Supporting Registration Decisions during 3D Medical Volume Reconstructions

Peter Bajcsy*, Sang-Chul Lee, David Clutter
1008 National Center for Supercomputing Applications, 1205 W. Clark St, Urbana, IL 61801

ABSTRACT

We propose a methodology for making optimal registration decisions during 3D volume reconstruction in terms of (a) anticipated accuracy of aligned images, (b) uncertainty of obtained results during the registration process, (c) algorithmic repeatability of alignment procedure, and (d) computational requirements. We researched and developed a web-enabled, web services based, data-driven, registration decision support system. The registration decisions include (1) image spatial size (image sub-area or entire image), (2) transformation model (e.g., rigid, affine or elastic), (3) invariant registration feature (intensity, morphology or a sequential combination of the two), (4) automation level (manual, semi-automated, or fully-automated), (5) evaluations of registration results (multiple metrics and methods for establishing ground truth), and (6) assessment of resources (computational resources and human expertise, geographically local or distributed). Our goal is to provide mechanisms for evaluating the tradeoffs of each registration decision in terms of the aforementioned impacts. First, we present a medical registration methodology for making registration decisions that lead to registration results with well-understood accuracy, uncertainty, consistency and computational complexity characteristics. Second, we have built software tools that enable geographically distributed researchers to optimize their data-driven registration decisions by using web services and supercomputing resources. The support developed for registration decisions about 3D volume reconstruction is available to the general community with the access to the NCSA supercomputing resources. We illustrate performance by considering 3D volume reconstruction of blood vessels in histological sections of uveal melanoma from serial fluorescent labeled paraffin sections labeled with antibodies to CD34 and laminin. The specimens are studied by fluorescence confocal laser scanning microscopy (CLSM) images.

Keywords: 3D volume reconstruction, registration, fluorescence confocal laser scanning microscopy (CLSM)

1. INTRODUCTION

The problems of 3D volume reconstruction and medical cross section registration have been approached by an overwhelming number of researchers over the past several decades [1,2,3,4,5,6] and remain still an open problem. There have been several survey papers about registration approaches that include selection of registration variables based on user decisions [7,8,9]. In this paper, we focus on optimal selections of registration variables that are inherent parts of 3D volume reconstruction process. In our work, the 3D volume reconstruction problem is defined as a registration problem without fiduciary markers[8]. The goal of 3D reconstruction is to form a high-resolution 3D volume with large spatial coverage from a set of spatial tiles (small spatial coverage and high-resolution 2D images or 3D cross section volumes). 3D volume data are acquired from multiple histological sections of a tissue specimen by (a) placing each section under a microscope, (b) moving the specimen spatially, and (c) changing the focal length to obtain a set of optical sections, e.g., image stack per histological section and per specimen location[1].

In order to demonstrate the issues related to selecting registration variables, we consider 3D volume reconstruction of blood vessels in histological sections of uveal melanoma[10] from paraffin-embedded serial sections labeled with antibodies to CD34 and laminin and studied by confocal laser scanning microscopy (CLSM) imagery[11,12]. The set of spatial tiles is acquired by CLSM and consists of images that came from (1) one cross section (same axial coordinate) in different lateral coordinates or (2) multiple cross sections of a 3D volume (different axial coordinates). The 3D volume reconstruction objectives are to (1) register (stitch together) spatial tiles that came from the same cross section, i.e., image mosaic, and (2) align spatial tiles from multiple cross sections with the end use for visual inspection or quantitative analysis[11]. An overview of the 3D volume reconstruction problem is illustrated in Figure 1.

* pbajcsy@ncsa.uiuc.edu, Phone: 217-265-5387, Fax: 217-244-7396

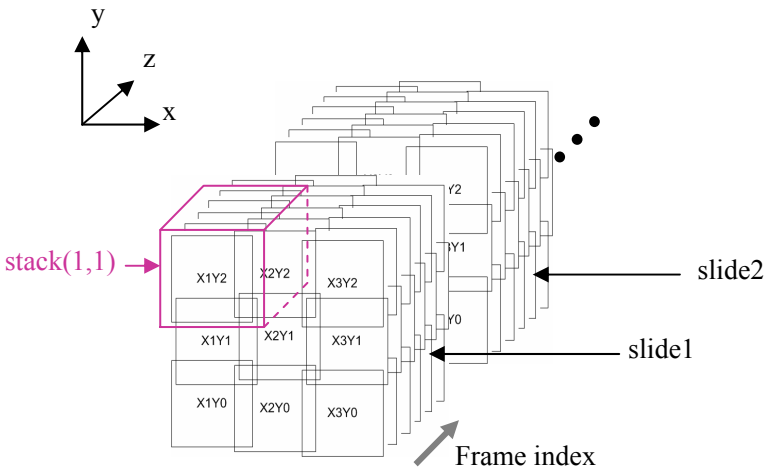


Figure 1. An overview of 3D volume reconstruction from fluorescent laser scanning confocal microscope images.

Based on our knowledge, there has been limited work on understanding accuracy, uncertainty, consistency and computational complexity characteristics of 3D volume reconstruction and their relationships to registration decisions. The past work usually addressed only certain aspects of registration decisions, for example the choice of transformation models[8], the combination of invariant registration features[13], the image data quality evaluation metrics[14], the choice of shape metrics[5], or the process of geometric (spatial registration related) and radiometric (intensity related) adjustments[15]. The past work has originated primarily from the computer vision community when tackling the problem of matching and alignment from points and frames while modeling rigid motion of objects. For example researchers, such as Pennec and Thirion[16,17], have developed a theoretical model defining the relationship between uncertainty of a rigid transformation applied to a set of 3D points or 2D frames and the registration accuracy. However, the model is defined for only a very small subset of typical registration decisions. A researcher performing 3D volume reconstruction is usually facing registration decisions about (1) image size used for registration, (2) transformation model (e.g., rigid, affine or elastic), (3) invariant registration feature (intensity, morphology or a combination of the two), (4) automation level (manual, semi-automated, or fully-automated), (5) evaluations of registration results (multiple metrics and methods for establishing ground truth), and (6) assessment of resources (geographically local or distributed computational resources and human expertise). Thus, there is a need to provide a mechanism for making optimal registration decisions, as well as to build a good understanding of the decision impacts on registration accuracy. Our work addresses this need by developing data-driven web-enabled analyses to support optimal registration decisions during 3D volume reconstruction. The analyses can be viewed as tradeoff studies of multiple registration decisions in terms of registration accuracy.

This paper describes data-driven optimization approaches to four registration decisions including image size selection, rigid or affine transformation model, intensity or morphological invariant feature selection, and manual (pixel-based) or semi-automated (centroid-based) automation level. The reason for using data-driven approaches lies in the large variability of objects of interest, specimen preparation, imaging modality, specific instrumentation characteristics, and so on, that is extremely difficult to model analytically with any generality whatsoever. We frame the problem first by presenting the links between 3D reconstruction accuracy and registration variables. Then, we lead to the data-driven methods for optimal choice of registration variables. The experimental results illustrate how the developed web-enabled software would be used by researchers to make the most optimal data-specific registration decisions.

The paper is organized as follows. First, we present the test material and CLSM imaging in Sections 2.1 and 2.2. Next, we describe registration decisions in Section 2.3 and their impacts on (a) anticipated accuracy of aligned images, (b) uncertainty of obtained results, (c) repeatability of alignment, and (d) computational requirements in Section 2.4 in order to provide better understanding of the relationships between registration variables and the quality of 3D reconstruction results. Third, we describe software algorithms in Section 3 that (a) can be used for making optimal registration decisions and (b) are available to geographically distributed researchers by using web services and high performance computing (HCP) resources[18]. The support developed for registration decisions about 3D volume reconstruction is available to the general community with the access to the NCSA supercomputing resources at

http://isda.ncsa.uiuc.edu/MedVolume/. Section 4 discusses the complexity of registration decisions during 3D medical volume reconstruction and Section 5 summarizes our work.

2. MATERIALS AND METHODS

2.1. Histological Materials

In general, registration decisions during 3D volume reconstruction are independent of image modality and histological materials. Nonetheless, the optimality of registration decisions and their impact on the resulting alignment are very much data dependent (and hence image modality and histological material dependent). The data used for explaining the use of registration decision software in Section 3 came from the University of Illinois at Chicago, Department of Pathology. The histological material represents formalin-fixed, paraffin-embedded uveal melanoma tissue samples sectioned at 4 μm thickness. The use of archival human tissue in this study was approved by the Institutional Review Board of the University of Illinois at Chicago.

Slides were deparaffinized in xylene and rehydrated through a decreasing ethanol gradient. Slides were rinsed in distilled water followed by antigen unmasking using Target Retrieval Solution 10X Concentrated (DAKO, Carpenteria, CA) according to the manufacturer's instructions and then rinsed in Phosphate Buffered Saline (PBS) for 5 minutes. Slides were incubated with monoclonal mouse anti-laminin antibody Sigma L8271, clone LAM 89 (Sigma, St. Louis, MO) at a dilution titer of 1:200 for 30 minutes at room temperature. Slides were rinsed in protein blocking solution (DAKO) for ten minutes followed by detection with Alexa Fluor 488 goat anti-mouse IgG (Molecular Probes, Eugene, OR) for 30 minutes at a dilution of 1:400. Slides were rinsed in buffer then mounted in Faramount Aqueous Mounting Medium (DAKO). For all staining procedures, secondary antibody was omitted in negative controls.

2.2. Confocal Laser Scanning Microscopy

All histological serial sections were acquired with a Leica SP2 confocal laser scanning microscope (Leica, Heidelberg, Germany) using the 20X objective. Images were stored in tagged information file format (TIFF). The experimental task was to perform 3D volume reconstruction of extravascular matrix patterns in primary human uveal melanoma tissue from 4 μm sections stained with laminin with signal detection by immunofluorescence as described above. We demonstrated the performance of our registration decision support system with one such 3D volume. The 3D volume consisted of four consecutive sub-volumes (histological sections) formed by 23, 23, 22 and 24 image frames (altogether 92 frames), where the term *image frames* refers to optical sections of a histological section. Each mosaicked frame is of type byte (8 bits per pixel) and of size 552 x 556 pixels. We analyzed each sub-volume separately to determine the most optimal set of registration parameters.

2.3. Registration Decisions

The overview of most common registration decisions is provided in Figure 2. The list of registration decisions was introduced in Section 1 (image size, transformation model, invariant feature, automation level, evaluations of registration results, and assessment of resources). In Figure 2, one should view black arrows as possible decision outcomes during the registration process and hence any possible combination of user decisions would characterize the obtained registration result. Some of the decisions could be easily expanded, e.g., other transformation models. Other decisions could be elaborated, such as methods for establishing ground truth could be classified into visual inspection, comparison with ground truth data, or measuring the degree of deviations from assumed data model. The purpose of Figure 2 is to present basic registration decisions rather than an exhaustive list of possible decision selections.

The user-driven registration decisions define the complexity of (a) registration model, (b) model parameter estimation, (c) registration computations to be performed and (d) evaluation strategy. For example, the case of a manual registration (alignment) of two image sub-areas containing a few features (visually salient pixel arrangements) using rigid transformation (rotation and translation) by overlaying two sub-areas and visually assessing the quality of alignment would be considered as a low complexity registration. It would use the simplest transformation model (rigid), subjective parameter estimation (visual), no computation (manual), and visual method for evaluating registration quality. In contrast, the case of a fully-automated alignment of two large images containing several millions of features using affine transformation (rotation, translation, scale and shear) by exhaustively evaluating the range of affine transformation parameters based on invariance of intensity (e.g., using normalized cross correlation or normalized mutual information)

would be considered as a high complexity registration[†]. In this case, the registration uses consistent parameter estimation by evaluating invariance of intensity, consuming significant computational resources and performing registration quality evaluations using mathematically defined metrics and based on a set of assumptions about data.

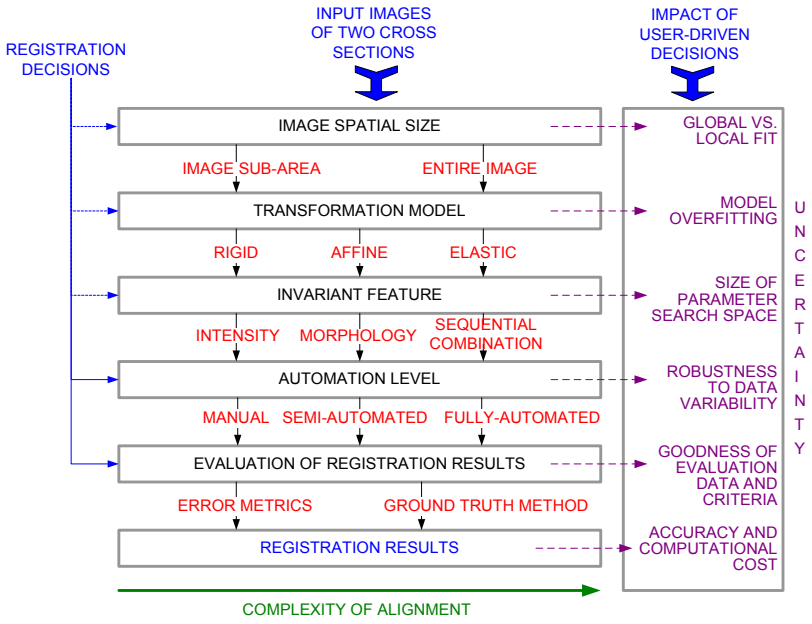


Figure 2: Registration decisions and their impact on registration results.

2.4. Impacts of Registration Decisions

Registration decisions during medical cross section alignment have a great impact on (1) anticipated accuracy of aligned images, (2) uncertainty of obtained results, (3) repeatability of alignment, and (4) computational requirements. As illustrated in Figure 2, the registration decisions affect (a) spatial distribution of registration error (global vs. local registration), (b) registration error composition (model overfitting, mosaicking x-y error, and alignment z error), (c) computational requirements on registration (invariance assumption and its degree of freedom, model complexity, search space to optimize parameters) and (d) validity of the obtained results (robustness of a registration model with respect to data deviations, quality of evaluation criteria).

Alignment Accuracy: The accuracy of aligned images can be measured either by visually inspecting anticipated structures or by defining quantitative metrics to evaluate accuracy with respect to ground truth data[21] (or a data model defined a priori). The analyses consist of (a) simulations with rigid and affine registration transformation, and (b) development of registration error models for better understanding of alignment results.

Uncertainty of Alignment Results: As illustrated in Figure 2, multiple registration decisions impact alignment uncertainty and include the decisions about (1) the tradeoffs between global or local registration fit (image spatial size decision), (2) the issues of transformation model overfitting (transformation model decision), (3) the degree of assumed intensity and morphological invariance (feature invariance decision), (4) the size of parameter search space and the algorithmic robustness to model deviations (automation level decision), and (5) the goodness of evaluation criteria for a registration problem. These uncertainties are very difficult to evaluate analytically and are very much data specific.

In order to understand registration uncertainties, data-driven evaluation software tools might provide insight about the anticipated result quality. Since some level of uncertainty is introduced to alignment results from every registration decision, we developed data-driven algorithms to calculate (1) normalized correlation based error as a function of image sub-area size, (2) error residuals as a function of rigid or affine transformation models (model complexity), (3) the degree

[†] Although there exist fast implementations of the intensity-based cross correlation methods, e,g., pyramid-based techniques, these methods cannot be applied in the case of CLSM due to large intensity heterogeneity.

of intensity and morphological invariance by de-coupling intensity and morphological variations, (4) the total search space size for semi-automated and fully-automated registration, and (5) error residuals as a function of evaluation metrics (normalized cross correlation or normalized mutual information[19]). Some of these algorithms are described in Section 2.5.

Repeatability of Alignment: Repeatability can be viewed as the consistency of alignment results obtained using multiple methods and processes[20]. The alignment processes usually include humans and computer algorithms. In general, alignment repeatability varies depending on the level of automation, registration methods and the complexity of human input. The higher automation level leads to better convergence of registration methods to global extreme[13]. In the semi-automated case, the less complex human input will lead to higher alignment accuracy.

The repeatability issue is usually addressed by performing studies using human subjects[21]. The studies are based on either a class of images acquired by the same imaging techniques and from similar specimens or a class of synthetic images that simulate different degrees of deviations from a registration model. Algorithmic repeatability is often evaluated together with its robustness to test that the algorithm avoids getting trapped in local minima, and can reliably find the best global minimum in complex landscapes defined by objective functions. Similarly, any measured repeatability due to automation is based on (a) making assumptions about acquired data, for instance, assuming feature invariance, (b) allowing only a subset of possible registration transformations (model constraints), or (c) searching only a subspace of possible transformation parameters. Thus, usually the robustness and accuracy of cross section alignment is decreasing with an increasing level of automation for data sets deviating from the automation model and violating the registration assumptions.

In order to understand repeatability issues, we developed web accessible registration tools that allow researchers to upload their data and collect measurements about repeatability as a function of (a) the complexity of human input (pixel or segment selection), (b) the level of automation (manual or semi-automated registration), (c) human expertise (experts and novices) and (d) data sets deviating from automation models.

Computational requirements: The computational requirements of alignment are directly proportional to the level of automation, the complexity of transformation model, and to the search space of transformation parameters. Furthermore, computational requirements for accommodating all researchers interested in using the same software with multiple registration tasks also have to be addressed.

In our prototype software, we utilize high performance computing resources at the National Center for Supercomputing Applications (NCSA) to guarantee sufficient computing power. The registration decision support software is also designed using web services technology that enables connecting multiple researchers at several geographical locations with computational resources at NCSA and algorithms developed by us[18].

2.5. Optimization Methods for Registration Decisions

Given a set of registration decisions and their impacts on accuracy, uncertainty, repeatability and computational requirements, there is a need for a set of methods and software tools to support optimal registration decisions. Our work aims at providing such a solution that can be described as a web-enabled, web services-based and data-driven registration decision support system. The system includes methods and software tools that provide data specific understanding of registration decision tradeoffs, such as (1) image sub-area size and location selections, (2) transformation model selection as a function of registration accuracy and possible image distortions (e.g., rigid or affine model), (3) registration feature selection (invariance of intensity or morphological features), and (4) choice of automation level (manual, semi-automated, or fully-automated). All of the above tradeoffs are evaluated as a function of registration accuracy. In addition, the system provides guidelines for an assessment of needed resources (geographically local or distributed computational resources and human expertise).

Specifically, the first method supporting registration decisions compares two images from a 3D volume, selecting a sub-area of interest, and then it reports a visualization of a correlation coefficient as a function of incrementally increasing sub-area size. The correlation coefficient is used as registration accuracy metric in this case. If two identical frames would be compared then the correlation metric value equals to one. The frames that would be completely dissimilar would lead to the value of zero.

The second method allows uploading two images to be registered, manually placing disks of variable size over visually matching features, and selecting rigid or affine transformation to compute the registered image pair. The reported registration transformation error and the level of distortion in the overlaid registered images with registration disks are used for aiding the choice of registration model.

The third method supporting registration decisions attempts to decouple intensity variations from morphological variations and then compute the degree of each variation. In order to decouple intensity variations from morphology variations, a 3D stack of images is thresholded by a chosen intensity to obtain a stack of binary images. The binary images are viewed as the data set with only morphological variations. The data set with only intensity variations is obtained by selecting one of the images in a 3D stack after thresholding, and re-populating the same thresholded image with intensities from each image in the 3D stack to form a new frame. In this way, the newly formed frames preserve intensities of the original frames without any morphological variations. The computation of the degree of intensity variation is achieved by using a correlation metric. The correlation metric reports statistical intensity deviations from two identical images. The degree of morphological variation is computed by defining a priori known 3D model (e.g., a cylinder) and calculating the deviations from the model. The aim is to support a decision based on the following rule: If *Degree*(Intensity variation) > *Degree*(Morphology variation) then use morphology, e.g., feature-based registration method, else use intensity to align images, e.g., intensity correlation-based registration method.

The fourth method provides an interface to pixel-based (manual) and segment centroid-based (semi-automated) registration methods. A user can choose any two images from a 3D stack of pre-segmented images, perform registration using manual or semi-automated registration and compare the registration accuracy[21]. If pre-segmented results are not available then segmentation and segment centroid extraction will be conducted.

3. RESULTS

This section describes how the set of developed methods was used for making optimal registration decisions during 3D volume reconstruction of the specimens imaged by CLSM and presented in Sections 2.1 and 2.2. We envision other researchers processing CLSM images to follow similar data-driven procedures to choose optimal 3D volume reconstruction parameters, such as registration area size, rigid or affine transformation model, intensity or morphology as invariant feature across cross sections, and pixel-based (manual) or centroid-based (semi-automated) alignment method.

3.1. Image Spatial Size

From a medical user view point, one would like to obtain 3D volume reconstructions over a large spatial area at a high spatial resolution, and with the best possible visual alignment of all salient image features. Let us assume that the visual alignment of salient image features is measured by a normalized correlation coefficient. Then, our goal is to provide data-driven analysis of input data to understand the tradeoffs between these two conflicting requirements, such as to maximize aligned image area and its measure of alignment goodness (e.g., normalized correlation coefficient) with other images.

In order to compute a data-driven dependency between normalized correlation metric and image sub-area size, at least two frames have to be aligned otherwise other registration unknowns would be inseparable from the variables under our scrutiny. In the case of CLSM, we can establish the data driven dependency by using any two frames from one stack of images. The assumptions are that these two selected frames are representatives of the two frames from two adjacent cross sections without any distortion due to specimen slicing, and the frame-to-frame morphology and intensity changes in the selected frames from one sub-volume are similar to those in the adjacent sub-volumes. Figure 3 illustrates our assumptions.

For the input data set with four sub-volumes, we varied the sub-area size of images shown in Figure 4 (a) from 24648 pixels (79x78) to 1236444 pixels (561x551) by increments of 2% of the entire number of pixels (approximately 25000 pixels). Figure 5 illustrates the dependency of the level of similarity measured by normalized correlation as a function of pixel count (sub-area size in pixels). The normalized correlation coefficients represent the maximum value obtained by optimizing spatial overlays using the affine transformation for every pair of image sub-areas. We can observe that the maximum similarity is obtained for the smallest experimental sub-area, and the minimum similarity is reported for the full image. Thus, alignment results for a sub-area that is approximately 37 times smaller than the full image size (915,788/24,648) will lead to 13% absolute improvement (0.85-0.72), and 46% relative improvement with respect to the best (1.0) and worst (0.72) alignment of the two frames (0.85-0.72)/(1-0.72). If one would like to explore the relationship without intensity variations then thresholded images can be evaluated the same way as. The thresholded images are shown in Figure 4 (b) and the dependency is shown in Figure 5.

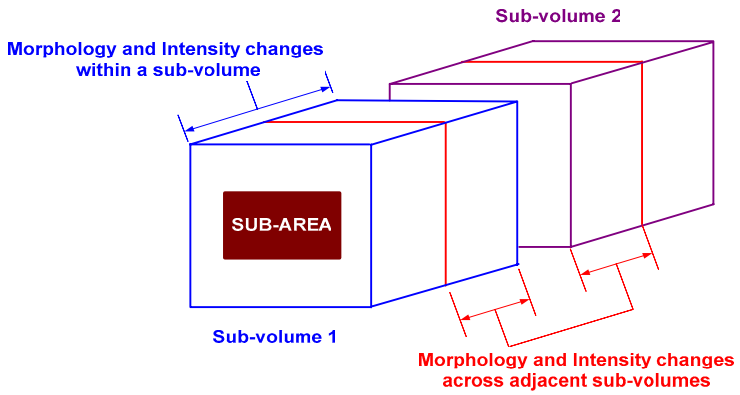
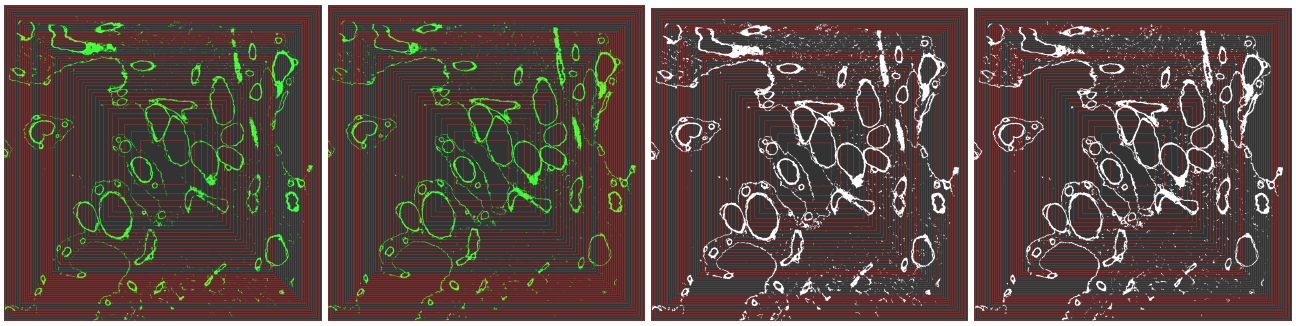


Figure 3: Assumptions about two aligned frames from one stack (sub-volume) for sub-area size evaluations. The illustration shows that processing end frames of one sub-volume (histological cross section) is similar to processing middle frames of two adjacent sub-volumes if morphological distortions due to preparation cut are neglected and an intensity profile is compromised. Based on these assumptions, one can select image sub-area size by analyzing two frames from one sub-volume.



(a) (b)

Figure 4: Pairs of images used for demonstrating the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision): (a) the image pair came from one stack of CLSM images and hence the ground truth alignment was known. The slides are 13 frames apart (1.7 microns distance along z-axis), and each pixel represent 0.478 microns (536 microns by 527 microns for the entire image) and (b) pair of binary images derived from the images in (a) by intensity thresholding to demonstrate the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision) for the case of zero intensity variation.

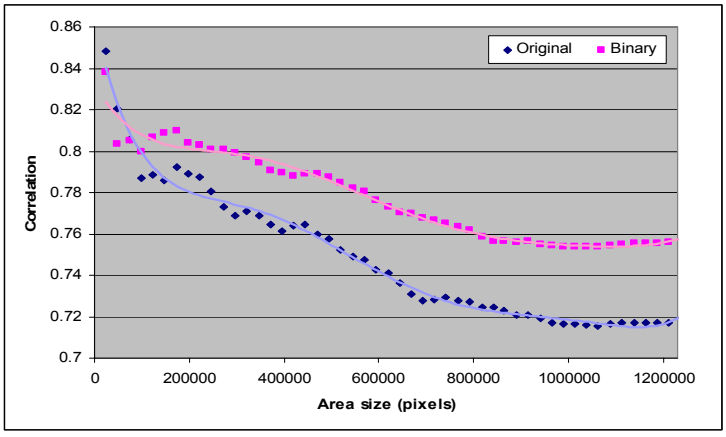


Figure 5: Normalized correlation coefficient as a function of image sub-area size in pixels for two perfectly aligned images that are 13 frames apart (1.7 microns apart), each image is of pixel size 1123 by 1104 corresponding to spatial resolution 536 microns by 527 microns (entire image). The two curves correspond to original image (blue) and thresholded image (pink). The images are shown in Figure 4. The normalized correlation coefficient represents a quantitative metric (similar to visual inspection) that ranges from [0, 1]. For our data set, we decided to use the entire image area equal to 561x551 pixels.

3.2. Transformation Model

Let us assume that a class of CLSM images acquired from similar tissues can be perfectly registered by an unknown transformation with N parameters. The N parameters define the order of transformation model. It is well known that a higher order transformation model than N would lead to data over-fitting while a lower order transformation model than N would lead to large registration errors. In our case, over-fitting would lead to distortions in the transformed image that could never happen in the original cross sections although the registration error would be small and indicate a good alignment. Using lower order transformation model might likely never satisfy the transformation error defined as user requirements.

The tradeoffs between transformation model complexity and registration error are usually resolved by practitioners based on a good understanding of medical specimen preparation. For example, for imaging cross sections of solid and hard specimens, rigid transformation might be appropriate (rotation and translation). If cutting the specimen introduces additional shear and scale changes, then affine transformation would be appropriate (rotation, translation, scale and shear). Our goal is to provide data-driven analysis of input data to understand the tradeoffs between these two conflicting requirements, such as to minimize transformation model complexity and registration error.

The task can be achieved by uploading two frames from two adjacent sub-volume images, manually placing disks of variable size over visually matching features, and selecting rigid or affine transformation to compute the registered image pair. We illustrate the matching feature selection for rigid and affine transformation models in Figure 6. The registration error for these two transformation models is summarized in Table 1. It provides a quantitative comparison of complexity versus error tradeoffs. For our data set, we decided to use affine transformation model.

Table 1: Summary of registration errors obtained using rigid and affine transformation models for the same set of four matching pairs of points. The residual errors X, Y and (X, Y) are computed as a sum of squared differences between the transformed coordinates of image 1 and the user chosen coordinates in image 2. The original residual error (X, Y) before any transformation was equal to 12.782.

Residual Error / Transformation Model	Error X	Error Y	Error (X,Y)
Rigid	1.528	1.802	1.225
Affine	0.927	1.581	0.935

3.3. Invariant Registration Feature

Registration features can span a wide range of image and volume attributes. For instance, image attributes would include pixels, edges, contours, image segments and their characteristics (centroids, sizes, etc.), and homogeneous or texture intensity profiles. Volume attributes would characterize 3D structures in each CLSM sub-volume, such as centroid trajectories of cylindrical structures, surface and volume descriptors, or intensity profiles of structures' surfaces and volumes. In general, these registration features can be divided into morphological (related to 2D or 3D shape) and intensity (related to fluorescent magnitude) attributes.

The goal of this step is to select either intensity or morphology as an invariant registration feature according to the

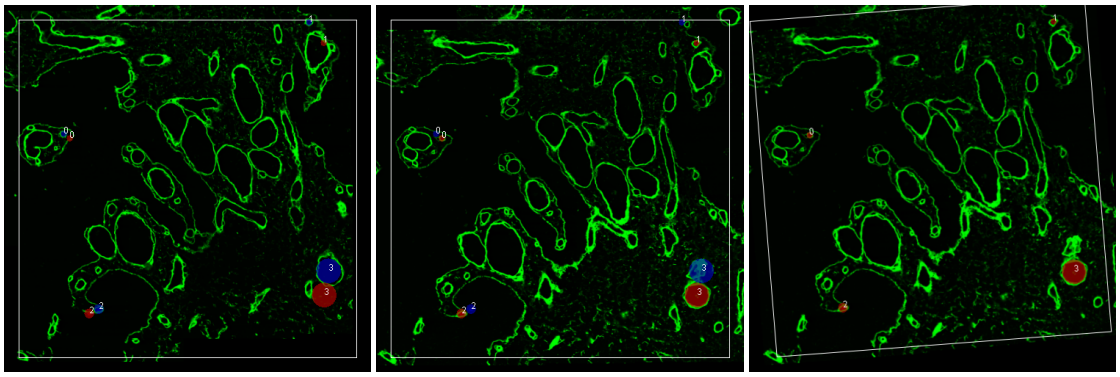


Figure 6: Transformation model selection. Target image with blue discs aligned with image features (left), source images (to be transformed) with red discs aligned with image features (middle) and transformed source image using an affine transformation model estimated based on the pairs of blue and red disc centroids (right).

following rule: If *Degree*(Intensity variation) > *Degree*(Morphology variation) then use morphology, else use intensity to align images. To follow the rule, one has to define and compute *Degree*(Intensity variation) and *Degree*(Morphology variation), as well as to compute the comparison of the degrees of intensity variation and morphological distortion in the input data. All three definitions and computations are still considered research topics on their own.

Due to the complexity of this problem, we introduced several assumptions. First, it is possible to de-couple intensity variation and morphological deviation within a 3D sub-volume. Second, morphological deviation and intensity variations can be measured by normalized correlation of frames from the de-coupled sub-volumes. While the computation to support the decision rule can be performed based on these two assumptions, it might be biased due to the normalized correlation metric. Thus, in our case, we decided to incorporate an additional evaluation of morphological deviation from a priori known straight cylindrical model of 3D structures.

Figure 7 and Figure 8 show examples from a CLSM sub-volume after de-coupling intensity variation and morphological deviations. The normalized correlation results for pairs of frames with an increasing axial distance or constant axial distance (adjacent frames) are reported in Figure 9. The average morphological deviation of a 3D segment from a cylindrical model was 0.21 pixels. The deviation was measured as a sum of Euclidean distances of segment centroid positions from a straight line along z-axis.

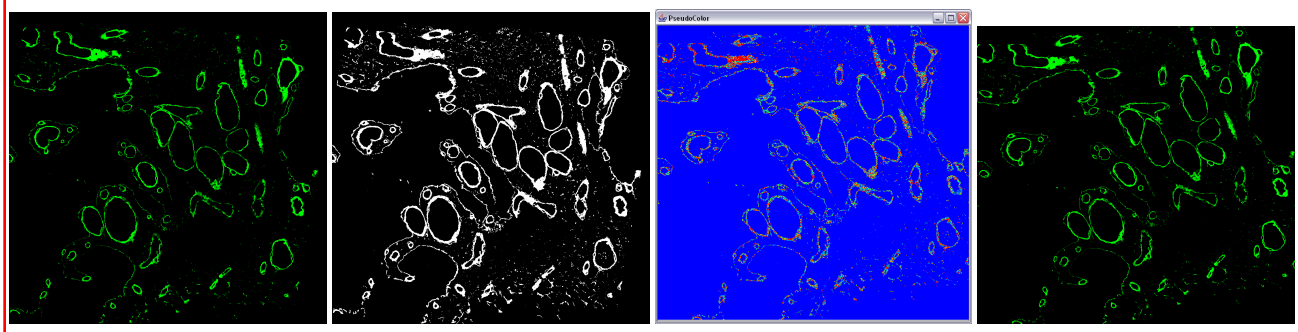


Figure 7: Simulation of intensity variations only. A template image (left), its intensity thresholded version for thresh = 2 (second left), and the pseudo-colored test frame (second right) generated by superimposing intensity variation of another frame (right). This type data preparation leads to a data set without morphological distortion and with the same intensity variation characteristics as found in the original stack of frames.

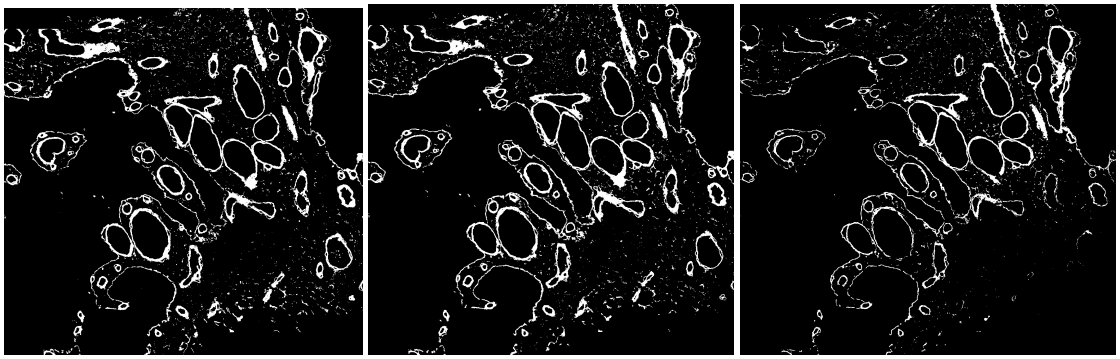


Figure 8: Simulation of morphological distortion only. Three frames from a stack of CLSM images that are 11 frames apart (frame index 0 (left), 11 (middle) and 22 (right)) and were intensity thresholded to remove intensity variations.

Based on our evaluations of the results presented in Figure 9, we concluded that the morphology is more invariant than intensity in our data sets. This conclusion is supported by (a) larger correlation value (better match) for the sub-volume with only morphological distortion (see Figure 9; left - all frame distances except from the last first and last ones, right – all adjacent frames), and (b) relatively small deviations of segment centroids from a straight trajectory.

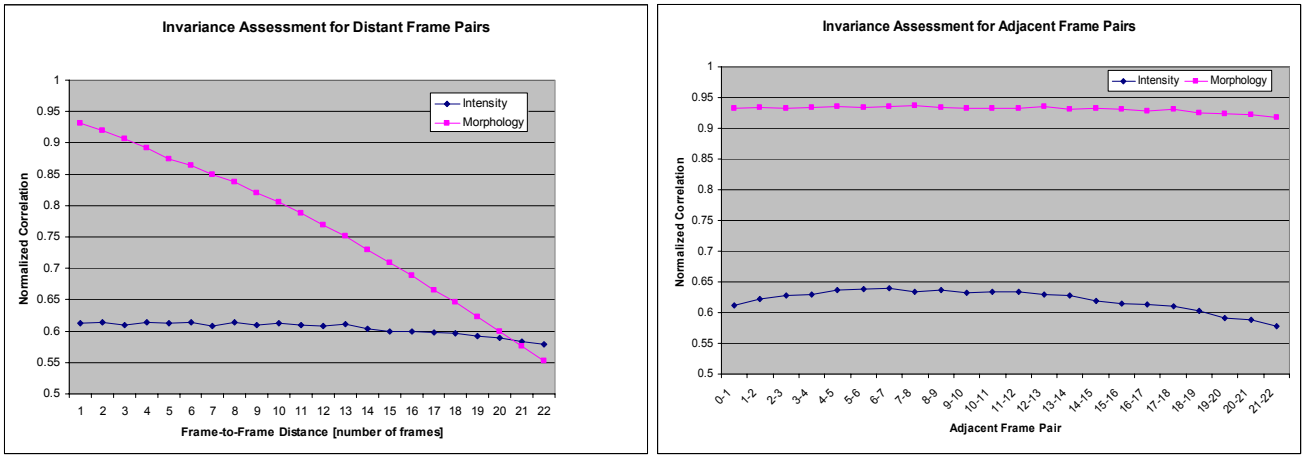


Figure 9: Evaluation of the degree of intensity variation and morphological distortion by a normalized correlation metric. Left: The correlation coefficients were computed for increasing frame-to-frame distances within one 3D stack of images (frames), where the 3D stack corresponds to 4 micrometers along axial coordinate and about 750x750 micrometers in lateral plane Right: The correlation coefficients were computed for adjacent frames for the same 3D stack of images (frames).

3.4. Automation Level

The goal of this step is to decide the level of automation based on the registration accuracy, computational resources, and geographic locations of needed 3D volume reconstruction expertise. We approached this problem as follows: The evaluations of registration accuracy at multiple levels of automation require involvement of human subjects and careful preparation of ground truth data. We have conducted a user study in the past using manual or semi-automated registration techniques[21] and decided to use the semi-automated (segment centroid based) automation level. By web-enabling the developed software, other researchers can perform similar studies to the study published in [21] and support their data-specific decision about the level of automation. The problem of available computational resources and geographically distributed expertise was solved by developing a web service-based mechanism for registration that provides access to image data at the location with 3D volume reconstruction expertise and performs computation at the location with computational resources.

4. DISCUSSION

Based on our experimental data-driven evaluations, we supported our registration decisions to choose (1) an entire image size for registration, (2) affine rather than rigid transformation model to compensate for observed warping, (3) morphology as the more invariant feature than intensity and (4) semi-automated centroid-based technique over pixel-based registration method. These registration decisions are optimal given the choice of accuracy quality metrics and algorithmic approaches. We provide next the discussion about (a) the need to have 3D volume reconstruction quality specifications, (b) the pros and cons of analytical and data-driven modeling approaches to registration decisions, (c) the challenges associated with computing and comparing the degrees of intensity variation and morphological distortions and (d) the difficulties with evaluations that support registration decisions.

3D volume reconstruction quality specifications: The outlined procedure for optimizing registration decisions brings up several issues related to well-defined 3D volume reconstruction requirements. In a real world scenario, input data are not ideal, and 3D volume reconstruction accuracy requirements are limited by the nature of specimen preparation, data acquisition and processing. We demonstrated that the requirements for the 3D volume reconstruction presented in this paper have to include several user-defined parameters. First, a user has to define the minimum area of the interesting region and the minimum intensity similarity, (e.g., normalized cross correlation coefficient) for accepting the alignment results (image sub-area size decision). Second, there is a need to specify the minimum registration error improvement due to increased transformation model complexity for adopting more complex transformation model (registration model decision). Third, one would specify the instrument noise threshold value to separate intensity variations from

morphological deviations, as well as the 3D model for segmented structures and the maximum estimated deviation from the model for accepting registration feature invariance (invariant feature decision). As for the automation level decision, a user has to decide what accuracy and repeatability of reconstructions would be acceptable as a function of automation.

Analytical and data-driven modeling approaches: Although we started the paper by considering the impacts of registration decisions on (1) anticipated accuracy of aligned images, (2) uncertainty of obtained results, (3) repeatability of alignment, and (4) computational requirements, we have not provided analytical models for all factors contributing to the quality of 3D volume reconstructions. Instead, we have offered data-driven models that facilitate our understanding of real data acquired from instrumentation. The reason for choosing data-driven models instead of analytical models lies in the large variability of objects of interest, specimen preparation, imaging modality, specific instrumentation characteristics, and so on. The use of analytical models might not be feasible (lack of knowledge) or introduce too many assumptions about data so that the generality of the models is very small. As illustrated in Section 3, the data-driven approach to support 3D volume reconstructions is more feasible and general. It is addressing several basic registration decision problems related to (a) what variables play major roles during 3D volume reconstruction, (b) how relevant 3D reconstruction variables (parameters) depend on each other and how they impact resulting 3D reconstruction, (c) how to optimize the choice of 3D reconstruction parameters, and (d) what solutions could accommodate computational and expertise requirements during 3D volume reconstruction.

Degrees of intensity variation and morphological distortions: Another discussion is about the challenge of assessing data in terms of the degrees of intensity variation and morphological distortions. The challenge lies in defining and comparing metrics evaluating the degrees of change. We chose the normalized correlation metric since it has been used frequently as a statistical measure of intensity similarity for two images of the same spatial size. The degree of morphological distortion is related to shape, and traditionally shape characteristics have been obtained (a) from local geometry called landmarks, (b) from a set of sampled boundary points, (c) by boundary-representing basis function coefficients, or (d) by hybrids of boundary and other curve loci with landmarks, and other sampling schemes, as overviewed in [5]. We incorporated the assessment of the degree of morphological distortion by computing the change of the first spatial moment, such as the centroid (0-th moment). However, it is not obvious how to compare a normalized correlation coefficient with an average change of centroid locations. Therefore, we opted for the normalized correlation coefficient as a measure of morphological distortion and based our decision on comparative results using normalized correlation and absolute morphological distortion assessment results using the centroid trajectory deviation from a priori known model. The current approach to choosing an invariant registration feature is just one solution that might be evaluated in the future with other possible approaches.

Evaluations supporting registration decisions: Ideally, one would like to perform data-driven evaluations with all possible registration accuracy metrics, all registration methods and approaches, and with unlimited computational resources. Each researcher would assess his/her data variability and quality with respect to the 3D reconstruction task under multiple modeling assumptions and without computational constraints. While this is currently unobtainable, our work presents a methodology how to obtain better 3D volume reconstruction understanding given a limited subset of metrics, methods and resources. In addition, the evaluation problems related to human subject selection, ground truth data preparation, and 3D volume validation (just to name a few), are beyond the scope of this paper but might be of interest to a user evaluating automation levels.

5. SUMMARY

In this paper, we addressed the problem of optimal registration decisions during 3D medical volume reconstruction. The registration decisions of interest included (1) image spatial size (image sub-area size), (2) transformation model (rigid or affine), (3) invariant registration feature (intensity or morphology), and (4) automation level (manual or semi-automated). We provided mechanisms for evaluating the tradeoffs of each registration decision in terms of 3D volume reconstruction accuracy, repeatability and computational requirements. Furthermore, we have built software tools for geographically distributed researchers to optimize their data-driven registration decisions by using web services and supercomputing resources. Finally, we presented an example of optimal registration decisions supported by the developed software and analyses. The example showed the case of 3D volume reconstruction of blood vessels in histological sections of uveal melanoma from serial fluorescent labeled paraffin sections imaged by fluorescence CLSM imaging.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Institute of Health under Grant No. R01 EY10457 and Core Grant EY01792. The on-going research is collaboration between the Department of Pathology, College of Medicine, University of Illinois at Chicago (UIC) and the Image Spatial Data Analysis Group, National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign (UIUC). We acknowledge NCSA/UIUC support of this work.

REFERENCE

1. G. Hermosillo and O. Faugeras, "Variational methods for multi-modal image matching", *Int. J. Comp. Vision,* 50, 329 – 343, 2002.
2. D. Hill, P. Batchelor, M. Holden and D. Hawkes, "Medical Image Registration." *Phys. Med. Biol.* 46 R1-R45, 2001.
3. Montgomery, K. and Ross, M.D., "Non-fiducial, shape-based registration of biological tissue." *3D Microsc.: Image Acquisition and Processing III. SPIE Electronic Imaging*, 1996.
4. C. Papadimitriou, C. Yapijakis and D. Davaki, "Use of truncated pyramid representation methodology in three-dimensional reconstruction: an example." *J. Microsc.*, 214, 70–75, 2004.
5. S. Pizer, D. Fritsch, P. Yushkevitch, V. Johnson and E. Chaney, "Segmentation, Registration, and Measurement of Shape Variation via Image Object Shape." *IEEE Trans.Med. Imaging*, 18, 851-865, 1996.
6. P. Viola and W. Wells, "Alignment by maximization of mutual information." *Int. Conf. on Computer Vision*, 16, 1995.
7. L. Brown, "A Survey of Image Registration Techniques." *ACM Comp. Surv.* 24, 325-376, 1992.
8. J. Maintz and M. Viergever, "A survey of medical image registration." *Med. Image Anal.,* 2, 1-36, 1998.
9. Zitov´a B. and Flusser, J., "Image registration methods: a survey." *Image and Vision Computing*, 21, 977–1000, 2003.
10. X. Chen, A. Maniotis, D. Majumdar and R. Folberg, "Uveal Melanoma Cell Staining for CD34 and Assessment of Tumor Vascularity." *Invest. Ophthal. Vis. Sci.* 43, 2533-2539, 2002.
11. H. Bornfleth, K. Satzler, R. Eils and C. Cremer, "High-precision distance measurements and volume-conserving segmentation of objects near and below the resolution limit in three-dimensional confocal fluorescence microscopy." *J. Microsc,* 189, 118–136, 1998.
12. J. Itoh, K. Yasumura, K. Ogawa, K. Kawai, A. Serizawa, Y. Yamamoto and Y. Osamura, "Three-dimensional (3D) imaging of tumor angiogenesis and its inhibition: Evaluation of tumor vascular-targeting agent efficacy in the DMBA-induced rat breast cancer model by confocal laser scanning microscopy (CLSM)." *Acta Histochem., Cytochem., Technical advancement,* 36, 27-36, 2003.
13. P. Cachier. et al., "Multisubject Non-rigid Registration of Brain MRI Using Intensity and Geometric Features." *LNCS* 2208, 734–742, 2001.
14. D. Wolin, K. Johnson and Y. Kipma, "The Importance of Objective Analysis in Image Quality Evaluation." *Imaging Sic. Tech. (IS and T)* 603-606, 1998.
15. R. Radke, S. Andra, O. Al-Kofahi and B. Roysam, "Image Change Detection Algorithms: A Systematic Survey." *IEEE Trans. Image Processing*, 14, 294-307, 2005.
16. X. Pennec and J-P. Thirion, A Framework for Uncertainty and Validation of 3D Registration Methods based on Points and Frames. *Int. J. of Computer Vision*, 25, 203-229, 1997.
17. X. Pennec, "Evaluation of the Uncertainty in Various Registration problems." *Meth. of Eval. Comp. Med. Imaging*, 15-16, 2001.
18. R. Kooper, A. Shirk, S-C. Lee, A. Lin, R. Folberg and P. Bajcsy, "3D Medical Volume Reconstruction Using Web Services," *IEEE Int. Conf. Web Services*, 2005.
19. R. Duda, P. Hart, and D. Stork, *Pattern classification* (Second Edition), 654, Wiley-Interscience, New York, N.Y., 2001.
20. C. Schmid, R. Mohr and C. Bauckhage, "Evaluation of Interest Point Detectors." *Int. J. Computer Vision*, 37, 151 – 172, 2000.
21. S-C. Lee and P. Bajcsy, "Feature based registration of fluorescent LSCM imagery using region centroids." *SPIE*, 5747, 170-181, 2005.