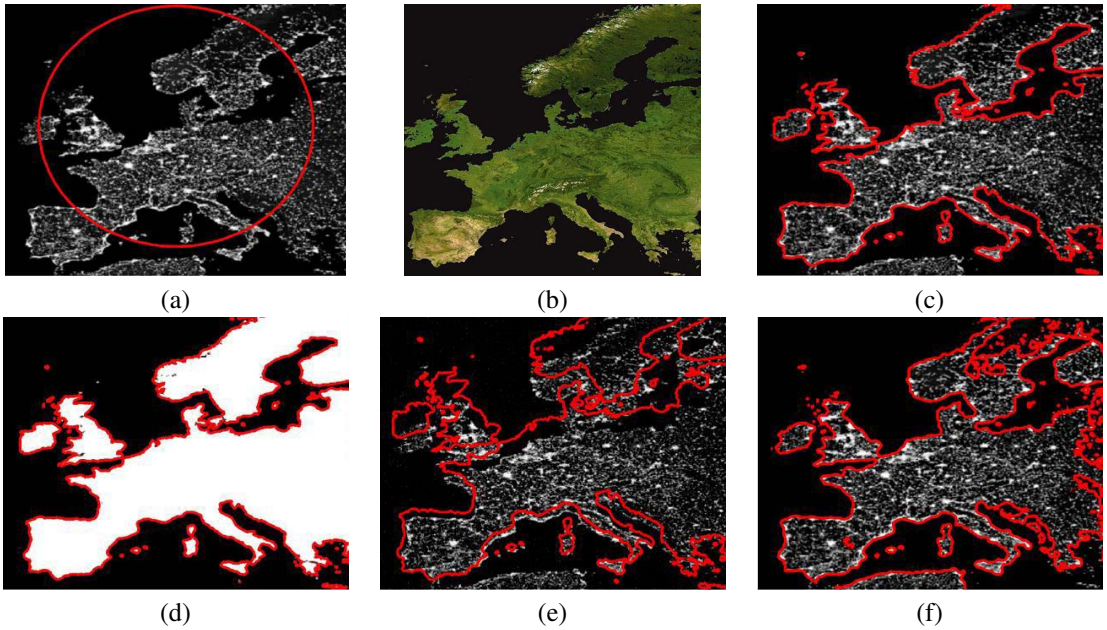


Figure 5. (a) Input image: Europe at night from a satellite. The initial contour (red) is drawn on the image. (b) Prior image. Fig.a-b were adapted from *earthobservatory.nasa.gov*. (c) Successful segmentation: the final contour (red) accurately traces the coast lines, despite the shaded parts of the continent. (d) Verification of the estimated transformation parameters: The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. (e) The unregistered prior shape contour (red) drawn on the input image. (f) Segmentation result without a prior. The dark areas of the eastern Adriatic shore and Scandinavia are not well segmented.

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**

Transformation parameters	α	β	γ	t_x	t_y	t_z	ψ	ξ
Proposed algorithm	0.069^0	0.075^0	7.4^0	-10	9.6	-0.1	0.086^0	0.14^0
Ground truth	0.075^0	0.075^0	7.5^0	-10	10	-0.1	0.075^0	0.15^0

Table 1. Comparison of the recovered and true transformation parameters, for the example in Fig. 2.

Homography matrix entries	h_{11}	h_{12}	h_{13}	h_{21}	h_{22}	h_{23}	h_{31}	h_{32}
Proposed algorithm	0.92	0.27	2.8	-0.30	0.70	33.35	-0.0003	0.0018
Point correspondence	0.91	0.28	3.2	-0.31	0.71	34.05	-0.0003	0.0019

Table 2. Comparison of the homography matrix entries obtained using the proposed algorithm and via manually selected corresponding point pairs, for the example shown in Fig. 3.

References

[1] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, 2002.

[2] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, Feb. 2001.

[3] Y. Chen, H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggs, and E.A. Geiser. Using prior shapes in geometric active contours in a variational framework. *IJCV*, 50(3):315–328, Dec. 2002.

[4] D. Cremers, T. Kohlberger, and C. Schnorr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, Sep. 2003.

[5] D. Cremers, N. Sochen, and C. Schnorr. Multiphase dynamic labeling for variational recognition-driven image segmentation. In *ECCV*, volume IV, pages 74–86, 2004.

[6] A. Duci, A. Yezzi, S. Mitter, and S. Soatto. Region matching with missing parts. In *ECCV*, volume 3, pages 48–64, May 2002.

[7] O. Faugeras, Q.T. Luong, and T. Papadopoulo. *The Geometry of Multiple Images*. MIT Press, 2001.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2003.

[9] X. Huang, Z. Li, and D. Metaxas. Learning coupled prior-shape and appearance models for segmentation. In *MICCAI*, volume I, pages 60–69, 2004.

[10] 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.

[11] M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, volume I, pages 316–323, 2000.

[12] Y. Ma, S. Soatto, J. Košecká, and S. S. Sastry. *An Invitation to 3-D Vision*. Springer-Verlag, 2003.

[13] 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.

[14] 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.

[15] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Unlevel-sets: Geometry and prior-based segmentation. In *ECCV*, volume 4, pages 50–61, 2004.

[16] M. Rousson and N. Paragios. Shape priors for level set representation. In *ECCV*, pages 78–92, 2002.

[17] C. E. Springer. *Geometry and Analysis of Projective Spaces*. Freeman, 1964.

[18] A. Tsai, A. Yezzi, Jr., W.M. Wells, III, C. Tempany, D. Tucker, A. Fan, W.E.L. Grimson, and A.S. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137–154, Feb. 2003.

Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)
1550-5499/05 $20.00 © 2005 **IEEE**