Digital Stereoscopic Imaging

A. Ravishankar Rao[a] and Alejandro Jaimes[b]

[a]IBM T.J. Watson Research
Center PO Box 218
Yorktown Heights NY 10598

[b]Columbia University
Department of Electrical Engineering
New York, NY 10027

ABSTRACT

The convergence of inexpensive digital cameras and cheap hardware for displaying stereoscopic images has created the right conditions for the proliferation of stereoscopic imaging applications. One application, which is of growing importance to museums and cultural institutions, consists of capturing and displaying 3D images of objects at multiple orientations. In this paper, we present our stereoscopic imaging system and methodology for semi-automatically capturing multiple orientation stereo views of objects in a studio setting, and demonstrate the superiority of using a high resolution, high fidelity digital color camera for stereoscopic object photography. We show the superior performance achieved with the IBM TDI-Pro 3000 digital camera developed at IBM Research. We examine various choices related to the camera parameters (aperture, focal length, focus), image capture geometry (distance of camera to the object, distance between the two camera positions, angle of elevation etc.) and suggest a range of optimum values that work well in practice. We also examine the effect of scene composition and background selection on the quality of the stereoscopic image display. We will demonstrate our technique with turntable views of objects from the IBM Corporate Archive.

Keywords: stereoscopic imaging, 3D imaging, 3D applications, digital camera, digital library.

1. INTRODUCTION

Stereoscopic photography has been used since the beginning of the century and yet, widespread application of 3D technology has been limited [Sawdai98]. One of the main reasons for this is the difficulty in providing the right viewing environment in terms of easily accessible 3D-display technology. A second reason is the fairly limited supply of high quality stereoscopic images. Recent advances in low cost computers, hardware for stereoscopic viewing [Halnon98, Qualman98] and digital image capture have created the right conditions for the proliferation of stereoscopic imaging applications on the personal computer platform.

The creation of high quality stereoscopic images is a challenging task for a variety of reasons. The viewer's perception of depth in a 3D application strongly depends on the image quality of the input capture device and the geometric accuracy with which the stereo images are captured. Image quality is typically evaluated in terms of color, resolution, contrast, and noise whereas geometric accuracy refers to the spatial alignment of the stereo images. If the objective of capturing stereoscopic images is the creation of a digital library for archival purposes or for scholarly study, then the amount of detail in the image is an important factor [Mintzer96]. Most of the research on image acquisition for digital libraries has been focused on issues such as faithful color reproduction and preservation of detail. The capture of stereoscopic images poses additional challenges especially when views of an object are desired at multiple orientations. A desirable setup is one that minimizes the

---

[a]A. Ravishankar Rao is with IBM TJ Watson Research Center. PO Box 218 Yorktown Heights, NY 10598. E-mail: ravirao@us.ibm.com
[b]Alejandro Jaimes is with the Image and Advanced TV Lab & Electrical Engineering Department, Columbia University, 1312 SW Mudd Mailcode 4712 Box No. F8, New York, NY 10027. E-mail: ajaimes@ee.columbia.edu Web: http://www.ee.columbia.edu/~ajaimes

amount of human error, and allows for semi-automatic capture of the images. Such a setup is particularly useful in applications in which large numbers of objects need to be scanned which are difficult to handle or are of great value, such as museum pieces or jewelry.

In this paper, we describe the superior performance achieved by using a high resolution, high fidelity digital camera and present our setup for a semi-automated stereoscopic capturing environment. This methodology is suitable for an application of growing importance to museums and cultural institutions in which an object is placed on a turntable and automatically scanned at multiple orientations. We examine in detail the various parameters that must be considered to successfully obtain stereoscopic images of objects and give guidelines for setting such parameters. These include choices related to the camera settings, such as aperture, focal length, etc., and image capture geometry such as distance of camera to the object, distance between the two camera positions, angle of elevation etc. We also examine the effect of scene composition and background selection on the quality of stereoscopic images.

For capturing the stereoscopic images we use the IBM TDI-Pro 3000 [Mintzer96, Giordano99] developed at IBM Research and a Kaidan computer-controlled turntable [www.kaidan.com] for automatically obtaining multiple views of the object. For viewing we use off-the-shelf hardware from NuVision Technologies [www.nuvision3d.com] on a standard CRT monitor.

In the following section, we discuss the advantages of using a high-end digital camera for stereoscopic applications. In Section 3 we describe the setup of our application and discuss the most important practical issues we have encountered. Section 4 describes issues related to the display of stereoscopic images. In Section 5 we discuss results and finally in Section 6 we discuss applications of our techniques.

2.   THE CASE FOR HIGH RESOLUTION DIGITAL CAMERAS

In this section we discuss the superiority of high-resolution digital cameras in comparison with film, and lower end digital cameras.

2.1 Film
It is widely accepted that digital photographic systems have definite advantages of convenience in the context of computer based environments. Digital cameras are advantageous in their instantaneous output compared to film, which takes considerable time to expose and print. It has been shown that digital acquisition systems offer fundamental technical advantages over silver-halide film based systems [Shaw98]. Though the resolution of film may exceed that of high-end digital cameras, digital images contain less noise and grain. This is due to the high signal-to-noise ratio in digital acquisition systems, for instance CCD based imaging systems. Film contains considerably more sources of noise, including spread of grain sizes, emulsion-layer grain shielding effects and so on [Shaw98]. Furthermore, digital cameras can be readily calibrated, giving rise to consistently accurate color reproduction. Other issues related to color reproduction and details about film/digital can be found in [Giorgianni98].

2.2 Digital
High-end digital cameras produce images that have superior image quality in terms of resolution, sharpness of detail and accuracy of color reproduction compared to lower-end digital cameras. Table 1 compares these two types of cameras in terms of resolution, cost, and performance. Further details about these cameras may be found in [Publish99, Output98]. For the IBM TDI-Pro 3000, which was used to obtain images in this paper, the resolution is 3000 by 4000 pixels. It's fidelity can be measured via the color error in the calibration of a Macbeth chart, which is typically less than two CIE delta E units per color square on the average [Rao98]. Although lower-end digital cameras have a pixel count that is approaching that of higher-end cameras, the image quality they produce is inferior. There are many reasons for this, such as the amount of interpolated data used, the type of color filters used, patterning of the color filters on the CCD chips, the response of the CCD pixels, noise sensitivity and so on.

Type of digital camera	Typical resolution	Typical cost	Image quality
Lower end digital camera: e.g. Nikon Coolpix 900, Olympus D-430L, Kodak DC210, Agfa ePhoto 1280.	1200x900 pixels (1 megapixels)	Less than $1000	Fair to acceptable (e.g. for web publishing).
Higher end digital camera with area CCD arrays e.g., Kodak DCS-520, Dicomed Bigshot, Leaf Oneshot, PhaseOne LightPhase, Megavision S3	3000x2000 pixels or higher	$10,000 - $50,000	Excellent, suitable for moving and still professional photography.
Higher end digital camera with linear CCD arrays, e.g., PhaseOne PowerPhase, BetterLight Model 6000, IBM-TDI Pro 3000	3000x4000 pixels or higher	$20,000 - $60,000	Excellent, suitable for still professional photography.

Table 1. A comparison of digital cameras.

Since the images are captured at high resolution, it is possible to magnify the displayed images considerably without losing detail. Furthermore, due to superior color fidelity, digital stereoscopic images reproduce a compelling likeness to the original objects.

3. ISSUES IN CAPTURING STEREOSCOPIC VIEWS

In this section, first we describe the basic setup of our system and then discuss the main issues regarding the capture of multiple view stereoscopic images in such an environment.

3.1 Setup

The physical setup of our stereoscopic capturing environment consists of the following objects: (1) computer controlled turntable, (2) camera tripod, (3) background, (4) lighting, and (5) high-end digital camera. Figure 1 illustrates the main elements. As will be shown later, each of the objects in the environment plays a crucial role in the successful acquisition of stereoscopic images. The parameters for each object must be carefully controlled (e.g., uneven lighting for the right and left views will result in ghosting effects and difficulty in fusing of the stereo images by the viewer).

When photographing an object, the objective is to obtain two images of the same object to emulate the images perceived by our left and right eyes, respectively. Once the proper parameters are set, the two images of the object are obtained. After capturing the first one (e.g., left view), the second one (e.g., right view) is obtained by horizontally *translating* the camera along a plane perpendicular to the first view's optical axis. The camera is translated so that optical axis of the second view is parallel to the optical axis of the first view, since this emulates our retinal images in the best way [Howard95]. The parallel camera configuration avoids depth plane curvature and keystone distortion [Woods93]. Capturing of other images of the same object at other orientations is performed automatically as the turntable rotates according to the parameters set by the operator.

In the next section we present the relevant definitions and parameters of our system.
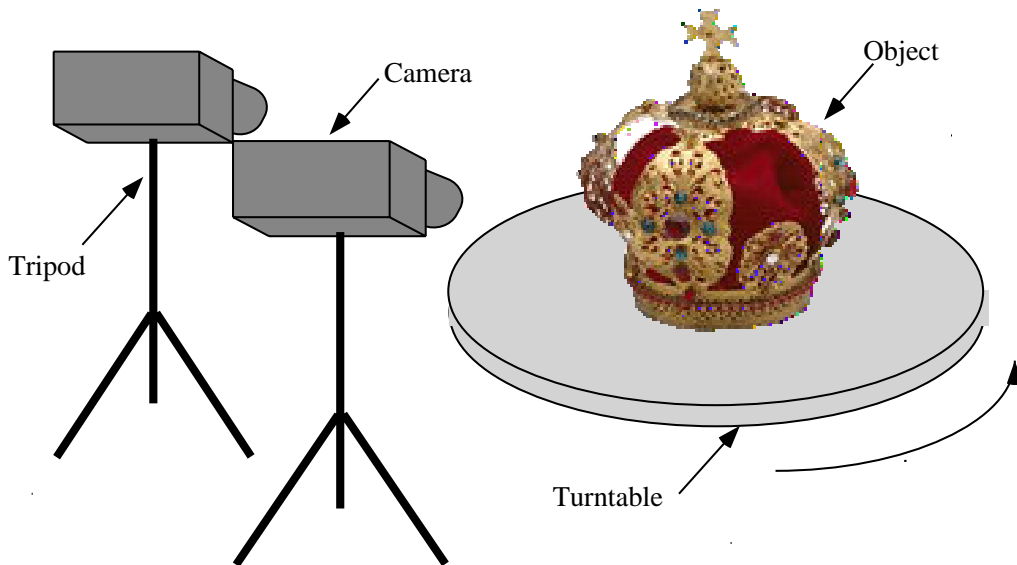
Figure 1. Basic Configuration

3.2 Environment definitions and parameters

In order to examine some of the important issues in a system like the one we are presenting, it is necessary to present only definitions and parameters that we consider of most importance. Further details can be found in [Howard95].

3.2.1 Basic definitions

Figures 2 and 3 show each of the elements that follow.
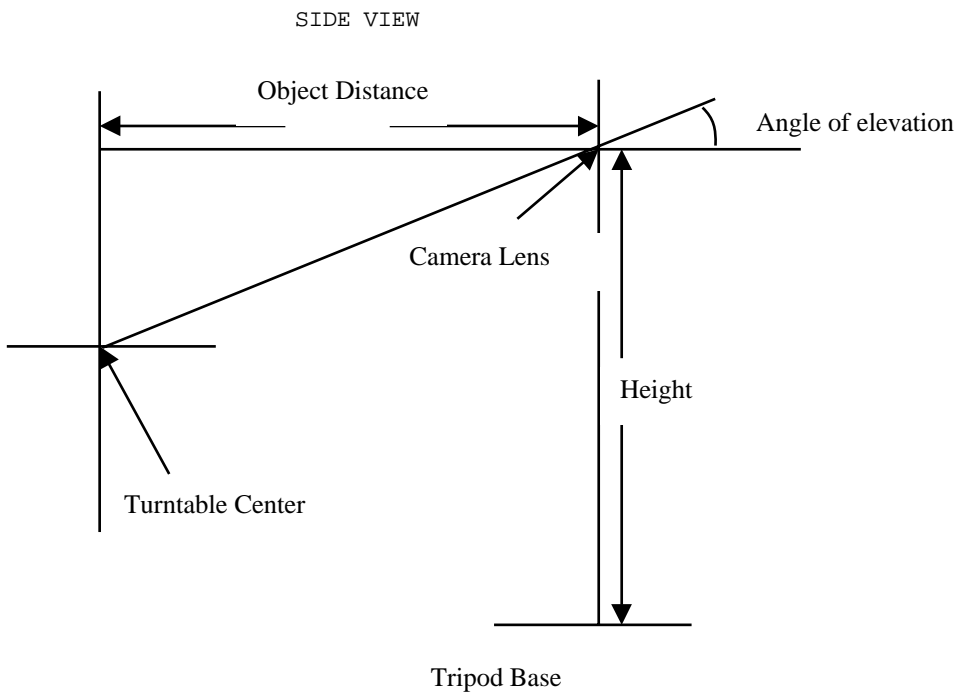
Figure 2: Side View.

Optical Axis: a line that passes through the center of the lens and is perpendicular to the image plane.

Baseline: a line between the lens centers of the left and right views. The baseline is perpendicular to the optical axes.

Turntable Center: point where the axis of rotation of the turntable and the plane of the table intersect.

Tripod Base: either the floor or any fixed point on the tripod from which camera height is measured.

3.2.2 Setup Parameters

Table 2 provides a template that can be used to record the scanning setup. Most of those parameters are described below:

Lens separation: amount by which the left and right views are separated horizontally; the length of the baseline.

Object Distance: distance from the baseline to the turntable center.

Height: vertical distance from the tripod base to the camera lens.

Angle: angle of elevation of the scanner. It is measured between a plane parallel to the turntable and the optical axis of the camera.

Figure 3. Top View
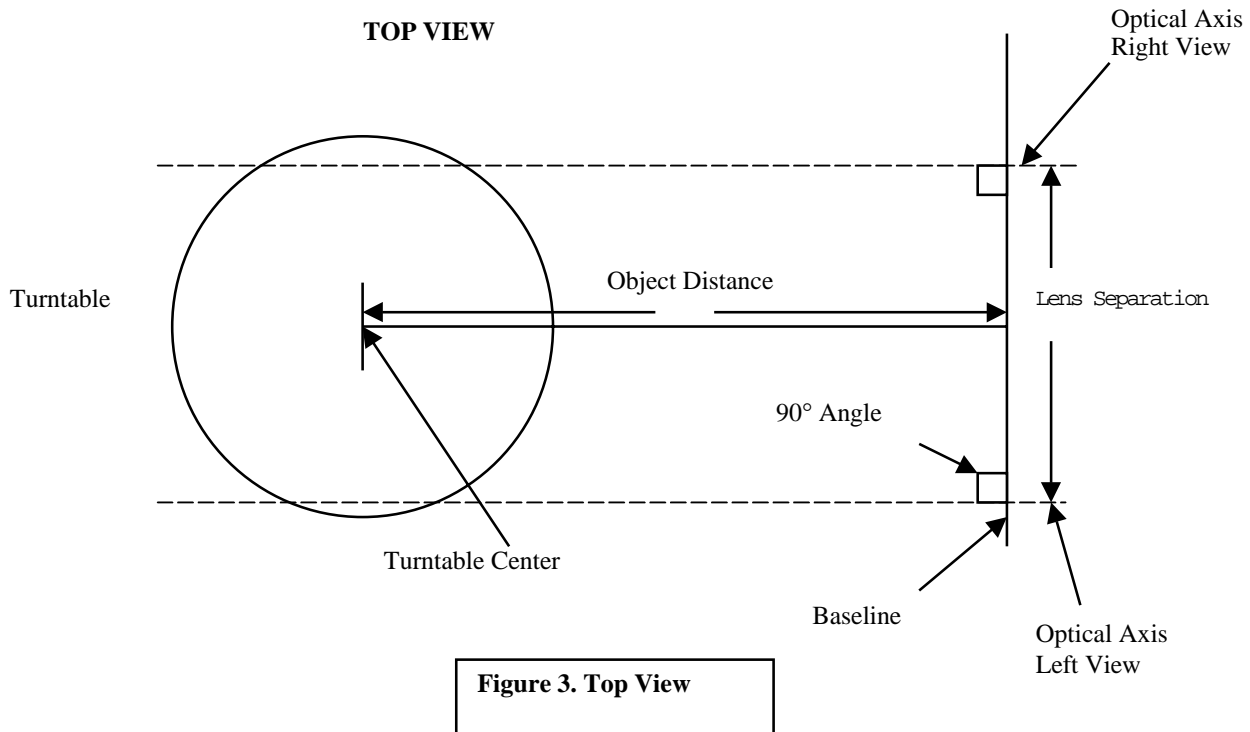
3.2.3 Camera Parameters

Aperture: the lens opening which determines amount of light that enters the camera and the depth of field. Measured in f-stops.

Focal Length: the distance between the optical center of a lens and the image plane. Determines how close the object appears in the scan. Measured in millimeters, depends on the lens used.

SETUP PARAMETERS AND SUGGESTED VALUES

Lens separation (cm)	6.5 cm
Object Distance (cm)	Usually between 4 and 10 feet.
Height (cm)	May vary depending on the object.
Scanner Angle (degrees)	May vary depending on the object.
Aperture (f-stops)	Small aperture gives larger depth of field
Focal Length (mm)	105 mm 1:4 lens (Rodenstock Apo Rodagon lens).
Focus (number)	Focus setting may be recorded. Useful for replicating scans.
Object Placement (description)	Important for replicating scans.
Initial Angle (degrees)	0
Rotation Angle (degrees)	30
Rotation Direction (increasing/decreasing)	increasing
Number of Views (1 to 360)	12 for a 30 degree rotation angle

Table 2: Chart to record setup parameters.

3.3 Important considerations and common problems

The purpose of this section is to emphasize some of the practical issues involved in obtaining stereoscopic views. These are to be viewed as guidelines which will enable someone with the right equipment to successfully obtain high quality stereoscopic images. These guidelines were formulated after considerable experimentation, and their observance removes common sources of human error.

Object placement: sequential viewing of the different views tends to be more pleasant when the object is placed exactly at the turntable center. This varies from object to object and is a subjective consideration (especially with objects that are not symmetrical).

Scanner position: special considerations should be taken when selecting the height, angle, and distance of the scanner. Different views of the object should be considered to determine these parameters. The best scanner position can be determined by finding the parameters where the least self-occlusion occurs (so that one part of the object does not hide another part) and most detail can be captured. Depending on the geometric complexity of the object, multiple angles of elevation may need to be used, but this adds to the capture time and increases storage requirements.

Containment of the object in the field of view: as the camera is translated from the left to right position, it is important to ensure that every view of the rotated object is contained in the field of view. We found it useful to spin the turntable while previewing the object. The object distance and scanner position can be varied to obtain appropriate containment of the object in the field of view.

Rotation angle: selection of the angle by which the turntable is rotated for each view is a subjective measure that depends on the object. It is important, however, to keep track of the angle being used and the direction of rotation.

Figure 4. IBM TDI-Pro 3000 scanner with an object in a Picture Magic Studio SL-36000.

Depth of field: it is preferable to use a smaller aperture to obtain a larger depth of field, so that the object remains focused at multiple orientations.

Alignment issues: it is crucial to make sure the object is not moved (relative to the turntable) during the process of rotation. If the turntable is covered by a material (such as paper), special care must be taken to ensure it does not slide during the process Fixing it to the table alleviates the problem. Similarly, other parameters, such as camera orientation and focus should be preserved.

Number of views: Typically 12 views of an object at 30 degree intervals are sufficient. However more complex objects may warrant additional views.

Lighting: We used four quartz halogen lamps (GE Quartzline, 250 W). Adequate lighting is important for accurate color reproduction and should be maintained constant for the left and right views and multiple object rotations.

3.4 Choice of background

In any image reproduction system, both the input requirements and output requirements must be examined in choosing the design parameters. *Page-flipped stereo* or field-sequential stereo (described in more detail in Section 4) is the preferred mode of display on a personal computer based CRT. In this mode, the composition of the input scene has an effect on the image quality of the final display. Specifically, one of the problems encountered in the display of page-flipped stereo is that of ghosting, where an after-image persists after being turned off [Lee97]. This is caused by the fact that the phosphors on a CRT have a finite decay time as the driving voltage is turned off. The ghosting effect is maximum across edges in the images that have the most contrast, i.e. possess large differences in luminance. Typically, the background chosen in object photography is white or black. However, these are inferior choices in the case of stereoscopic object photography, as they tend to increase the ghosting effect. Ideally, each background should be customized for a given object. However, it is easier to choose a fixed background. One method for reducing the ghosting effect at edges with a fixed background is to choose a neutral gray background, say of 128 on a scale of 0-255. This way, the maximum contrast, i.e. 128, is halved relative to a background of 0 or 255, where the contrast could be 255.

Since the ghosting problem does not occur in some other forms of display, such as autostereoscopic displays, or in stereoscopic prints, any background can be chosen in these cases without affecting the output image quality.

3.5 Apparatus used for stereoscopic object photography

The setup we described above has been automated to allow rapid capture of object sequences. We use a Kaidan computer-controlled turntable, the Meridian TM-400 [http://www.kaidan.com]. The turntable is placed in a Picture Magic lighting studio, which provides a uniform lighting environment. Quartz halogen lamps were used to provide illumination. We used the IBM TDI-Pro 3000 Scanner as the high-end digital camera mounted on a Bogen camera stand to allow the camera to be moved laterally. Figure 4 shows the setup.

4. ISSUES IN DISPLAYING STEREOSCOPIC VIEWS

4.1 Stereoscopic Display

The most common type of electronic stereoscopic viewing devices available are Liquid Crystal shutter glasses, such as those made by StereoGraphics Corporation [Halnon98] or NuVision Technologies [Qualman98]. These glasses work by alternating the left and right eye shutters in sync with a left and right display field. Thus, the shutters flip alternately between opaque and clear so that only the left eye sees the screen when the left image is displayed, and only the right eye sees the screen when the right image is displayed. This method is known as page-flipped stereo. Sawdai provides an overview of other low-cost PC stereoscopic display technologies such as polarized displays and autostereoscopic displays [Sawdai98].

We used the shuttered LCD glasses made by Nuvision Technologies. These are inexpensive and work with a variety of graphics cards. This was run on an IBM PC. With sufficiently high refresh rates, say in excess of 75 Hz, the flicker on the display is less noticeable.

4.2 Viewing conditions

The expected viewing environment is a CRT display attached to a PC. If the graphics card can support sufficiently high refresh rates, flicker should not be a problem. In any case, it is preferable to view the images in dim lighting or with monitor shades as this reduces the problem of flicker. This also helps to avoid color adaptation effects where the eye may adapt to different colored surrounds, affecting the perceived color of the object. In case accurate color reproduction is desired, it may be necessary to use standardized viewing conditions, such as D50 illumination and color-calibrate the display. These aspects are beyond the scope of this paper and further details may be found in [Hunt95]. New methods of display are evolving, such as stereographic prints on polarized media [http://www.slidefactory.com]. In this case, flicker is not an issue.

4.3 Preparation of images

The images were converted to a standard SMPTE color space [Hunt95]. Two sets of images were prepared for each object: one with a VGA resolution of 800x600 pixels and the other with the maximum resolution at capture time (typically close to the maximum resolution of the digital camera, which was 2Kx3K in our case). The VGA resolution images can be viewed as a sequence with a smaller delay between successive views. The higher resolution stereo pairs preserve detailed information such as writing or inscriptions on the object, and are suitable for scholarly study.

5. RESULTS

We successfully used our setup and methodology to obtain multiple orientation stereoscopic views of several objects. Some objects are from the IBM Corporate Archive, such as an early hard disk drive and one of the first card punching machines. We also scanned some sculptures from the Smithsonian Museum consisting of pottery dating from 3000 B.C.

Figures 5 and 6 show the resulting stereoscopic views. The object shown is an early diskdrive from the IBM Corporate Archives in Kingston, NY. In Figure 5 we show the sequence of twelve views of the rotated object. This represents the left view of the object. In Figure 6 we show one of the stereo pairs.
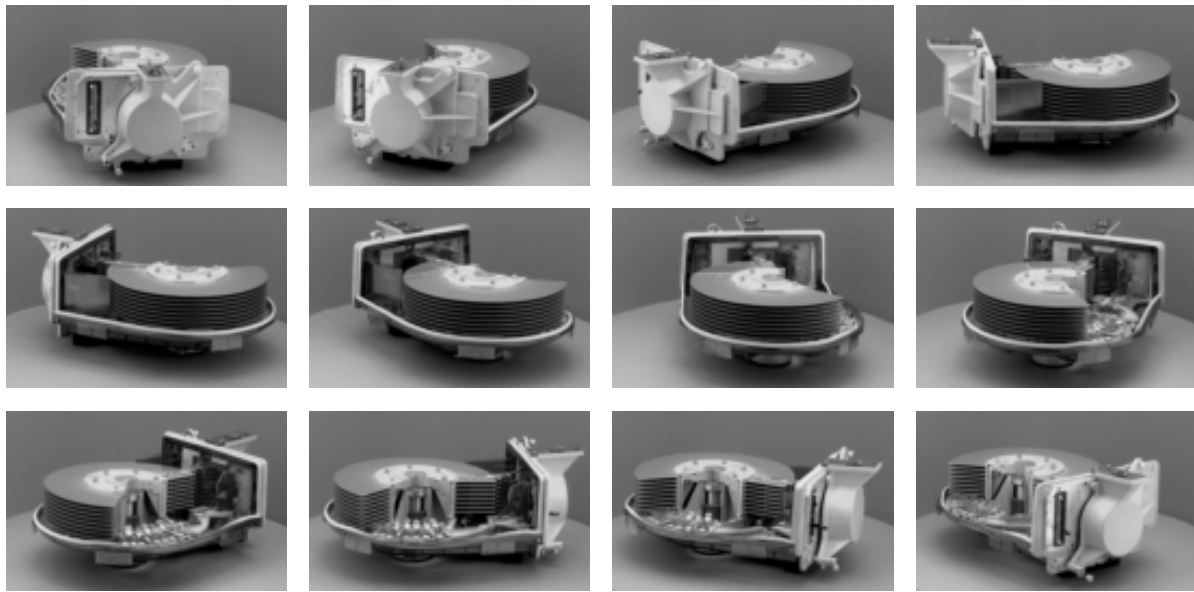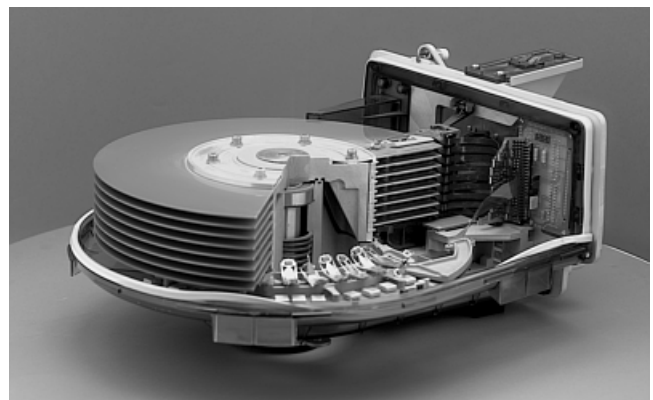
Figure 5. A sequence of 12 images obtained by rotating the turntable at 30 degree intervals. The left views are shown here.
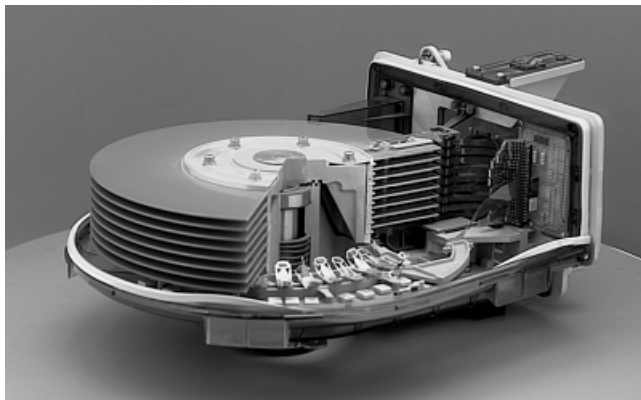
Right View   **Left View**

Figure 6. One of the twelve stereo pairs obtained from the sequence.

6. APPLICATIONS

Some of the projected application areas we see for stereoscopic object photography include cultural institutions, museums, universities, and merchandise catalogs. We are currently involved in a project with the Smithsonian, Washington DC where we are digitizing fossils and ancient pottery dating to 3000 BC. Since these are rare objects, access is limited even to scholars. Transportation of these objects is costly and potentially damaging. One way to increase access to these materials is through stereoscopic object photography. This provides stereoscopic views at multiple viewing angles, so that the object is experienced in its entirety. Such a technique could also be useful in teaching art classes, where sculptures could be viewed realistically by a large number of students. A few universities have expressed interest in this possibility. Stereoscopic object photography can also prove useful in merchandise catalog applications such as those involving jewelry or other expensive and difficult-to-transport objects. Eventually this could become a popular way to view object merchandise on the web.

7. CONCLUSIONS AND FUTURE WORK

In this paper we presented our stereoscopic imaging system and methodology for semi-automatically capturing multiple orientation stereo views of objects in a studio setting. We demonstrated the superiority of using a high resolution, high fidelity digital color camera for stereoscopic object photography and outlined the most important parameters in the capture process. This methodology should benefit those interested in replicating our results. With the decreasing cost of high-resolution digital cameras, stereoscopic object photography is becoming increasingly viable. We expect several applications of this technique to emerge in the future. An important application is the capture of art objects for scholarly study. Other applications include the capture of objects that are rare, difficult to transport, or expensive, such as museum sculptures or jewelry. The multiple views allow the user to view the object from a desired vantage point, and the stereoscopic capability allows the third dimension of depth to be perceived realistically. A high-end digital camera provides accurate color reproduction. Based on these advantages of using digital stereoscopic imaging to capture objects, we expect its usage to grow significantly as a media form in digital libraries. In the future, the approach we have described can be extended to stereoscopic panoramic photography.

We have not investigated issues related to the compression of 3D stereoscopic image sequences in this paper, but this has been done elsewhere [Siegel97]. Further efforts in this direction will lead to acceptable quality while reducing storage and transmission requirements.

The setup we used is partially automated, where the camera movement on the tripod is performed manually. Complete automation is preferable, though it comes at additional cost. Furthermore, parameters related to the object capture, such as distance to the object, scanner elevation angle, focus etc. can be automatically measured and stored in the image header.

It has been observed [Sawdai98] that stereoscopy has not yet penetrated the PC market to any appreciable extent. Part of the reason is the lack of standards for stereoscopic drivers. Another reason is the lack of compelling digital stereoscopic content for the PC environment. It is hoped that the techniques demonstrated in this paper will spur the growth of such content.

REFERENCES

[Giordano99] F. Giordano *et al*, "Evolution of a high-quality digital imaging system," Proceedings of the SPIE, Vol. 3650, **Sensors, Cameras and Applications for Digital Photography**, San Jose, 1999.

[Giorgianni98] E.J. Giorgianni and T.E. Madden, "Digital color management: encoding solutions," Addison-Wesley, New York, 1998.

[Halnon98] J. Halnon and D. Milici, "Solving the interface problem for Windows stereo applications," Proceedings of the SPIE, Vol. 3295, **Stereoscopic Displays and Applications IX**, pp. 12-21, 1998.

[Howard95] I.P. Howard and B.J. Rogers, "Binocular vision and stereopsis," Oxford University Press, 1995.

[Hunt95] R.W.G. Hunt, "The reproduction of color," Fountain Press, England, 1995.

[Lee97] B. Lee and M. Katafiaz, "Evaluation of a 3D autostereoscopic display for telerobotic operations," Proceedings of the SPIE, Vol. 3012, **Stereoscopic Displays and Virtual Reality Systems IV**, pp. 48-58, 1997.

[Mintzer96] F.C. Mintzer *et al,* "Toward on-line worldwide access to Vatican Library materials," *IBM Journal of Research and Development,* Vol 40 No. 2, March 1996.

[Output98] "Digital Cameras," pg. 22-29, **Digital Output***,* Vol. IV, No. 12, December 1998.

[Publish99] "Camera backs for the digital studio," **Publish**, pg. 32, January 1999.

[Qualman98] D. Qualman, "Development of stereoscopic software tools for Windows95 and WindowsNT computer applications," Proceedings of the SPIE, Vol. 3295, **Stereoscopic Displays and Applications IX**, pp. 35-44, 1998.

[Rao98] A.R. Rao and F.C. Mintzer, "Color calibration of a colorimetric scanner using non-linear least squares," IS&T PICS98 Conference, Portland, Oregon, pp. 118-120, May 1998.

[Sawdai98] D. Sawdai, G. Hamlin and D. Swift, "Software issues for PC-based stereoscopic displays: how to make PC users see stereo", Proceedings of the SPIE, Vol. 3295, **Stereoscopic Displays and Applications IX**, pp. 23-34, 1998.

[Shaw98] Rodney Shaw, "Quantum efficiency considerations in the comparison of analog and digital photography," IS&T PICS'98 Conference, Portland, Oregon, pp. 165-68, May 1998.

[Siegel97] M. Siegel et al, "Compression and interpolation of 3D-stereoscopic and multi-view video," Proceedings of the SPIE, Vol. 3012, **Stereoscopic Displays and Virtual Reality Systems IV**, pp. 227-238, 1997.

[Woods93] A.J. Woods, T.M. Docherty, R. Koch, "Image distortions in stereoscopic video systems," Proceedings of the SPIE, Vol. 1915, **Stereoscopic Displays and Applications IV**, pp. 227-238, 1993.