

# Scanning and Processing 3D Objects for Web Display

Mohamed Farouk ‡ Ibrahim El-Rifai ‡ Shady El-Tayar ‡ Hisham El-Shishiny †  
Mohamed Hosny † Mohamed El-Rayes † Jose Gomes † Frank Giordano †  
Holly Rushmeier † Fausto Bernardini † Karen Magerlein †

†IBM Corporation  
P.O. Box 704, Yorktown Heights, NY 10598  
fausto@watson.ibm.com

‡Center for Documentation of Cultural & Natural Heritage  
Zamalek, Cairo, Egypt  
cultnat@mcit.gov.eg

## Abstract

*We present a case study of scanning 3D objects for the purposes of education and public information. We begin by describing the original design of a 3D scanning system now in use in Cairo's Egyptian Museum. The system captures both the geometry and surface color and detail of museum artifacts. We report on the experience using the system in the museum setting, and how practical problems with the system were addressed. We present samples of how the processed 3D data will be used on a web site designed to communicate Egyptian culture.*

## 1. Introduction

One role that 3D scanning has in the area of cultural heritage is to produce content that can be used for the education of students and for informing the general public. This area was pioneered by the National Research Council of Canada in such projects as the on-line presentation of Inuit sculpture in a virtual museum [1]. We have recently installed and begun using a 3D scanning system for capturing artifacts in the Egyptian Museum in Cairo for a similar project. Our goal is to further the use of 3D objects in web communication by using the objects in novel virtual restorations, and integrating the visual presentations with extended textual description by subject experts.

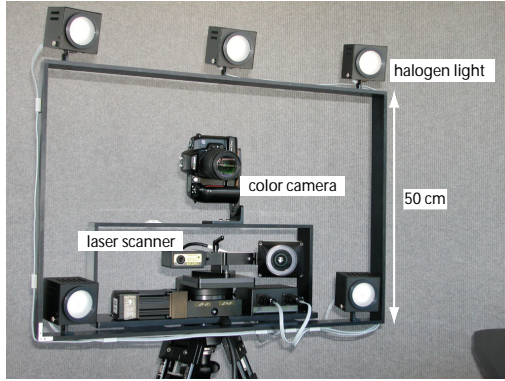
This system is part of a joint project between the Egyptian National Center for Documentation of Cultural and Natural Heritage (CultNat) and IBM to develop a web site to communicate Egyptian culture to a world-wide audience. Many important artifacts from Egyptian history can be effectively displayed using photographs, or series of photographs combined to form object movies or panoramas.

We plan to use 3D scanned objects to present artifacts when they are changed in a manner that cannot be physically photographed. Changes requiring a digital model include placement in virtual environments and virtual restoration. While many image-based techniques have been developed for interactive viewing of objects, making changes in the object's shape or surface finish requires a complete geometric model and map of surface attributes.

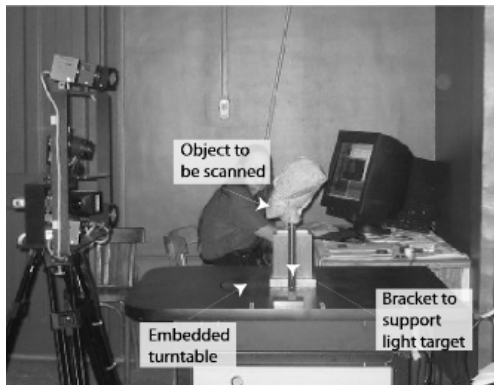
Working with experts in Egyptology, we selected a set of objects to scan. Our selection criteria include: objects which need restoration but cannot be restored physically; objects with an interesting history; objects of research interest in establishing the identity of the person being represented; objects and tools for which a simulation would be helpful to present their function.

Our goal is to display the scanned museum objects with high visual quality. Early attempts at web presentation of 3D objects used relatively crude representations with the expectation that users would be satisfied with interactivity in place of visual quality. This type of trade-off is effective in applications such as computer games. However, studies of cultural web sites show that users are interested in viewing material more passively, and not having a high level of interaction [2], an observation referred to as "less clicking, more watching." The use of 3D models on the Egyptian culture web site will be to generate high quality images and animations illustrating the history of object in ways only possible with a 3D digital model, rather than attempting to give the user full 3D game-style interactivity.

Acquiring high quality data is necessary, but not sufficient. The data must be processed into a form that can be edited to simulate virtual restorations or imported into virtual environments. We describe our processing pipeline. The pipeline is still far from fully automatic, and we describe the human intervention required to obtain models.



**Figure 1. The scanning system.**



**Figure 2. The system in the museum.**

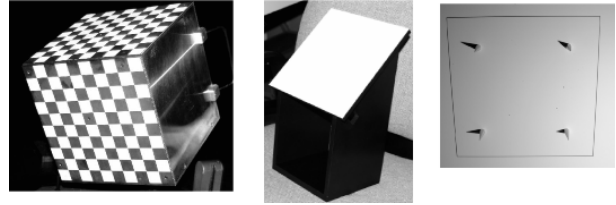
We illustrate difficulties in editing the processed models, and present some preliminary results.

## 2. The Capture System

In this section we give an overview of the hardware and calibration systems, with an emphasis on the user interfaces to the system. Additional detail on the system design and implementation can be found in a companion paper [3].

### 2.1 Hardware

We used a capture approach employed by researchers in previous cultural heritage projects to [4, 5]. The basic concept is to couple a device for capturing a geometric range image with a photometric stereo capture system. The system, shown in Fig. 1, consists of a ShapeGrabber laser range scanner, a Fuji FinePix S1 Pro digital color camera, and five halogen light sources mounted on a 100 cm by 50 cm aluminum rack. As shown in Fig. 2, the system is used with a scanning table that contains an embedded Kaidan turntable. The camera has an adjustable lens so that the user can adjust



**Figure 3. Calibration targets for (left to right) the camera, turntable and lights.**

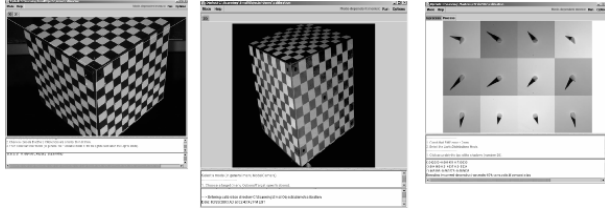
the zoom to accommodate scanning objects of various sizes. The camera can also be moved and tilted with respect to the range scanner to adjust for different types of objects. All of the system elements are under the control of the scanning workstation, an IBM Intellistation MPro with a Pentium IV and NVIDIA Quadro2 graphics.

The ShapeGrabber comes with a target and software procedure to maintain the calibration between the scanning head and the panning stage that moves it across the object. The 3D point values the ShapeGrabber produces establish the base coordinate system. All of the other system components, the digital camera, lights and turntable, must be calibrated in terms of this base coordinate system. To perform these additional calibrations we designed and constructed a series of calibration targets, Fig. 3. We developed a software system we call “Djehuti”, Fig. 4 that guides the user through acquiring and processing the necessary calibration data.

### 2.2. Calibration

We calibrate the digital camera with respect to the laser scanner coordinate system by taking a digital image and a 3D scan of a cube calibration object, shown in the left of Fig. 3. We have cube objects of different sizes, so that the size of the calibration volume can be tailored to the size of the object to be scanned. We use a precise checkerboard pattern on each of three adjacent sides of the cube to establish the three dimensional positions of the two dimensional features detected in the image from the camera.

For the calibration cube placed in the standard position, automated methods can be used to find the checkerboard corners. For operation in the museum, however, an automated method that might on some occasion fail for unforeseen reasons is not acceptable. The Djehuti interface presents the digital camera image to the operator, as shown in on the left of Fig. 4, so the operator can assist the calibration by clicking key points, and can inspect the calibration result to verify success. The camera coordinate system needs to be expressed in terms of the coordinate system of the laser scanner. A laser scan is taken of the cube, with the



**Figure 4. “Djehuti” graphical interface for calibration.**

laser intensity values presented to the user (center of Fig. 4) to verify a successful scan. Using operator input, the three planes of the front of the cube are computed from the laser scan. Knowing the coordinates of these planes in the laser system allows the transformation of the camera calibration results into the laser coordinate system.

The lights must be calibrated both for position and for directional distribution across the scan volume. The right image of Fig. 3 shows the image captured for calibrating the position of one of the lights. The position of the board is found by identifying the corners of the rectangular pattern surrounding the four conical pins. Djehuti presents the shadows of the pins detected to the operator as a mosaic, as shown on the right of Fig. 4. Given the positions of the sharp tips of each of the four shadows cast by the conical pins, the position of each light is computed by finding the least squares solution to the intersection of the four rays defined by shadow and cone tips. Note the variation of the light distribution in the rightmost image of Fig. 3, which falls off from very bright for positions away from the lower right hand corner. The image of a planar white board is obtained for each source to account for this non-uniformity in the lighting.

The turntable axis is calibrated by placing the pin registered 45 degree plane onto the turntable. The target is scanned at three different turntable positions. The axis is then found by finding the intersection of the three planes and the center of a circle swept out by the normal to the planes.

The Djehuti interface is designed to guide the user through all of the steps, with (referring to Fig. 4) the top pane of each window displaying an image of the calibration data, the center pane giving tips on what to do next, and the lower pane reporting the progress and accuracy of the calibration. The calibration procedures still require considerable skill on the part of the operator to properly position calibration targets, acquire data in the correct order, and interpret the calibration reports of results. One way procedures could be simplified would be to reduce the flexibility of the system, i.e. to fix the camera position and focal length and so

to eliminate the need for frequent camera calibration. In our application where the objects range in size from a few centimeters in height to up to a half meter, we decided this would restrict the level of detail and efficiency of scanning. Another way to simplify procedures would be to rely completely on automated detection of features in the images acquired for calibration. In using calibration systems that are completely automated, we found great frustration when the automated system failed and there was no opportunity to manually assist the system. We decided that it would be more time efficient to have the user “in the loop” assisting the detection process.

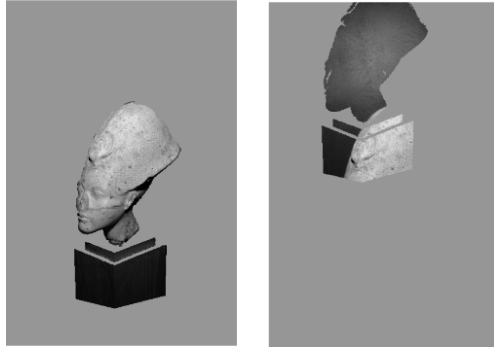
### 3. In-Museum Scanning

In-museum scanning operations have two phases – weekly calibrations and day to day scans. The ShapeGrabber is calibrated weekly using the calibration target and software provided by the manufacturer. After the laser scanner is recalibrated, the light sources are recalibrated in the new scanner coordinates.

Figure 2 shows the scanning system as it is installed in the museum. Scanning can only be performed during the hours the museum is open to the public. Each object is brought to the scanning area by a curator (who is the only person who can handle the object) and a guard. Because the scanning requires the time of the museum staff, and because the object is not available for public display during the scan, there is a premium on keeping the scanning time as short as possible.

One method to minimize scanning time is to plan the scanning strategy to the greatest extent possible before the object is brought to the scanning room. Each object to be scanned is quite different and so significant time has been required prior to scanning each object to plan for handling, scanning and processing. We measure the object and draw plans for how we will position the object for scanning. To understand how long the scanning will take we consider the mix between dark and shiny materials, how rounded the object is, whether there are hidden corners, whether the object is semi-transparent, or whether the object needs special planning because it is very large or small, and the size of the details that need to be captured. We prepare the scanning device to be as close as possible to the right position before the object is brought for scanning. The scanning device is calibrated with the correct sized calibration cube and with the correct zoom level anticipated for the object. Only then is the object placed on the turntable.

Prior to the actual scan, preliminary range scans and color images are obtained at a variety of exposures. The operator then selects the exposures to use for the object. Ideally, images at many different exposures would be used automatically for each object. However, the time to ac-

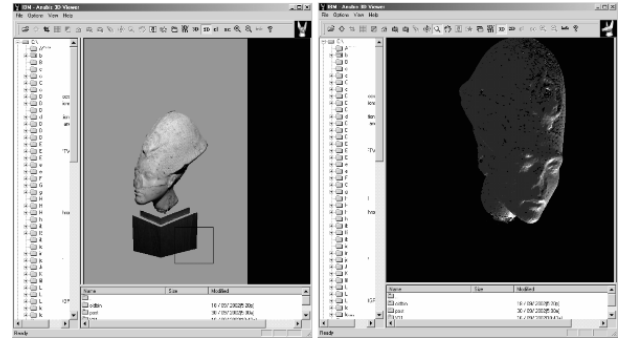


**Figure 5. The images used for image-based scan cleaning and checking the camera calibration.**

quire and transmit the data to the computer – a little over two minutes for each range scan and set of five images – makes this unacceptable in the museum environment. Alternatively an exposure would be selected automatically, but experiments with using the automatic camera exposure feature of the digital camera gave unsatisfactory results. The selection of the correct exposures therefore currently relies on the judgement and experience of the scanner operator.

After the object is placed on the turntable and exposure levels have been set, the scanner is run from the workstation. A graphical user interface (GUI) guides the user through the scanning process. We have found that all of the in-museum operations need to be integrated into this single interface for efficient scanning. To begin, the operator identifies the appropriate directories for reading the relevant calibration files and for writing the scan data. The first actual data acquisition is an automated sequence of eight scans, one for each 45 degree increment of the turntable. A set of eight scans takes approximately 20 minutes.

After the first eight scans, the first estimate of the digital model is needed to plan additional views that are needed to completely cover the object. The processing of the initial data requires user input to “clean” the scans to eliminate the geometry of support fixtures or cushions from the scan of the actual object. An image-based system is used to perform this cleaning. In the image-based system, the captured geometry for each scan is projected into a color image of the object from the same view using the camera calibration data. The result of this is shown on the left of Fig. 5. Besides being used for the cleaning, this image is useful as an additional check of the camera calibration. If the wrong calibration file is used, or the file is corrupted, the geometry will not be projected into the correct portion of the image, as shown on the right of Fig. 5. Since the data will be useless if the calibrations are not correct, it is critical to detect any

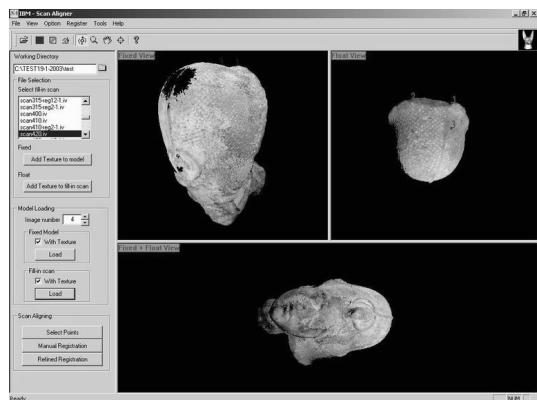


**Figure 6. The operator cleans the scans by editing an image of the captured points. After cleaning the user is presented with a crude of the model.**

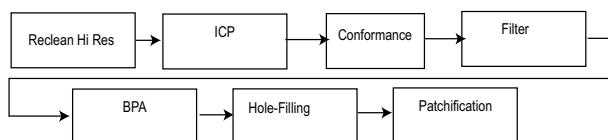
problems at this stage, suspend scanning, and recalibrate. Given the correct image, the scanning GUI presents the image for editing as shown on the left of Fig. 6. While any image editor could be used to paint the unwanted geometry green, we found that the scanning process is smoother using a simple image editor embedded into the scanning GUI.

After the images are processed to paint out unwanted support geometry, a simplified version of each scan is generated using every fourth point in the horizontal and vertical directions, and omitting any points that project into the background area. A crude version of the object is then formed by applying rigid transformations computed from the turntable calibration, and integrating the eight scans into a single mesh. The operation to form the low resolution model takes a couple of minutes. The operator can view this simplified model from within the GUI, as shown on the right of Fig. 6. Any 3D viewer could be used to view the geometry, but again we found that a simple viewer embedded in the GUI streamlines the scanning process. The 3D view allows the operator to plan the next scanning view needed. The GUI representation of the next view can be shown to the curator who must physically reposition the object.

Scan cleaning to build initial models cannot be avoided, since safe support of the objects frequently requires including additional material in the scanner field of view. Preparing support materials, such as foam, to be all black is impractical, and also ineffective when very dark objects are scanned. Image-based cleaning of the scan has the advantage that time is not wasted during scanning operations loading and manipulating each scan in a 3D editor. Furthermore, the cleaning is recorded in the image. The same image can be used to form a high resolution scan in the post-processing phase without the operator having to repeat work. This also means that the operator who collected the



**Figure 7. A graphical interface for the “fill-in” scan alignment.**



**Figure 8. Pipeline of geometric post-processing.**

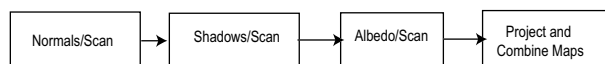
data does not have to be involved in the post-processing phase to identify which portion of geometry was a part of the object, and which is extraneous support.

Several additional scans are generally necessary after the initial eight turntable scans. In our original system, the additional “fill-in” scans were positioned relative to the crude initial model by using a third party software package to import the current model and the new scan into a 3D viewer and select three pairs of matching points to position the new scan. This proved to be an unsatisfactory disruption to the work flow, and was the step most prone to failure. We have developed our own graphical interface for alignment, shown in Fig. 7, so that this step can be smoothly integrated into the museum scanning interface.

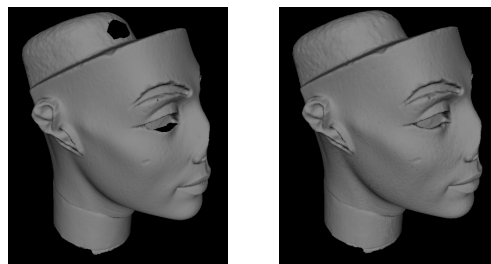
#### 4. Postprocessing

Postprocessing is a longer process that we perform after leaving the museum. There are two major phases in the postprocessing to obtain a model from the scanned data: geometric processing (Fig. 8) and texture processing (Fig. 9).

A GUI is provided to guide the operator through the geometric post-processing to build the high resolution model. The processing begins with forming high resolution scans



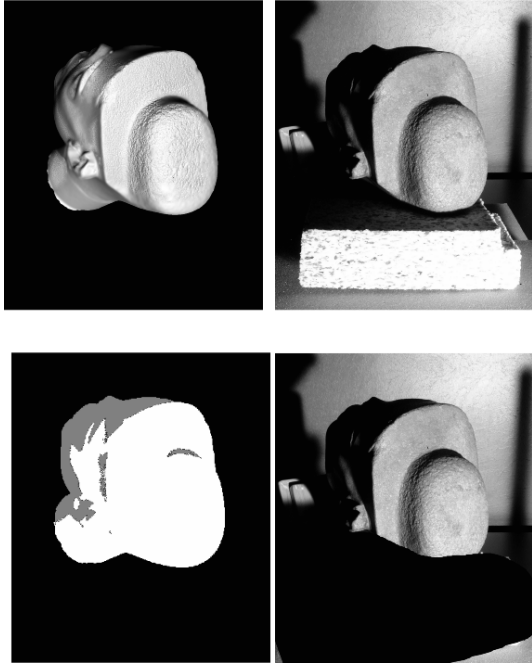
**Figure 9. Pipeline of texture post-processing.**



**Figure 10. Automated hole-filling.**

from the green-background cleaned images. The cleaning on these images can be refined if necessary using any image editor. The iterative closet point algorithm (ICP) [6] is then run to refine the initial transformations computed in the museum. The conformance process iteratively deforms the scans relative to one another to remove the noise resulting from line-of-sight errors in the scans. A filter is used to reduce the high number of points resulting from multiple overlapping scans to a point that is consistent with the scanner sampling resolution (usually 1 mm sampling distance for the scanner placed the “standard” distance from the scanned object.) The ball pivoting algorithm (BPA) [7] is used to integrate all the scans into a full mesh. Because presenting an object with data holes is unacceptable on the web site, automated hole filling as illustrated in Fig. 10 is included. The GUI is useful in this step, since in some cases the standard parameters for the hole-filling process need to be adjusted. The model can be viewed before and after the hole-filling phase so that the user can validate that the holes have been filled in a manner consistent with the captured images of the object. Finally the model is patchified, i.e. partitioned into non-overlapping height fields, to prepare the model for texture mapping. The full resolution geometry pipeline takes on the order of one to two hours to complete.

The geometry is then used in computing the surface details and unshaded color as diagrammed in Fig. 9. For each scan, detailed normals and an unshaded image (albedo map) are computed from the five color images, along with an image of weights that encode the confidence in the data at each pixel. These calculations may require user intervention as shown in Fig. 11. The geometry for which texture is being calculated is shown in the upper left of Fig. 11, and one of the lit images is shown in the upper right. The geometry



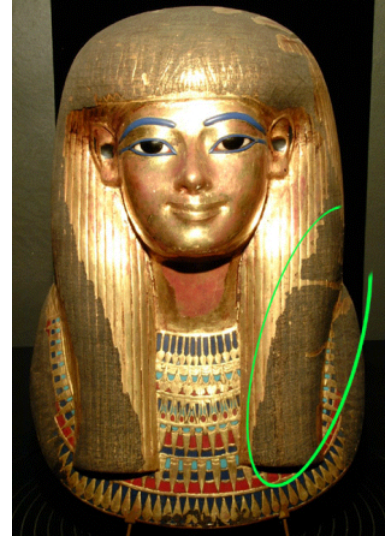
**Figure 11. Upper left: geometric model; Upper right: image of lit object; Lower left: predicted lit (white) and shadow(grey) areas; Lower right: lit image edited before texture processing.**

of the object is used to compute the sections of the object that are lit from this viewpoint and for this light source, as shown in the lower left. However, to obtain the image in the upper right, the object had to rest on a foam support. This support obscures the view of part of the object, and casts a shadow on the object that cannot be predicted. The user must manually black out the foam and the shadow it casts so that this part of the image will not be incorrectly used as input for computing the surface albedo.

The normal and albedo maps are projected on to the partitioned surface and combined to form seamless surface textures using the methods described in [4]. Once the lit images are adjusted for the types of problems illustrated in Fig. 11 the calculation of the detailed maps and reprojecting them on the model takes on the order of a couple hours on a Pentium IV processor. At the end of the processing, there may still be holes in the texture. We plan to extend the processing with texture hole-filling.

## 5. Results

The scanning system is currently in use in the museum by the CultNat scanning team. Several objects have been



**Figure 12. Photograph of the mask of Thuya. Circled area to be virtually restored.**

scanned, and early studies in virtual restoration have been performed. Here we present some initial results and work in progress.

### 5.1 The Mask of Thuya

Figure 12 shows a picture of the funerary mask of Thuya, a queen of the 18th Dynasty. The mask is 40 cm in height and is composed of cartonnage, gold leaf, glass paste, and alabaster. The objective in scanning this object is surface restoration. In ancient times the mask was covered with a black cloth which over time adhered to the surface of the mask. Efforts to remove the black cloth succeeded in many areas but parts could not be removed. Additionally, parts of the gold leaf covering are missing, and other parts need polishing. The gold leaf is fragile and too sensitive for more hand restoration. Using the laser scanning and virtual restoration is essential in this case.

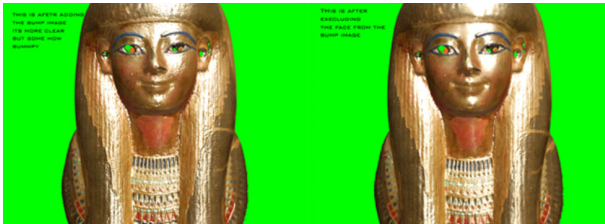
The major challenge in data acquisition for this object was setting the parameters of the laser scanner to acquire the shape. It was hard to set the laser intensity to this object as it has two extremes., the shiny material of gold and the dark material of the black cloth. We had to set the laser intensity to an intermediate level in order to avoid losing data from either area. We had to do many trials to find the correct value.

Geometric processing presented two challenges – filling the holes where there are black areas, and smoothing the surface. This object is a case where fine tuning of the hole-filling parameters is necessary. Smoothing is often needed





**Figure 13. Thuya, with recovered Lambertian reflectance (left) and with estimated specular reflectance (right).**



**Figure 14. Thuya, with detailed bump maps (left) and bumps excluded from face (right.)**

in shiny areas because of the noise in the acquired data, but uniform smoothing of the object is not desirable.

Our texture processing pipeline currently extracts colored albedo only, as shown in the image of the left of Fig. 13. While the shininess of gold is not captured, enough data is present to map out the different materials on the mask's surface. To give the illusion of gold, the image on the right of Fig. 13 shows the model with a value of specularity to match the appearance of the original object added manually. Clearly a lot of detailed further editing is needed to exclude this shininess from the areas of the mask that are not gold. The mask has a lot of fine detail that is not visible in the model. In Fig. 14 we show experiments with adding additional detail.

To restore the object, much detailed editing needs to be done on the surface description. This editing can be performed using the model "flattened" into a single vml file and the textures packed into a single image. In this form, the model can be imported into a 3D paint package and painted in a "projection paint" mode. Even with a 3D paint package, editing surface detail is a tedious process. Unfortunately, virtual restoration is not a "one-shot" process – the graphics specialist who can perform the editing does not have the subject knowledge or vision that the Egyptologist has of what the correct alteration is. Many versions of the model will be needed as ideas for the restoration are pre-



**Figure 15. Thuya, after initial virtual restoration. Image used to start discussion of alterations with experts.**

sented and discussed. To assist with this iterative process quick methods of viewing possible alterations are needed. Figure 15 shows an example of a quick initial restoration to be used for discussion with subject experts. An image of a completely gold version of the geometric model was generated. This was then composited with an image of the scanned model, using only the colored necklace and eyes.

## 5.2 The Head of a Queen

Figure 16 shows the results of scanning a sculpture that is the head of a queen, believed to be of Nefertiti. This model is to be used for comparison and for virtual restoration. Figure 17 shows a scan of a sculpture known to be of Nefertiti. These two digital models will be used to explore whether similarities of shape and proportion are observed that aren't possible to see using just the physical objects.

The head of the queen is damaged, and unlike Thuya, geometric as well as texture modifications are necessary to make restoration. Again a quick method of illustrating possibilities is needed to begin a dialog regarding the correct modifications. In this case models of possible ear and nose replacements are positioned on the scanned model, as shown on the upper left of Fig. 18. These are then used to generate images that serve as a starting point for discussion with the Egyptologists on the most appropriate restoration.

Any virtual restoration is speculative. Images of the objects as they are today will always be shown along



**Figure 16. Four views of a digital model of a head of a queen, believed to be of Nefertiti.**



**Figure 17. Four views of a digital model of a head of a queen, known to be of Nefertiti.**

with the altered models. The results will appear on [www.eternegypt.org](http://www.eternegypt.org), to be launched in late 2003.

## References

- [1] G. Godin, J.-A. Beraldin, J. Taylor, L. Cournoyer, M. Rioux, S. El-Hakim, R. Baribeau, F. Blais, P. Boulanger, J. Domey, and M. Picard, "Active optical 3D imaging for heritage applications," *IEEE Computer Graphics and Applications*, vol. 22, no. 5, pp. 24–35, 2002.
- [2] C.-M. Karat, C. Pinhanez, J. Karat, R. Arora, and J. Vergo, "Less clicking, more watching: Results of the iterative design and evaluation of entertaining web experiences," in *Proc. of Interact'2001, Tokyo*, July 2001.
- [3] J. Gomes, F. P. Giordano, H. El-Shishiny, H. Rushmeier, K. Magerlein, and F. Bernardini, "Design and use of an in-museum system for artifact capture," in *IEEE Workshop on Applications of Computer Vision in Archaeology*, June 2003.
- [4] F. Bernardini, H. Rushmeier, I. M. Martin, J. Mittleman, and G. Taubin, "Building a digital model of Michelangelo's Florentine Pietà," *IEEE Computer Graphics and Applications*, vol. 22, no. 1, pp. 59–67, Jan /Feb 2002.
- [5] C. Rocchini, P. Cignoni, C. Montani, and R. Scopigno, "Acquiring, stitching and blending appearance attributes in 3D models," *The Visual Computer*, vol. 18, no. 3, pp. 186–204, 2002.
- [6] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
- [7] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin, "The ball-pivoting algorithm for surface reconstruction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 349–359, 1999.



**Figure 18. A possible restoration.**