Exploring Approaches to Layered Image Registration

Ping-Feng Chen
Hamid Krim
Olga L. Mendoza*

North Carolina State University
Department of Electrical and Computer Engineering
2099 EBII, Campus Box 7911
Raleigh, NC 27695-7911

Air Force Research Lab*
Sensors Directorate

pchen@ncsu.edu, ahk@ncsu.edu

ABSTRACT

Image registration is an important fundamental process, which has many useful applications to Tracking, Automatic Target Recognition (ATR), and Sensor Fusion to name a few. Recent publications illustrate the Air Force's desire to exploit the benefits of a layered (altitude) sensing environment. In such an environment, data from different sensors, of different modalities, at different elevations of the same scene are fused to gain a better understanding of the operational environment. This research, sponsored by the Air Force Research Lab (AFRL), builds on classical registration techniques to explore novel registration algorithms applied to data under the new layered sensing environment. Our main focus, herein, is to register large-scale aerial Electro Optical (EO) images collected from cameras at different altitudes. Particularly we propose a method to jointly segment and register the same object in two layered images. A multiphase, region-based active contour method,[1, 2] together with an adapted joint segmentation-registration technique,[3] are combined to provide a way of registering layered sensing images. This method provides a *Level 0* mechanism in data fusion hierarchy to preprocess (segment) and align data for future fusion stages. In this paper, theories of multi-view geometry is reviewed for understanding the possible form of registration, our proposed method is illustrated and substantiating examples and results are provided as well.

Keywords: Image Registration, ATR, segmentation-registration, region-based active contour, multiphase level set, level set method

1. INTRODUCTION

Image registration has often been confused with data fusion. Image registration is an enabling technology for data fusion but it should not be confused with data fusion. Specifically, by registering images, the fusion of multimodality information becomes possible.[4] That is something we explain below.

Multisensor data fusion is an evolving technology, concerning the problem of how to fuse data from multiple sensors in order to make a more accurate estimation of the area of interest.[5] Applications of data fusion include environmental monitoring, ATR, tracking, battlefield surveillance, remote sensing, and global awareness. The applications are usually time-critical, cover a large geographical area, and require accurate information. Data fusion incorporates a well-established taxonomy of "levels" that group different iterative processes of differing maturity levels. Level 1 is "object refinement", which is the process of fusing data to determine the identity of entities. Level 1 is usually partitioned into four functions: data alignment, association, tracking and identification.[6] Level 2 performs "situation refinement", which is the process of fusing the spatial and temporal relationships between entities to group them together and form an abstracted interpretation of the patterns in the order of battle data. The product from this level is called the situation assessment.[7] Level 3 performs

Intelligent Computing: Theory and Applications VI, edited by Kevin L. Priddy, Emre Ertin
Proc. of SPIE Vol. 6961, 69610J, (2008) · 0277-786X/08/$18 · doi: 10.1117/12.777408

"threat refinement", which is the process of fusing the combined activity and capability of enemy forces to infer their intentions and assess the threat that they pose. The product from this level is called the threat assessment. Level 4 performs "process refinement", which is an ongoing monitoring and assessment of the fusion process to refine the process itself and to regulate the acquisition of data to achieve optimal results.[8] Level 4 interacts with each of the other levels. Due to the confusion of image registration and data fusion, a new level was introduced in 1998, Level 0, "data alignment".[7] This is the level that prepares the data for the other fusion levels, and includes data preprocessing, image registration, and image geo-registration. The product from this level is well aligned images that includes coordinate information for every pixel. This is then the input for tradational 'data fusion' algorithms, described as Level 1 - Level 4.

In this paper we propose a novel method to process data in Level 0, to segment object and align (register) data in two layered images. Particularly we use a method that combines joint segmentation-registration[3] along with multiphase active contours.[1,9] Our method first approximates the object surface that co-exists in both layered images by several planes, and then evolves a corresponding number of multiphase active contour in a variational framework. The result of this method will not only segment the object of interest but also register the object in two images. This paper is organized as follows: section 2 gives a background knowledge about the requirement of homography between two images with the same scene, section 3 first reviews the theory of joint segmentation and registration and then shows how we adapt it to our application, section 4 illustrates multiphase active contour and shows how we exploit it to jointly segment and register images as arising in our case. Section 5 provides some concluding remarks.

2. HOMOGRAPHIES BETWEEN MULTI-VIEWS

An image is a 2-D projection of a 3-D world, where one dimension is reduced and some relationship is carried out to this mapping. It is this relation that even with a reduced 2-D representation, it still truly reflects the world for human perception. Multi-view Geometry is the study of these relationships and is able to achieve several goals. For example, by picking some corresponding points in different images, the distortion of objects within the images may be rectified, or the 3-D scene may be constructed by two or three views up to a certain level of ambiguity. Multi-view geometry may assist us in the registration problem by first understand the transformation between two views.

It is known that when the data in the 3-D world is not lying on the same plane, the images taken by different cameras do not relate to each other by a single transformation, or homography; it is only when the data being projected on images reside on a plane, a homography exists between one image and another. In the following we review the camera model and bring in the idea of our registration method. The traditional camera model is a *pinhole model* where the camera is modeled as a point and the 2-D projections of world data are the intersection between the image planes (or retina plane) and the rays illuminated from the camera to the 3-D data (Fig. 1). The projection can then be described as a mapping, or a matrix multiplication by a camera matrix P, from the

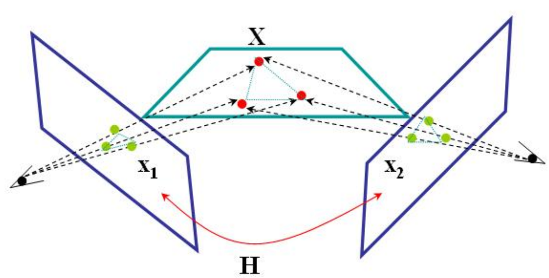


Figure 1. A *pinhole* camera model where the 2-D projections of world data are the intersections between the image planes and the rays illuminated from the camera to the 3-D data.

3-D coordinate to 2-D image. In order to talk about this model, a homogeneous coordinate system has to be introduced first. In traditional coordinate system, a point at infinity is just a mathematical idea, denoted as ∞, for help of our understanding. It helps to explain things such as the place where parallel lines intersect, and the treatment of infinite points are separate from from other ordinary points. However, after the projection of 3-D world to 2-D image, the points at infinity are tangible in the image, such as the horizon (line of infinity), or the intersection of rail ways which are no longer parallel after a perspective transformation; these points should be treated the same as every other point in the image. Therefore homogeneous coordinate was introduced where one additional dimension is added to the system. For example, an ordinary point in 2-D image would receive a 1 to have its homogeneous coordinate as $(x, y, 1)$ ($(x, y, z, 1)$ for 3-D), while points at infinity would receive a 0 as $(x, y, 0)$($(x, y, z, 0)$ for 3-D). The advantage of this representation is that now the manipulation of points at infinity can be treated as ordinary points.

The projection of 3-D world to 2-D image can now be modeled as the multiplication

$$\mathbf{x} = \mathrm{P}\mathbf{X}, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{X}$ are 3×1 and 4×1 vectors denoting 2-D and 3-D points in homogeneous coordinate and P is the 3×4 camera matrix. Other properties and important tools such as epipolar geometry and fundamental matrix in multi-view geometry are developed from this simple model.

To relate this model to the registration problem, imagine two candidate images to be registered are taken by two cameras at different positions. Registration is then a correspondence problem that one seeks the corresponding points within these two images. If there exists a simple transformation that can align the corresponding points, then this transformation is the goal. Now suppose with this pinhole model, we have two images $\mathbf{x_1} = \mathrm{P}_1\mathbf{X}$ and $\mathbf{x_2} = \mathrm{P}_2\mathbf{X}$ of the same scene $\mathbf{X}$, and we hope to find the relationship between $\mathbf{x_1}$ and $\mathbf{x_2}$. It is known that only when 3-D points $\mathbf{X}$ lie on a plane, a single homography exists between $\mathbf{x_1}$ and $\mathbf{x_2}$, and the relation can be written as

$$\mathbf{x_2} = \mathrm{H}\mathbf{x_1}, \tag{2}$$

where $\mathrm{H} = \mathrm{P}_2\mathrm{P}_1^+$ is a 3×3 transformation matrix and P^+ represents the pseudo inverse of matrix P. In general, it is shown[10] that the homography between two images of the 3-D points lying on a plane follows a *projective transformation*. Projective transformation has 8 unknown parameters, and with four-point correspondence in two images, it can be fully determined.

An interesting case and also the one we took as our assumption is that when the camera is sitting at infinity, the camera matrix P degenerates to be an affine camera[10] matrix

$$\mathrm{P}_A = \begin{bmatrix} m_{11} & m_{12} & m_{13} & t1 \\ m_{21} & m_{22} & m_{23} & t2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

with 8 free real parameters. We claim this assumption is valid since our data are images taken by cameras at very high attitude (airborne scan) and it is very close to the assumption of infinite camera. The homography induced by this assumption between two images becomes an affine transform

$$\mathbf{x_2} = \mathrm{A}\mathbf{x_1}, \tag{4}$$

where A is an 3×3 affine matrix with 6 parameters, and can be reduced to a 2×2 matrix with $\mathbf{x_1}$ and $\mathbf{x_2}$ expressing in homogeneous coordinate system. This relationship will assist us later in the technique of joint segmentation-registration.

Our strategy is then to first assume the 3-D object surface consists of a limited number of planes. This assumption is of course unrealistic for natural object, however it is very practical for artificial objects, say buildings, vehicles, or machines, which are of special interest of ATR. With this assumption, the 2-D images taken by two different cameras will have a homography relationship, particularly affine transformation, for the two projected images restricted on each corresponding plane. Secondly we would like to search the transformation parameters of each corresponding plane by evolving multiphase active contours on both images, and at the same time enforcing a relationship between each pair of corresponding active contours. The advantage of this strategy is two folded: 1.) the homography between each pair of the corresponding planes is obtained at the end of contour evolution, and 2.) the segmentation of object is at the same time being achieved. In the following section we will illustrate the joint segmentation-registration technique and then describe how we adapt it to our case.

3. JOINT SEGMENTATION-REGISTRATION

The technique of joint segmentation-registration technique was proposed by Yezzi *et al.*,[3] and the motivation is that segmentation and registration processes depends on one another. With good performance of one, the other is improved as well. The technique is carried out in a variational framework, particularly region-based active contour method. In traditional region-based active contour methods, an energy functional was proposed, and by minimizing the functional, the active contour is driven to attach on the boundaries of objects. Now suppose we have two images that are to be registered. If we evolve two contours individually, each corresponding to one image with its own energy functional, then we simply segment two images with no relationship to each other. To couple registration and segmentation, one then device the energy as the sum of the two energy functionals and enforce a relationship, in our case a homography, between the contours in two images. One then evolves the contours and registration parameters along the energy gradient flow and joint segmentation and registration will be achieved.

Below we review the theory in Ref. 3 for understanding our adaptation to layered imaging. Suppose $I : \Omega \subset \mathbf{R}^2 \rightarrow \mathbf{R}$ and $\hat{I} : \hat{\Omega} \subset \mathbf{R}^2 \rightarrow \mathbf{R}$ denote two images containing the same object to be segmented and registered. Using Chan and Vese's[9] region-based energy functional for each of the images, we have

$$E_1(C) = \int_{C_{in}} (I - u)^2(\mathbf{x})d\mathbf{x} + \int_{C_{out}} (I - v)^2(\mathbf{x})d\mathbf{x}, \tag{5}$$

$$E_2(\hat{C}) = \int_{\hat{C}_{in}} (\hat{I} - \hat{u})^2(\mathbf{x})d\mathbf{x} + \int_{\hat{C}_{out}} (\hat{I} - \hat{v})^2(\mathbf{x})d\mathbf{x}, \tag{6}$$

where C and $\hat{C}$ represent the evolving contours in two separate images, $C_{in} \subset \Omega$ and $C_{out} \subset \Omega$ ($\hat{C}_{in} \subset \hat{\Omega}$ and $\hat{C}_{out} \subset \hat{\Omega}$) represent the regions inside and outside the contour C ($\hat{C}$) separately, and $u \in \mathbf{R}$ and $v \in \mathbf{R}$ ($\hat{u}$ and $\hat{v}$) represent the mean of I ($\hat{I}$) inside and outside the contour C ($\hat{C}$). The energy models the image as the simplest case, piecewise constant function, and the edges correspond to the place where the constants change. This energy functional performs well for more complicated images or images with noise and not necessarily has to satisfy the piecewise constant condition. Evolving the contours C and $\hat{C}$ by minimizing E_1 and E_2 separately, we end up with two gradient flows

$$\frac{\partial C}{\partial t} = \left((I-u)^2 - (I-v)^2\right) N \quad \text{and} \quad \frac{\partial \hat{C}}{\partial t} = \left((\hat{I}-\hat{u})^2 - (\hat{I}-\hat{v})^2\right) \hat{N}, \tag{7}$$

where t is an artificial time parameter and N ($\hat{N}$) is the unit outward normal to the contour C ($\hat{C}$).

To couple segmentation and registration, we then enforce a relationship between the two contours $\hat{C} = g(C)$, where g is some transformation with parameters g_i, $i = 1, \ldots, n$, and then sum the energy functionals (5) and (6). This transformation g is then the registration parameter we would like to find out

$$\begin{aligned} E(g, C) &= E_1(C) + E_2(g(C)) \\ &= \left(\int_{C_{in}} f_{in}(\mathbf{x})d\mathbf{x} + \int_{C_{out}} f_{out}(\mathbf{x})d\mathbf{x}\right) + \left(\int_{\hat{C}_{in}} \hat{f}_{in}(\mathbf{x})d\mathbf{x} + \int_{\hat{C}_{out}} \hat{f}_{out}(\mathbf{x})d\mathbf{x}\right), \end{aligned}$$

where we have replaced $(I-u)^2$ by f_{in} and $(I-v)^2$ by f_{out} ($(\hat{I}-\hat{u})^2$ by $\hat{f}_{in}$ and $(\hat{I}-\hat{v})^2$ by $\hat{f}_{out}$) for brevity. This energy functional may be re-expressed only within the domain Ω where the first image resides as

$$E(g, C) = \int_{C_{in}} (f_{in} + |g'|\hat{f}_{in} \circ g)(\mathbf{x})d\mathbf{x} + \int_{C_{out}} (f_{out} + |g'|\hat{f}_{out} \circ g)(\mathbf{x})d\mathbf{x}, \tag{8}$$

where g' denotes the Jacobian of g, $|\cdot|$ is the determinant, and $\circ$ represents a function composition.

Segmentation and registration are then successfully coupled together and by minimizing the energy (8). The gradient flow for the contour and registration parameters are shown to be

$$\frac{\partial C}{\partial t} = (f + |g'|\hat{f} \circ g)N, \quad \text{where} \quad f = (f_{in} - f_{out}) \quad \text{and} \quad \hat{f} = (\hat{f}_{in} - \hat{f}_{out}), \tag{9}$$

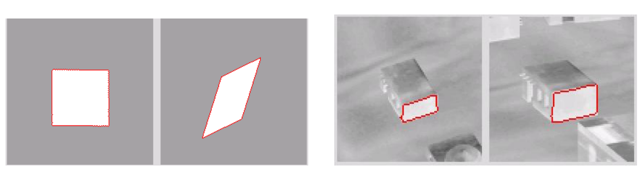
$$\frac{\partial g_i}{\partial t} = \int_C \hat{f}(g(\mathbf{x})) \left\langle \frac{\partial}{\partial g_i} g(\mathbf{x}), ((g')^{-1}|g'|)^T N \right\rangle ds. \tag{10}$$

To adapt this technique to our application, we recall that in the previous section we illustrate that homography exists between two images only if the 3-D points to be imaged lie in a plane. Suppose the object which exists in both images indeed consists of only a plane (of course this is an unrealistic assumption, and thus in the next section we illustrate the multi-phase active contour in combination with joint segmentation-registration to provide a solution), then the homography between two images for the interior region of the object complies a projective transformation. We can then enforce the relationship g, as a projective transformation, between the evolving contours C and $\hat{C}$ within the two images. However, a projective transformation has 8 degrees of freedom, and every element in the projection matrix in Eq. (2) is a free parameter except for a scaling factor.[10] Enforcing a projective transformation is equivalent to not restricting any relationship between contours, and it is the same as evolving two contours separately. Therefore we upgrade the transformation to be affine transformation with the assumption illustrated in previous section that the camera is sitting at very high attitude and the illuminating rays are parallel. An affine transformation can be written as the composition of two rotations, a non-isotropic scaling and a translation, with 6 free parameters

$$\begin{aligned} g(\mathbf{x}) &= R(\theta)R(-\phi)S(\lambda_1, \lambda_2)R(\phi)\mathbf{x} + D(x_0, y_0) \\ &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \end{aligned} \tag{11}$$

where this transformation $g(\mathbf{x})$ corresponds to the affine matrix A in Eq. (4). An intuitive interpretation of this affine transform is to rotate $\mathbf{x}$ with an angle ϕ and stretch it in x and y direction with different amount, then rotate it back and then rotate another angle θ, and finally add a translation x_0 and y_0. The evolution of the contour and each of the 6 registration parameters will then follow the gradient flows in Eq. (9) and (10). One advantage of this method we would like to point out is that even though the transformation g is only restricting on the two contours C and $\hat{C}$, it actually represents the transformation for the whole interior regions of the two. In other words, once the evolution is done, we may apply the resulting transformation on the whole interior region of one contour C in alignment to the interior of another. Thus a Level 0 alignment process is achieved, and further fusion step may be carried out from there on.

Fig. 2(a) shows the joint segmentation-registration result for simulated images, and Fig. 2(b) shows the result of segmenting and registering one side plane of a building. The results successfully segmented the images and captured the registration parameters. In the next section we will review the theory of multi-phase active contour method and show how to combine the two to jointly segment and register an object which consists of more than one plane.



(a) (b)

Figure 2. (a.) Simulated images with a square on the left and its affine transformation on the right. The contours successfully segmented both images and captured the registration parameters, (b.) Joint segmentation-registration result for a layered EO image (grey scale image). It captured the side plane of a building.

4. MULTI-PHASE ACTIVE CONTOUR WITH JOINT SEGMENTATION-REGISTRATION

Multi-phase active contour method has been proposed in several literatures using different analysis.[1, 2, 11, 12] The main difference between them is that if the contours intersect each other at the end of evolution. Ref. 2 allows the intersections of contours, so for example, two contours would end up with four regions (suppose the two contours are C_u and C_v, then four regions: $C_{u_{in}} \cap C_{v_{out}}$, $C_{v_{in}} \cap C_{u_{out}}$, $C_{u_{in}} \cap C_{v_{in}}$, and $C_{u_{out}} \cap C_{v_{out}}$), Ref. 11 does not allow the intersections of contours, thus resulting in three regions ($C_{u_{in}} \cap C_{v_{in}} = \varnothing$), and Ref. 1 devised the energy functional by a model with the assumption of no intersections of contours. We would not like to have intersections between contours because of the following reason: since we will use joint segmentation-registration technique combining multi-phase active contour, each contour in one image is enforced by a transformation to another in the second image; furthermore, the transformation represents a mapping for the *whole interior region* of the contour, therefore we should not allow intersections for the resulting contours. The theory in Ref. 1 is elegant and easier to implement, therefore we first review the theory and then adapt it to our application.

We look at a ternary flow as an example first, and then a general multi-phase flow can be derived similarly. Suppose an image contains two disjoint, simply connected foreground regions, R_a and R_b, and a background region R_c, with their intensities I^a, I^b and I^c respectively. An initial closed contour C_u generally encloses some portion of each region, and the mean intensity u inside the contour C_u can be written as a convex combination of I^a, I^b and I^c, i.e. $u = \alpha I^a + \beta I^b + \gamma I^c$, where $0 \leq \alpha, \beta, \gamma \leq 1$ and $\alpha + \beta + \gamma = 1$. However, if $I \in \mathbf{R}$, this combination is not unique, and only if $I \in \mathbf{R}^n$, where $n \geq 2$, the combination has a unique solution. Therefore the image has to be a multichannel data, say a RGB colorful image, and the intensities representing each region become $I^a = (I^a_1, I^a_2, \ldots, I^a_n)$, $I^b = (I^b_1, I^b_2, \ldots, I^b_n)$, and $I^c = (I^c_1, I^c_2, \ldots, I^c_n)$. The mean intensity of the interior of the first and the second contour, C_u and C_v, has the same vector representation $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$, and the third vector $w = (w_1, w_2, \ldots, w_n)$ denotes the mean intensity of the mutual exterior region of C_u and C_v. The intensity of the three regions R_a, R_b and R_c forms a triangle T_{abc} because of their geometrical independence, and the intensity of u, v and w forms another triangle T_{uvw} lying within T_{abc}. It is only when the two contours C_u and C_v correctly capture the boundary ∂R_a and ∂R_b that the triangle T_{uvw} coincides with T_{abc}. It is easy to see that when pulling the mutual distances between u, v and w most apart, the segmentation is achieved. In other words, segmentation is driven by maximizing the area expended by vertices u, v and w, or the area of triangle T_{uvw}.

The energy functional can then be constructed as the negative area of triangle T_{uvw}, plus a traditional regularizing term to penalize excessive length of the two contours

$$E = -2\text{area}^2(T_{uvw}) + \alpha \left(\int_{C_u} ds + \int_{C_v} ds \right), \tag{12}$$

where $\alpha \in \mathbf{R}$, and $\text{area}^2(T_{uvw}) = (||u-w||^2||v-w||^2 - ((u-w)\cdot(v-w))^2)/4$. The gradient flows for this energy are shown to be[1]

$$\frac{\partial C_u}{\partial t} = \left\{ \sum_{i=1}^{n} \left(\bar{w}_i \frac{I_i - u_i}{A_u} - \bar{v}_i (1-\chi_v) \frac{I_i - w_i}{A_w} \right) - \alpha\kappa_u \right\} N_u, \tag{13}$$

$$\frac{\partial C_v}{\partial t} = \left\{ \sum_{i=1}^{n} \left(\bar{u}_i \frac{I_i - v_i}{A_v} - \bar{v}_i (1-\chi_u) \frac{I_i - w_i}{A_w} \right) - \alpha\kappa_v \right\} N_v, \tag{14}$$

where A_u, A_v and A_w represent the area of the region interior of C_u, C_v and the mutual exterior of the two, χ_u and χ_v are the characteristic function of the interior of C_u and C_v, κ_u and κ_v represents the curvatures, N_u and N_v the outward normal to C_u and C_v, and $\bar{u}$, $\bar{v}$, $\bar{w}$ are defined as follows

$$\begin{array}{lll} \bar{u} = \tilde{u} - \tilde{v} & \tilde{u} = \hat{u}(\hat{v}\cdot\hat{w}) & \hat{u} = u - v \\ \bar{v} = \tilde{v} - \tilde{w} & \tilde{v} = \hat{v}(\hat{w}\cdot\hat{u}) & \hat{v} = v - w \\ \bar{w} = \tilde{w} - \tilde{u} & \tilde{w} = \hat{w}(\hat{u}\cdot\hat{v}) & \hat{w} = w - u. \end{array} \tag{15}$$

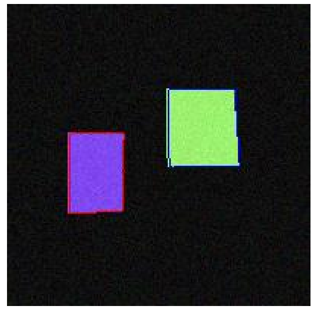


Figure 3. 3-phase active contour segmentation result for a simulated RBG image.

To generalize this formulation, a multi-phase flow will then have m contours, and the mean intensities for each of the m interior regions form a polygon. The segmentation will be achieved by maximizing the area of the polygon expended by m vertices. A 3-phase segmentation result for a simulated RGB image is shown in Fig. 3.

To combine multi-phase flow method with the joint segmentation-registration technique, similar to the rationale in previous section, all we have to do is to enforce a relationship between each pair of m active contours and then sum the energy functionals of the two images. Take ternary flows for example, we observe that the energy functional corresponding to Eq. (13) and (14) can also be written as (ignoring the penalizing term for brevity)

$$E_1(C_u, C_v) = \int_\Omega \{f_1(C_u, C_v)\chi_u + f_2(C_u, C_v)(1-\chi_v)(1-\chi_u) + f_3(C_u, C_v)\chi_v\}(\mathbf{x})d\mathbf{x}, \tag{16}$$

$$E_2(\hat{C}_u, \hat{C}_v) = \int_{\hat{\Omega}} \left\{\hat{f}_1(\hat{C}_u, \hat{C}_v)\hat{\chi}_u + \hat{f}_2(\hat{C}_u, \hat{C}_v)(1-\hat{\chi}_v)(1-\hat{\chi}_u) + \hat{f}_3(\hat{C}_u, \hat{C}_v)\hat{\chi}_v\right\}(\mathbf{x})d\mathbf{x}, \tag{17}$$

where $\Omega \subset \mathbf{R}^2$ ($\hat{\Omega} \subset \mathbf{R}^2$) is the space where the first (second) image resides, χ_u and χ_v ($\hat{\chi}_u$ and $\hat{\chi}_v$) are the characteristic functions of the interior of C_u and C_v (the interior of $\hat{C}_u$ and $\hat{C}_v$), and f_1, f_2 and f_3 ($\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$) are defined as

$$\begin{array}{ll} f_1 = \sum_{i=1}^n \left(\bar{w}_i \frac{I_i - u_i}{A_u}\right), & \hat{f}_1 = \sum_{i=1}^n \left(\bar{w}_i' \frac{I_i' - u_i'}{A_u'}\right), \\ f_2 = \sum_{i=1}^n \left(\bar{v}_i \frac{I_i - w_i}{A_w}\right), & \hat{f}_2 = \sum_{i=1}^n \left(\bar{v}_i' \frac{I_i' - w_i'}{A_w'}\right), \\ f_3 = \sum_{i=1}^n \left(\bar{u}_i \frac{I_i - v_i}{A_v}\right), & \hat{f}_3 = \sum_{i=1}^n \left(\bar{u}_i' \frac{I_i' - v_i'}{A_v'}\right), \end{array} \tag{18}$$

where $\bar{u}$, $\bar{v}$, $\bar{w}$, A_u, A_v and A_w are as defined in Eq. (13) - (15) and $\bar{u}'$, $\bar{v}'$, $\bar{w}'$, A_u', A_v' and A_w' are similarly defined for the second image.

Now to couple two ternary flows in two images, we restrict a relation between each pair of contours, $\hat{C}_u = g_1(C_u)$ and $\hat{C}_v = g_2(C_v)$, and sum the energies in Eq. (16) and (17)

$$\begin{aligned} E(g_1, g_2, C_u, C_v) &= E_1(C_u, C_v) + E_2(g_1(C_u), g_2(C_v)) \\ &= \int_\Omega \Big\{\Big(f_1 + |g_1'|\hat{f}_1 \circ g_1\Big)\chi_u + \Big(f_2 + |g_1'||g_2'|\hat{f}_2 \circ (g_1, g_2)\Big)(1-\chi_u)(1-\chi_v) \\ &\quad + \Big(f_3 + |g_2'|\hat{f}_3 \circ g_2\Big)\chi_v\Big\}(\mathbf{x})d\mathbf{x}, \end{aligned} \tag{19}$$

where g_1' and g_2' denote the Jacobian of transformations g_1 and g_2, $|\cdot|$ represents the determinant and we assume $C_{u_{in}} \cap C_{v_{in}} = \varnothing$, otherwise the second term involves nonlinear operation and it will be hard to express it in an analytic form. The evolutions of contours C_u and C_v can then be shown as

$$\frac{\partial C_u}{\partial t} = \left\{\left(f_1 + |g_1'|\hat{f}_1 \circ g_1\right) - \left(f_2 + |g_1'||g_2'|\hat{f}_2 \circ (g_1, g_2)\right)(1-\chi_v) - \alpha\kappa_u\right\} N_u, \tag{20}$$

$$\frac{\partial C_v}{\partial t} = \left\{\left(f_3 + |g_2'|\hat{f}_3 \circ g_2\right) - \left(f_2 + |g_1'||g_2'|\hat{f}_2 \circ (g_1, g_2)\right)(1-\chi_u) - \alpha\kappa_v\right\} N_v, \tag{21}$$

and the evolution of registration parameters g_{1_i} and g_{2_i} are

$$\frac{\partial g_{1_i}}{\partial t} = \int_{C_u} (\hat{f}_1 - \hat{f}_2) \left\langle \frac{\partial}{\partial g_i} g_1(\mathbf{x}), \left((g_1')^{-1}|g_1'|\right)^T N_u \right\rangle ds, \tag{22}$$

$$\frac{\partial g_{2_i}}{\partial t} = \int_{C_v} (\hat{f}_3 - \hat{f}_2) \left\langle \frac{\partial}{\partial g_i} g_2(\mathbf{x}), \left((g_2')^{-1}|g_2'|\right)^T N_v \right\rangle ds, \tag{23}$$

where in Eq. (20) and (21) we include the curvature penalizing term.

Now with the adapted technique, the whole procedure to joint segment and register two layered image is as follows: 1.) approximate the revealing part of the object surface in both images by m planes; this knowledge is assumed as *a priori* in this paper, 2.) apply m-phase active contours on both images, each with its own energy functional analogous to Eq. (16) and (17), 3.) constrain an affine transformation relationship to each pair of contours, sum the two energy functionals, and evolve the contours and affine transformation parameters by minimizing the summed energy functional.

Fig. 4 shows the 3-phase active contour combining joint segmentation-registration result of a pair of simulated RGB images. It shows that the contours successfully captured the object boundaries in both images and the registration parameters. Fig. 5(a) shows a pair of layered EO RGB images, where the revealing part of the object consists of only 2 planes (the roof and the front side), and our methods works well even though the segmentation results has a little error of protrusion. Fig. 5(b) shows another pair of layered EO RGB images where the revealing part of object consists of 3 planes. Theoretically we should use 4-phase active contour to segment it, however a 3-phase active contour can capture 2 planes of the object (roof and side plane) and the registration parameters as well.

5. DISCUSSION AND CONCLUSION

In this paper we proposed a novel registration method in Level 0 of data fusion hierarchy. This method is a combination of multi-phase active contour and joint segmentation-registration techniques, where we relate each pair of contours in two layered images by an affine transformation, and evolve the contours in both images to segment the object. The co-existing part of the object has to be first approximated by a *a priori* number of planes. There are also some restrictions of our method. It the revealing part of the object in two images consists of different number of planes, our method will not be suitable or should be modified accordingly (ex: evolve only the co-existed number of contours to segment partial object planes). Moreover, if the plane in one image is blocked by other scene but it's not blocked in another image, then our method will not work. In conclusion, our method successfully segments each revealing plane of the object and capturs the registration parameters. It not only registers by also segments the object in layered images, hence allowing one to align the data taken by different sensors and any fusion technique may be studied and carried out hereafter in the future.

REFERENCES

1. A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *Journal of Visual Communication and Image Representation* **13**, 2002.
2. L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision* **50**, pp. 271–293, Dec. 2002.
3. A. Yezzi, L. Zollei, and T. Kapur, "A variational framework for integrating segmentation and registration through active contours," *Medical Image Analysis* **7**, pp. 171–185, 2003.
4. A. A. Goshtasby, *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*, Wiley-Interscience, 2005.

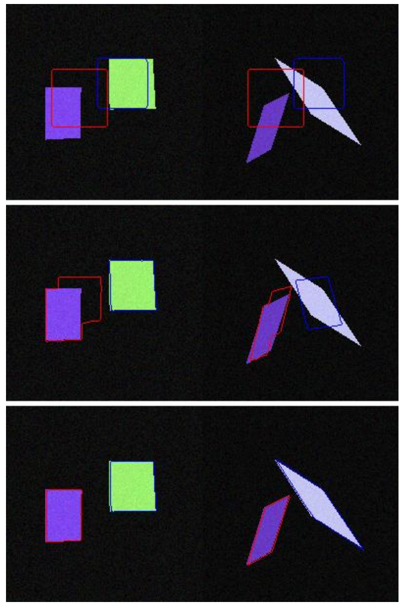
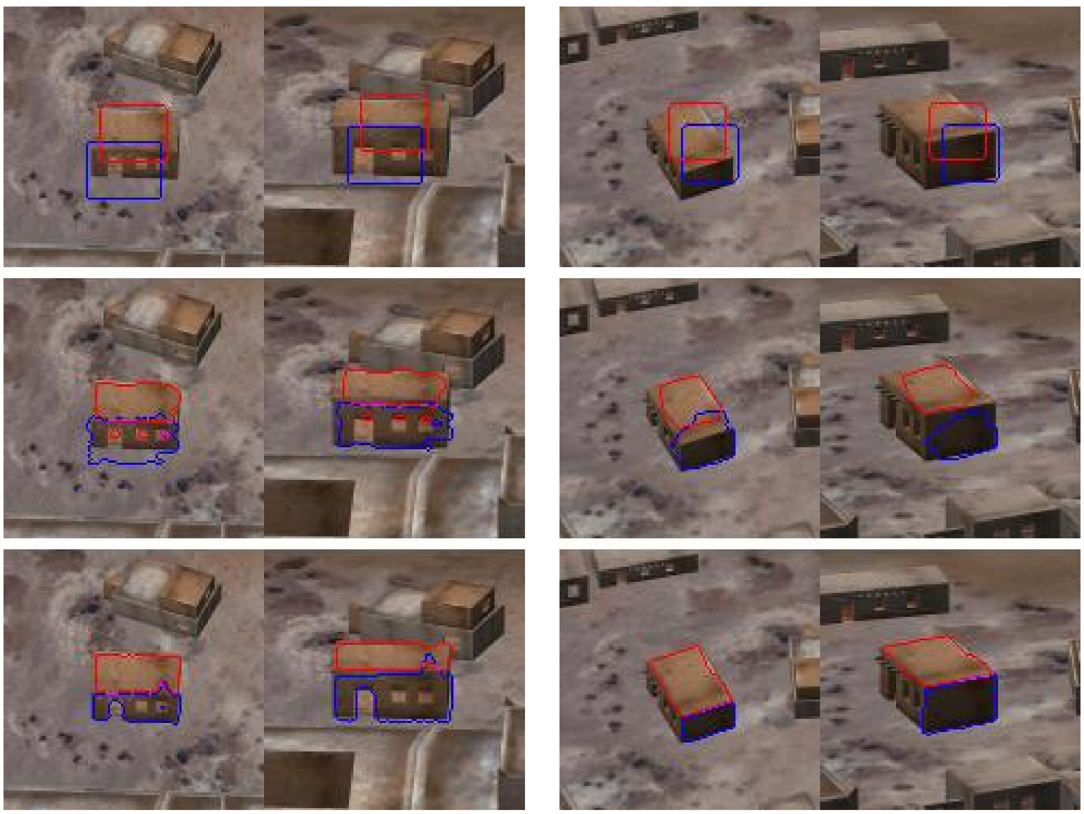


Figure 4. 3-phase active contour combining joint segmentation-registration result for a pair of simulated RGB images. On the left are the original images and on the right are affine transformed ones.

5. H. Qi, X. Wang, S. Iyengar, and K. Chakrabarty, "Multisensor data fusion in distributed sensor networks using mobile agents," in *International Conference on Information Fusion*, pp. pp. 11–16, August 2001.
6. D. L. Hall, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Inc., Norwood, MA, USA, 1992.
7. D. D. Science and T. Organization), "Data fusion lexicon," in *DSTO (Defence Science and Technology Organization) Data Fusion Special Interest Group Department of Defence*, p. pp. 21, September 1994.
8. L. A. Klein, *Sensor and Data Fusion Concepts and Applications*, Society of Photo-Optical Instrumentation Engineers (SPIE), Bellingham, WA, USA, 1993.
9. T. Chan and L. Vese, "Active contours without edges," *Image Processing, IEEE Transactions on* **10**(2), pp. 266–277, Feb 2001.
10. R. Hartley and A. Zisserman, *Multi-View Geometry in Computer Vision*, Cambridge Univ. Press, 2000.
11. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *J. Comput. Phys.* **127**(1), pp. 179–195, 1996.
12. G. Unal, A. Yezzi, and H. Krim, "Information-theoretic active polygons for unsupervised texture segmentation," *International Journal of Computer Vision* **62**, pp. 199–220, 2005.



(a) (b)

Figure 5. (a.) (b.) the result for two pairs of layered EO images (RGB image).