Facial Reconstruction Using Volumetric Data

Mark W. Jones

University of Wales Swansea
Singleton Park
Swansea, SA2 8PP, UK
Email: m.w.jones@swan.ac.uk

Abstract

In this paper we demonstrate computer aided techniques in the area of the reconstruction of facial features from skeletal remains. We outline the desired features of a system for skull identification, and demonstrate our progress in the area which builds upon our previous work on distance fields for volumetric data. In addition to the application we demonstrate an accurate 3D closing operator in practise on large data sets, cross correlation for automatic feature detection, and the current output from the technique which uses a 2D warp of the 3D distance field.

1 Introduction

Identification of human remains is a challenging area of forensic science. In cases where there is some evidence that the remains are linked to a known person (perhaps a missing person in the local area), reliable techniques such as using dental and medical records may be used to identify the person. This particular application is concerned with cases where there is no likely candidate for the remains, and where the remains are disfigured, either through burning, decomposition or intentionally. In such cases the skeletal remains are all that are available from which a picture of that person should be created. Such a picture may be useful in triggering the memories of people who can then come forward with other information leading to the use of medical records for accurate identification.

Similar techniques are also used in archaeology – in the UK television programmes such as BBC's *Meet the Ancestors* have recreated the facial features of archealogical skeletal remains, and set them in historical context. Such methods were being used as early as 1895 when W. His reconstructed the head of Johann Sebastian Bach using tissue depths collected from cadavers.

This paper outlines a method for the fast reconstruction of facial features from discovered remains by acquiring volume data of both the remains and a reference head.

The traditional manual method for reconstructing a face is to reconstruct the head using a cast of the discovered skull as a basis [1]. This method requires an expert knowledge of anatomy aswell as artistic skills. Markers are placed at known reference points on the skull and cut to the correct depth using the data for each reference point. The face is built up using the appropriate muscles and overlayed with skin tissue using clay, and takes several days. The expert skill and amount of time required have motivated researchers to try to computerise the technique. The abovementioned reference points are collected at a limited number of feature points on a cadaver using a variety techniques (see Rhine and Campbell [2] for examples of tables). The principal technique is to use a needle to measure the depth at each point. The measurement process can be negatively effected by a number of factors such as the deformation of the surface during measurement, dehydration of tissue which starts immediately upon death, and the skill of the collector to judge the feature points and insertion angles. The majority of data has been acquired manually using the needle techique. Newer approaches [3] have used CT and ultrasound to collect depths more accurately on live subjects, but still suffer from the second main disadvantageous factor which is the data is only collected at a limited number (usually around 20-30) feature points. For example the *Manchester protocol* indicates 32 points, of which only 3 points are collected on each cheek. Almost all current computer assisted methods simulate the manual method by placing patches over the skull using the depths at feature points. The patches are textured using *average* faces obtained from image scans of various peoples faces. Quatrehomme et. al. [4] use CT scans of the discovered skull, and reference skull and a cast of the reference head tissue (to reduce radiation exposure). After registration the reference skull mesh is deformed parametrically to the discovered skull, but only on a subset of features (known as crest lines). When applying the same mapping to the face, the reconstructed face should be the candidate for the discovered person. In this paper we propose a pipeline that avoids the limited feature point table data sets by using the original depths from the CT data. We are able to achieve this through the use of our fast distance field computation algorithm (section 2.2.2).

In this paper we introduce the stages of a pipeline for the computer aided reconstruction of facial features. Section 2 will introduce the overall process and indicate our progress on each stage, including automatic feature correspondence (section 2.1), facial mapping using distance fields and closed skulls (section 2.2), and the distance field warp function and results (section 2.3). As this work is *in progress* we give details of some of the problems we face and suggestions for their solution in section 3.

2 Proposed Process

Figure 1 shows the main steps in the proposed process. The discovered skull is to be scanned using a CT scanner to obtain volumetric data. A reference head is chosen that has the same sex, racial and age characteristics as the discovered skull. A correspondence is created between the two heads, and using this correspondence the soft tissue from the reference head is mapped onto the discovered skull giving a candidate face for the unknown person. The problems that arise in and some potential solutions for each stage of the process are discussed in the following sections.

2.1 Feature Correspondence

In order to map the tissue depths from the reference head onto the discovered skull, a mapping between each point on the discovered skull onto the reference head must be calculated. The corresponding tissue depth can be looked up for each point on the discovered skull in order to render the new face.

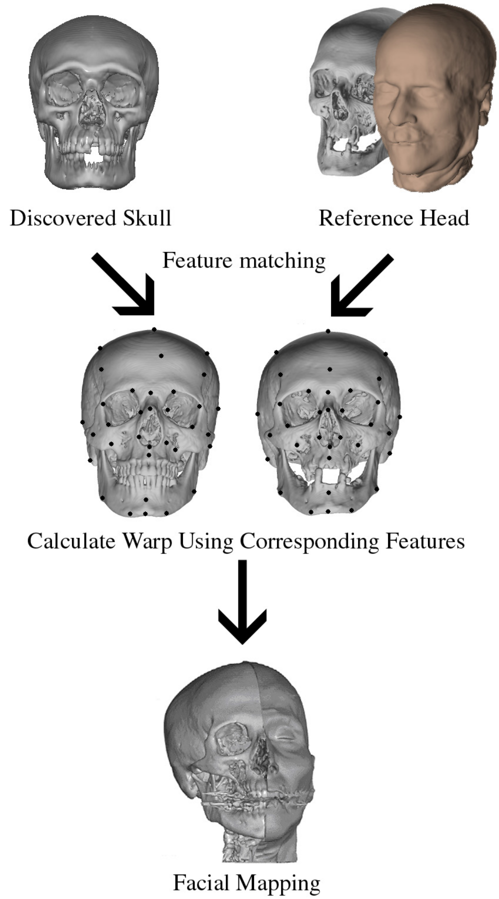

Figure 1: Reconstruction process

One of the desireable features of a complete system would be to allow the automatic correspondence of features between the reference head and the discovered skull. Feature detection is an active research area, but as the data is of a certain type it may be possible to use some prior knowledge to aid feature detection. In our case we have used cross-correlation to some success to enable automatic feature detection.

2.1.1 Correlation

Correlation [5] is the process whereby a filter can be convolved with an image in an attempt to identify regions of similarity in the image. It assumes that the image is approximately aligned with the filter. The application of the filter can be expressed as follows:

$$g(x,y) = \sum_{j=-q}^{q} \sum_{i=-p}^{p} h(i,j) f(x+i, y+j) \quad (1)$$

where p is half of the size of the filter in the x dimension, and q similarly for the y dimension.

The conventional application of this filter will not only give high values where the filter matches the image, but will also will give high values in regions of the image that are generally brighter than the rest. For this reason a normalised convolution is used:

$$g'(x,y) = \frac{g(x,y)}{\sum_{j=-q}^{q} \sum_{i=-p}^{p} f(x+i, y+j)} \quad (2)$$

We found that this formulation still presented a problem in identifying the feature points, so we used the following modification (to give image difference):

$$g(x,y) = \sum_{j=-q}^{q} \sum_{i=-p}^{p} |h(i,j) - f(x+i, y+j)| \quad (3)$$

and used the minimum as the candidate position for a match. Overall this obtained very good results. Figure 2 shows the use of our correlation function with a window size of 20×20 and 34 feature points that are useful for determining the construction of the skull. The left image shows the reference skull (in this case) with the feature points that are to be identified. The right image shows the discovered skull with the automatically detected feature points. In general the placement is good, but there are some points for which a useful position is not returned. We have further refined the algorithm by observing that the skull can be constrained to a similar view, and by adding a term accounting for the distance of the current point from the reference point:

$$\begin{gathered} g(x,y) = |(x,y) - (s_x, s_y)|^2 + \\ \sum_{j=-q}^{q} \sum_{i=-p}^{p} |h(i,j) - f(x+i, y+j)| \end{gathered} \quad (4)$$

where (s_x, s_y) is the position of the reference point. This yields the correlation given by Figure 3. It can be observed that the points which have failed are due to the very different bone construction behind the holes in the cheek and eye socket. We give further suggestions for dealing with this problem in Section 3. Currently the feature location algorithm searches for the match in the whole image, a further refinement would be to search in an area local to the reference point which should further remove problem cases. The algorithm currently finds the 34 feature points in a few seconds (on a 1.2GHz Athlon), so the time reduction by restricting the search area will not be significant.

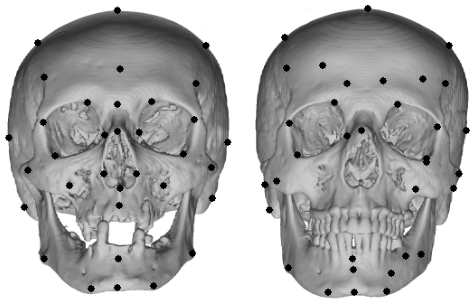

Figure 2: Correlation using equation 3

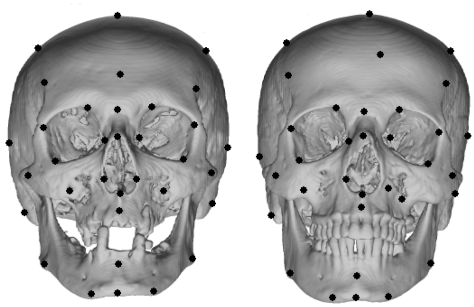

Figure 3: Correlation using equation 4

2.2 Facial Mapping

After the correspondence of features some computerised methods will take an average image of a range of people and map this texture onto the skull using the corresponding points and the appropriate depths at those points. The inbetween points are interpolated from the known points. This method of smoothing is similar to the manual method known as *strip plastic facial reconstruction* [6].

Previously Nelson and Michael [6] have suggested volume deformation of the reference head using the correspondence points to control it. The result is that the reference skull contained in the volume data is deformed to the same shape as the discovered skull, and as a volume deformation is carried out, the skin data within the data set should now correspond to the face of the discovered person. This method produced good results but is quite complex in terms of specifying the deformation reliably and in terms of computation. Also it does not refer to the tissue depths directly which diverges from the current well researched manual methods.

The approach we follow here is to choose a less computationally complex method using the correspondence features to direct a warped mapping between the skulls. For each point in the discovered skull the warp defines a corresponding point in the reference skull. We can look up the tissue depth at that point, and map the depth back onto the discovered skull. Rendering the new tissue should give the unknown face.

One of the main problems identified with previous manual and computerised techniques is that they use depths at limited positions around the skull. A requirement of the collection process is that the depths are to be collected at a certain angle. From a computer graphics viewpoint we could regard this as being *normal* to the surface, although some depths are collected in the direction of features rather than normal to the surface (for the ease of collection). All points inbetween the 20 or so feature points are interpolated from the known feature points, and this process has been identified as being a major source of innacuracy.

Our method essentially solves this problem by using the depths at all points on the skull. For example rather than just three points being known on a cheek and all others being interpolated from these, in our case each point on the skull in the cheek area will be mapped back to the reference head, and the depth of the corresponding point will be measured from the CT volume data set and used as the tissue depth for the discovered skull.

One problem with this method is for example the mapping of tissue onto the eye. The closest point may be the eye socket, but intuitively we would like to map the eye onto a closed surface at the eye socket. For this reason we decided to compute on a data set formed by a closing operation on the skull.

2.2.1 Mathematical Morphology

The mathematical morphological operation of erosion [5] will remove external parts of an object (depending upon the structuring element). Dilation will add parts to the boundary of the object. Opening will enlarge cracks and cavities, and closing will remove cavities and smooth spikes. Definitions for these operations are given below.

Given a structuring element B, erosion of an object X is defined as:

$$X \ominus B = \{x | B_x \subset X\} \quad (5)$$

and dilation as:

$$X \oplus B = \{x | B_x \cap X \neq \emptyset\} \quad (6)$$

Closing is a dilation followed by an erosion:

$$X \bullet B = (X \oplus B) \ominus B \quad (7)$$

and openning is an erosion followed by a dilation:

$$X \circ B = (X \ominus B) \oplus B \quad (8)$$

If B is defined as a ball of radius r:

$$B = \{b | d(b, (0,0)) \leq r\} \quad (9)$$

where d is the distance between two points, then erosion can be defined as:

$$X \ominus B = \{x | sgn(x) \cdot dist(x) \leq -r\} \quad (10)$$

and dilation as:

$$X \oplus B = \{x | sgn(x) \cdot dist(x) \leq r\} \quad (11)$$

where $dist$ is the distance between x and the boundary of an object X, and sgn is -1 inside the object, and $+1$ outside the object.

This gives us a practical algorithm for finding a closed skull (i.e. one in which the cavities such as the eye sockets have been closed). First we calculate the distance field D for our surface as in equation 12. The surface, S, representing a dilation with a ball of radius r is equivalent to $S = \{q : D(q) = r\}$ where $q \in \mathbb{R}^3$. To calculate the erosion with a ball of radius r for this surface S, we calculate a new distance field, D' based upon distances from surface S using the techniques described earlier (measuring to the triangulation of the isosurface). The surface S^{C_r} which has been closed with a ball of radius r through a dilation of degree

r followed by an erosion of degree r is given by $S^{C_r} = \{q : D'(q) = -r\}$.

This method has not found widespread use for large three-dimensional data sets because it is so expensive to compute the distance field function. Morphological operators have been carried out on binary segmented data. Höhne [7] has reported the use of three-dimensional erosions and dilations as a useful application for the extraction of homogenous regions in body scans. An alternative to erosions and dilations upon binary segmented data, is to increase or decrease the grey-level voxel values themselves. For example, if the template is a $3 \times 3 \times 3$ matrix where each value in the matrix is 1, for a dilation from a binary segmentation f (equation 13), each voxel v where $f(v) = 1$ will be included in the new surface along with all of its 26 neighbours. Using the original voxel values, a grey-level dilation will set a voxel to the value of its maximum neighbour (plus 1). This does not seem to lend itself to the intuitive view that a dilation or erosion of 5mm should be an offset surface 5mm outside or inside the object.

2.2.2 Distance Field Calculation

At last year's VMV conference we presented distance fields as an approach to modelling objects. Distance fields can be calculated efficiently for volume data using the following method [8, 9].

To calculate the distance field in the required form:

$$\begin{aligned} & D(p) = sgn(p) \cdot min\{|p - q| : q \in S\} \\ & sgn(p) = \begin{cases} -1 & \text{if p inside} \\ +1 & \text{if p outside} \end{cases} \\ & \text{where } || \text{ is the Euclidean norm} \end{aligned} \quad (12)$$

we first of all calculated the segmentation function f as:

$$f(v) = \begin{cases} 1 & \text{if } v \text{ is inside the surface} \\ 0 & \text{otherwise} \end{cases} \qquad \text{where } v \in \mathbb{Z}^3 \quad (13)$$

Then for each voxel, v, we add v and v_{26} (the 26 neighbours of v) to S_v, when $f(v) = 1$ and $\exists p$ such that $f(p) = 0$ where $p \in v_{26}$. Then for each voxel $v \in S_v$ we calculate the distance from that voxel to the closest point on the surface of the skull. This is achieved by organising the space using an octree so that parts not containing the surface can be rejected along with nodes that are further away than the current minimum distance already computed for the voxel v. As the maximum distance of any point in the shell from the surface is less than $2\sqrt{3}$ we can reject many nodes of the octree structure. To measure the distance to the surface we perform a triangulation of each transverse voxel [10], and measure the distance to the triangulation. This produces a shell of distance values around the skull in about 2 minutes on a 1.2GHz Athlon. This method can also be implemented in parallel as each voxel is independent of each other voxel, and only requires a copy of the CT volume data set. At the moment we calculate each slice in parallel, taking about 2 seconds per slice.

We now take this *sub-voxel accurate distance shell* and propogate it using our vector city vector distance transform (VCVDT) [9], which produces a very low error approximation to the true distance field. (Previous methods have operated on binary segmented data rather than the sub-voxel accurate distance shell, and have employed worse chamfer or vector distance transforms). The propogation process takes about 7 seconds on a 1.2GHz Athlon, and the result is a complete distance field of the desired format upon which the morphological closing operator can be carried out. The total running time of about 2 minutes (on one processor) to obtain a distance field, compares to 2 hours if the distance is calculated accurately from every voxel rather than the voxel shell. For this approximation the average error per voxel is 0.0097 (with 72% of voxels being incorrect) which is caused by vectors pointing to the closest point for voxels in the shell, and not the correct closest point for the current voxel. A recalculation of the distances using these vectors to indicate which voxel to use to calculate the surface reduces the average error to 0.0013 and produces a data set in which 93% of voxels are correct. This adds an additional 2 minutes to the running time. We are still making progress towards a 100% accurate method using the vector information and by keeping a list of *tie* situations.

2.2.3 3D Closing

The 3D closure of the skull, S^{C_n}, (Section 2.2.1) can be calculated using the following process. The original data can be converted into a distance field, another distance field can be calculated for the dilated surface, and a third distance field can be calculated for the erosion of the dilated surface to create the closed surface. Previously this computation would not have been considered due to the sheer amount of time, or will have been computed inaccurately (by using binary segmented data). We have shown that it is possible to calculate accurately for large data volumes.

Figure 4 demonstrates accurate 3D closing operations of 20, 10 and 5 on the eye socket and cheek of the UNC CThead.

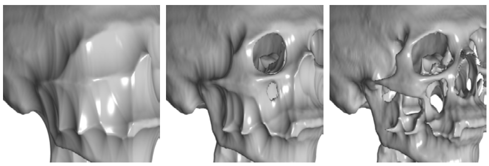

Figure 4: True 3D distance closure of 20, 10 and 5 voxels

The use of the closed skull (for both the discovered and reference data sets) facilitates the mapping of the tissue depths from the reference head onto the discovered skull, and it is such closed data sets that we use in our facial mapping process.

2.3 Warping

The assumption behind the process of creating data sets of depth measurements and using these to create a face from a discovered skull is the fact that the tissue depth should not differ greatly between skulls. This is assumed to be true for the manual methods, and is subject to choosing the correct data set according to the sex, racial type, origin, and build of the discovered skull. For our process we make a similar assumption – the reference CT data set has been chosen so that it matches the discovered skull in terms of the above factors.

The manual reconstruction methods will take the known data and map them on to the appropriate positions on the discovered skull, and then use interpolation or anatomical knowledge to fill in the areas in between those sparse points. For our method we suggest that a warp based on the correlated points should be used which defines a deformation from the discovered skull to the reference skull. It is then possible to map each point from the discovered skull onto the reference skull, and determine

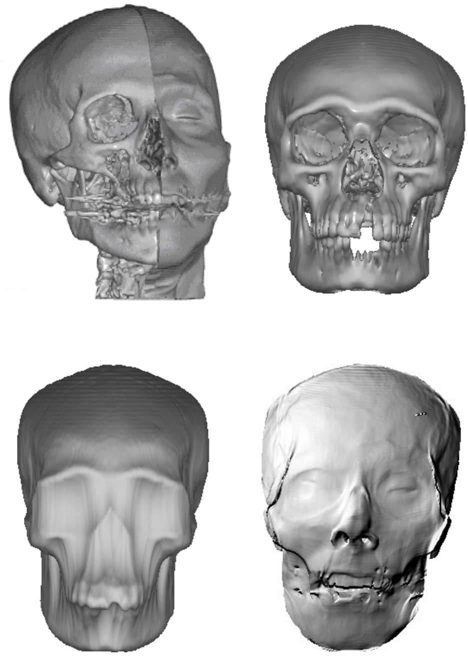

Figure 5: (a) Reference head (skull and soft tissue – the noise around the mouth is due to X-ray scattering off the metal fillings) (b) Discovered skull (c) Discovered skull closed (the reference head is also closed in the process) (d) Reference face mapped onto discovered skull

the distance from the skin from that point using the precomputed distance data of section 2.2.2.

As an initial test of the method we have used a 2D warp of the front view of the skull. This gives the result that for each point on the forward facing skull we are able to determine the corresponding point on the reference head, measure the depth and map it back onto the discovered skull. We tested two warping techniques for this part of the process – field warping and mesh warping [11].

Figure 5 shows the source heads, intermediate closed skull (for the discovered skull) and the result of the process. It was found that the two warping methods produce very similar results. In the case of the field warping method, the field lines can be defined once for the reference head using the desired feature points, and this connectivity is translated to the corresponding points on the discovered skull (once the feature correspondence has been found). This leads to an almost automatic algorithm – at the moment only the few points that are misidentified need to be moved before the final skin mapping can take place. In the case of the mesh warping method, a mesh can be created using the feature points, and as for the field warping technique, this can be mapped to the feature points on the discovered skull, again leading to an automatic algorithm (assuming the feature points are identified correctly). At the moment it takes about 20 seconds to automatically detect feature points (using correlation) and correct the field lines or mesh by moving the incorrect feature points.

Once the input to the warp method has been specified, the warp looks up each point for the discovered skull in the reference head data set and finds the corresponding tissue depth at that point. At the present time it takes about 20 seconds to calculate the warp and present an image. This demonstrates the power of the method when compared to the few days a forensic artist requires to produce a head.

As can be seen in figure 5 there are still some problems with the method. First and foremost is the fact that as this method is intended to be a *proof of concept* and therefore uses just a 2D frontal warp, it is not possible to map skin where there is no bone behind. Looking at a head, this corresponds to the ears and tissue around the neck. At the moment the mapping is restricted to the discovered skull, and so the outline is still skeletal in appearance. We discuss this problem further in the next section.

The second noticeable artifact is the apparent skin tear along the sphenoid bone (either side of the forehead). This may be due to the fact that we used a caucasian reference head to reconstruct a negroid skull which has a more prominent brow, and that there may be difference in orientation. More experiments will need to be made, in particular to orient skulls in the Frankfurt Horizontal position (the line between the external auditory meatus (centre of ear) and the point on the lower orbit (bottom of the eye socket) is made horizontal).

The success of using a closed data set can be clearly seen in the reconstruction. The eye tissue and nose (neither of which had any basis for mapping in the original discovered skull data set) have been mapped successfully.The mouth also has a satisfactory map, even though some of the teeth are missing from the discovered skull, and the reference head had some scattering of the x-rays around the mouth due to metal fillings. Large areas of the head, such as the forehead and chin have also mapped successfully.

3 Problems and Future Work

This is a work that is very much *in progress*. We have solved some issues with our first attempt, but we still have many problems which this section will cover along with our suggestions for solution.

3.1 Removing the View Restriction

As this work was carried out to assess the method, several shortcuts were taken, the most significant being the fact that the reconstruction is based upon one view in the direction of the viewing angle, which causes some tissue such as the ears not be be reconstructed. Removing this restriction has implications for several stages of the pipeline:

3.1.1 Correlation

In this case we now must correlate points at all parts of the skull and not just from the fixed front view. We have already made progress towards achieving this by projecting viewing rays from a sphere enclosing the skull, rather than from a fixed viewing plane. This creates images such as the one seen in Figure 6. Correlation has been used on this image to automatically identify points. At this current time we would like to use all of the standard Manchester feature points (most of which we use at the moment), along with some other anatomical features. Some of these are well defined in the medical literature, and it should be possible to incorporate some of those definitions into our work.

3.1.2 Depth Measurements

At the moment the tissue depths are available from the closest point on the skull, but are mapped as a depth along the ray path back to the viewing plane. In future we would like to map the tissue depths into the volume data according to the surface normal of the skull. This should remove current problem that depths along the viewing rays are only displayed where continuing the viewing ray results in an intersection with bone.

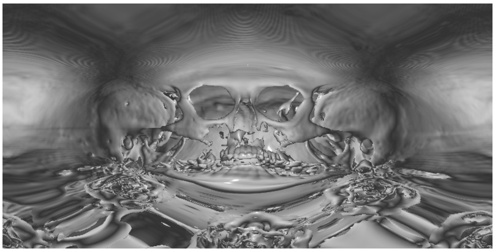

Figure 6: Projecting view rays from a sphere rather than plane

3.1.3 Orientation

For these experiments the skulls were oriented *by eye*. In future we would like to use a more consistent approach by using feature points to align the skulls.

3.2 Warping

We have demonstrated that a simple 2D warp produces very good results. We anticipate that using the spherical view of the data sets to obtain a warp should increase its accuracy. We will also consider 3D warping techniques suited to distance fields [12].

4 Conclusions

In this paper we have demonstrated a process for the reconstruction of facial features from discovered skulls. Depths are available at all points on a scanned reference head, and by corresponding each point on the discovered skull with a point on the reference head it is possible to map the tissue depths onto the discovered skull. We have demonstrated the progress made towards this goal, in particular within the areas of automatic detection of feature points, the closure of the skull in order for a *good* surface for mapping, the warping process, and the results of a facial map. We have also given our suggestions on how to improve this work.

Acknowledgements

I would like to thank Chris Whyley and Vincent Wu for their implementation of the mesh and field warps respectively. This work has been funded by EPSRC UK under grant GR/R11186.

References

[1] J. Prag and R. A. H. Neave. *Making Faces Using Forensic and Archaeological Evidence*. British Museum Press, 1997.

[2] J. S. Rhine and H. R. Campbell. Thickness of facial tissues in American Blacks. *Journal of Forensic Science*, 25:847–858, 1980.

[3] V.M. Phillips and N.A. Smuts. Facial reconstruction: Utilization of computerized tomography to measure facial tissue thickness in a mixed racial population. *Forensic Sci. Int.*, 83:51–59, 1996.

[4] G. Quatrehomme, S. Cotin, G. Subsol, H. Delingette, Y. Garidel, G. Grevin, M. Fidrich, P. Bailet, and A. Ollier. A fully three-dimensional method for facial reconstruction based on deformable models. *J. Forensic Sci.*, 42(4):649–652, 1997.

[5] R. Gonzalez and R. Woods. *Digital Image Processing*. Addison-Wesley, 1993.

[6] L.A. Nelson and S.D. Michael. The application of volume deformation to three-dimensional facial reconstruction: A comparison with previous techniques. *Forensic Sci. Int.*, 94(3):167–181, 1998.

[7] K. H. Höhne and W. A. Hanson. Interactive 3D segmentation of MRI and CT volumes using morphological operations. *Journal of Computer Assisted Tomography*, 16(2):285–294, 1992.

[8] M. W. Jones and R. Satherley. Voxelisation: Modelling for volume graphics. In B. Girod, G. Greiner, H. Niemann, and H. P. Seidel, editors, *Vision, Modeling, and Visualisation 2000*, pages 319–326. IOS Press, 2000.

[9] R. Satherley and M. W. Jones. Vector-city vector distance transform. *(Accepted for) Computer Vision and Image Understanding*, 2001.

[10] B. A. Payne and A. W. Toga. Distance field manipulation of surface models. *IEEE Computer Graphics and Applications*, 12(1):65–71, 1992.

[11] G. Wolberg. Image morphing: A survey. *The Visual Computer*, 14(8/9):360–372, 1998.

[12] D. E. Breen and R. T. Whitaker. A level-set approach for the metamorphosis of solid models. *(To appear) IEEE Transactions on Visualization and Computer Graphics*, 2001.