Brett A. Nickerson,[1] *M.Sc.; Patrick A. Fitzhorn,*[1] *Ph.D.; Stephen K. Koch,*[1] *M.Sc., and Michael Charney,*[2] *Ph.D., D.A.B.F.A.*

A Methodology for Near-Optimal Computational Superimposition of Two-Dimensional Digital Facial Photographs and Three-Dimensional Cranial Surface Meshes

REFERENCE: Nickerson, B. A., Fitzhorn, P. A., Koch, S. K., and Charney, M., **"A Methodology for Near-Optimal, Computational Superimposition of Two-Dimensional Digital Facial Photographs and Three-Dimensional Cranial Surface Meshes,"** *Journal of Forensic Sciences*, JFSCA, Vol. 36, No. 2, March 1991, pp. 480–500.

ABSTRACT: The authors present a methodology for human identification based on digital superimposition techniques. This methodology computes a fast, near optimal fit between a three-dimensional skull surface mesh and a two-dimensional digitized facial photograph. Since this is done digitally, (1) the photograph can be enhanced to reduce or eliminate motion blur, overexposure or underexposure, and out-of-focus distortions; (2) previous problems with skull/photograph scaling and alignment are minimized or eliminated; and (3) the photograph and skull can be numerically correlated. Two of several test cases produced from an implementation of this methodology are also presented.

KEYWORDS: physical anthropology, human identification, superimposition, photography, photographic superimposition, digitizing, computer graphics, image processing, nonlinear optimization

Historical Background

The successful comparison of human skeletal remains with artistic or photographic replicas has been achieved many times, beginning scientifically with Welcker. In 1867, he compared known skull measurements of Dante (whose tomb was renovated in 1865 on the 600th anniversary of his birth) with a reputed death mask. Welcker concluded that the mask did indeed belong to Dante's skull [*1,2*]. He continued this work with the skull of Kant (1724–1804) and a death mask, and remarked that the comparison was so striking that the match could have been found by picking from hundreds of different skulls and masks [*3,4*]. In the comparison of a supposed skull cast of Schiller (1564–1616) and two death masks in 1883, Welcker doubted the authenticity of both the cast and the masks [*3,5*]. In 1884, Welcker and Schaaffhausen combined their efforts in skull and

Received for publication 22 Dec. 1989; revised manuscript received 30 May 1990; accepted for publication 5 June 1990.

[1]Graduate research assistant, associate professor, and graduate research assistant, respectively, Department of Mechanical Engineering, Colorado State University, Fort Collins, CO.

[2]Director, Forensic Science Laboratory, Colorado State University, Fort Collins, CO.

portrait comparison to measure the accuracy of three profile portraits of Raphael (1483–1520) in relation to his known cast [6,7]. Work continued in 1895 with His, who took facial-tissue thickness measurements from 24 male and 4 female cadavers. These measurements were then used to construct a bust of Johann Sebastian Bach, producing an extremely lifelike resemblance [8]. In 1898, Kollmann and Büchly added facial-tissue thickness measurements for 45 male and 8 female cadavers [9]. These measurements were separated into four distinctive body condition categories to increase the accuracy of these facial-tissue depths further. Skull and portrait comparison continued to develop, and in 1905, the skeletal contents of a leaden coffin, discovered after being misplaced for 113 years, were likened to the great American admiral John Paul Jones [10]. These early cranial comparisons were possible because there existed sculpted busts and paintings of the deceased. In 1918, Lander compared the skull of an individual of known age, race, and sex with a photograph of the individual and concluded, "The contrast between the narrow appearance of the latter [the skull] and the broad fleshly face is striking. It seems improbable that anyone examining the skull would postulate a type of face similar to that seen in the photograph" [11]. The first documented use of photographic superimposition appears to have been by Pearson, Morant, and Derry, who in 1934 demonstrated that tracings taken from the photographs of the head and macerated skull of an executed Egyptian criminal could be superimposed to produce a good fit [12].

Despite some reluctance to use superimposition for positive identification, most researchers agree that this technique provides good corroborative evidence to reinforce a probable identification and, in some cases, good evidence for exclusion of a suggested identification [13]. Improved techniques in photography, scale-matching, and orientation-matching were most prevalent up to the 1980s [14–17], at which time there was a shifted emphasis towards quick comparison using video cameras, split screens, and adjustable skull mounts [18–21]. Currently, when given the task of identifying unknown skeletal remains, the forensic scientist has three superimposition methods to choose from: (1) photographs of the skull and face can be photographically superimposed; (2) a transparency of the facial photograph can be superimposed on the cranium; or (3) a video camera method can be used. However, with all three methods there is still difficulty in finding the correct angulation and scaling for a good fit. Even when a reasonable, qualitative fit is found, it is difficult to guarantee that the angulation and scale truly represent that of the facial image in the photograph. These problems force examination of the methods for, and attributes of, the superimposition methods currently in use.

Superimposition Methods

Static Photographic Superimposition Method

The first and classical superimposition method is static photographic superimposition. In this methodology, the term "photographic" generally applies to any type of pre or antemortem record. This could be an actual photograph, an X-ray, a portrait, or a painting. As was first done in a criminal case in 1937 [15], the investigator compares the antemortem photograph with a photograph taken of the skull. The two photographs are then photographically superimposed, typically using a negative of the skull and a positive of the face. The resulting photograph can then be studied to determine the validity of the skull/face fit. Obviously, bony landmarks, (orbits, nasal cavity, and dentition) and general size and shape characteristics of the skull are indicators in the evaluation of the fit.

Several serious problems are associated with this method:

- *angular positioning:* the skull and facial regions in the photographs must have a reasonably equivalent spatial positioning through six degrees of freedom;

- *scaling factors:* similarly, the areas of interest must be differentially scaled so that they appear equivalently close to the observer;
- *camera lens incongruities:* differing camera lens focal lengths can easily distort the contours of regions of high curvature; and
- *poor initial photography:* poor camera or film characteristics can easily lead to noisy or grainy exposures, overexposure or underexposure, blurring, and other photographic enigmas.

The first two of the three superimposition methods, which typically require manual dexterity, are an attempt to generate a set of rigid body transformations or functions which map a perspective projection of the skull onto a perspective projection of the facial region. The most recent techniques of photographic superimposition employ a more dynamic method [*22*]. The images of the face and skull are projected into the camera using a beam splitter and mirror. The skull can be dynamically angled with the help of a suitable tripod stand to approximate the best fit. If known objects in the facial photograph are present, then predetermined scaling factors can still be utilized with this method. Again, the final fit is approximated by the forensic science specialist.

Dynamic Video Camera Superimposition Method

A similar approach is video camera superimposition. The apparatus is used to superimpose a facial photograph (negative) in front of the skull. A "best-fit" approach is taken by physically angulating the skull. Problems with facial photograph enlargements are similar to those for the other methods of photographic superimposition. Another video camera method has also been used, in which two video images are mixed into one composite image. The mixing unit provides convenient fade-in and fade-out of either image as well as composite slicing of each image. This method does decrease the number of associated problems in superimposition. First, the number of manual manipulations and the length of time required per iteration are cut dramatically. Second, since the skull is used rather than a photograph, one can judiciously choose an appropriate lens, and there is obviously no problem with poor skull photography. However, it has been argued that new sources of optic and electronic error are added when using this method [*22*].

An interesting combination of video camera superimposition and facial reconstruction was used in the case of "Mr. X" [*18*]. In this case, the facial features of the left side of the skull were reconstructed using the technique described by Krogman [*23*], and three video cameras were used to obtain a three-way superimposition. One camera was directed at an antemortem photograph, a second camera was focused on an X-ray of the skull, and the remaining camera was focused on the partly reconstructed skull. A special effects generator provided easy, flexible comparison by providing fade-in, fade-out, and sweep (vertically or horizontally) mechanisms in which the different images could be progressively superimposed upon each other.

Problems with Calculation of the Proportions of the Antemortem Photograph—In the identification process, establishing the correct enlargement of the photograph of the skull is generally considered to be critical. This enlargement factor has been based on linear measurements of items within the antemortem photograph, such as fabric [*16*] and other objects of known geometry [*15,17*]. One of the earliest and most widely recognized applications of this method was made in 1935 by Glaister and Brash in their study of the Ruxton Case [*15*]. For the first time, careful attempts to scale the photographic records to life size were made using the known dimensions of objects present in snapshots, such as a tiara headdress, the outline of the neckline of a frock, and the heights of a door gate and wall. In one instance, the known focal length of the camera was used to establish a magnification factor. In other cases, the linear pattern of a tie [*16*], the pattern on the

border of a sari [*17*], and the diameter of a button on a sweater [*24*] were used to find a correct photographic enlargement to yield successful superimpositions and probable identifications. Measurements of a wooden chair present in a photographic portrait also yielded a magnification factor from which an accurate enlargement was made [*17*].

In the absence of objects of known size in antemortem photographs, several researchers have combined the use of anatomical landmarks and anthropometric measurements of the facial skeleton with the existing data for soft tissue thickness to estimate a magnification factor [*13,25,26*]. Sekharan suggested a standard interpupillary distance of 6 cm as the measurement from which life-size enlargements of antemortem photographs could be made [*17*]. McKenna et al. tested this claim by measuring interpupillary distance on 75 Chinese subjects between the ages of 19 and 22 years and concluded that there was too much variance in this measure to use it as an average figure for the basis of photographic enlargements [*14*]. This group investigated a series of possible cranial reference points for calibrating measurements but finally concluded that none of their trials using nondental landmarks and measurements proved satisfactory. In their research, McKenna et al. state that their most satisfactory method of enlargement was to use the dimensions of the anterior teeth in portraits. With a magnification factor established from the anterior teeth, life-size transparencies of the photograph are superimposed on the skull's dental landmarks, after which the skeletal and facial features are compared. This technique requires that the anterior teeth be clearly visible in the antemortem photograph. More recently, videotaped recordings of mounted skulls and antemortem photographs have been made using a best-fit approach in which enlargement of the photograph is determined by visual assessment rather than through the use of a positive magnification factor [*20*]. In discussing this technique for video superimposition, Bastiaan et al. state:

> Inherent in all superimposition procedures are assumptions and estimations that have to be made of the bony anatomical features on the antemortem photograph. The average thicknesses of tissue over bone have been recorded, and therefore, a calculation can be made of the soft tissue outlines on the skull [*20*].

In this statement, Bastiaan et al. suggest that statistics drawn from flesh thickness data can be used to interpret a best-fit superimposition. They also state, however, that establishing the correct enlargement using dentition measurements is the more quantitatively accurate superimposition technique when dentition is present. In their concluding remarks concerning the use of anterior dental landmarks to scale antemortem photographs, McKenna et al. suggest that extensive records of smiling people could be compiled with relative ease from passport photographs because of their size requirements.

Unfortunately, however, the slightest misangulation when using the dentition method can yield erroneous photographic scaling factors. At this time, a system of identification based on dentition has not been endorsed by medicolegal authorities. A system of identification supplemental to fingerprinting would be highly desirable, especially in cases in which the soft tissues have been mutilated, destroyed by fire, or simply lost through decomposition.

Problems with Calculation of the Angulation of the Antemortem Photograph—The difficulty of establishing correct enlargements of photographs is compounded by the need for establishing a correct angulation of the skull to the photograph. Full facial photographs are not always available, and any variation in facial angulation from the frontal view must be carefully estimated before the skull can be oriented properly for comparison and scale matching. Typically, flexible mounts are used to position the skull while it is being viewed through a camera. Older techniques using still photography suffered in that the face and skull could not be compared directly while the skull was being positioned for a photograph. As a result, many photographs had to be taken of the skull in varying orientations. The technique of video superimposition overcomes the protracted time

involved with photographic superimposition. In their paper describing a procedure for video superimposition, Bastiaan et al. concluded that "video superimposition will provide quicker and greater flexibility in . . . the enlargement of the antemortem photograph and the alignment of the skull to the photograph" [*20*]. This process is still quite subjective, however, relying on the skill and dexterity of the operator. Bastiaan also observes that a great advantage of the video technique is the operator's ability to fade either the skull or the antemortem photograph in and out of the video screen, which allows careful overall assessment of how well the two images match.

Problems with Lack of Quantitative Data on the Fit—Even though superimposition identifications continue to appear in the literature, the value of superimposition has been challenged on the basis that alignment and enlargement factors are too variable. Forensic scientists, judges, attorneys, and juries often view superimposition with skepticism. The statement made by Justice Wrottesley while concluding the *Rex v. Dobkin* (Baptist Church Cellar Murder) case is typical: "These photographs do not disagree, and if you take my advice you will not regard that as evidence more compelling than that" [*13*]. According to DeVore, a specialist in forensic dentistry,

> It must be strongly emphasized that this method is for general information only and cannot be used for positive identification since the magnification and angulation of the original . . . picture are unknown. . . . This method of photographic superimposition is of more value in exclusion than identification [*27*].

It should be noted that Australian courts now accept video superimposition as an identification tool [*28*]. Its value is highlighted when other methods of identification are not possible or reliable. Also, most recently photographic superimpositions were accepted as probable identifications by counsel for the defense and prosecution in the High Court of Hong Kong [*14*]. These identifications were also accepted by the families of the victims.

Recently, a computer program and video camera system has been developed to improve the reliability of video superimposition and to reduce the problem of alignment and enlargement [*21*]. This system uses a special optical device to project a grid onto the superimposed image. A computer software package is used to evaluate the fit of the superimposition to yield analytical descriptions by kth-order polynomial equations and Fourier harmonic analysis. This system still requires manual repositioning of the specimens, however. Thus, the technique generates localized fit evaluations, each based on a manual repositioning of the specimens. One must then manually iterate through a number of positions to determine the closest. This is obviously tedious, with no guarantee that the closest fit will always be found.

Digital Photographic Superimposition Method

In order to reduce the abberant and nonquantifiable behavior of the previous techniques, the methodology developed here relies on stricter generation, control, and filtering of data before the superimposition process. Then, graphical superimposition is accomplished through the use of a near-optimal, computational transformation. The methodology (illustrated in Fig. 1) includes the following:

- two-dimensional (2-D) digitization of an antemortem facial photograph,
- three-dimensional (3-D) digitization of the surface mesh of a skull,
- application of digital filtering techniques to either or both models to reduce or eliminate systematic error,
- selection of four landmark points on the digital facial image and four equivalent noncoplanar landmarks on the skull surface mesh,

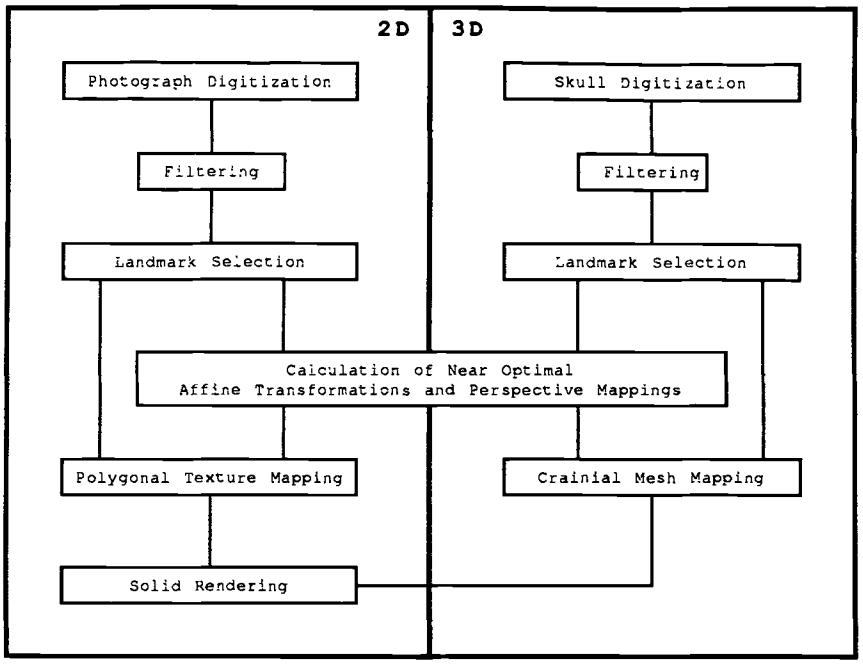

FIG. 1—*Digital photographic/surface mesh superimposition.*

- calculation of the near-optimal affine and perspective transformations required to map the skull surface mesh into two dimensions and onto the face, and
- joint solid rendering of the digital facial photograph and transformed skull surface mesh for visual analysis.

Given (1) a single polygon texture mapped with the digital facial image, and (2) a three-dimensional polygonal mesh of the skull, one can then digitally superimpose them, producing a solid rendering of skull and image. Given the appropriate set of rigid body transformations and a perspective projection for the skull polygon mesh and the texture mapped image, the skull mesh is then rendered with a semitransparent polygon image. Thus the viewer looks through the texture map at the skull. This gives a very apparent superimpositioning of the skull and facial image. Of the many benefits accrued with this technique, one of the most important is the quantitative fit between the skull surface mesh and the digital facial photograph. This fit may provide additional insight to the forensic scientist for making an informed identification decision. The remainder of this paper discusses each of the steps outlined in Fig. 1.

Near-Optimal Skull/Antemortem Facial Photograph Superimposition

Two-Dimensional Antemortem Facial Photograph Digitization

The 2-D digitizing system used for photograph measurement is a digital image processing (DIP) system. A digital image denotes an image which has been decomposed into a discrete 2-D matrix. Each of these picture elements (pixels) is a 3-tuple, $p[x, y, i]$, where $\{x, y\}$ is the matrix position of p, and i is the integer light intensity value, proportional to the brightness (or gray level) of the image at that point. That is, a digital image is an image that has been mathematically discretized in both spatial coordinates

and brightness. Resolutions of digital cameras range up to 1024 × 1024 pixels depending on price. Similarly, color cameras can be used as well. If the camera is color, the pixel becomes $P[x, y, c]$, where c is a vector whose direction is color hue in a {red, green, blue} additive system, and whose magnitude is intensity. One of the first applications of DIP was in the transmission of newspaper pictures between London and New York [*29*], dropping the time for transmission of a picture across the Atlantic from more than a week to less than 3 h. More recently, DIP has been utilized to transmit space images in the Surveyor, Ranger 7, Mariner, and Apollo space missions. One area of current emphasis is in the enhancement of pictorial information for human interpretation and analysis, which can, of course, be used in digital superimposition. This digitizing process can be accomplished by a number of currently available hardware and software digitizing systems. The system used in this research is the TRAPIX image processing hardware and RTIPS (real-time image processing software) [*30*]. When using the system, one places a photograph in the view area of the digital camera and then focuses and adjusts the intensity and gain for optimal picture quality. The photograph is then discretized and stored for later processing. The RTIPS software is capable of enlarging selected areas of the image, as well as performing a number of potentially important enhancements, as detailed below.

Processing the Facial Image—In superimposition, the quality of the facial superimposition is heavily dependent on the quality of the photograph. Thus, a photograph which is overexposed or underexposed, blurry, or noisy will result in a low-quality superimposition. In addition, if valuable edges such as dentition and feature profiles are vague because of the poor quality of the photograph, then the final fit will also be vague. The goals of image enhancement for superimposition can be enumerated as follows:

- enlargement of the facial region of interest, manually eliminating or blocking background features not germane to the superimposition; and
- filtering the image to produce as much useful image information, contrast, and clarity as possible.

The first task in this section is to isolate the facial region digitally. This is accomplished through selective enlargements, rotations if necessary, and replacement of noisy backgrounds with a neutral color or grey shade. Although not an absolutely necessary step, this isolation reduces the complexity of the remainder of the filtering. This step is typically easily accomplished with appropriate software contained in most digital image processing packages. In addition, many filtering algorithms are also contained in such packages. Techniques that have been used to advantage in the test cases outlined later include the following:

- *median filtering*—used to remove photographic noise;
- *histogram equalization*—used when a photograph is overexposed or underexposed or has specular highlights or distinct shadows; and
- *Wiener filtering*—used to remove linear uniform blurring of an image.

These techniques are of particular interest for photographic enhancement and restoration; however, they are by no means representative of all the enhancement and restoration techniques available. For a more comprehensive study of image enhancement and restoration techniques in general, see Ref *29*. After this step is complete, the facial image should be digitized and enhanced as much as is feasible.

After any needed digital isolation and restoration of the image, the next step in the digital fit process is the selection of landmark points. This selection process in photosuperimposition has been studied by Chai et al. [*31*] with the specification of 34 landmarks along 8 reference lines (also called marking or determining lines) on the face and skull. The landmarks used here consist of the (1) *ectocanthion* points at the left and right outer

orbits along the ectocanthion determining line and the (2) *glabella*, (3) *nasion*, and (4) *subnasal* landmarks along the front central determining line. These points and their determining lines, shown in Fig. 2, were shown by Chai et al. to have the highest probability of superimposition. If available on the image, dentition landmarks (one of which is also shown in Fig. 2) have been used here as well, although only on upper mandibular teeth. Landmark points on the lower mandible are not suitable because of unknown angulations of the jaw in the facial image. Of the six landmark points presented, one overriding concern is the necessity that the set of four chosen be noncoplanar.[3] If the landmark set is planar, the application of rigid body transformations and projections of the cranial mesh will result either in aberrant fits with the facial image or in numerical problems with the fit algorithms themselves. In addition, as the spatial separation between landmarks grows, the time to convergence will decrease and, in general, better fits will result. The test cases outlined in a later section typically used either the glabella or nasion landmark, the two ectocanthion landmarks, and an upper mandibular dentition point if present or the subnasal point. If the subject's facial image does not present one or more of these landmarks, others can be chosen from the list presented by Chai et al. in their paper.

Since the rendering methods of most common computer graphics algorithms require polygons, it is useful now to map the discretized and possibly restored and enhanced image (and its landmarks) onto a polygon. Intuitively, when rendered as a solid model,

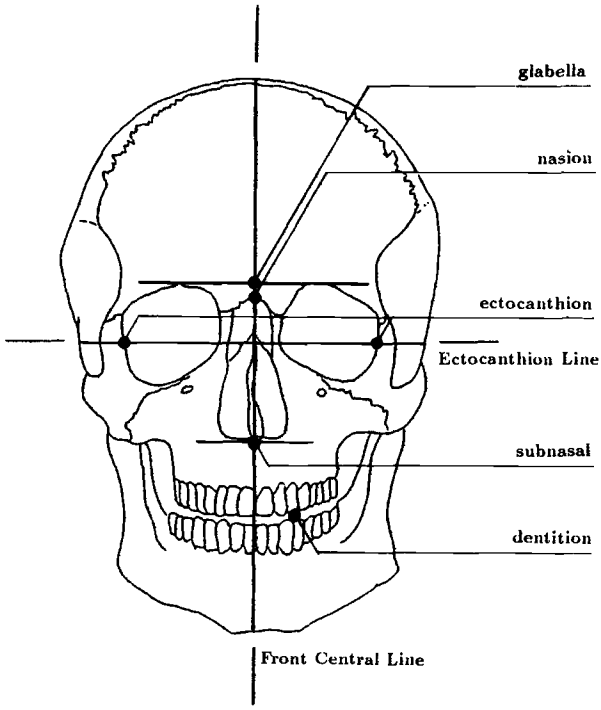

FIG. 2—*Landmark selection and the determining lines.*

[3]The reason four landmark points are selected will be discussed in the section entitled Calculation of Near-Optimal Affine and Perspective Mapping.

the polygon will contain the digital image. Given the discretized image as $p[x, y, i]$ and function f spatially mapping the image onto the polygon, as the polygon point $\{x_i, y_i\}$ is rendered, the intensity of that point is calculated from $p[f(x_i, y_i), i]$. This process of mapping images to polygons is known as texture mapping. Texture mapping has been extensively used in computer imaging for such varying texture simulations as the texture of a brick wall to a flower pattern on a teapot [*32*]. Extremely complex patterns and textures are easily simulated using this method. After texture mapping, note that the polygon can be uniformly or nonuniformly scaled about three axes, rotated, translated, skewed, and so forth, with a corresponding distortion in the resulting image. The texture-mapped photograph is now a distinctive, unique object and can be manipulated using computer graphics techniques as easily as the skull model. Algorithms for texture mapping are discussed in Ref *32*.

Three-Dimensional Skull Surface Mesh Digitization

The ideal situation in superimposition would be to compare a 3-D skull with a 3-D head or face. Unfortunately, photographs provide only a planar representation of the face. Here, a 3-D representation of the skull is used for iteration through different scales and angulations until an optimal or near-optimal fit is achieved. With this in mind for digital superimposition, the skull must be both digitized and represented in three dimensions. The digitizer used in this research was developed concurrently with the superimposition and is known as the Digibot [*33*]. It is an automated 3-D range finding system capable of measuring surface contours of an object with a resolution better than 100 surface points per square centimetre. The measuring principle used in the Digibot is triangulation, as demonstrated in Fig. 3. A spot of light produced from a laser is projected onto the object being measured. Then, a sensor methodically searches for this spot and when found records its position with reference to the laser and object. When $\angle B = 90°$

$$\frac{\bar{D}}{\bar{R}} = \tan A \quad \text{and} \quad \bar{D} = \bar{R} \tan A$$

with a single point being digitized in the laser-sensor reference frame. The object is measured through a 360° rotation, producing a 2-D horizontal slice of the object. This process takes place at multiple vertical levels, and the resulting combination of slices

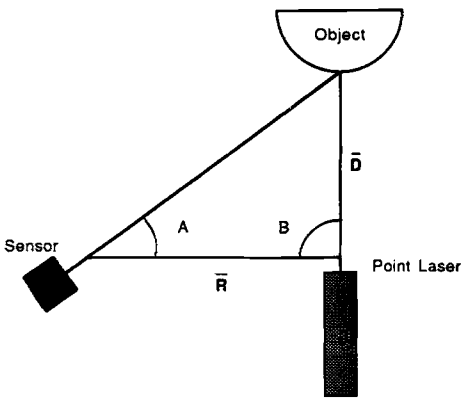

FIG. 3—*Triangulation theory utilized by the Digibot.*

defines the object in three dimensions. The output from the Digibot is a slice, or level-organized list of cartesian triples, $\{x, y, z\}$, measured with respect to an origin in a skull reference frame. These horizontal levels are then radially joined, producing a 3-D mesh of the object. Figure 4*b* shows a computer representation of the skull mesh consisting of nearly 30 000 points. In digitizing the skull, such regions as the eye orbits, nasal cavity, and temple regions are not digitized to highlight the bony edges for superimposition. The digitizing time requirements, design methodology and optimization, and error analyses of the Digibot in general and as optimized for skull digitization are outlined in Ref *34*. In addition, other 3-D digitizers and range-finding equipment are now available in the marketplace.

Polygon Filtering—Skull mesh generation is a necessary but not entirely satisfactory step since it does not provide an accurate, visual depiction of the skull. To generate a more succinct feeling of reality, the skull is rendered as a solid model rather than a wire-frame. Actually, the correct term for this type of graphics rendering is "surface" modeling. However, to differentiate between the "flat" photograph and the 3-D skull, this type of modeling will be referred to here as solid modeling. In solid modeling, the vectors defining the object are combined forming planar polygons. Using computer graphics lighting models and polygon smoothing algorithms, polygonal models can be smoothly shaded, yielding extremely life-like resemblances, such as is shown in Fig. 4*a*. Solid modeling of objects yields a realistic rendering; however, this method becomes computationally intensive for complex objects. The algorithms involved in solid renderings are well known and can be found in any good text on computer graphics (see Ref *35* as an example).

In an effort to reduce the computational and memory complexities inherent in solid modeling, a polygon reduction algorithm was developed. The wire-mesh skull shown in Fig. 4*b* results in over 30 000 polygons. Polygonal reduction is quite desirable to provide a reasonable computational load in the rendering phase on small desktop computers. However, a high resolution of the skull provided by the original model is desired. Note that most information occurs where adjacent polygon curvature is high. Areas such as

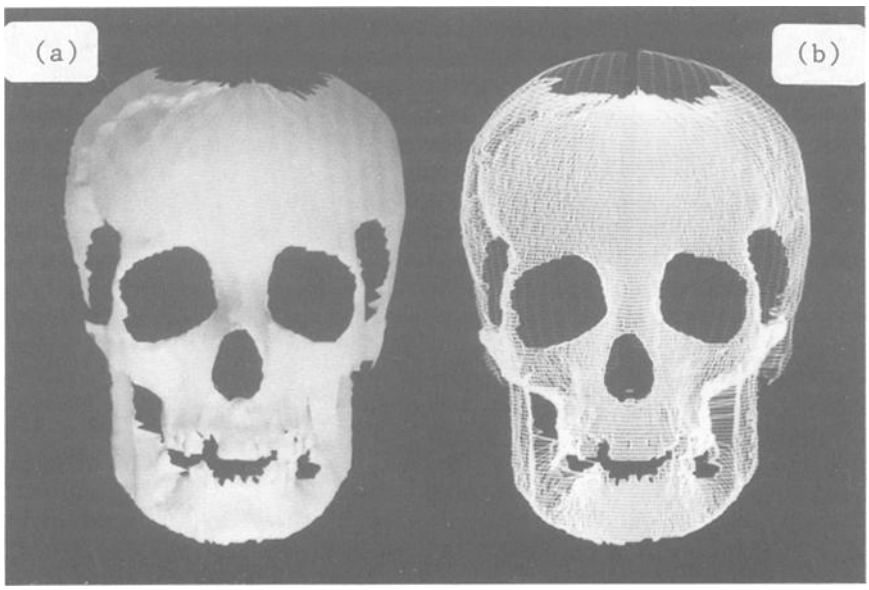

FIG. 4—*Models of the skull:* (a) *solid and* (b) *wire-frame.*

the forehead, top, sides, and rear of the skull do not have high rates of curvature and thus may be combined into larger polygons without loss of desired high-frequency data.

The polygon reduction scheme uses polygon normals computed by taking the cross product of two sides of the triangular polygon. Then, the dot product of the polygon surface normals between adjacent polygons is computed to measure the surface curvature. When the change in angle between the surface normals of adjacent polygons reaches a user-specified limit, these polygons are replaced with a single, larger polygon. Using this scheme, contiguous sets of polygons which do not uniquely define the model to some small error are replaced by larger polygons. Polygons which define sharp changes (high-frequency objects) are left alone, retaining original characteristics of the object at those points. The algorithm, fully detailed in Ref *34*, was implemented using a successive pass method where each pass further decreases the number of polygons. After reduction of the polygons, the model is then ready for solid rendering. Using this reduction technique, the mesh in Fig. 4*b* was reduced by over 75%, from 30 000 polygons to nearly 7000. Since many solid-rendering algorithms require computation time on the order of the square of the number of polygons, such reduction can theoretically drop execution time to 1/16th of that required for the nonreduced polygonal data. As an example then, if an original mesh requires 1 h of rendering time, the reduced mesh will be rendered in less than 4 min. Many IBM personal computer (PC) class machines, as well as mainframes and engineering workstations, support solid-rendering software and can take excellent advantage of the speedup.

Calculation of Near-Optimal Affine and Perspective Mappings

At this point, the digital image of the face should be texture mapped onto a polygon. Also, this facial polygon should be in some standard, predefined position in two-space, with the $\{x, y\}$ positions of the landmark points recorded. In addition, the skull mesh should be filtered, with the $\{x, y, z\}$ positions of its landmark points recorded. The final step before rendering the superimposition of both models is the calculation of the necessary affine transformations and perspective projection that will map the 3-D skull mesh and its landmark points as closely as possible to the 2-D landmark points on the facial image. In general, the mappings are developed from sets of rigid body transformations and a perspective projection. These transformations, listed in the order in which they are performed, consist of the following steps:

1. general rotation of the skull mesh by θ deg about an arbitrary three-space point and rotation axis, followed by
2. uniform scaling of the mesh, followed by
3. arbitrary translation of the mesh in three-space, followed by
4. mirroring the mesh about the z-axis, followed by
5. perspective projection of the mesh from three-space to two-space as a function of angle of view.

Uniform scaling coupled with perspective mapping of the cranial mesh is required to reduce or eliminate aberrant, nonproportional distortions of the mesh. A mirroring operation is performed, since data for the cranial mesh are gathered in a right-handed coordinate system but graphics devices and software usually operate in a left-handed system. The perspective distortion then maps three-space points into two-space such that objects further from the projection point become smaller. The angle of view of the projection is computable from the focal length of the camera lens used in the facial photograph if known. Here, however, it is left as an additional variable to be solved simultaneously with the remainder of the unknowns. This mapping problem can be generally specified as a set of eight equations in twelve unknowns. The unknowns are

(r_x, r_y, r_z) = origin of the general three-space rotation,
(d_x, d_y, d_z) = direction cosines of the axis of rotation,
θ = general angle of rotation of the cranial mesh,
s = constant three-space mesh scaling factor,
(t_x, t_y, t_z) = general three-space mesh translation vector, and
ϕ = the perspective angle of view,

with the known quantities of the system being the set of for landmark points from which to compute the projection:

$(x_{fi}, y_{fi}, 1)$ = facial landmark points on the $z = 1$ viewplane, and
(x_{ci}, y_{ci}, z_{ci}) = cranial landmarks.

The entire system of equations, then, in homogeneous coordinates [*36*] and in matrix form is shown in the Appendix. Given the twelve unknowns, four sets of landmark points from the digital image and cranial mesh then yield the resulting equations. Intuitively, four landmark points in three-space yield twelve knowns as well, which would then yield a deterministic system of equations (twelve unknowns and twelve knowns). However, remember that the facial image landmarks are all on the plane, so their z-axis coordinates are equal and dependent (that is, movement of one of the z-axis coordinates of a facial landmark equivalently moves the remainder). The actual system of equations then is eight equations in twelve unknowns. The other four equations are equal to zero however, since the z depth of the cranial mesh projection plane is necessarily equal to the z depth of the two-space facial plane. This system of equations allows total repositioning of the cranial mesh in three-space, including angulation, proportioning, and linear movement, as necessary. After repositioning, the perspective distortion is performed to map the three-space points into two-space. This distortion mimics the perspective result of photography, including the usage of camera lens focal length and format size, if known. If the camera lens focal length is known, one can find the appropriate viewing angle from charts available in most photography manuals. Obviously, this would reduce the number of unknowns by one. Three different methods to calculate these transformations were evaluated. The first was a heuristic method that exactly fit two landmark points and then iterated to find the closest fit to a third. The second was the classical approach to nonlinear optimization, and the third was a newer approach to numerical solutions that uses "learning" algorithms from artificial intelligence. This last fit technique has been found to be the most promising in terms of robustness, minimal user interaction, and calculation of the optimum.

Classical nonlinear optimization methods require a set of constraint equations along with an objective function to maximize or minimize. For the problem presented here, the constraint equations are the set of eight equations in twelve unknowns. The objective of the optimization is then to minimize the sum of the squared two-space Euclidean distances between the facial landmarks and the transformed and projected cranial landmarks. The amount of computation required to evaluate the constraints and objective function is certainly not trivial. Performing the matrix multiplications alone requires on the order of 500 floating point multiplications and 375 floating point additions. In addition, symbolically expanding the matrix arithmetic into eight algebraic equalities reveals that they are non-convex, ill-behaved, and nonlinear. Intuitively, there are an infinite number of "optimal" solutions to the problem down to the resolution of floating point numbers on the particular computer used to solve the problem. Unfortunately, these types of problems are some of the most difficult, if not the most difficult, to solve using classical nonlinear optimization techniques. Even so, potential (although certainly not guaranteed) solutions do exist, such as with the SUMT or penalty methods [*37*]. Both of these methods require the user to have some insight into either potential solutions or the weighted importance of one or more of the unknowns. That is, the classical solution to this problem

(if one can be found) requires continuing sophistication and insight into the problem for each imposition accomplished. In order to gain more insight into the problem, one can symbolically expand the objective function. This only compounds the problem, however. In fact, reducing the problem by assuming less general rotations and symbolically reducing the problem size as much as possible before numerical computation still yield an objective function that contains over 1000 floating point multiplications and nearly 600 transcendental functions.

Newer, nonclassical approaches to optimization of this type of problem may prove useful. For example, genetic algorithms [*38*] are now approaching the speed and generality to compute optimal or near-optimal solutions without needing the constraint equations. Specifically, they build a solution lattice on a hyperplane whose dimension is equal to the number of unknowns in the system. Generally then, the method iteratively chooses a lattice point and computes a value for the objective function. Since the algorithms deal with a hyperplane lattice, the unknowns must be discretized; furthermore, the integer ranges must be in powers of two. This poses no problem here. Consider a skull mesh whose convex hull measures 200 by 200 by 200 mm. Optimization with genetic algorithms could then work to an accuracy of at least 0.25 mm (which is no doubt accurate enough, considering that the result is visually analyzed) by discretizing each of the unknowns with the following mappings:

$$-2^{10} \leq \lceil \{t_x, t_y, t_z\} \cdot 4 \rceil \leq 2^{10}$$

$$-2^{10} \leq \lceil \{(r_x, r_y, r_z\} \cdot 4 \rceil \leq 2^{10}$$

$$0 \leq \left\lceil \frac{\{d_x, d_y, d_z\}}{0.072} \right\rceil \leq 2^{13}$$

$$0 \leq \left\lceil \frac{\theta}{0.072} \right\rceil \leq 2^{13}$$

$$0 \leq \left\lceil \frac{\phi}{0.072} \right\rceil \leq 2^{13}$$

$$0 \leq \left\lceil \frac{1024 \cdot (s - 1)}{-5} \right\rceil \leq 2^{10}$$

For the direction cosines that describe the axis of rotation as well as the rotation angle θ, a 0.072° angle change yields a maximum 0.25-mm step change at a 100-mm radius. At values of (θ/0.072) between 5000 and 2^{13} 8192, these angles overshoot 360°, which is obviously of no consequence. The perspective angle ϕ needs only the range over 180° (which would correspond to an extremely wide angle lens), but it must cover twice the radius, that is, the full convex hull. The discretized mapping of s is simply the inverse of a slope-intercept form of the line connecting the points (0, 1) and (1024, 0). The y value of the line then corresponds to a normalized scaling ratio whose smallest value is slightly less than 0.25 mm. The sum of the total number of bits required to contain the twelve discretized unknowns is then 141 ($3 \cdot 2^{11} + 3 \cdot 2^{11} + 3 \cdot 2^{13} + 2^{13} + 2^{13} + 2^{10}$). Bit sizes of 100 and over are generally thought to be "very large" problems for genetic algorithms.

Given the objective function to minimize, then, along with discretized unknowns, the method "learns" how to traverse the hyperplane to yield relatively better and better solutions (that is, evaluations of the objective function that are closer and closer to zero).

This learning process is at the heart of the genetic algorithm. Since this technique does not require continual sophisticated analysis by the user, it is deemed more appropriate for the optimization problem presented here. In addition, genetic algorithms return near-optimal solutions rather than failure where appropriate. This solution typically takes on the order of 10^4 iterations to converge to a minimum. On a SUN-4 Microsystems computer, this requires in excess of 4 to 6 h. These fit transformations are then used for final solid rendering.

Results from Digital Superimposition

Approximately twelve case studies have been superimposed using variants of the methodology discussed here. These included known matches and mismatches, several contrived mismatches, both sexes, and several age groups and races. Three of these were field cases to aid authorities in the identification of unknown human remains. Two of these resulted in a probable identification, one of which was later confirmed through other evidence. The other two have been neither confirmed nor denied. The identification of the remainder of these cases was known a priori; they were used for initial testing, debugging, and evaluation. Several of the case studies are discussed here.

Case Study 1

The skeletal remains of Case Study 1 were recently returned from the Vietnam War with an identification. The family was requesting further verification of this identification. Figure 5 shows one of the final probable identifications. Notice that the absence of temple regions in the skull improves eye and eye orbit comparison. Another photograph was used for a dentition superimposition (Fig. 6). The superimposition results of this case study were by far the most convincing of all the cases conducted thus far. In addition, the superimpositions were accepted as circumstantial identification by the family of the deceased.

FIG. 5—*Case 1: skull and face superimposition.*

FIG. 6—*Case 1: dentition superimposition.*

Case Study 2

Two superimpositions of this case study were conducted as shown in Figs. 7 and 8. The near-optimal fit method produced acceptable results, as is obvious from these superimpositions. Three different eye orbit points (at each eye orbit) were taken from the skull. Use of the lower point resulted in the most successful superimpositions. The two sets of skull and face control points varied less than 1%, based on the distance between the eye points for the automatic superimposition. In addition, the probable identifications

FIG. 7—*Case 2: skull and face superimposition.*

FIG. 8—*Case 2: skull and face superimposition.*

from these superimpositions were achieved quickly and accurately with no human intervention.

Mismatch Cases in Digital Superimposition

Figures 9 and 10 demonstrate a mismatched skull and photograph. The skull was taken from one case study and the photograph from another. Both subjects in these cases were

FIG. 9—*Mismatch 1: mismatched skull and face superimposition.*

FIG. 10—*Mismatch 2: mismatched skull and face superimposition.*

of similar height, weight, and sex. In Fig. 9, the width of the nose was used to scale and position the skull in front of the photograph. The nose appears to fit; however, the rest of the skull obviously does not match. Figure 10 demonstrates a match of the eyes/eye orbits and nose/nose cavity. Again, the skull does not fit the face: the skull is somewhat wide and the chin extends too far below that shown in the photograph, even though the teeth in the skull were clinched when digitized. Other mismatch studies have been done, but as of yet, not enough data have been gathered to compute mismatch results. We conclude that an expert forensic scientist must still view the results and pronounce whether or not they, in fact, match.

Conclusions

When the first steps in a photographic superimposition for human identification consist of digitizing the facial image and cranium in their respective spaces, a useful methodology can be wielded that reduces or eliminates major sources of error inherent in other techniques. The methodology developed here has been tested with success on known matches and mismatches, as well as on several in-progress identifications. Although requiring more sophisticated computing hardware and peripherals (typical two- and three-space digitizers will interface with PC level computer equipment) than other techniques, the gains in terms of image restoration and enhancement and near-optimal superimposition appear to outweigh these disadvantages. Currently, the time requirements for setup, digitization, and fit are several days. This can be reduced dramatically by integrating the digitizers with a color computer workstation. The skull typically requires 2 to 4 h to digitize, and the computation of an optimal mapping of skull onto facial image requires nearly 4 h more. The remainder of the processes typically take from several minutes up to 30 min of computation. The remainder of the time is spent transferring and translating data from one system to another. If the digitizers are attached as compatible peripherals to a single computer system, the entire process could require less than 8 h.

Further work is necessary to determine the continuing feasibility of nonlinear optimization of the transformation and mapping constraints. One new open question raised

by this technique is whether or not a correlation exists between the objective function minimization and match or mismatch between the facial image and the cranium. Obviously, the stochastic solution to this question will require digital superimposition of a large number of known cases. Such a solution will not replace the forensic scientist, of course, but it may provide strong guidance and support for legal question of identification decisions.

Appendix

The Matrix Transformations and Perspective Projection

Derivations of these transformations can be found in Refs *36* and *39*. The affine transformations and perspective mapping consists of sets of 4 by 4 matrices as follows:

$$[F] = \begin{bmatrix} x_{f1} & y_{f1} & 1 & 1 \\ x_{f2} & y_{f2} & 1 & 1 \\ x_{f3} & y_{f3} & 1 & 1 \\ x_{f4} & y_{f4} & 1 & 1 \end{bmatrix}$$

$$[C] = \begin{bmatrix} x_{c1} & y_{c1} & z_{c1} & 1 \\ x_{c2} & y_{c2} & z_{c2} & 1 \\ x_{c3} & y_{c3} & z_{c3} & 1 \\ x_{c4} & y_{c4} & z_{c4} & 1 \end{bmatrix}$$

$$[A] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -r_1 & -r_2 & -r_3 & 1 \end{bmatrix}$$

$$[D_1] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & d_2/v & d_1/v & 0 \\ 0 & -d_1/v & d_2/v & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[D_2] = \begin{bmatrix} v & 0 & d_1 & 0 \\ 0 & 1 & 0 & 0 \\ -d_1 & 0 & v & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[R] = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[S] = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[T] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ t_x & t_y & t_z & 1 \end{bmatrix}$$

$$[M] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[P] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \tan(\phi/2) & \tan(\phi/2) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where

$[F]$ = four facial landmark points mapped onto the polygon,
$[C]$ = four cranial landmark points in three-space,
$[A]$ = translation of rotation origin to three-space origin,
$[D_1]$ = initial setup rotation of the cranial mesh,
$[D_2]$ = secondary setup rotation of the cranial mesh,
$[R]$ = rotation θ deg about the axis of rotation,
$[S]$ = general constant scaling of the mesh,
$[T]$ = general three-space translation of the mesh,
$[M]$ = mirroring about the z-axis,
$[P]$ = perspective projection, and
$v = \sqrt{d_x^2 + d_z^2}$

Then the resulting system can then be computed as

$$[F] = [C] \cdot ([A] \cdot [D_1] \cdot [D_2] \cdot [R] \cdot [D_2]^{-1} \cdot [D_1]^{-1} \cdot [A]^{-1}) \cdot [S] \cdot [T] \cdot [M] \cdot [P]$$

where the parenthesized matrix operations perform the general rotation.

References

[1] Welcker, H., "Der Schädel Dante's," *Jahrbuch der Deutschen Dante-Gesellschaft*, Vol. 1, 1867, p. 35.

[2] Welcker, H., "On the Skull of Dante: A Letter from Hermann Welcker, Professor of Anatomy, Halle, Hon. Fellow, A.S.L., Corr. Memb., E.S.L., to Dr. J. Barnard Davis," *Anthropological Review*, Vol. 4, 1867, p. 56.

[3] Welcker, H., *Schiller's Schädel und Todtenmaske, nebst Mittheilumgen über Schüdel und Todtenmaske Kant's*, Fr. Vieweg und Sohn, Berlin, 1883.

[4] Kupffer, C. and Hagen, F., "Der Schädel Immanuel Kant's," *Archiv für Anthropologie*, Vol. 13, 1881, p. 359.

[5] Welcker, H., "Zur Kritik des Schiller-Schädels," *Archiv für Anthropologie*, Vol. 17, 1887, p. 19.

[6] Welcker, H., "Der Schädel Rafael's und die Rafaelportrats," *Archiv für Anthropologie*, Vol. 15, 1884, p. 417.

[7] Schaaffhausen, H., *Der Schädel Raphael's: Zur 400 Jährigen Geburtstagsfeier Raphael Santi's*, Max Cohen und Sohn, Bonn, Germany, 1883.

[8] His, W., "Anatomische Forschungen über Johann Sebastian Bach's Gebeine and Antlitz Nebst Bemerkungen über dessen Bilder," *Abhandlung Durch Mathematik und Physik*, Vol. 22, 1895, pp. 380–420.

[9] Kollman, J. and Büchly, W., "Die Persistenz der Rassen und die Reconstruction der Physiognomie Prähistorischer Schädel," *Archiv für Anthropologie*, Vol. 25, 1898, pp. 329–359.
[10] Stewart, C., "Committee on Printing, Superintendent Library and Naval War Records," in *John Paul Jones Commemoration at Annapolis*, U.S. Government Printing Office, Washington, DC, 24 April 1906.
[11] Lander, K., "The Examination of a Skeleton of Known Age, Race, and Sex," *Journal of Anatomy*, Vol. 52, 1918, pp. 282–291.
[12] Pearson, K. and Morant, G., "The Wilkinson Head of Oliver Cromwell and Its Relationship to Busts, Masks and Painted Portraits," *Biometrika*, Vol. 26, No. 3, Dec. 1934, pp. 1–116.
[13] Simpson, K., "*Rex v. Dobkin:* The Baptist Church Cellar Murder," *Medico-legal Review*, Vol. 11, 1943, pp. 132–145.
[14] McKenna, J., Jablonski, N., and Fearnhead, R., "A Method of Matching Skulls with Photographic Portraits Using Landmarks and Measurements of the Dentition," *Journal of Forensic Sciences*, Vol. 29, No. 3, July 1984, pp. 787–797.
[15] Glaister, J. and Brash, J., *Medico-Legal Aspects of the Ruxton Case*," E. & S. Livingstone, Edinburgh, United Kingdom, 1937.
[16] Gordon, I. and Drennan, M., "Medico-legal Aspects of the Wolkersdorfer Case," *South African Medical Journal*, Vol. 22, Sept. 1948, pp. 543–549.
[17] Sekharan, P., "A Revised Superimposition Technique for Identification of the Individual from the Skull and Photograph," *Journal of Criminal Law, Criminology, and Police Science*, Vol. 62, No. 1, 1971, pp. 107–113.
[18] Koelmeyer, T., "Videocamera Superimposition and Facial Reconstruction as an Aid to Identification," *American Journal of Forensic Medicine and Pathology*, Vol. 3, No. 1, March 1982, pp. 45–48.
[19] Iten, P., "Identification of Skulls by Video Superimposition," *Journal of Forensic Sciences*, Vol. 32, No. 1, Jan. 1987, pp. 173–188.
[20] Bastiaan, R., Dalitz, G., and Woodward, C., "Video Superimposition of Skulls and Photographic Portraits—A New Aid to Identification," *Journal of Forensic Sciences*, Vol. 31, No. 4, Oct. 1986, pp. 1373–1379.
[21] Delfino, V., Colonna, M., Vacca, E., Potente, F., and Introna, F., Jr., "Computer-Aided Skull/Face Superimposition," *American Journal of Forensic Medicine and Pathology*, Vol. 7, No. 3, 1986, pp. 201–212.
[22] Dorion, R., "Photographic Superimposition," *Journal of Forensic Sciences*, Vol. 28, No. 3, July 1983, pp. 724–734.
[23] Krogman, W., *The Human Skeleton in Forensic Medicine*, Charles C Thomas, Springfield, IL, 1962.
[24] Janssens, P., Hänsch, C., and Voorhamme, L., "Identity Determination by Superimposition with Anthropological Cranium Adjustment," *Ossa*, Vol. 5, 1978, pp. 109–122.
[25] Prinsloo, I., "The Identification of Skeletal Remains in Regina Versus K and Another: The Howick Falls Murder Case," *Journal of Forensic Medicine*, Vol. 1, No. 1, July/Sept. 1953, pp. 11–17.
[26] Sivaram, S. and Wadhera, C., "Identity from Skeleton—A Case Study," *International Criminal Police Review*, Vol. 32, 1977, pp. 158–160.
[27] DeVore, T., "Radiology and Photography in Forensic Dentistry," *Dental Clinics of North America*, Vol. 21, No. 1, Jan. 1977, p. 81.
[28] Brown, K., Clarke, B., Hollamby, C., and Congdon, I., "Identification in the Truro Murders," presented at the 7th Australian International Symposium on the Forensic Sciences, Sydney, Australia, March 1981.
[29] Gonzalez, R. and Wintz, P., *Digital Image Processing*, Addison-Wesley, 2nd ed., Reading, MA, May 1987.
[30] *Command Language User Guide, Real-Time Image Processing Software*, TAU Corp., Los Gatos, CA, 1987.
[31] Chai, D., Yu-Wen, L., Tao, C., Gui, R., Mu, Y., Feng, J., Wang, W., and Zhu, J., "A Study on the Standard for Forensic Anthropologic Identification of Skull-Image Superimposition," *Journal of Forensic Sciences*, Vol. 34, No. 6, Nov. 1989, pp. 1343–1356.
[32] Blinn, J. and Newell, M., "Texture and Reflection in Computer Generated Images," *Communications of the ACM*, Vol. 19, No. 10, Oct. 1976.
[33] *DIGIBOT Product Literature*, DigiBotics Inc., Ft. Collins, CO, 1989.
[34] Nickerson, B., "Automated Digital Skull-Face Superimposition in Human Identification," Master's thesis, Department of Mechanical Engineering, Colorado State University, Ft. Collins, CO, 1988.
[35] Rogers, D., *Procedural Elements for Computer Graphics*, McGraw Hill, New York, 1985.
[36] Newman, W. and Sproull, R., *Principles of Interactive Computer Graphics*, McGraw-Hill, New York, 1979.

[37] Himmelblau, D., *Applied Nonlinear Programming*, McGraw-Hill, New York, 1972.
[38] Fitzpatrick, J. and Grefenstette, J., "Genetic Algorithms in Noisy Environments," *Machine Learning*, Vol. 3, Nos. 2–3, 1988.
[39] Foley, J. and Van Dam, A., *Fundamentals of Interactive Computer Graphics*, Addison-Wesley, Reading, MA, 1982.

Address requests for reprints or additional information to
Dr. Patrick A. Fitzhorn
Department of Mechanical Engineering
Colorado State University
Fort Collins, CO 80523