Real-Time Three-Dimensional Visualization of Standard Light Microscope Image Sequences for Microassembly

Michael A. Greminger

Department of Mechanical Engineering
University of Minnesota
Minneapolis, MN

James F. (Red) Jones

Intelligent Systems and Robotics Center (ISRC)
Sandia National Laboratories
Albuquerque, NM

Abstract

In this paper we present a novel method to visualize a sequence of standard light microscope images in 3D for teleoperated microassembly. This method increases the depth of field of the microscope images over 60 times. A volume rendering approach is used to construct the images by making use of the functionality built into current 3D graphics hardware. Through the use of this hardware, the views can be constructed at interactive rates. Two views can be constructed simultaneously to construct stereo pairs for a 3D display. This method greatly assists the operator of a microassembly system by providing the same type of visual feedback that is available during macroassembly.

1 Introduction

One of the main challenges for in developing microassembly systems is providing the operator with the same visual cues that are available when performing a macroassembly task. There are two aspects of microscope optics that make this goal difficult. First, high resolution microscope images have a very low depth of field. Because of this, the operator can only see a very narrow slice of the scene at any given time. Second, high resolution stereo image pairs cannot be provided by an optical microscope. Stereo vision is necessary in order to locate objects relative to each other in a complex scene.

This paper will describe a method to visualize a sequence of standard light microscope images in order to provide depth information while maintaining high resolution. Not only does the technique presented here increase the depth of field, it also allows the scene to be viewed from angles other than that used to obtain the original image sequence. It will be shown that this 3D visualization of microscope image sequences can been performed at interactive rates by making use of modern 3D graphics hardware. By rendering multiple views at once, stereo pairs can be generated so that the operator gains a sense of depth that will aid in the microassembly procedure. The increased depth of field and the stereo image pairs provided by this approach supply the operator of a teleoperated microassembly system with the same type of visual cues available during macroassembly.

2 The Limits of Microscope Optics

As mentioned previously, there are two limitations inherent to high resolution microscope optics that limit their usefulness for teleoperated microassembly. The first is the small depth of field and the second is the difficulty in obtaining high resolution stereo image pairs. Both of these limitations are described in this section.

2.1 The Low Depth-of-Field Problem

The relationship between microscope resolution, R, and numerical aperture (NA) is [3]:

$$R = \frac{0.61\lambda}{\text{NA}} \qquad (1)$$

where λ is the wavelength of light. The numerical aperture is defined as $\sin(\alpha)$ where α is a measure of the size of the light gathering cone of the objective lens (see Figure 1).The numerical aperture also determines the depth of field, D, of the microscope system. The relationship is [3]:

$$D = \frac{n\lambda}{(\text{NA})^2} + \frac{ne}{(M)(\text{NA})} \qquad (2)$$

where n is the refractive index of the medium (air in this case), M is the magnification of the image as it appears on the image sensor and e is the smallest distance that can be resolved on the image sensor. It can be seen that a large numerical aperture leads to a high resolution system with a small depth of field. Because of this small depth of field,
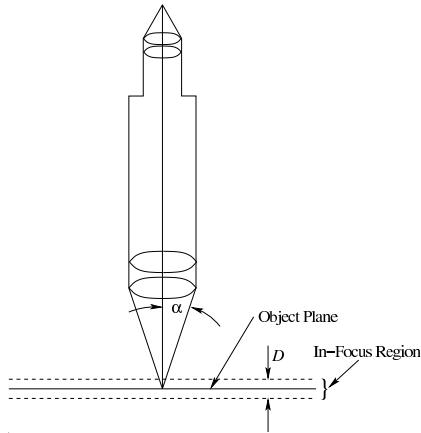
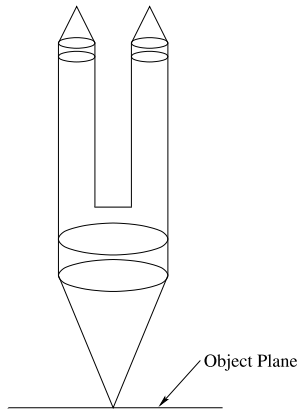Figure 1: Microscope depth of field.

Figure 2: The common main objective (CMO) stereo microscope configuration.

very little depth information is available in a single microscope image.

It will be shown that this low depth of field can be used to our advantage if a sequence of microscope images is captured with each image at a different focus plane.

2.2 The High Resolution Stereo Problem

The need for high numerical aperture optics also limits the ability to obtain stereo image pairs. Figure 2 shows a stereo microscope configuration called the common main objective (CMO) stereo microscope. Half of the light that enters the objective lens is directed to each eye-piece. Since only half of the light cone entering the objective lens reaches each eye, the effective numerical aperature of the system will be reduced when compared to a non-stereo system. Also, because of the low depth of field of high NA objective lenses, only a single plane will be in focus which will drasticly reduce 3D effect obtained by the user. These factors limit the performance of an optical stereo microscope system.

3 Previous Work

There have been two dominant methods to increase the depth of field of the conventional light microscope. The first method is to take multiple microscope images at different positions along the z-axis of the microscope (the z-axis is defined to align with the optical axis of the system). These images are obtained by actuating the objective lens of the microscope as shown in Figure 3. The in-focus regions of each image are then combined to create one large depth of field image. The increase of depth of field depends on the number of images used. This approach was used by Widjanarko et al. [11]. Because all of the original images are combined into one final image, the depth information is lost in this approach. Ohba et al. [8] preserve the depth information separately from the increased depth of field image but this depth information must be rendered separately and can not be rendered at interactive rates.

Nayar et al. [7] and Subbarao et al. [9] attempted to retain the depth information contained with the image stack by using a technique called shape from focus. In this technique, surfaces are fit to the in-focus regions of the images in an attempt to reconstruct the geometry in the scene. This approach is computationally expensive so it cannot be performed at interactive rates and the assumption is made that only one surface exists in a scene. If one object lies in front of another, one surface will be fit to both objects.

The second dominant approach to increasing depth of field is known as wave-front coding and makes use of a special lens placed after the objective lens of an optical system. This method was developed by Dowski and Cathey [1]. This lens uniformly blurs the resultant image in such a way that, upon a post processing deconvolution step, an image is obtained with a depth of field increased by up to 20 times. This method does not provide the user with any depth information about the scene.

The method presented here is able to generate a large depth of view image from an arbitrary view point allowing for stereoscopic display of the scene. This is done at interactive rates so it is useful as a feedback mechanism for microassembly.

4 System Overview

The method presented here first acquires a series of images along the optical axis in a manner similar to the approach used in the shape from focus method, but rather than fit surfaces to the data, the data is rendered directly using a technique known as volume rendering [6]. The stack
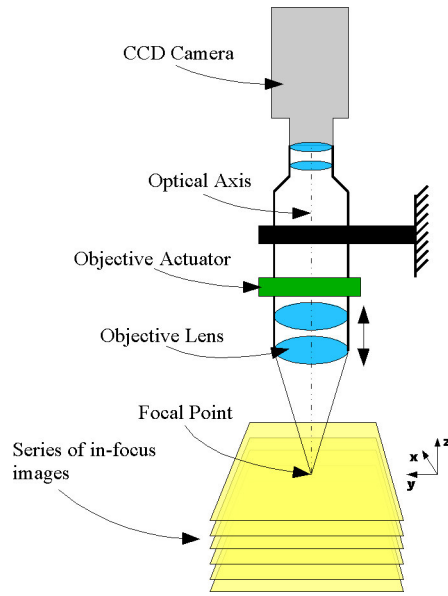
Figure 3: Z-axis scanning approach to increase depth of field.

Figure 4: Steps to render a microscope z-volume.

of images is stored as a volume that will be referred to as the microscope z-volume. By storing the stack of images as a volume, the depth information contained within the images is preserved. After some preprocessing of the volume data using a process known as classification, this volume can be rendered. Volume rendering allows the volume to be viewed from any vantage point. Figure 4 shows the steps involved in volume rendering a microscope z-volume. The process of volume rendering and classification will be described in the following sections. In a similar fashion, two separate views can be generated simultaneously in order to display the scene to the viewer in stereo through the use of shutter glasses or some other stereoscopic display technology.

5 Volume Rendering by the Ray Casting Technique

Volume rendering is a technique by which three-dimensional volume data can be rendered to a two-dimensional image for display. Volume data consists of samples of a function defined in three spatial dimensions. Volume data is commonly rendered in the field of medical imaging where there are technologies, including MRI, that provide volume data directly. Scientific visualization is another example where volume rendering is used to visualize three-dimensional fields such as fluid flow. Each sample in the volume is called a voxel. In the case of a microscope z-volume, each sample represents a color value, or an intensity value for grey-scale images, that was acquired from a microscope. Before volume data can be rendered prop-
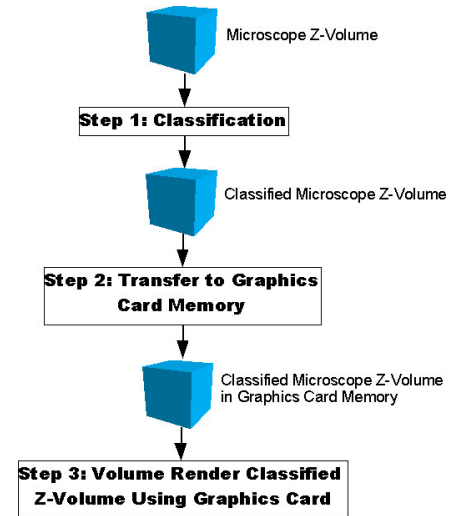
erly, the data must be preprocessed to define an opacity and color for each voxel of the original volume data. The process of mapping from raw volume data to opacities and colors is known as classification. A radiologist may, for example, classify skin tissue in MRI volume data to have a low opacity so that internal organs are visible in the final volume rendered image. For the microscope z-volume, opacity will be based on a focus measure for each layer of the volume data. The in-focus portion of each image slice will be set to a high opacity while the out-of-focus regions of each slice will be set to a low opacity so they do not appear in the final image. The focus measure used will be discussed in the next section. Assigning a color to each voxel from the microscope z-volume is straight forward because the data in the volume is already composed of the color information present in the microscope images that form each layer.

Once the volume data has been assigned opacity values the final image will be rendered using the ray casting approach as proposed by Levoy [6]. For each pixel in the image plane, a ray is shot into the volume and samples are taken at regular intervals along the ray. If these ray samples do not fall at a voxel location, tri-linear interpolation is used to compute the ray sample's opacity and color from the eight nearest voxels. The ray samples are combined by the following compositing equation to determine the final grayscale intensity C for the pixel from which the ray originates:

$$C = \sum_{i=0}^{N} \left[C_i \alpha_i \prod_{j=i+1}^{N} (1 - \alpha_j) \right] \qquad (3)$$

where C_i is the grayscale intensity of the ith ray sample (numbered from furthest away to closest), α_i is the opacity
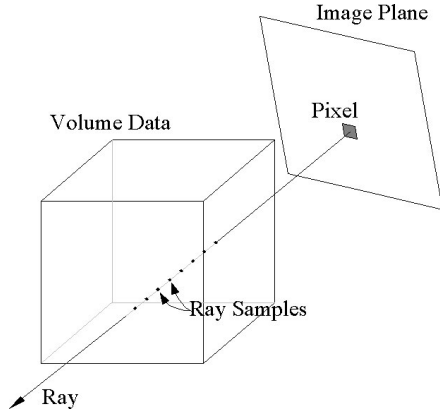
Figure 5: Ray casting approach to volume rendering.

of the ith ray sample, C_0 is the grayscale intensity of the background and $\alpha_0 = 1$. If the full color information is available for each voxel, the above equation is evaluated for each color component in the volume data: red, green and blue.

Ray casting is computationally expensive and generally cannot be performed in real time with software. Current graphics hardware can perform the calculations for ray casting in hardware through the use of an extension of OpenGL called 3D textures[1] [2]. Using this hardware, the volume rendering can be done in real-time on a PC.

6 Classification

As was mentioned in the previous section, every voxel in the microscope z-volume must be assigned an opacity value. The areas of each layer in the z-volume that are in focus will be assigned high opacities and the areas that are not in focus will be assigned low opacities to make those regions transparent.

A focus measure provides a means to measure the level of focus in a region of an image. There are many available focus measures [5][10]. Two of the most commonly used are the magnitude of gradient:

$$G(x,y) = \sqrt{\frac{\partial I}{\partial x}^2 + \frac{\partial I}{\partial y}^2} \qquad (4)$$

and the Laplacian:

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \qquad (5)$$

where I is the image grayscale intensity of the original images that make up the microscope z-volume. If the image

---

[1]3D textures became an official part of OpenGL in the OpenGL 1.2 specification and some consumer level graphics cards now support 3d textures.
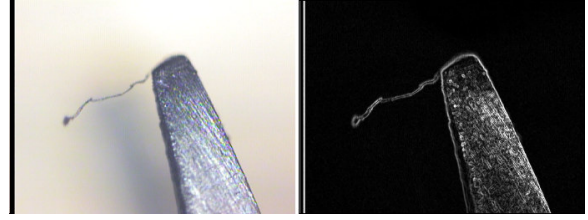

Figure 6: Original microscope image on left and opacity image on right (white indicates high opacity).

is in color, it must be converted to grayscale before computing the Laplacian or the gradient magnitude. Typically the magnitude of the gradient is approximated by:

$$G(x,y) = \left|\frac{\partial I}{\partial x}\right| + \left|\frac{\partial I}{\partial y}\right| \qquad (6)$$

to avoid the expensive square root calculation. Since the Laplacian focus measure requires second derivatives, it introduces more noise to the focus measure than the magnitude of gradient. For this reason, we will use (6) as the focus measure. Figure 6 shows an opacity image calculated by this method.

Determining voxel color is straight forward because the z-volume data contains color information. Therefore, in the case of grayscale images, the grayscale intensity C and opacity α values for each voxel are computed using the following formulas:

$$C(x,y,z) = I(x,y,z) \qquad (7)$$

$$\alpha(x,y,z) = \left|\frac{\partial I(x,y,z)}{\partial x}\right| + \left|\frac{\partial I(x,y,z)}{\partial y}\right| \qquad (8)$$

If the original microscope z-volume consisted of full color images, (7) would consist of three components for red, green and blue. Once a color and opacity has been assigned to each voxel in the microscope z-volume, the ray-casting operation can be performed as described in the previous section.

7 Local Maximum Filtering

As the microscope sweeps along its z-axis, features will gradually come into focus and then go out of focus. It is desirable to only have the most in-focus layer of a feature appear in the final volume rendered image. To insure this, each voxel's opacity is compared the opacity of neighboring voxels in the plus and minus z-direction. If the voxel's opacity is not greater then its neighbor's opacity, there is a more in focus representation of that feature and the voxel's opacity is set to zero. This filter only allows the local maximum in-focus feature appear in the final volume rendering

of the scene. For the results presented in this paper each voxel's opacity is compared to the three voxels in front of and behind it in the z-direction. Also, if a voxel's opacity is above a predetermined threshold, it is not compared to its neighbors because there is a high probability that it represents a true feature.

8 Results

The microscope z-volume was obtained by manually moving the microscope by 100 μm increments for each layer of the z-volume. Each layer of the z-volume has the dimensions of 512x384 and 25 layers were acquired. The microscope used to obtain the images has a numerical aperture of 0.14 and the size of each pixel in object space is 3.227 μm (therefore, $e/M = 3.227$ μm where e and M are defined as in (2)). Assuming that the wavelength of light is 500 nm and that the index of refraction of air is 1.0, the resolution of the system is 2.2 μm and the depth of field is 37 μm using (1) and (2) respectively.

The first microscope z-volume that was rendered is a set of images of a pair of tweezers. Two slices of the z-volume are shown in Figure 7a and b. It can be seen from the figures that one jaw of the tweezers is nearer the microscope then the other. Figure 7c and d show the volume rendering of the tweezers images. Two different views of the tweezers are shown. The depth of field is increased by a factor of over 60 times.

A second, more complex, z-volume was also rendered. This is shown in Figure 8. Figures 8a-e show some of the 25 slices, Figure 8f shows a volume rendering recreating the original view with increased depth of field and Figures 8g - h show two additional views of the z-volume. These images show a pair of tweezers in front of a LIGA micropart (for a description of the assembly procedure used with this micropart, see [4]). It can be seen that, even though the tweezers occlude a portion of the micropart, the occluded portions of the micropart are rendered correctly in the volume renderings. This is because, as can be seen in Figure 8c, the tweezers are out of focus when the micropart is in focus so the micropart is assigned the correct opacity values in the occluded region. Existing methods of three-dimensional reconstruction such as shape-from-focus are unable to handle this situation correctly because these methods assume that there is only a single surface in the image and that no surface occludes another.

The previous two cases demonstrate the method for a single z-volume. For a practical microscope system it is desirable to have the microscope z-volume updated in real-time to give the user feedback for micromanipulation or microassembly operations. In order to test how quickly the
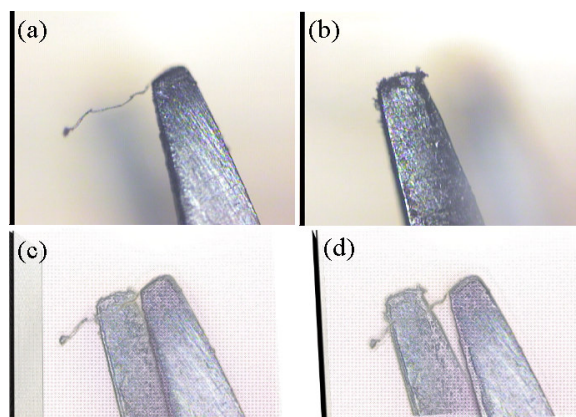

Figure 7: Layers 3 (a) and 22 (b) of the z-volume. Two volume rendered views of the z-volume: (c) and (d).

z-volumes can be continuously sent to the graphics hardware and rendered, 5 different z-volumes were loaded into computer RAM as if they were just obtained from a microscope. The z-volumes are then classified, sent to video RAM and rendered one-by-one in a continuous loop (the opacity values are not stored but are recalculated each time a z-volume is redisplayed). With each z-volume consisting of 25, 512x384 pixel, grayscale images, the display could be updated at 12 fps (on a Pentium 4, 3 GHz, with a Radeon 9600 Pro 128 MB graphics card). In order to obtain this rate, local-maximum-filtering was disabled and the image was rendered in grayscale rather than in full color. Figure 8i shows the quality of image that was rendered at this rate.

9 Conclusions

We have presented a method for creating high quality three-dimensional views of standard light microscope image sequences in real-time. This is made possible by making use of the 3D texture mapping features available in current graphics hardware. Multiple views can be generated simultaneously to create a stereoscopic view of the scene. These capabilities give the user the same type of visual cues that are available when performing manipulation tasks at the macroscale.
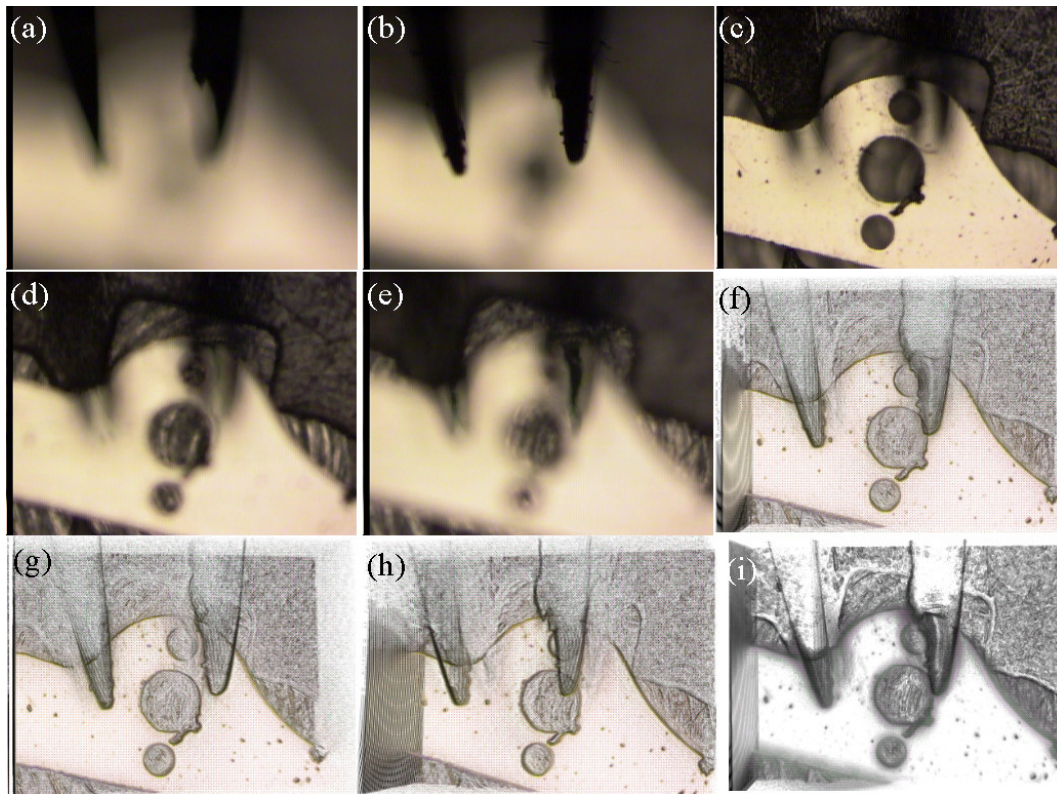
Acknowledgments

Figure 8: Selected layers from the microscope z-volume (a) - (e) (layers 1, 8, 19, 21 and 24 respectively). Volume rendering of original view (f) and of two additional views (g) and (h). Volume rendering of original view without applying local maximum filter (i).

References

[1] Dowski, E.R., Cathey, W.T., "Extended depth of field through wave-front coding," *Applied Optics*, Vol. 36, No. 35, Dec. 1997.

[2] Gelder, V., Kim, K., "Direct volume rendering with shading via three-dimensional textures," *Volume Visualization Symposium Proceedings*, pp. 23-30, San Francisco, CA, October 1996.

[3] Inoue, S., Spring, K.R., *Video Microscopy*, Plenum Press, New York, 1997.

[4] Jones, J.F., Kozlowski, D.M., Trinkle, J.C., "Microscale Force-fit Insertion," *Journal of Micromechatronics*, Vol. 2, No. 3-4, pp. 185-200, 2004.

[5] Kautsky, J., Flusser, J., Zitova, B., Simbuerova, S., "A new wavelet-based measure of image focus," *Pattern Recognition Letters*, Vol. 23, No. 14, pp. 1785-1794, December 2002.

[6] Levoy, M., "A Hybrid Ray Tracer for Rendering Polygon and Volume Data," *IEEE Computer Graphics and Applications*, pp. 33-40, March 1990.

[7] Nayar, S.K., Nakagawa, Y., "Shape from Focus," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 16, No. 8, pp. 824-831, August 1994.

[8] Ohba, K., Ortega, J.C.P., Tanie, K., Tsuji, M., Yamada, S., "Microscopic vision system with all-in-focus and depth images," *Machine Vision and Applications*, Vol. 15, pp. 55-62, Dec., 2003.

[9] Subbarao, M., Choi, T., "Accurate Recovery of Three-Dimensional Shape from Image Focus," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 17, No. 3, pp. 266-274, March 1995.

[10] Subbarao, M., Tyan, J.K., "Selecting the Optimal Focus Measure for Autofocusing and Depth-From-Focus," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 20, No. 8, pp. 864-870, August 1998.

[11] Widjanarko, T., Hardie, R.C., "A post-processing technique for extending depth of focus in conventional optical microscopy," *Optics & Laser Technology*, Vol. 34, No. 4, June 2002.