Assembly of Large Three-Dimensional Volumes from Serial-Section Transmission Electron Microscopy

Pavel Koshevoy[1], Tolga Tasdizen[1], Ross Whitaker[1], Bryan Jones[2] and Robert Marc[2]

[1]Scientific Computing and Imaging Institute, University of Utah, USA
[2]Moran Eye Center, University of Utah, USA

***Abstract*— Serial-section transmission electron microscopy (TEM) is an important imaging modality for studying neuronal connectivity patterns. However, before serial-section TEM images can be used to reconstruct connectivities of neurons, several image registration problems must be addressed. The first problem arises due to the large sample size and limited field of view: each section must be assembled from many overlapping tiles, a process also referred to as mosaicing. The second problem is the co-registration of slice mosaics into a single three-dimensional volume. In both problems, non-linear distortions of individual images must be corrected. We present a carefully engineered solution to these problems making by modifying existing registration paradigms to better fit the requirements of serial-section TEM images.**

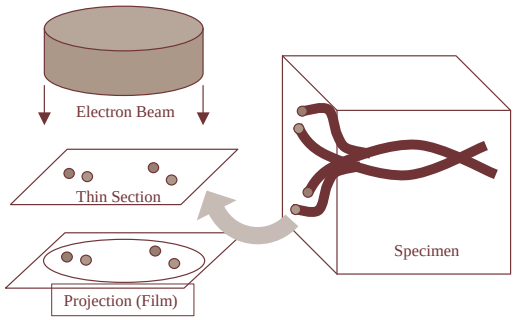

Fig. 1. Ultra-thin serial-section TEM.

I. Introduction

Detailed, data-driven descriptions of microscopic structures are very important in neurobiology. While neural modeling is critical to our understanding of the central nervous system, state-of-the-art models are relatively unconstrained by anatomical data. Very little is known about the physical organization and connectivities of neurons. Motivated by this lack of data, a number of researchers [1], [2], [3] have undertaken extensive imaging projects, to produce detailed maps of neuronal structure and connectivity. Magnetic resonance imaging (MRI) can only provide information at a macroscopic level. Many imaging needs can be met with traditional light microscopy (LM) and confocal microscopes, but resolution issues limit the ability to resolve smaller structures such as synapses. For these problems, electron microscopy (EM) approaches can help to define structures smaller than the wavelength of white light and as such, EM is the primary method for resolving sub cellular anatomy and cellular connectivity.

A. Electron Microscopy

Electron microscopy can produce 3D data in several different ways. The method underlying the proposed work is *ultra-thin serial-sectioning* which removes 40-90 nanometer (nm) slices, one at a time, from a solid specimen. There are two methods for forming and imaging the slices. The first method is serial block-face scanning EM [9]. In this method, each slice is cut away and discarded, and the electron beam is scanned over the remaining block face to produce electron "reflectance" images. The solid block is dimensionally stable hence no slice-to-slice registration is required in three-dimensional (3D) volume assembly. However, the image resolution is not adequate for certain problems such as ganglion-cell connectivity reconstruction, which is discussed in Section I-B. The image acquisition method we address in this paper, transmission electron microscopy (TEM) [10], [11], [3], captures the slice itself by suspending it in an electron beam to obtain a single projection of electron dense and transparent structures transverse to the plane of the slice, see Figure 1. TEM offers good resolution (< 1 nm in plane) and relatively high signal-to-noise ratio.

An alternative to serial-sectioning is electron microscope tomography (EMT) [6], [7]. The resolution of this method is similar to TEM, but reconstructions often suffer from artifacts due to limited acquisition angles and low signal-to-noise ratios.

B. Ganglion-cell connectivity in the retina

The driving application for our work is the reconstruction of ganglion-cell connectivity in the retina. The mammalian retina contains at least 55-80 classes of neurons [4], [5]. Photoreceptors drive horizontal cells and bipolar cells (BCs). BCs drive amacrine cells (ACs) and 15-20 ganglion cell (GC) classes that project to the brain. Ganglion cells play a central role in visual perception. The connectivities of most AC and GC classes are inferred (incompletely) from optically acquired stereomorphology or limited physiology. To fully understand the role of GCs, one needs to reconstruct a complete map of identified synaptic inputs to each GC class in the mammalian retina. This requires the tracking of individual retinal neurons in EM images.

Rabbit retinal tissue are serially sectioned parallel to the retinal plane (90 nm slice thickness) through the inner plexiform layer (0.025 mm depth). To reconstruct the connectivities

Corresponding author: Tolga Tasdizen, email: tolga@sci.utah.edu

of all GC types, an area no smaller than 0.2 x 0.2 mm has to be imaged to ensure that a sufficient number of samples of each GC will be present in the dataset [4]. Each slice is film-imaged as 270–330 overlapping tiles on a TEM at a true magnification of 5300×, and 8-bit scanned at 1200 dpi, which yields adequate resolution to identify synapses and gap junctions.

C. Image registration problems in serial-section TEM

Several types of image registration problems are encountered in the application described in Section I-B. The first is the **mosaicing** of a large number of individual images (tiles). Due to the large field of view (0.2mm×0.2 mm), and the high-magnification requirements, each slice is composed of a large number of overlapping tiles (270–330). The same problem is also encountered in CCD camera based microscopy image acquisition. The limited field of view imposed by CCD chips tend to constrain what can be seen in any one digital frame versus what can be seen with the eye and this limitation is only made more severe at higher magnifications. The solution to this has been the implementation of image mosaicing approaches which stitch or tile adjacent images together to create a seamless complete image. There are a variety of methods for mosaicing images ranging from encoding stage information and syncing image capture with stage position to coarse optical matching of intensities to tile images, but these approaches tend to work best with confocal data that has broad contrast and loose z-axis focus requirements. For microscopy requiring high contrast work with narrowly constrained focal planes, current approaches tend to fail at image edges where errors in the range of 100-250 nm corrupt data making some projects designed to reconstruct entire networks of neurons impossible.

However, regardless of whether a pure digital work-flow or a hybrid film/digital imagery work-flow is followed, there are inherent problems with mosaicing EM images in that each image has within it a microscope specific image aberration that typically presents itself as a polynomial warp from the center to the periphery of the resulting image. Any one image is perfectly interpretable, but when one attempts to mosaic several images lining up structures becomes a difficult task. Efforts attempting to define complete neuronal circuits through the use of hundreds of image mosaics in numerous layers are made impossible unless images can be warped with a polynomial transformation to match one another. Furthermore, due to the large number of tiles per slice, user interaction is not feasible at this stage.

The second registration problem is that each section is a physically separate section that maintains for the most part, its dimensional stability. There are in reality, subtle deformations on the nanometer scale that are introduced through the preparation of these sections and these deformations also have to be taken into account when digitally reconstructing serial sections. This problem is also further complicated by the change in the visible structures going from one 2D slice to the next. In this paper, we describe a carefully engineered solution to these problems.

The fields of image processing and computer vision have made significant progress in the quantitative analysis of biomedical images such as MRI and CT over the last 20 years; however, progress has been much slower in EM image processing. In this methods paper, we outline a complete algorithm for assembling 3D volumes from stacks of serial-section EM data. This is an essential first step in automated processing. Large parts of our solution relies on existing methods from the image processing and computer vision literature; however, we focus on the specific properties and unique requirements of very large datasets of serial-section EM images.

II. Related Work

Medical image registration is a very active research area, and an extensive review is beyond the scope of this paper. For an in-depth survey, the reader is referred to [13], [14]. Methods can be classified according to three criteria: the image-matching metric, the type of coordinate transformation, and the optimization procedure. For instance, intensity-based methods compute transformations using image intensity information [15], [13] while landmark-based methods match a set of fiducial points between images [16], [17], [18], [19], [20], [21]. Fiducial points can be anatomical or geometrical in nature and are either automatically detected or entered manually by a user. More complex landmarks, such as contours and surfaces have also been used for registration [22], [23], [13]. A more recent research direction is to combine both landmark and intensity based approaches [24]. The range of allowed transformations can be rigid, affine, polynomial, thin-plate splines or large deformations [25], [13], [26], [27], [20].

Serial-section EM registration methods in the literature primarily have been manual or semi-automatic [28]. In an early work, Carlbom *et al.* propose a manual method [29]. Fiala and Harris also propose a manual method which estimates a polynomial transformation from fiducial points entered by a user [30]. Randall *et al.* propose an automatic method for coregistering EM images acquired with a CCD camera [31]. However, they assume that the successive slices are very similar and use rigid transformations.

III. Image mosaicing

Each section is imaged in overlapping tiles in scan-line ordering. Digital electron microscopes can provide precise information about location of individual tiles which can be used in the mosaicing process; however, analog electron microscopes can provide only coarse positioning information. On the other hand, digital EM imaging has poor intrinsic resolution defined by the density of optical fiber bundles; hence, the method of choice is to use analog film processing followed by digital film scanning at high resolution allowing full resolution to be preserved. During analog film capture, an operator manually adjusts the position of each tile aiming for a 20% overlap between adjacent tiles. Due to the imprecise nature of this operation, exact position information for each tile is not available. The mosaicing algorithm addresses this problem. Furthermore, every section undergoes an unpredictable rotation when placed in the microscope; hence, the number of

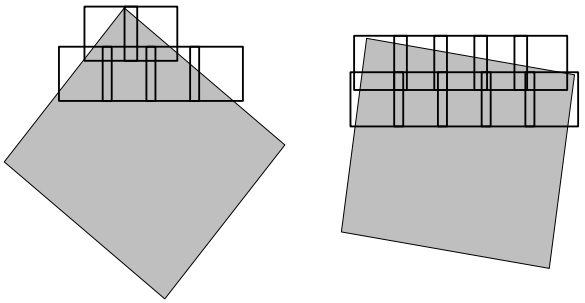

Fig. 2. First two rows of tiles in imaging two section with different orientations. The number of tiles in each row depend on the unpredictable orientation of the slice in the electron microscope.

tiles in each scan-line will differ, see Figure 2. Therefore, even though the order in which tiles were imaged is known , we do not know which tiles are neighbors in the tile collection unless the microscope operator records this information. To make our algorithm as general as possible, we propose an algorithm which deduces the tile ordering automatically. Hence the problem can be stated as: Given a large number of tiles specified in no particular order, a mosaic must be constructed and individual tiles must be corrected for radial distortion. The problem that can be split up into more manageable sub-problems:

A) Find pairs of overlapping tiles and estimate the relative displacement between them.
B) Deduce the tile ordering and build a rough estimate of the mosaic without radial distortion correction.
C) Iteratively refine the mosaic by alternating the refinement of the radial distortion correction and position of each tile in the mosaic.

A. Matching pairs of tiles

The first sub-problem is to find pairs of overlapping tiles. The main constraint at this stage of the algorithm is computational complexity because this procedure is applied to approximately n^2 tile pairs (number of potential matching pairs) where n is the total number of tiles in a slice. Hence, methods that incrementally improve cross-correlation or similar measures by iterative optimization are not feasible.

If we restrict the class of allowed coordinate transformations between pairs of tiles to only translation, a closed-form solution exists [32]. Let F_i denote the Fourier transform of image S_i. The cross-correlation Φ_{ij} between S_i and S_j is calculated as

$$\Phi_{ij} = F_i F_j^*,$$

where F_j^* denotes the complex conjugate of F_j. Similarly the auto-correlation terms (the power spectral density) is computed as $\Phi_{ii} = F_i F_i^*$. Then using the Fourier transform property $F[S(x-x_o)] = e^{-jwx_o} F[S(x)]$, a probability density function (PDF) for the displacement vector between images S_0 and S_1 is calculated as

$$PDF\,(x,y) = Real\left[F^{-1}\left[\frac{\Phi_{10}}{\sqrt{\Phi_{00}\Phi_{11}}}\right]\right] \qquad (1)$$

where F^{-1} denotes the inverse Fourier transform [32]. Notice that, under assumptions of periodicity, if the image pair differs exactly by a translation then the PDF computed by (1) will have non-zero probability for a single displacement vector. In practice, finding the maximum of the displacement PDF is non-trivial. The PDF is very noisy due to the textured nature of electron microscopy images. Also, the PDF of two non-overlapping images may contain several maxima, or none at all. These problems are not addressed in [32].

Four main steps are necessary to identify the location of the correct maxima in the PDF. The first step is to pre-smooth the images to reduce the amount of noise. The second step is to select and apply a threshold to the PDF image to isolate global peaks. We choose a threshold so that 1% of the total pixels are considered. In the third step, we look for a cluster of at least five 8-connected pixels that indicate a strong maximum. If the pixels are scattered across the PDF, it is likely the PDF does not have a strong maximum. Once all of the clusters have been identified, the clusters that are broken up across the PDF boundary are merged together. This step is required because the Discrete Fourier Transform assumes that the signal is periodic; therefore, the PDF is also periodic. The coordinates of the PDF maxima are calculated as the centers of mass of the corresponding clusters. The final step is to verify which, if any, of the maxima found in the previous step is the true displacement between the image pair. Non-overlapping image pairs typically produce a PDF with several maxima points at roughly the same value, while the PDF of two overlapping tiles produces one maximum significantly higher than the rest. If the strongest maxima is at least twice as strong as the rest, it is marked as a good match; otherwise, we determine that the tiles do not overlap.

In order to find the displacement vector, it is not enough to simply find the maximum of the displacement PDF. The coordinates (x_{max}, y_{max}) are always positive, yet the displacement vector may very well have negative coordinates. As mentioned earlier, the Discrete Fourier Transform assumes that the signal is periodic whereas the actual images are not. Therefore, once the coordinates of the maximum (x_{max}, y_{max}) are known, there are four possible permutations of the displacement vector that could produce the corresponding high cross-correlation between the tiles. The permutations are (x_{max}, y_{max}), $(x_{max} - w, y_{max})$, $(x_{max}, y_{max} - h)$ and $(x_{max} - w, y_{max} - h)$ where w and h denote the image width and height, respectively. We choose the best permutation as the one with the minimum squared image difference normalized by the overlap area. Also, displacement vectors resulting in less than 5% of overlap are discarded without further consideration. This decision is based on the fact that the microscope operator aims for approximately 20% overlap along the edges of the tile, and approximately 10% overlap at the corners. These overlap amounts provide a near optimal trade-off between matching algorithm robustness vs. redundancy in images, wasted storage space and extra microscope operation time. Matching becomes unreliable for image pairs with less than 10% overlap. Successful results of our image matching on two tiles with approximately 10% overlap is demonstrated in Figure 3.

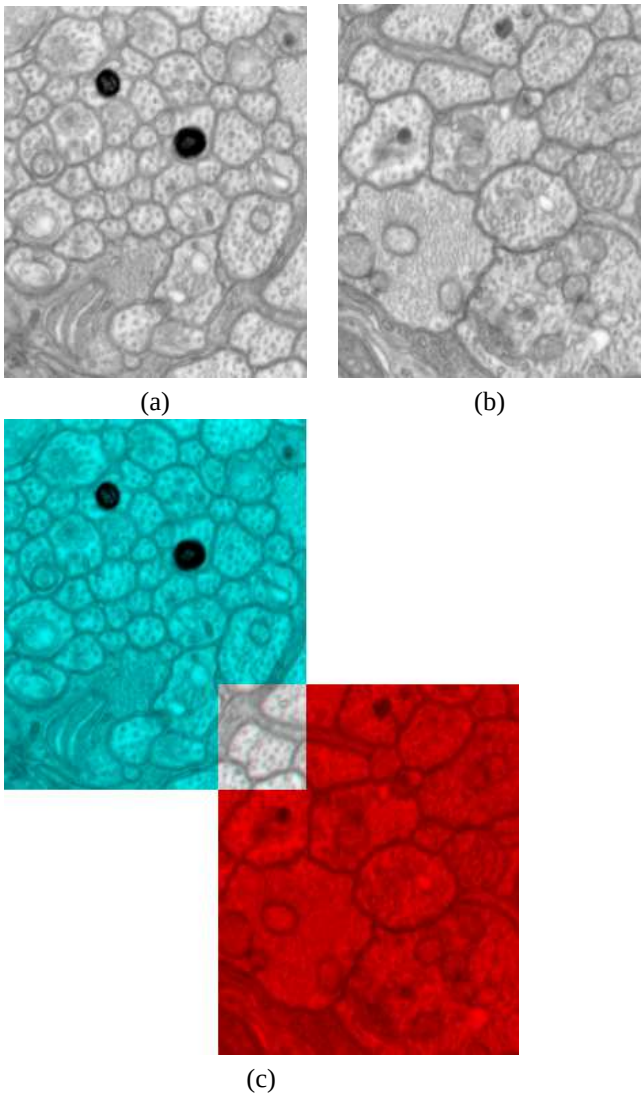

Fig. 3. [COLOR image] (a,b) Two overlapping images, (c) result of matching.

B. Deducing the tile ordering

Matches between tiles found by the algorithm described in Section III-A are used as the basis for deducing the tile ordering. The mappings between the tiles found with the method described in Section III-A are stored as connections in a graph of tiles. Each mapping (connection) is weighed according to the normalized squared image differences metric mentioned earlier.

After selection of an anchor tile, our algorithm tries to find the best possible mapping from the image space of any tile into the image space of the anchor tile, S_{target}. This is accomplished by cascading the mappings via intermediate tiles. For example, there may exist a mapping $S_0 : S_1$ between tiles S_0 and S_1, and another mapping $S_1 : S_{target}$ between tiles S_1 and S_{target}. A mapping $S_0 : S_1 : S_{target}$ between tiles S_0 and S_{target} can be created via the intermediate tile S_1. The mapping with the least cost (normalized square intensity difference) is preferred even when it has greater cascade length. Redundant mappings allow us to select the best mapping possible and avoid problems due to occasional bad matches which might result from the previous stage. It is also worth mentioning that any tile can be chosen as the anchor for the mosaic without significantly changing the final performance. The gross tile placements in the mosaick are used as the initialization to the iterative non-linear refinement discussed next.

C. Nonlinear distortion correction

As mentioned earlier, each tile undergoes a nonlinear warping due to bending of the electron beam. The warping is well approximated by a radial distortion if the image coordinates of the center of the microscope is known, which is not necessarily the center of the image itself. Unfortunately, this information is not available and can not easily be inferred from the image. Therefore, we choose to model the warping with bivariate Legendre polynomials. This is a more general class of transformations than radial distortion and the unknown center of the electron beam can be handled through the additional degrees of freedom.

Correcting the radial distortion is important because it can amount to mismatches upto tens of pixels between overlapping tiles in areas far from the electron beam center. If uncorrected, this large mismatch is likely to cause neuron segmentation and tracking algorithms to fail. We have found that a bivariate cubic Legendre polynomial provides very high quality matches. The transformation from the mosaic image coordinates (u, v) to the coordinate frame of tile k is given by

$$x_k(u,v) = X_k \sum_{i=0}^{N} \sum_{j=0}^{i} a_{k,j,i-j} P_j\left(\frac{u-\bar{u}_k}{X_k}\right) P_{i-j}\left(\frac{v-\bar{v}_k}{Y_k}\right) \tag{2}$$

$$y_k(u,v) = Y_k \sum_{i=0}^{N} \sum_{j=0}^{i} b_{k,j,i-j} P_j\left(\frac{u-\bar{u}_k}{X_k}\right) P_{i-j}\left(\frac{v-\bar{v}_k}{Y_k}\right), \tag{3}$$

where P are the Legendre polynomial basis functions. The transform for the k'th tile is parameterized by polynomial coefficients $a_{k,i,i-j}$ and $b_{k,i,i-j}$. The normalization constants X_k and Y_k correspond to the half-width and half-height of the k'th tile. The image center of the k'th tile in the mosaic coordinate frame is located at $(\bar{u}_k, \bar{v}_k)$; this center is given by the displacement vectors and the tile ordering computed previously. Note that the zero'th degree coefficients of the transformation allow further shifts in the tile displacements as necessary. Also, if the center of electron beam is not at the image center, the radial distortion can still be modeled due to the larger degree of freedom available.

The polynomial coefficients are found iteratively using the ITK [33] optimization framework, specifically the *itk::RegularStepGradientDescentOptimizer* class. Since more than one tile may overlap the same pixel, the average intensity variance within overlapping regions was chosen as the tile mismatch metric:

$$V = \frac{1}{A} \sum_{u=0}^{W-1} \sum_{v=0}^{H-1} \Delta^2(u,v) \tag{4}$$

where W and H are the dimensions of the entire mosaic and A is the area of the overlapping regions. The function $\Delta^2(u,v)$

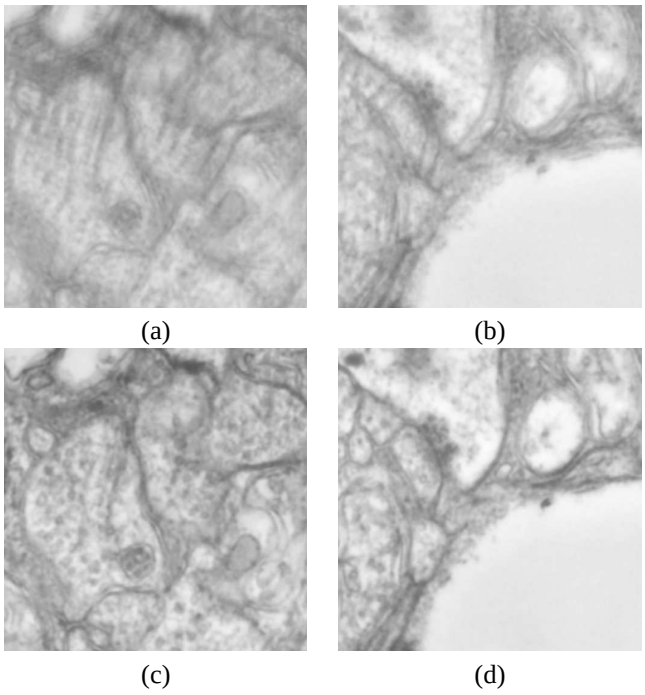

Fig. 4. Overlapping tiles before (a,b) and after (c,d) nonlinear distortion correction.

is the variance at mosaic pixel location (u, v):

$$\Delta^2(u,v) = \frac{1}{N(u,v)} \sum_{k=0}^{N(u,v)-1} \left(S_k(x_k, y_k) - \mu(u,v)\right)^2 \quad (5)$$

$N(u,v)$ is the number of tiles overlapping a pixel at the given mosaic coordinates (u, v). The functions x_k and y_k are as defined in equations (2) and (3). The mean intensity value at the specified mosaic coordinates is computed as

$$\mu(u,v) = \frac{1}{N(u,v)} \sum_{k=0}^{N(u,v)-1} S_k(x_k, y_k) \quad (6)$$

The optimization proceeds in two stages. First, we optimize a shared set of transform parameters $\vec{a}$ and $\vec{b}$ for all tiles. Note that, $\bar{u}_k$ and $\bar{v}_k$ which define the gross tile placement are still unique to each tile and are treated as fixed. Since the properties of the electron beam remains constant, the distortions of the various tiles are similar. Hence, this stage compensates for large scale radial distortion common to all tiles. The remaining variance in the mosaic is due to unique distortions present in each tile such as small shifts in the center of the electron in terms of image coordinates. Therefore, in the second stage, we restart the optimization with the initialization set to the common parameters, but optimize tile transforms without sharing the parameters. This produces the unique transform parameters for each tile and is found to be more robust than optimizing unique parameters directly. The variance minimization iterates until it converges or exceeds the maximum number of iterations.

Figure 4 illustrates the importance of nonlinear distortion correction in EM image mosaicing. Figures 4 (a) and (b) show close-up views of two areas of the section where two tiles were matched using only global displacement vectors. Figures 4 (c) and (d) show the same areas after applying the nonlinear distortion correction algorithm outlined in this section. Notice that the blurring in the overlapping region due to non-linear coordinate distortions are not present in Figures 4 (c) and (d). Two mosaicked sections (including nonlinear distortion correction) are shown in Figure 5. Both of these sections were assembled from 12 tiles. The size of the mosaicked sections are approximately $10,000 \times 11,000$ pixels. The computation of each mosaic takes approximately 2 hours on a high-end PC. However, it is important to note that the computation time effectively scales linearly with the number of tiles in the slice n. The only portion of the algorithm that scales with $O(n^2)$ (image pair matching) is computationally very cheap; this portion takes only 3 minutes of the 2 hour computation required to assemble the slice.

IV. Slice-to-slice matching

After mosaicing individual tiles into slices, the goal is to find the transformation between consecutive slices to assemble a registered stack of images, or in other words a 3D volume. Note that the ordering of the slices in the stack is known. The transformation between adjacent slices is composed of several factors: (i) nonlinear warping due to the physical sectioning process, (ii) the physical changes in the structure being imaged due to the 90nm gap between consecutive slices and (iii) an unknown rotation and displacement when the section is placed in the field of view of the microscope. We use a fiducial-point based approach as outlined below:

A) For each slice, a gradient vector image pyramid and a Difference-of-Gaussian (DoG) image pyramid is constructed. The extrema points of the DoG pyramid are determined.
B) The dominant gradient vector orientations in the neighborhood of each extrema point are detected. A descriptor for every detected gradient vector orientation of the extrema point is generated.
C) For each pair of adjacent slices, matching descriptors are found.
D) Given the matching descriptors, a transform that best maps the extrema points from the image space of slice A into the image space of slice B is calculated.

A. Detecting extrema points

The specifics of the construction of the image pyramids are thoroughly covered by David G. Lowe[34] and will not be repeated here. Suffice it to say, that a pyramid is a collection of octaves, where each octave represents a reduction of image resolution by a factor of 2. Each octave is partitioned into a set of scales where each successive image is convolved with a Gaussian filter of increasing sigma value. The extrema points are the local minima and maxima points of the DoG image pyramids. Lowe[34] proposed looking for an extrema point in a $3 \times 3 \times 3$ neighborhood within a DoG pyramid.

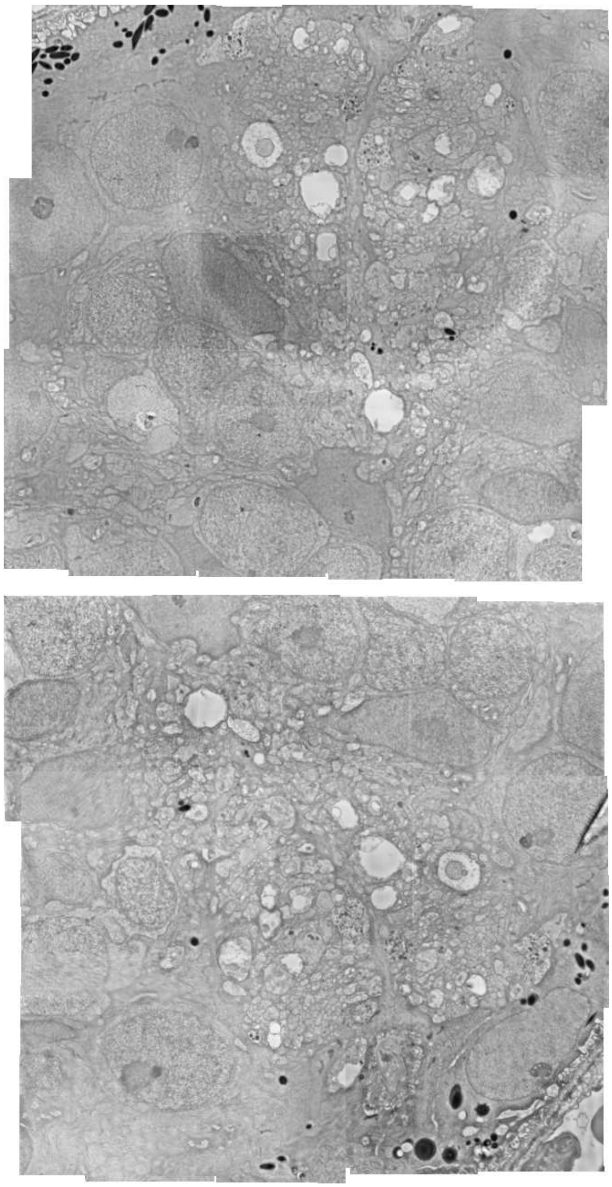

Fig. 5. Two consecutive slices. Each slice was mosaicked from 12 high resolution Transmission Electron Microscopy images of a rabbit retina. Each image is approximately $10,000 \times 11,000$ pixels.

B. Descriptors

The descriptors are based on neighborhood properties of the gradient image at the extrema points. The descriptors have to be rotationally invariant, therefore it is necessary to select a consistent frame of reference for sampling the neighborhood around each extrema point. The neighborhood gradient orientation angles are accumulated into a 1D histogram. Each contribution is weighed by the gradient magnitude and a 2D Gaussian weighting function centered at the extrema point. The peaks of the histogram define the feature vector orientation angles [34]. After determining the angle, the neighborhood gradient orientations are sampled within a local coordinate system based on the descriptor orientation angle. It is important to note that the radius of the sampling window has to be large enough (in pixels) to capture the neighborhood properties. We partition the neighborhood into a set of concentric annuli, where each annulus is partitioned into a set of cells of equal area. This strategy is found to perform slightly better for EM images than the rectangular grid used for sampling in [34]. Each histogram cell holds a gradient vector orientation histogram.

C. Descriptor matching

The matching process is slightly different from the one outlined in [34]. That work addresses a more general computer vision problem, where detection of the same object at different scales is important. The electron transmission microscopy images are taken at the same scale, and undergo minor deformation on the global scale, making the scale invariant feature matching unnecessary. Therefore, for the purposes of TEM image registration, the descriptors are matched against other descriptors selected from the same octave and scale of the pyramid. We use an optimized kd-tree [35] with a best-bin-first nearest neighbor search algorithm [36].

The number of mismatched descriptors can be reduced effectively for TEM images based on the ratio of the distance (in image space) between nearest extrema points in the two sections. Since the scales are the same, the distance between nearest neighbors in one image and the matching image should be nearly identical. If the ratio of the two distances deviates significantly from 1.0, it can be assumed that one of the matches is wrong. Figure 6 shows the descriptors remaining after this filtering for matching the two sections shown in Figure 5. In this case, 4601 possible matches are filtered down to 459 matches.

D. Estimating the transform

The remaining set of matches may still contain some errors. Brown and Lowe [37] propose the use of RANSAC [38] to select a set of matches that define a consistent transform. Essentially, a few matches are selected at random to solve for the transform parameters. The number of initially selected matches depends on the number of transform parameters. For example, a 2nd order (linear) bi-variate Legendre polynomial transform has 6 parameters, it requires 3 distinct matches. A 4th-order (cubic) bi-variate Legendre polynomial transform has 20 parameters, it requires 10 distinct matches.

Once a transform has been estimated, the rest of the matches are verified as inliers or outliers. For each match point pair, the point expressed in the space of image S_i is mapped via the transform into the space of image S_{i+1}. The distance of the mapped point to its match is used to classify the match as an inlier or an outlier based on some threshold. The inliers and the original set of matches are then used to re-estimate the transform. This can be an iterative process, where at each iteration the matches are classified as inliers and outliers, until convergence or a maximum number of iterations is reached. Since the goal is to maximize the number of inliers, the process is repeated with a new set of initial random matches, and the best results are kept. As shown in Figure 7, this process further reduces the number of potentially matching descriptors. In this case 165 matching descriptors remain for the final transform

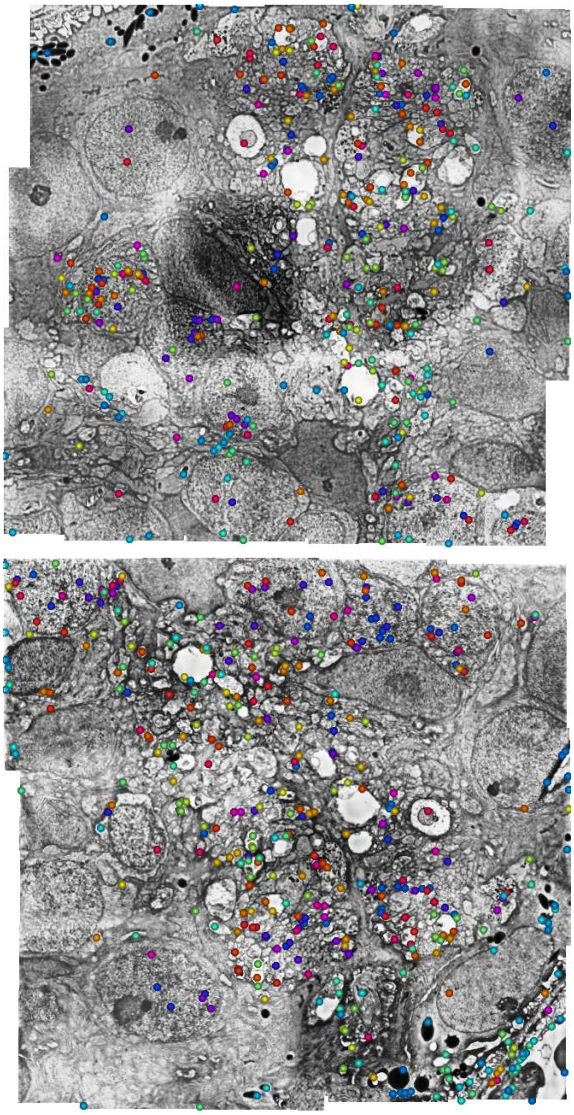

Fig. 6. [COLOR image] Remaining descriptors after distance-ratio filtering to reduce the number of mismatches. Descriptor locations are shown as colored dots; matching descriptors share the same color in both images.

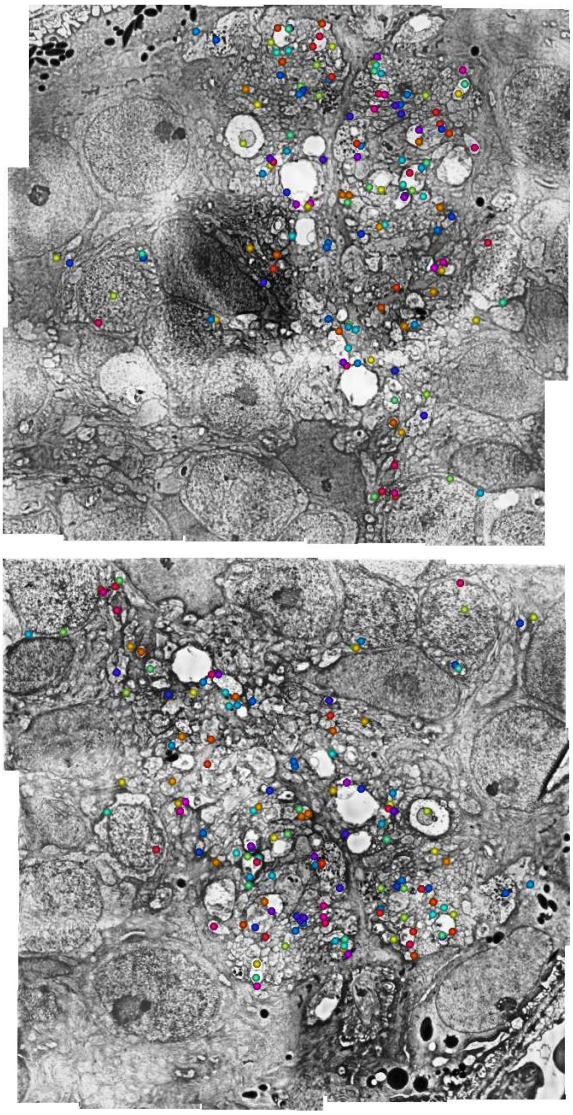

Fig. 7. [COLOR image] Final matching descriptors after application of RANSAC. Descriptor locations are shown as colored dots; matching descriptors share the same color in both images.

estimation. Finally, figure 8 illustrates the results of registering the two sections.

While the section mosaicking stage of the algorithm has been extensively tested, more testing is still needed to validate the slice-to-slice matching. Certain slices which have very different contrast properties than other slices in the stack and slices which have major artifacts (tears, folds) have been found to cause the slice-to-slice matching stage to fail. These issues will be addressed in future work.

V. Conclusion

Microscopy is undergoing a new renaissance as molecular biology and genetics begin to require and rediscover LM and EM based histological and anatomical approaches to validate and inform their science and methodologies. The advent of affordable CCD cameras combined with easy to implement interfaces such as Firewire have resulted in wider accessibility, higher throughput and easier interpretation and visualization of digital image data to bioscience laboratories. While digital image based microscopy has been almost a universally appreciated convenience, for some users pushing the boundaries of projects requiring imaging, there are some unique challenges. The method described in this paper address two of these challenges: section assembly from thin-sections captured in tiles with high-resolution analog TEM and section-to-section matching. Future work will focus on tracking neurons in the assembled 3D volumes. While the methods described in this paper are fine-tuned for the specific requirements of serial-section TEM imaging, the high-level approach, in principal, is applicable to any imaging method that captures a volume serially as a stack of 2D images.

Acknowledgments

This work was supported in part by NIH RO1 EB005832-01, NIH EY0015128, EY002576 and NEI Vision Core EY014800.

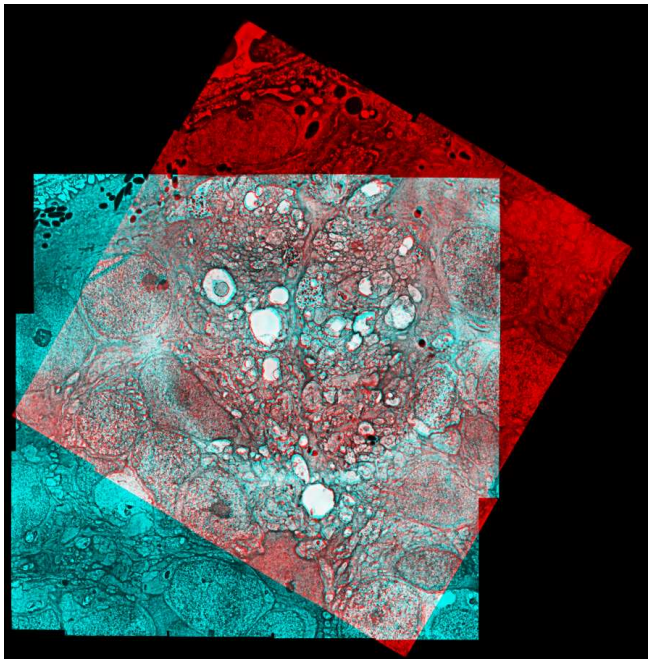

Fig. 8. [COLOR figure] Visualization of the slice-to-slice registration results.

REFERENCES

[1] K. M. Harris, J. C. Fiala, and L. Ostroff, "Structural changes at dendritic spine synapses during long-term potentiation," *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, vol. 358, no. 1432, pp. 745–748, 2003.

[2] E. D. Cohen and P. Sterling, "Parallel circuits from cones to the on-beta ganglion cell," *European Journal of Neuroscience*, vol. 4, pp. 506–520, 1992.

[3] R. F. Dacheux, M. F. Chimento, and F. R. Amthor, "Synaptic input to the on-off directionally selective ganglion cell in the rabbit retina," *Journal of Comparative Neurology*, vol. 456, no. 3, pp. 267–278, 2003.

[4] R. E. Marc and B.W. Jones, "Molecular phenotyping of retinal ganglion cells," *Journal of Neuroscience*, vol. 22, pp. 412–427, 2002.

[5] R. H. Masland, "Neuronal diversity in the retina," *Current Opinion in Neurobiology*, vol. 11, no. 4, pp. 431–6, 2001.

[6] M. E. Martone, T. J. Deerinck, N. Yamada, E. Bushong, and M. H. Ellisman, "Correlated 3d light and electron microscopy: Use of high voltage electron microscopy and electron tomography for imaging large biological structures," *Journal of Histotechnology*, vol. 23, no. 3, pp. 261–270, 2000.

[7] M. E. Martone, B. R. Hu, and M. H. Ellisman, "Alterations of hippocampal postsynaptic densities following transient ischemia," *Hippocampus*, vol. 10, no. 5, pp. 610–616, 2000.

[8] B.F. McEwen and M. Marko, "The emergence of electron tomography as an important tool for investigating cellular ultrastructure," *J Histochem Cytochem*, vol. 49, pp. 553–563, 2001.

[9] W. Denk and H. Horstmann, "Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure," *PLoS Biol*, vol. 2, no. 11, pp. e329, 2004, 1545-7885 Journal Article.

[10] E.V. Jr. Famiglietti and H. Kolb, "A bistratified amacrine cell and synaptic circuitry in the inner plexiform layer of the retina," *Brain Research*, vol. 84, pp. 293–300, 1975.

[11] D. J. Calkins, Y. Tsukamoto, and P. Sterling, "Microcircuitry and mosaic of a blue-yellow ganglion cell in the primate retina," *Journal of Neuroscience*, vol. 18, no. 9, pp. 3373–3385, 1998.

[12] D. I. Vaney, "A quantitative comparison between the ganglion cell populations and axonal outflows of the visual streak and periphery of the rabbit retina," *Journal of Comparative Neurology*, vol. 189, pp. 215–233, 1980.

[13] A. Toga, Ed., *Brain Warping*, Wiley-Interscience, New York, 1999.

[14] J. B. A. Maintz and M. A. Viergever, *Navigated Brain Surgery*, chapter A Survey of medical image registration, Oxford University Press, 1999.

[15] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision Graphics and Image Processing*, vol. 46, no. 1, pp. 1–21, 1989.

[16] A. C. Evans, C. Beil, S. Marret C. J. Thompson, and A. Hakim, "Anatomical-functional correlation using an adjustable mri-based region of interest with positron emission tomography," *Journal of Cerebral Blood Flow and Metabolism*, vol. 8, pp. 513–530, 1988.

[17] J. Thirion, "Extremal points: definition and application to 3d image registration," in *Proceedings of Computer Vision and Pattern Recognition*, 1994, pp. 587–592.

[18] J. Thirion, "New feature points based on geometric invariants for 3d image registration," *International Journal of Computer Vision*, vol. 18, no. 2, pp. 121–137, 1996.

[19] F. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, 1997.

[20] Karl Rohr, Mike Fornefett, and H. Siegfried Stiehl, "Approximating thin-plate splines for elastic registration: Integration of landmark errors and orientation attributes," *Lecture Notes in Computer Science*, vol. 1613, pp. 252–265, 1999.

[21] K. Rohr, M. Fornefett, and H. S. Stiehl, "Spline-based elastic image registration: integration of landmark errors and orientation attributes," *Comput. Vis. Image Underst.*, vol. 90, no. 2, pp. 153–168, 2003.

[22] C. A. Davatzikos, J. L. Prince, and R. N. Bryan, "Image registration based on boundary mapping," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 112–115, 1996.

[23] P. M. Thompson and A. W. Toga, "A surface-based technique for warping three-dimensional images of the brain," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 1–16, 1996.

[24] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 450–461, 2002.

[25] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[26] G. Christensen, R. Rabbit, and M. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[27] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-d image matching," *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 317–328, 1997.

[28] D. P Huijsmans, W. H. Lamers, J.A. Los, and J. Strackee, "Toward computerized morphometric facilities: A review of 58 software packages for computer-aided three-dimensional reconstruction, quantification, and picture generation from parallel serial sections," *Anat. Record*, vol. 216, 1986.

[29] I. Carlbom, D. Terzopoulos, and K. M. Harris, "Computer-assieted registration, segmentation, and 3d reconstruction from images of neuronal tissue sections," *IEEE Transactions on Medical Imaging*, vol. 13, no. 2, June 1994.

[30] J. C. Fiala and K. M. Harris, "Extending unbiased stereology of brain ultrastructure to three-dimensional volumes," *Journal of the American Medical Informatics Association*, vol. 8, no. 1, pp. 1–16, 2001.

[31] G. Randall, A. Fernandez, O. Trujillo-Cenoz, G. Apelbaum, M. Bertalmio, L. Vazquez, F. Malmierca, and P. Morelli, "Neuro3d: an interactive 3d reconstruction system of serial sections using automatic registration," in *SPIE Proceedings Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing V*, January 1998, vol. 3261.

[32] B. Girod and D. Kuo, "Direct estimation of displacement histograms," in *Proceedings of the Optical Society of America Meeting on Understanding and Machine Vision*, 1989, pp. 73–76.

[33] "Insight segmentation and registration toolkit," www.itk.org.

[34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 2004.

[35] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Transactions on Mathematical Software*, vol. 3, no. 3, pp. 209–226, 1977.

[36] J. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1997, pp. 1000–1006.

[37] M. Brown and D. G. Lowe, "Invariant features from interest point groups," in *British Machine Vision Conference*, 2002, pp. 656–665.

[38] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.