USSR: A Unified Framework for Simultaneous Smoothing, Segmentation, and Registration of Multiple Images

Nicholas A. Lord  Jeffrey Ho  Baba C. Vemuri
Department of Computer and Information Sciences and Engineering
University of Florida
Gainesville, FL 32611 USA
email:  nal,jho,vemuri@cise.ufl.edu

Abstract

Image smoothing, segmentation and registration are three key processing steps in many computer vision applications. In this paper, we present a novel framework for achieving all three seemingly disparate goals simultaneously across multiple images in a unified framework via a single variational principle. The proposed method ensures that the estimated registration is unbiased and all compositions of registration maps are compatible. The solution to the variational problem is achieved efficiently by solving a coupled system of partial differential equations over the common domain on which the registration maps are defined. The effectiveness of the proposed framework is demonstrated on sets of real images.

1. Introduction and Related Work

Smoothing, segmentation, and registration are the most important image processing operations (IPOs) in many computer vision applications, particularly medical imaging. Accordingly, an inscrutably vast body of literature now exists for each of the three subfields: some survey papers include [5, 6, 15]. Throughout their development, these problems have traditionally been considered independently of one another. Yet, the idea that the solution of one might benefit from coupling to the solution of another is nothing new: the natural interdependence of segmentation and smoothing was elucidated early on in the seminal paper by Mumford and Shah[7], wherein an optimal image reconstruction is defined in terms of a segmenting curve set specifying the boundaries of the smooth regions. In the curve evolution implementation of minimization of the Mumford-Shah functional by Tsai et al.[11], the mutual assistance of segmentation and smoothing is made explicit: a preliminary solution to each alternately serves as input to the other. As such, tentative smooth reconstructions enable the curve to find object boundaries, and the identification of object boundaries by the curve enables the piecewise smooth reconstruction to remain maximally faithful to the initial data.

In recent years, we have seen a similar framework arise for segmentation and registration. Yezzi et al. [1] were the first to motivate an explicit interdependence: registration methods can make use of the feature detection inherent in segmentation, while segmentation can utilize the redundancy provided by correctly registered images of the same structures to segment those structures jointly more robustly than individually. This is the relationship exploited by model-based segmentation schemes, e.g. [13, 3]. This proposal involved the unification of mutual information(MI)-based flat 2D rigid image registration with a level set implementation of a piecewise constant segmentation scheme (e.g., [2]) through a variational principle. Since then, alternate approaches and extensions to the cases of nonrigid registration, nonflat surfaces, and/or multiple (≥ 2) input images have included Wyatt et al. [16], Xiaohua et al. [17], Unal et al. [12], F. Wang et al. [14], and Young et al. [18].

We have already noted the prior motivation of integrating smoothing with segmentation, and segmentation with registration. But smoothing is a preprocessing step for other IPOs, with registration being no exception. Synthesizing all of these observations, we produce the system relationship depicted in Fig. 1. All dependences are direct with the exception of that of smoothing on registration. However, considered in the context of simultaneous segmentation, this relationship clearly emerges. At this point, then, it is natural to inquire into the possibility of unifying all three image processing operations under a single framework that achieves all goals simultaneously while exploiting their interdependence. Yet, proposals of this nature are recent and limited in number: references which could be considered to fall into this category are Vemuri et al.[13], Pohl et al.[9], and Unal et al.[12]. Moreover, each of the works just cited contains at least one of the following restrictions, simplifi-
cations, or limitations: (1) reliance on explicit atlasing, (2) low-dimensional registration parametrization (i.e. rigid or affine), (3) lack of guaranteed symmetry/unbiasedness with respect to input arguments or consistency (through composition) of registration maps, (4) reliance on piecewise constant reconstruction model, and/or (5) restriction to image pairs (rather than larger sets). In this paper, we present a unified variational framework (dubbed USSR) for simultaneous segmentation, (piecewise) smoothing, and (nonrigid) registration of image suites ($K \geq 2$) which eliminates all of the aforementioned limitations. The energy functional of the proposed method is the combination of two concept components (besides regularization): (1) an intensity-based matching penalty on the evolving registration maps, where the intensities are drawn from Mumford-Shah based reconstruction estimates for the input images, and (2) a sum of Mumford-Shah functionals on each of the input images, where the segmenting curve in the functional is carried between the images by the current estimates of the registration maps.

The simultaneity and interdependence are clearly present in the proposed scheme. However, the allowance of multiple input images raises the additional question of registration representation and constraint. As all pairs will be brought into correspondence, it is natural to require that all of these pairwise registrations be compatible (i.e. consistent with respect to composition). It is also natural to require that the registration be unbiased with respect to the input images: permutation of the inputs should not alter the result. Note that this requirement is violated by any method which chooses an atlas from the input set. In fact, with sporadic exceptions such as the “geometrized” approach of [10], theoretical symmetry of registration algorithms is an understudied issue, even for image pairs. Through appropriate casting of the problem (onto a single shared canonical domain $D$), we are able not only to theoretically ensure a result which is symmetric, unbiased, and compatible, but to solve a single joint smoothing and segmentation problem on the common domain, rather than $K$ different problems on the input image domains.

The “triangular” relationship between smoothing, segmentation, and registration.

Figure 1: The “triangular” relationship between smoothing, segmentation, and registration.

2. Variational Framework

In this section, we present the USSR framework: a unified variational framework for simultaneous smoothing, segmentation and registration of multiple images. Let $\{I_1, \ldots, I_K\}$ denote the set of $K$ input images. The outputs of the USSR algorithm are

1. A collection $\Phi$ of registration maps (diffeomorphisms) $\Phi_{ij}$ between $I_j$ and $I_i$, $1 \leq i, j < K$: $\Phi_{ij} : I_j \rightarrow I_i$ and $\Phi_{ji}^{-1} = \Phi_{ji}$.
2. A collection of smoothed images $\hat{I}_i$, $1 \leq i \leq K$.
3. A collection of segmenting contours $C_1, \cdots, C_K$ on $I_1, \cdots, I_K$, respectively.

The outputs are subject to the following compatibility constraints,

1. $C_i = \Phi_{ij}(C_j)$, for $1 \leq i, j \leq K$.
2. $\Phi_{ij} = \Phi_{ik} \Phi_{kj}$.
3. $\Phi_{ii} = \text{id}$ is the identity map for all $i$.

The compatibility constraints ensure the consistency among all $K$ images of the results of all three operations given the registration maps $\Phi_{ij}$ and the segmenting contour $C_i$. A general form of the energy functional can be written down immediately as

$$E(\Phi, C, \hat{I}) = \sum_{i=1}^{K} \int_{I_i} |I_i - \hat{I}_i|^2 dA + \sum_{i=1}^{K} \int_{I_i \setminus C_i} |\nabla \hat{I}_i|^2 dA + \sum_{i} \text{Length}(C_i) + \text{MatchingCost}(\Phi, \hat{I}) + \text{Regularization}(\Phi).$$

The first three (summation) terms represent the combination of Mumford-Shah functionals of each of the input images. As discussed previously, these terms represent an integrated approach to assessing the quality of smoothing and segmentation on each image. The last two terms measure the quality of registration and its smoothness. The interaction between the registration and segmentation is explicit in the compatibility constraint above, while the interaction between the registration and smoothing is given by $\text{MatchingCost}(\Phi, \hat{I})$, which measures the quality of the registration using the smoothed versions of the input images. The interdependence of the three operations thus appears naturally.

A direct approach to solving this constrained variational problem is difficult. However, a slight adjustment in perspective allows for a much more efficient computational strategy (Figure 2). Instead of computing the $K^2$ registration maps $\Phi_{ij}$ and $K$ contours $C_i$, we will compute $K$...
diffeomorphisms $\Phi_i$ between $I_i$ and a canonical domain $D$ and a segmenting contour $C$ on $D$. This simplifies the computation enormously while simultaneously satisfying all of the constraints mentioned above: the registration $\Phi_{ij}$ is defined as $\Phi_i \circ \Phi_{ij}^{-1}$ and the segmenting contour on $I_i$ is defined as $C_i = \Phi_i(C)$.

Figure 2: Unbiasedness and compatibility can be guaranteed by performing all computations on a canonical domain $D$.

Mathematically, the canonical domain $D$ and the $K$ registration maps provide parameterizations for the $K$ input images with a common domain. Computations on each individual image can now be covariantly formulated on $D$ using the concept of pullbacks. For instance, the integral on the left
\[
\int_{I_i \cap C_i} |\nabla I_i|^2 dx dy = \int_D |\nabla I_i(\Phi_i)|^2 J_{\Phi_i} dx dy
\]

is the pullback of the Euclidean metric $g$ as the images. For the matching cost term. Finally, the length of the contour $C_i$ is given as $\Phi_i(C)$ can be computed directly on $D$. Let $\gamma(t)$ denote a parametrization of the contour $C$: the length of the contour $C_i = \Phi_i(C)$ is then given by the integral
\[
\int_C ds_{\Phi_i} = \int \sqrt{<d\Phi_i(\gamma), d\Phi_i(\gamma)>} dt,
\]

where $d\Phi_i$ is the Jacobian of $\Phi_i$. The full energy function is thus given as
\[
E(\Phi, C) = \sum_{i=1}^{K} \int_D |I_i - \hat{I}_i|^2 J_{\Phi_i} dx + \beta_1 \sum_{i=1}^{K} \int_D |\nabla I_i(\Phi_i(x))|^2 J_{\Phi_i} dx + \beta_2 \sum_{i=1}^{K} \int_D |\hat{I}_i - \frac{\sum_{j=1}^{K} \hat{I}_j}{K}|^2 J_{\Phi_i} dx + \beta_3 \sum_{i=1}^{K} \int_C ds_{\Phi_i} + \beta_4 \text{Reg}(\Phi).
\]

3. Computational Framework Details

Working in the canonical domain $D$ ensures the satisfaction of all compatibility constraints, and the absence of bias towards any image or images. To solve the variational problem, we follow the standard approach of using appropriate gradient flows. In the following, $g_i$ will denote the Riemannian metric on $D$ given as the pullback of the Euclidean metric on $I_i$ by the diffeomorphism $\Phi_i$. The gradient flow with respect to the contour is given by
\[
\frac{\partial C}{\partial t} = \frac{1}{2} \sum_{i=1}^{K} \left( (I_i - \hat{I}_i)^{out} - (I_i - \hat{I}_i)^{in} \right)^2 J_{\Phi_i} \hat{N}
\]
\[
+ \frac{\beta_1}{2} \sum_{i=1}^{K} \left( \|\nabla g_i \hat{I}_i^{out} \|_{g_i}^2 - \|\nabla g_i \hat{I}_i^{in} \|_{g_i}^2 \right) J_{\Phi_i} \hat{N}
\]
\[
+ \frac{\beta_2}{2} \sum_{i=1}^{K} \left( \|\hat{I}_i^{out} - \frac{\sum_{j=1}^{K} \hat{I}_j^{out}}{K} \|_{\hat{N}}^2 - \|\hat{I}_i^{in} - \frac{\sum_{j=1}^{K} \hat{I}_j^{in}}{K} \|_{\hat{N}}^2 \right) J_{\Phi_i} \hat{N},
\]

where $\hat{N}$ is the unit normal field of the contour, and $\hat{I}_i^{in}, \hat{I}_i^{out}$ are the smoothed images inside and outside of the contour $C$, respectively. In the above, $\kappa_i$ is the curvature of $C$ with respect to the Riemannian metric $g_i$, and $\eta_i$ is a function that captures (details omitted) the difference between the metric $g_i$ and the standard Euclidean metric on $D$, with respect to which the contour will be evolving (i.e., the normals of the contour $C$ are defined by the Euclidean metric). Both quantities $\kappa_i, \eta_i$ can be easily incorporated into a level set framework. In particular, the curvature $\kappa_i$ with respect to the metric $g$ is given as $<\phi$ is the level-set

1In practice, $D$ is taken to be a rectangular domain with the same size as the images.

2We refer the reader to [8] for the definition of pull-back metric $g_i$. Briefly, the metric $g_i$ is the distorted version of the standard Euclidean metric on $C_i$ under the nonlinear map $\Phi_i$.
Solve the nonlinear elliptic PDEs in Equation 6. This given by, the chain rule, resulting in a formula con-
The last summand can be computed in closed form using the curve length can be derived using Equation 3:

where \( g^{-1} \) are the components of the local metric tensor.

Next, we derive the gradient for the deformation parameter \( \mu \) of the thin-plate spline basis function

The latter integral can be evaluated on the contour \( C \) for each \( \Phi_i \). Finally, the Euler-Lagrange Equation for \( \hat{I}_i \) is

where \( \omega(\hat{J}_i, \hat{I}) \) is a term involving the smoothed images \( \hat{I}_i \) and the determinants of the Jacobians \( J_{\Phi_i} \), and \( \Delta g_i \) is the Beltrami-Laplacian operator for the metric \( g_i \). The above equations give a system of nonlinear second-order elliptic equations, which can be solved using an iterative method, such as quasi-Newton [4]. The entire algorithm is summarized below.

Algorithm Summary
Given a collection of \( K \) images, \( \{I_1, \ldots, I_K\} \), the output of the algorithm is a collection of registration maps \( \Phi_i \) between image \( I_i \) and the domain \( D \), and a closed contour \( C \).

1. Optimize motion parameters \( \mu \) using gradient descent with gradient given by Equation 5. Update the deformation fields \( \Phi_i \).

2. Evolve the level set function \( \phi \) using Equation 4. The contour is updated as the zero level set of \( \phi \).

3. Solve the nonlinear elliptic PDEs in Equation 6. This updates the smoothed images \( \hat{I}_i \).

4. If the difference between consecutive iterates is below a pre-chosen tolerance, stop, else go to Step 1.

4. Results

We here present visual examples of simultaneous smoothing, segmentation, and registration, obtained through a level set contour evolution and thin-plate spline implementation. In Figure 3, the inputs are trios of hands making similar gestures, but in visibly different places and manners³. The segmentation, smoothing, and registration results are all contained directly in the displayed outputs: recall as well that the segmentation curves correspond across all images in a set through the registration maps. The examples shown here therefore demonstrate the success of the registration in matching even such fine features as single fingers, when those fingers are undergoing the fairly complex motion of spreading apart from one another. The (correct) identification of the fingers in the fourth column of Figure 3 is possible only because of the coupling of the segmentation of that image to the segmentation of the other two (in which the gaps are more readily distinguished).

Of course, dense correspondence is also implicitly being obtained between the hand features (e.g. the crease of the palm) as part of this process. The third sample case makes this correspondence explicit: in Fig. 4, we see the isolation of brain ventricles in human MRI cross sections. Note that ventricle identification in the second (middle) image is far more difficult than in the first or third, and yet an excellent result is achieved, thanks to the process coupling. In Fig. 5, two alternate methods of visualizing the dense registration alignment are provided. On the left, we have the effect of each \( \Phi_i \) on a set of concentric circles resident in shared domain \( D \). Since \( \Phi_{ij} = \Phi_i \circ \Phi_j^{-1} \) for all \( i \) and \( j \), this display provides a visual intuition of how structures in one image are carried onto another. On the right, we have the initial and final average image as seen from domain \( D \) (i.e. the average pullback). As \( D \) is the domain of correspondence, note that the sharpness of the pullback is a visual method of assessing the success of the registration in matching intensity values. As we are here segmenting the ventricles, it is the improvement in clarity and definition of the ventricles with which we are concerned. This effect is clearly evidenced. (Note that the unusual warping of the ventricles as seen from this perspective is of no significance. The “implicit atlas” which resides in the domain is merely a correspondent intermediary between input images, and need not represent a realistic input, merely a smooth deformation of a realistic input.)

³Thanks to M.R.T. Tarrosa for acquiring the input images.
5. Conclusion

We have presented both a general variational method for simultaneous smoothing, segmentation, and (nonrigid) registration of multiple images, and an implementation of one instantiation of the method. The chosen computational framework provides not only for the efficient solution of the variational problem, but inherently enforces symmetry, consistency, and unbiasedness across all processes without introducing any additional complexity (rather, the opposite). The mutually assistive coupling between the processes allows for the obtainment of solutions to all three, including in cases for which the nature of the data is such that the solution to any one might be difficult or impossible. This is precisely the motivation of, for instance, atlas-based registration-assisted segmentation, yet here no explicit atlas input or construction is ever required. An immediate extension of this work would be the incorporation of an information theoretic registration match metric in the energy functional, to extend functionality to input images of differing modalities.

References


---

Figure 3: **Left Three Columns:** Processing of variations on a hand position. Each column represents the evolution for a single input image in the trio, with the original input on top, and the final smooth reconstruction at bottom. The evolving segmentation and registration are depicted through the red curve and blue grid, respectively. Note the success in corresponding the fingers between all images. **Right Three Columns:** Processing of various openings of the hand, organized as in the example on the left. Note the identification of the fingers in the first column despite their near closure in the input image, under heavy noise. This owes to the “implicit atlassing” inherent in the method.
Figure 4: Identification of the ventricles of the brains of three different patients (in respective rows) in cross-sectional MRI scans. Note the successful identification in the case of the second input, despite the relative fineness of the structure. The final smoothed images are displayed in the last column.

Figure 5: **Left:** Concentric circles as warped by the final $\Phi_i$ for each $i$ (row) in Fig. 4, $i$ increasing from left to right. This is one method of visualizing the consistent correspondences between all three pairs. **Right:** An alternate visualization of registration success, as the sharpness of the average intensity pullback of the structure being segmented. On the left is the initial average (under identity mappings), on the right the converged result. The sharpening of the ventricle is dramatic.


