

Combined volume registration and visualization

Arlie G. Capps,^{1,2,3} Robert J. Zawadzki,² John S. Werner,² and Bernd Hamann³

¹Lawrence Livermore National Laboratory, U.S.A.

²Vision Science and Advanced Retinal Imaging Lab, Department of Ophthalmology, UC Davis, U.S.A.

³Institute for Data Analysis and Visualization, Department of Computer Science, UC Davis, U.S.A.

Abstract

We describe a method for combining and visualizing a set of overlapping volume images with high resolution but limited spatial extent. Our system combines the calculation of a registration metric with ray casting for direct volume rendering into a combined operation performed on the graphics processing unit (GPU). The combined calculation reduces memory traffic, increases rendering frame rate, and makes possible interactive-speed, user-supervised, semi-automatic combination of many component volume images. For volumes that do not overlap any other imaged volume, the system uses contextual information provided in the form of an overall 2D background image to calculate a registration metric.

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Image Processing and Computer Vision]: Enhancement—Registration

1. Introduction

Volumetric imaging modalities, such as magnetic resonance imaging (MRI), ultrasound (US), computed tomography (CT), confocal scanning optical microscopy and optical coherence tomography (OCT) have become increasingly important in biological and medical applications. Each of these modalities produces a “volume image” composed of samples or voxels commonly arranged in a 3D Cartesian grid of resolution $K \times L \times M$, having a uniform distance between samples in each dimension. Continual advances in technology and engineering have resulted in steady refinement of image resolution, increase in data acquisition speed, and increase in data size and dimensionality. However, with some clinical *in vivo* imaging techniques such as OCT the coverage provided by individual high resolution volume images is small. Multiple volume images that partially overlap must be captured, registered and combined (“stitched”) to allow a single larger region to be analyzed as one data set. Gaps or holes, areas of the larger region not covered by any volume image, must be easily identifiable as such by the user. Overlaps, areas covered by more than one volume image, must contain enough information for the stitching operation to unambiguously determine the volume images’ correct relative position.

We have developed methods and a prototype system

for simultaneously stitching and visualizing multiple high-resolution volume data sets. Our system allows a user to place individual volume images into rough alignment, then automatically refines the placement of the component images. In addition to calculating registration values based on overlapping 3D regions of the individual volumes, we calculate registration to 2D background reference images which the volumes may overlay, allowing users to properly position volumes which have little or no mutual overlap. Our system also supports combination of scanned data sets in which correction for motion artifacts has changed a deformed, regular image into a correct image with sampling irregularities and even gaps in coverage. Successive motion-corrected scans of the same region contain different motion and hence different sampling gaps, which are filled in when the scans are properly registered and combined.

The procedures for visualizing multiple overlapping data sets and calculating a metric for registration are closely linked. Both tasks require sampling throughout the image space overlapped by multiple volumes followed by combination of the sample values to a summary, either to display as an image or to report as a metric for goodness-of-fit. Since both tasks require intensive and wide-ranging data sampling, an efficient data access pattern will directly benefit the speed of the system. We combine computation of the registration

metric with the volume ray casting algorithm running on a workstation GPU.

We developed our algorithm for use with non-invasive *in vivo* cellular resolution human retinal imaging modalities. Adaptive optics scanning laser ophthalmoscopy (AO-SLO) [WH81, ZJP*11] captures a sequence of 2D images at 27 Hz with roughly $3\mu\text{m}$ lateral resolution which form a movie of the scanned retinal patch. Adaptive optics optical coherence tomography (AO-OCT) [WLK*02, ZCF*09] provides 3D images which are acquired by scanning a series of consecutive 2D cross-sectional tomograms. AO-OCT images with axial resolution of $3\mu\text{m}$ and lateral resolution of $3.5\mu\text{m}$ show subcellular detail of structures in the human retina, including microcapillaries and the photoreceptor mosaic, extending over a retinal patch of $300\mu\text{m} \times 250\mu\text{m}$. When researchers simultaneously acquire AO-SLO and AO-OCT, eye motion recovered from the AO-SLO movie can be removed from the AO-OCT volume, and the AO-SLO movie frames can be combined to form a high-quality 2D image of the scanned retinal region. The key contributions made in this paper are: (i) an algorithm combining visualization and computation of a registration metric for overlapping volumes, and (ii) a system using that algorithm to assist users in stitching multiple retinal OCT volumes to extend high-resolution coverage.

2. Related work

Several systems for registration and display of volume data sets have been published recently. In [KYWH12], the authors present a system they call “magic lens visualization,” providing overall context and specialized rendering at the focus of attention. A similar system is described in [DPLR10]. Both of these systems support surgery planning, paying careful heed to appropriate visualization of diverse, complementary data modalities. Accurate data registration is a prerequisite for these systems, in contrast to ours. In [HBJP12] the authors report a system using GPU ray casting to visualize petascale electron microscopy volume data. The authors acknowledge the need for volume registration and stitching and treat it as an external process. [BSS*12] describes a complementary system for stitching 3D confocal ultramicroscopy (CU) data sets, with no function for visualization. Unlike OCT, the CU data sets are sparse; like OCT, subvolume relative positions (starting points for registration) are approximately known when starting registration. The system described in [DHA10] uses a multi-step process using recorded sensor position, a wavelet transform for feature extraction and finishing with intensity-based volume registration. The system is focused on noninteractive 3D registration of ultrasound data sets and contains no visualization component. We note that the character of the data in retinal OCT volumes includes a complex network of sometimes poorly resolved vessels and other structures, in between structures visible as bright bands (on slices) or sheets (in volumes)

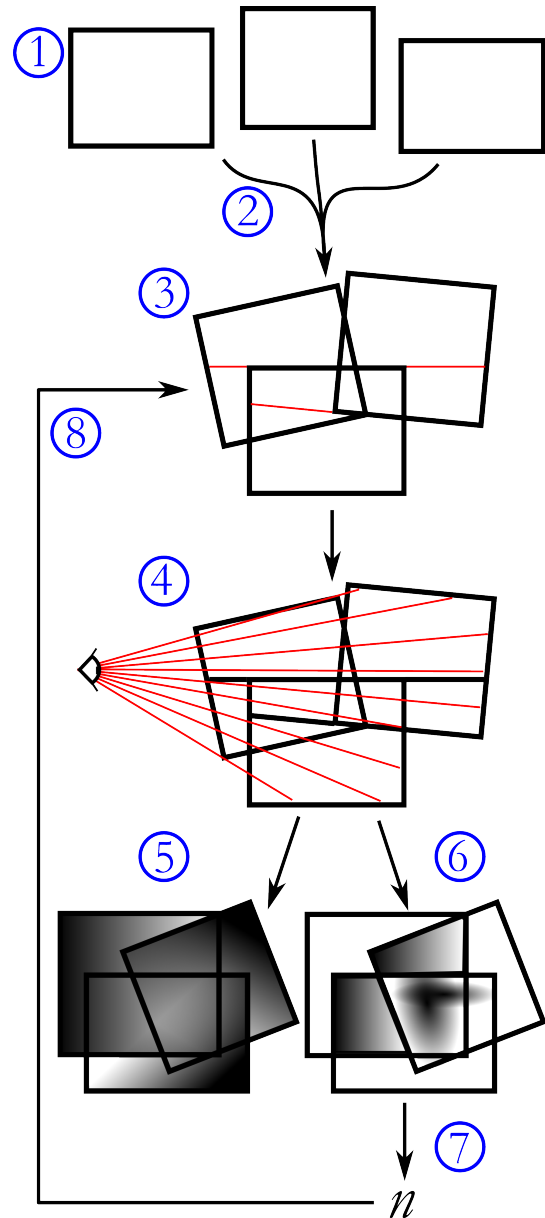


Figure 1: Overview of the algorithm. (1) Input volumes. (2) User adjusts position. (3) Partition into binary space partition (BSP) tree. (4) Ray casting. (5) Each ray produces a color value for its pixel. The resulting image is displayed on the screen. (6) Each ray that intersects a volume overlap also produces an intermediate metric result. The resulting image is stored in a texture. (7) The metric image is reduced to a single value. (8) An optimization step adjusts volume positions. Note that the image texture and the metric intermediate values shown at (5) and (6) represent the view from the eye point, a quarter turn from the representation at (4).

which are fundamentally different from the images of large bones used for successful registration by the authors. Specific to OCT data sets, [ZFC*08, ZCF*09] present a system that lets the user stitch multiple high-resolution AO-OCT data sets, allowing axis-aligned translation in whole-voxel increments. Image combination was accomplished by choosing the maximum intensity of overlapping voxels. We found [LLHY09] to be invaluable, and based our visualization procedure on the described algorithm.

3. Methods

Our method is summarized in Figure 1. The user supplies several input volumes (1) and adjusts their positions (2). The superimposed input volumes are divided using a binary space partitioning (BSP) tree (3) on the CPU, using the boundaries of the input volumes as partition planes. At the end of the partitioning, all the BSP tree leaf volumes (“cells”) intersect a constant number of input volumes. In other words, no cell is cut through by the boundary of an input volume. The ray casting algorithm is executed on the GPU through the BSP tree cells (4). The output of ray casting is an image (5), displayed on the screen, and a 2D set of intermediate metric results (6), one per ray, stored in a graphics texture. The metric image is reduced to a single metric value (7) and used as input to an optimizer which adjusts the input volume positions (8). After the user specifies an initial placement, an optimizer thread repeatedly runs steps (3) through (8), using the simulated annealing algorithm [KGJV83] to choose successive relative positions. We use simulated annealing because it does not need a gradient estimate, only the function value at any given point, and because it deals well with the high amounts of noise we find in OCT volumes. However, other optimizers could also be successfully used. As with all registration optimization problems, this system is susceptible to initial placement of component volumes which fall outside the capture region of the optimizer.

With changing eye position and distance to the displayed volumes, the same alignment may result in different values being reported by the metric-calculating part of the raycasting operation. Thus, only metric values are computed from the same eye point are comparable. Practically, this means the simulated annealing registration must restart when a user moves a volume or the global point of view. When an overlap region is affected by the adjustment of a volume’s position while not in view, we compute a metric using the same ray casting algorithm, but we do not compute a color value for on-screen display. Finally, we must deal with “island” volumes that do not overlap with any other volume. In this case, a registration metric cannot be calculated using ray casting. If the user provides a shared 2D background image, we calculate a metric value based on the correlation of the island volume’s 2D projection with the background image. Lack-

ing even such a background image, the island volume has no quantifiable position so our algorithm, by default, ignores it.

We use the cells of the BSP tree to render overlapping volumes. Rendering rear-facing cell boundary polygons provides us with the position each ray will exit the BSP necessary for both accurate volume rendering and metric calculation. Rendering front-facing polygons triggers the ray casting algorithm for each pixel, as well as metric calculation for cells intersecting multiple volumes. Our algorithm, outlined below, is adapted from [LLHY09] by the addition of the metric calculation.

1. For each BSP cell, ordered front to back,
 - a. Render rear-facing polygons; record fragment depth to texture.
 - b. Render front-facing polygons to trigger ray cast.
 - Use depth texture to determine ray exit points
 - At each sample along ray, accumulate color
 - If the current BSP cell intersects more than one volume, accumulate cross-correlation.
 - c. Render front-facing polygons once more, to combine accumulated color with that from previous cells.
2. Send color to screen and metric (if any) to metric texture.
3. Reduce metric texture with GPU code to one value

The registration metric m that we calculate is the normalized cross-correlation, shown in Equation 1. Here, S denotes the set of all ray casting sample positions, s denotes a particular ray casting sample position, v_i denotes the i th input volume (of n overlapping volumes), and \bar{v}_i and σ_{v_i} denote the mean and standard deviation of the values composing v_i .

$$m = \frac{1}{|S|} \sum_S \prod_{i=1}^n \frac{(v_i(s) - \bar{v}_i)}{\sigma_{v_i}} \quad (1)$$

Simply stated, at each ray casting sample position we take the product of all the normalized image values at that position. When a ray terminates, it stores a color value for its pixel in the image buffer, the number of sample points the ray sampled to an auxiliary buffer, and the sum of the value products at each of the ray’s sample points to a second auxiliary buffer. The auxiliary buffers are summed to produce the count of all raycast sample sites $|S|$ and the sum of all cross-correlation components, which are used in the final metric calculation.

The common technique of early ray termination cannot be applied to our combined metric and visualization method. Even though a ray may accumulate enough opacity that its color will not change over further sample steps, metric calculation must still be done at those sample steps, so all rays must run to volume exit. However, the ray may still skip empty space since it lacks volume data on which to compute a metric.

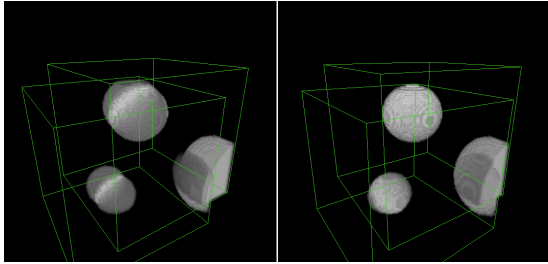


Figure 2: A registration example using a simple analytical data set. Two volumes were sampled from a function which gave a small, bright ellipsoid and a larger, dimmer ellipsoid. Left: an approximate registration done by the user. Right: registration refined by our system to a closer alignment.

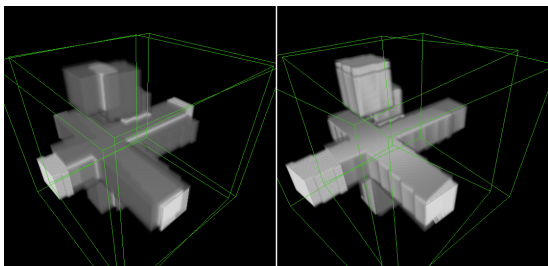


Figure 3: A second example of three orthogonal rectangular bars, with a gap in the Z-axis bar and an extra blob indicating first quadrant. Left: an approximate registration done by the user. Right: registration refined by our system to a closer alignment. The perturbed volume was rotated as well as translated, which our system was able to correct.

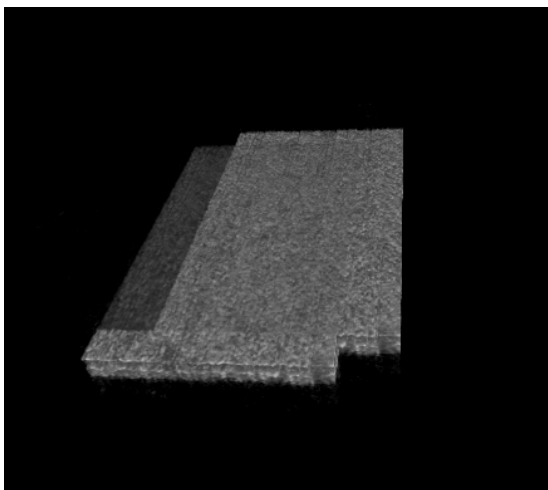


Figure 4: A pair of OCT scans registered and combined using our system. Work is ongoing to tune the metric calculation to deal with the noise and speckle patterns characteristic of retinal images.

Data set	Data set size	fps	Final error
Fig. 2	$64 \times 64 \times 64$	24	1.673, 0.0742
Fig. 3	$64 \times 64 \times 64$	20	6.089, 0.0257
Fig. 4	$256 \times 256 \times 256$	15	

Table 1: Data set size, average frames displayed per second, and final position error attained performing the registrations presented in each Figure. Final position error is presented for the generated data sets, and consists of the norm of the displacement in voxels and the magnitude of the aggregate rotational error in radians.

4. Results

Figures 2 through 4 show some preliminary results from our system. Table 1 reports timing and error results from the registrations shown. The analytical data set examples work well. A rough registration accomplished by the user is easy and quick to accomplish, especially with simple analytical data sets, and the improvement in position given by the automatic registration is substantial, as shown by the figures for final registration error. Although we continue to tune the system for use with AO-OCT data sets (the primary anticipated use for the technique at this point) we are able to achieve good registration with interactive frame rates for data set with sizes typical of our main use case. We expect to integrate this system with our current visualization application and make semi-automatic guided volume stitching part of our normal work flow for more informative clinical evaluation.

References

- [BSS*12] BRI A., SILVESTRI L., SACCONI L., PAVONE F., IANNELLO G.: Stitching terabyte-sized 3D images acquired in confocal ultramicroscopy. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* (May 2012), pp. 1659–1662. doi:10.1109/ISBI.2012.6235896. 2
- [DHA10] DALVI R., HACIHALILOGLU I., ABUGHARBIEH R.: 3D ultrasound volume stitching using phase symmetry and Harris corner detection for orthopaedic applications. *Proc. SPIE 7623* (2010), 762330–1. doi:10.1117/12.844608. 2
- [DPLR10] DIEPENBROCK S., PRASSNI J.-S., LINDEMANN F., ROPINSKI T.: Pre-operative planning of brain tumor resections, 2010. 2
- [HBJP12] HADWIGER M., BEYER J., JEONG W.-K., PFISTER H.: Interactive volume exploration of petascale microscopy data streams using a visualization-driven virtual memory approach. *Visualization and Computer Graphics, IEEE Transactions on* 18, 12 (Dec. 2012), 2285–2294. doi:10.1109/TVCG.2012.240. 2
- [KGJV83] KIRKPATRICK S., GELATT JR. C. D., VECCHI M. P.: Optimization by simulated annealing. *Science* 220, 4598 (May 1983), 671–680. 3
- [KYWH12] KIRMIZIBAYRAK C., YIM Y., WAKID M., HAHN J.: Interactive visualization and analysis of multimodal datasets for surgical applications. *Journal of Digital Imaging* 25 (2012), 792–801. doi:10.1007/s10278-012-9461-y. 2

- [LLHY09] LINDHOLM S., LJUNG P., HADWIGER M., YNNERMAN A.: Fused multi-volume DVR using binary space partitioning. *Computer Graphics Forum* 28, 3 (2009), 847–854. doi:10.1111/j.1467-8659.2009.01465.x. 3
- [WH81] WEBB R. H., HUGHES G. W.: Scanning laser ophthalmoscope. *Biomedical Engineering, IEEE Transactions on BME-28*, 7 (July 1981), 488–492. doi:10.1109/TBME.1981.324734. 2
- [WLK*02] WOJTKOWSKI M., LEITGEB R., KOWALCZYK A., BAJRASZEWSKI T., FERCHER A. F.: *In vivo* human retinal imaging by Fourier domain optical coherence tomography. *J. Biomed. Opt.* 7 (2002), 457–463. doi:10.1117/1.1482379. 2
- [ZCF*09] ZAWADZKI R. J., CHOI S. S., FULLER A. R., EVANS J. W., HAMANN B., WERNER J. S.: Cellular resolution volumetric *in vivo* retinal imaging with adaptive optics–optical coherence tomography. *Opt. Express* 17, 5 (Mar 2009), 4084–4094. doi:10.1364/OE.17.004084. 2, 3
- [ZFC*08] ZAWADZKI R. J., FULLER A. R., CHOI S. S., WILEY D. F., HAMANN B., WERNER J. S.: Improved representation of retinal data acquired with volumetric Fd-OCT: co-registration, visualization, and reconstruction of a large field of view. In *Proceedings of SPIE, the International Society for Optical Engineering* (2008), Society of Photo-Optical Instrumentation Engineers, pp. 68440C–1. 3
- [ZJP*11] ZAWADZKI R. J., JONES S. M., PILLI S., BALDERAS-MATA S., KIM D. Y., OLIVIER S. S., WERNER J. S.: Integrated adaptive optics optical coherence tomography and adaptive optics scanning laser ophthalmoscope system for simultaneous cellular resolution *in vivo* retinal imaging. *Biomedical Optics Express* 2, 6 (June 2011), 1674–1686. 2