Max–Planck–Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 49

As we get older, do we get more distinct?

Alice J. O'Toole[1],Thomas Vetter[2],Harald Volz[3] & Elizabeth M. Salter[4]

March 1997

[1] The University of Texas at Dallas, E–mail: otoole@utdallas.edu
[2] AG Bülthoff, E–mail: vetter@mpik-tueb.mpg.de
[3] AG Bülthoff, E–mail: volz@mpik-tueb.mpg.de
[4] The University of Texas at Dallas

This document is available as /pub/mpi-memos/TR-49.ps via anonymous ftp from ftp.mpik-tueb.mpg.de or from the World Wide Web, http://www.mpik-tueb.mpg.de/bu.html.

As we get older, do we get more distinct?

Alice J. O'Toole, Thomas Vetter, Harald Volz & Elizabeth M. Salter

Abstract. We applied a standard facial caricaturing algorithm to a three-dimensional representation of human heads. This algorithm sometimes produced heads that appeared "caricatured". More commonly, however, exaggerating the distinctive three-dimensional information in a face seemed to produce an increase in the apparent age of the face — both at a local level, by exaggerating small facial creases into wrinkles, and at a more global level via changes that seemed to make the underlying structure of the skull more evident. Concomitantly, de-emphasis of the distinctive three-dimensional information in a face made it appear relatively younger than the veridical and caricatured faces. More formally, face age judgements made by human observers were ordered according to the level of caricature, with anti-caricatures judged younger than veridical faces, and veridical faces judged younger than caricatured faces. We discuss these results in terms of the importance of the nature of the features made more distinct by a caricaturing algorithm and the nature of human representation(s) of faces.

1 Introduction

Facial caricatures are used commonly by artists to accentuate or exaggerate the distinctive information in individual faces. The automated production of caricatures has been possible for many years due to the relative simplicity of the algorithms needed to make them (e.g., Brennan, 1985). Typically, such algorithms operate as follows. First, a measure of the average value of a set of "features" across a large number of faces is computed. These features are defined, usually, as a set of facial landmark locations (e.g., corners of the eye and other points that are reasonably easy to localize/match on all faces)[1] in the two-dimensional image. Next, to create a caricature of an individual face, a measure of the deviation of the face from the average two-dimensional configuration is computed. Finally, "distinctive" or unusual features of the face are exaggerated to produce the caricature.

This generic algorithm has been applied both to line drawings of faces (e.g., Brennan, 1985; Rhodes, Brennan, & Carey, 1987; Benson & Perrett, 1994) and to photographic quality images (e.g., Benson & Perrett, 1991); both representations yield perceptually compelling caricatures. More formally, there is evidence from psychological studies that caricatures of faces are recognized more quickly and accurately than veridical images of faces (Benson & Perrett, 1994; Mauro & Kubovy, 1992; Stevenage, 1995) and further, are rated as "better likenesses" of individuals than veridical images (Benson & Perrett, 1994).

Face distinctiveness effects have figured prominently in many theoretical accounts of human face processing (e.g., Bruce & Young, 1986; Goldstein & Chance, 1980; Morton & Johnson, 1991; Valentine, 1991). Several studies have demonstrated that faces rated as distinct by observers are better recognized than faces rated as typical (e.g., Light, Kayra-Stuart, & Hollander, 1979). Computer-generated caricatures have been of interest to psychologists because they provide a direct method for testing the role of distinctiveness in face perception and recognition tasks (see Stevenage, 1995; O'Toole & Edelman, 1996). This direct control of distinctiveness is an important advantage of caricature studies by comparsion to rating studies. Studies in which the distinctiveness of individual faces has been manipulated directly via caricaturing lend more support to the claim that face distinctiveness is an important factor for human face perception and recognition (Perrett and Benson, 1991; 1994; Rhodes et al. 1987, Stevenage, 1995).

The primary difference between a caricature approach and assessing the relationship between rating and performance data (e.g., Light et al., 1979) is the explicitness of the definition of "distinctiveness" required in the caricature approach. Specifically, to apply a caricature algorithm to faces, one must first operationally define the information in faces to be exaggerated or caricatured. This is

[1] Though see also Burt & Perrett, 1995 for an application of this technique to image intensities.

done, typically, in terms of a set of landmark facial "features", e.g., corners of the eyes, tip of nose, that can be located in the image and altered. As noted, although caricature algorithms have been applied to different qualities of image data (line drawings versus photographs), they have not been applied generally to features other than those based on the two-dimensional configural structure of faces (though we discuss an exception, Burt & Perrett, 1995, in Section 4).

The possiblility of exploring the perception of caricatures made from an inherently three-dimensional representation of faces is interesting for several reasons. First, the question of the extent to which human representations of objects and faces encode two- versus three-dimensional "features" has been very much-debated in the psychology literature since the important papers of Biederman (1987) and Bülthoff and Edelman (1992). Second, these representation questions are now being investigated actively by neuroscientists interested in the neurophysiological substrates of object and face recognition (Logothetis, Pauls, Poggio, 1995; Perrett, Hietanen, Oram, & Benson, 1992). A computationally based approach for defining and altering face distinctiveness, in combination with a human perceptual assessment of the results of such alterations, is potentially very informative about these issues because it mandates a reasonably precise definition of the face representation. As such, the ways in which individual faces can be considered distinct within these different representational contexts can be specified, and the perceptual consequences of these different specifications can be examined empirically. As we will illustrate shortly, "distinctiveness", defined with respect to the three-dimensional information in faces, captures at least one *perceptual* dimension that is not similarly present in face representations defined only on two-dimensional data.

In the present study, we applied a standard caricature algorithm to a representation of the three-dimensional structure of the head. To our surprise, we were unable, for the most part, to produce good caricatures of faces that did not, also, appear to have been aged – often considerably. Because this finding was unexpected, the present paper is organized in an unusual way. We first demonstrate the application of the caricature procedure to a number of male and female heads, using a range of caricature distortion values. Next, we present an experiment in which human observers judged the age of four versions of the heads, including an anti-caricature and two levels of caricature. These data indicated that the caricature manipulation increased the apparent age of the face. Finally, we discuss the primary human face and skull changes that occur with age and speculate about the cues that seem to be captured in the present manipulation. We argue that these cues form a useful complement to the relatively narrow repertoire of facial age cues that has been considered thus far in psychological studies.

2 Caricature Procedures

2.1 Description of Laser Scan Head Stimuli.

Laser scans (CyberwareTM) of 100 heads of young adults (50 male and 50 female) were used as stimuli. The mean age of faces in the data base was 26.9 years (standard deviation = 4.7 years), with a minimum age of 18.1 years and a maximum of 45.8. The distribution was slightly skewed toward the younger ages — e.g., we had only four faces older than 35 years.

The subjects were scanned wearing bathing caps, which were removed digitally. Additionally, further pre-processing of the heads was done by making a vertical cut behind the ears, and a horizontal cut to remove the shoulders.

The laser scans provide head structure data consisting of the lengths of 512×512 radii from a vertical axis centered in the middle of the subject's head to "sample" points on the surface of the head. This is a cylindrical representation of the head surface, with surface points sampled at 512 equally-spaced angles around the circular slices of the cylinder, and at 512 equally spaced vertical distances along the long axis of the cylinder. The radii in our data set varied from a minimum of 4 cm (neck area) to a maximum of 12 cm (nose area). The resolution of the scanner was 16 microns (0.0016 cm).

2.2 The Correspondence Problem

One basic requirement for any automated caricaturing or morphing algorithm is that the example faces be set into some kind of "correspondence". This is necessary regardless of whether the caricaturing or morphing operation is applied to two-dimensional images or to three-dimensional head models. Correspondence refers to an alignment of the important parts of the faces so that these face parts overlap or coincide in a meaningful way in the representational code. In most implemented morphing or caricaturing programs, correspondence is achieved via a non-automated prerequisite process in which a human operater locates and "marks" a number (usually between 30-300)[2] of facial "landmarks" in each face image. Landmarks refer to facial structures that are easy to locate on all faces, e.g., the corners of the eyes or tip of the nose.[3] The set of locations of these points in each face forms the basis of a standardised two-dimensional shape code that can be compared meaningfully across all faces images. Additionally, the areas between the distinct feature points are often triangulated and linearly matched (Beier & Neely 1992, Craw & Cameron 1991). Applied to the problems of morphing and caricaturing, this correspondence-based code can be manipulated, averaged, and warped without incurring the difficulties encountered applying these kinds of operations to raw image codes (see Pittenger, 1991 for a discussion of the problem).

Although this sort of alignment or face correspondence is absolutely necessary for caricaturing or morphing, the hand placement of hundreds of landmark points on hundreds of faces is extremely time consuming. On the other hand, several decades of work in computer vision can attest to the fact that automating the correspondence-finding process is far from trivial. The correspondence problem for human faces is functionally equivalent to the better-known correspondence problems in structure-from-stereopsis (Marr & Poggio, 1976) and structure-from-motion (Ullman, 1981). In stereopsis, the left and right retinal images must be brought into correspondence, whereas, in motion, images taken at time t and at time $t+1$ must be made to correspond. Likewise, for faces, the problem is to match the corresponding "features" (i.e., facial landmarks) in the two images. In this broader context, it is worth pointing out that although functionally equivalent to the stereopsis and motion problems, the "correspondence" problem for human face morphing and caricaturing has been defined in a more limited way as the matching of *only* a subset of points, facial landmarks, in the image.

In the present work, we cast the face correspondence problem into the more general case where one attempts to match all of the data points, rather than just a subset of landmark points. In the three-dimensional case, these data points comprise the 512×512 matrix of radii just described. This means that we have to find, for every radius sample point on the face, (e.g., the radius specifying the right corner of the right eye), the corresponding radius location at the same feature in the other face. While this problem is far from solved in a perfectly general form, a great deal of progress has been made recently on the problem with faces. Specifically, several methods have been applied successfully to the task of automating a correspondence finding procedure for images of human faces (Beymer, Shashua & Poggio, 1993; Cootes, Taylor, Cooper & Graham, 1995; Lanitis, Taylor, Cootes & Ahmad, 1995; Vetter & Poggio, 1996). The automation of a pixelwise correspondence finding procedure using a gradient-based optical flow algorithm (Bergen & Hingorani, 1990) has been one promising approach (Beymer, Shashua & Poggio, 1993; Vetter & Poggio, 1996). For each pixel in one image, these algorithms compute the spatial distance and direction to the corresponding pixel in the other image. The gradient-based optical flow algorithm, as applied to the problem of pixelwise correspondence for human faces, is well-described elsewhere (see especially Beymer & Poggio, 1996; and Vetter, 1996),[4] and so we describe only the basics of the computational procedure here, placing greater emphasis on the principles of what is being computed and why.

For present purposes, computing the correspondence between three-dimensional models of human heads is a straightforward extension of the approaches used successfully with images. Whereas in the two-dimensional case, pixels are matched, in the three-dimensional case, radius values are matched. The adaptation of the algorithm to the three-dimensional head models is computationally transparent due to the fact that the cylindrical representation of a head model, as described earlier, is perfectly analogous to an image. Specifically, the head structure can be represented in the same format as an image, where the 512×512 matrix of radii replaces the image intensities. The coordinates (i.e., x, y for a pixel value in an im-

[2] These numbers vary widely due to the different uses to which they have been put. For example, high quality human face morphs generally require over 100 points (e.g., cf., Burt & Perrett (1995) who use 208 points), whereas, in psychological studies, often the important point is to separate the contributions of shape and image-based codes, and less points are adequate for the task (e.g., cf., Hancock, Burton & Bruce (1996) who use 34 points).

[3] In practice, these easily defined points are often supplemented by set numbers of intervening points. For example, the human operator may decide that 6 equally spaced points between the left and right corner of the eye should define each eyelid.

[4] This latter reference addresses both the optic flow application and also the problem of dealing with non-matchable surface markings, e.g. moles, a critical shortcoming of pure optic flow approach.

age) of a particular radius are given by the angle α (the position on the circle of the cylinder) and the z-coordinate (the height on the cylinder).

The basic principal is to match each individual face to the average face. First, the radii coding corresponding facial surface points between each individual face and the average face are located. Once matched, the result is a representation of the spatial distance and directional offset of a face's "features" *relative* to the "features" of the average face. This is simply a very high resolution representation of how the three-dimensional structure of a particular head differs from the average head. In the motion literature, this representation is referred to as an "optic flow field" — here we will refer to it as a "correspondence field". With this representation, caricatures can be made using completely corresponded three-dimensional faces.

2.2.1 Computing the average head

. In this correspondence-based representation, a boot-strapping method is required to compute the average of the heads. Specifically, this average was obtained using an iterative procedure by which all heads were first set in correspondence with an arbitrary head. "Successful matches" among these were selected and an average of these was computed in the standard way. Successful matches were defined as those for which the correspondence algorithm made no obvious mismatch mistakes. These failures were easy to detect by eye, and ususally involved problems like matching the corners of the eye to the eyebrow corners, or matching the mouth corners to dimples. The average of the successful matches then served as the reference face to which all heads were once again matched and an average of these matches was computed. The process was repeated until the obtained average was stable. This occurred after three iterations of the procedure.

2.2.2 Correspondence algorithm definition.

The basic principal of algorithm we used can be explained as follows. The algorithm is designed to compare scalar functions of two variables, which in our case, represent the surface of the two head models to compare. Each scalar function $r(\alpha, z)$ is given by the radii r parameterized through the angle α (the position on the circle of the cylinder) and the z-coordinate (the height on the cylinder). The algorithm can then be applied to the "image" given by the radius values, considered as a function of α and z, in the same way as it has been applied to the image intensities at the x and y positions in the two-dimensional image case.

To compare two heads in this representation, we used a coarse-to-fine gradient-based procedure (Adelson and Bergen, 1986), following an implementation described in detail in Bergen and Hingorani (1990). Beginning with the lowest level of a resolution pyramid, for every vertex (α, z) in the model, the error term $E = ((\delta r_1/\delta\alpha)\Delta\alpha + (\delta r_1/\delta z)\Delta z - \Delta r_{1,2})^2$ was minimized for $\Delta\alpha$ and Δz. r_1 denotes the radius value at (α, z) on the first model and $\Delta r_{1,2}$ stands for the difference of the radii in the two models. This error term is nothing more than a linearization of the surface, which is assumed to equal the difference of the radii values at a vertex, multiplied by the correct spatial displacement. In more intuitive terms, the error term penalises large spatial displacements, but does so as a function of the quality of the local surface gradient information. Thus, as for all optic flow problems, the regions of the image that are most useful for matching are those for which the image intensities, or head surface values, are changing, i.e. places where the spatial derivatives are non-zero.

The resulting vector field $(\Delta\alpha, \Delta z)$ was then smoothed and the procedure was iterated through all levels of the resolution pyramid. As a result we obtained the correspondence field capturing the spatial displacement for each point between the two head models.

This algorithm yielded successful correspondence for the internal parts of nearly all of the head models. Correspondence errors at the edges of the head, and in the ear regions occurred most frequently. Approximately 15 of the 100 head models showed these kinds of errors and were eliminated from consideration in the psychological experiment. This elimination was done largely by eye, because correspondence errors produced noticeable distortions of the head surfaces, and hence were very easy to detect.

In summary, the head structure data provide an object-centered, view-independent representation of faces that can be thought of as the deformation of the three-dimensional head sample points from their *corresponding* sample points on the average head.

2.3 Description of the Caricature Algorithm

The caricaturing was performed after transforming the heads from cylindrical coordinates, on which all correspondence computations were made, into standard Cartesian coordinates, X, Y

and Z. Each face was represented as a vector containing the deformations or deviations of each head sample point, ΔX, ΔY, ΔZ, from its corresponding point in the average head. Caricatures of varying degrees were made by multiplying this vector by a scalar and adding it to the average head. In terms of a face space framework (Valentine, 1991), the procedure can be described as follows. With this representation, each face, including the average, can be thought of as a point in high dimensional space. Each element of the vector contains the face's coordinate on a particular "feature" axis. The caricature algorithm operates simply by drawing a line between the average face and the face to be caricatured. The caricatured face is made by creating a face at a new position along the line, but at a greater distance from the average. An anti-caricature is made in the same way, but at a lesser distance from the average. This new face is created with feature values, (in this case, the head deformation values, ΔX, ΔY, ΔZ), specified by coordinates of the (anti-) caricature.

In most past studies, caricatures of varying degrees have been made by multiplying a face vector, consisting of the original face "features" by a set factor (e.g., 1.5). We altered this strategy in two ways. First, we modeled our face space with a principal component analysis (PCA) representation, which allows us to work in a much smaller dimensional space than the original feature space. A PCA was applied to the head deformation data from the 100 faces, yielding 99 principal components. Each face was then represented as a 99-dimensional vector consisting of its normalised projection coordinates on each of the principal components. In other words, each element of the vector contained the z-score of the projection coefficient, with respect to the entire set of faces.

The second way in which we altered the procedure concerns our definition of caricature levels – previously defined in terms of the degree of amplification, i.e., the multiplication factor applied to the face vector. As Benson and Perrett (1991) and Stevenage (1995) have noted, individual faces vary *a priori* in their distinctiveness — or, from the point of view of caricature algorithms, in their distance from the average face. Additionally, Benson and Perrett (1994) found an inverse correlation between the rated distinctiveness of a face and the degree of caricature human observers chose as the "best-likeness" of the face. More distinctive faces needed to be less-caricatured than less distinctive faces to achieve the same "best-likeness" preference (cf., also Rhodes & McLean, 1990). Accordingly, we defined particular caricature levels as points equidistant from the average face, using the normalized coefficient vectors. Thus, for a particular caricature level, all faces had the same distance from the mean. This distance is also called the Mahalanobis distance (Duda & Hart, 1973) and is simply a Euclidean distance weighted by the covariance matrix of the sample data. In terms more familiar to psychologists, in the present case, because the face projections on the eigenvectors were expressed in z-score units, the face space was structured so that the average face was at the origin. Individual faces varied in their distance and direction from this average. Thus, face vectors in each caricature level were on the surface of a hypersphere. This equidistance approach seemed to produce more satisfactory results, in terms of the evenness of exaggeration, than applying the same degree of amplification to all faces regardless of their base level distinctiveness.

2.4 Illustration of Caricature Algorithm

In Figures 1 and 2 we present four caricature levels of each of four male and female faces. The mean distance of the heads from the average in this 99-dimensional space was 9.9. The caricature levels were set to be at distances of 6.5 (an anticaricature), 10.00 (an estimate of the original head),[5] 13.5, and 17.0 (both caricatures). For reference, the four male subjects in Figure 1 rows 1-4 were ages 26.0 years, 32.6 years, 23.0 years, 26.8 years, respectively. The four female subjects in Figure 2 rows 1-4 were ages 28.0 years, 37.6 years, 24.3 years, 26 years, respectively. The application of this simple caricature algorithm to the three-dimensional data from human heads seemed to us to have affected the apparent age of the head. We wished to confirm this observation more formally with a psychological experiment.

3 Experiment : Facial Age Estimates

3.1 Observers. Ten volunteers from the University of Texas at Dallas community participated in the experiment. Some of the participants were

[5] It should be noted that due to the fixed distance measure of distinctiveness, and the overall face distance mean of 9.9, this alteration can produce slight (anti-)caricaturing for any given face. The distances used for the anticaricature and caricature manipulations, however, were very much larger than the small offset between the veridical face distance mean and the distance of individual faces.

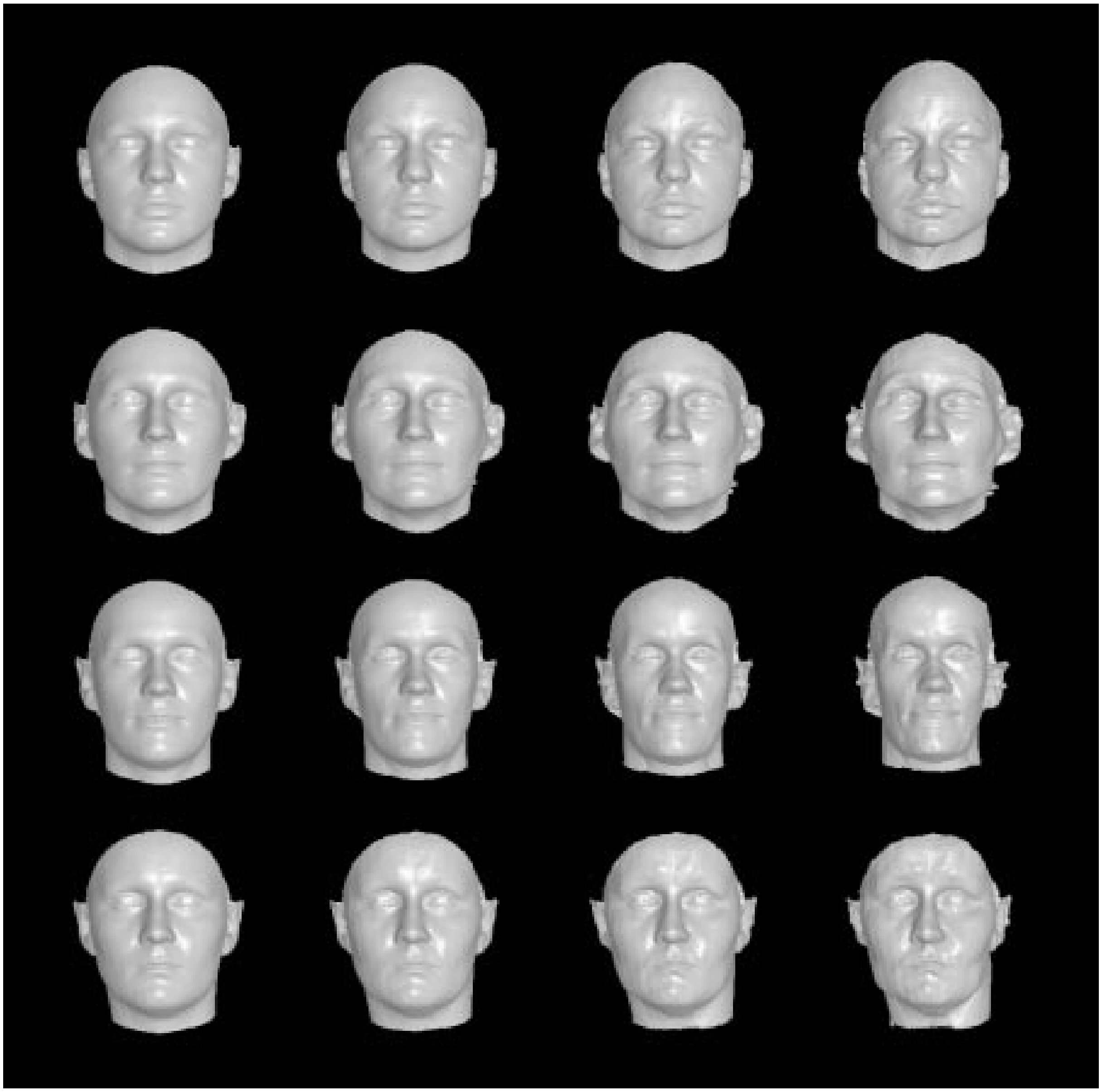

Figure 1: Four caricature levels (columns) of four female heads (rows). Caricature levels from left to right are made at distances of 6.5 (anti-caricature), 10 (approximately veridical), 13.5 and 17 (both caricatures).

.

students who were compensated with a research credit in a core course in the psychology program.

3.2 Apparatus. All experimental events were controlled by a Macintosh computer programmed with PsyScope (Cohen, McWhinney, Flatt, Provost, 1993).

3.3 Stimuli. Of the approximately 85 usable heads, 30 male and 30 female heads (pseudo-randomly chosen) were caricatured to 4 distances from the mean (6.5, 10.0, 13.5, and 17) and were rendered at a viewing angle of 30 degrees. We divided these faces into two groups (15 male and 15 female faces) and presented the different sets of faces to separate groups of observers. This separation was done so that the experiment would not be too time-consuming for the observers. The face sets were counterbalanced across the observers with five observers viewing each set of faces.

3.4 Procedure. Observers were asked simply to make their "best guess" of the age of the faces.

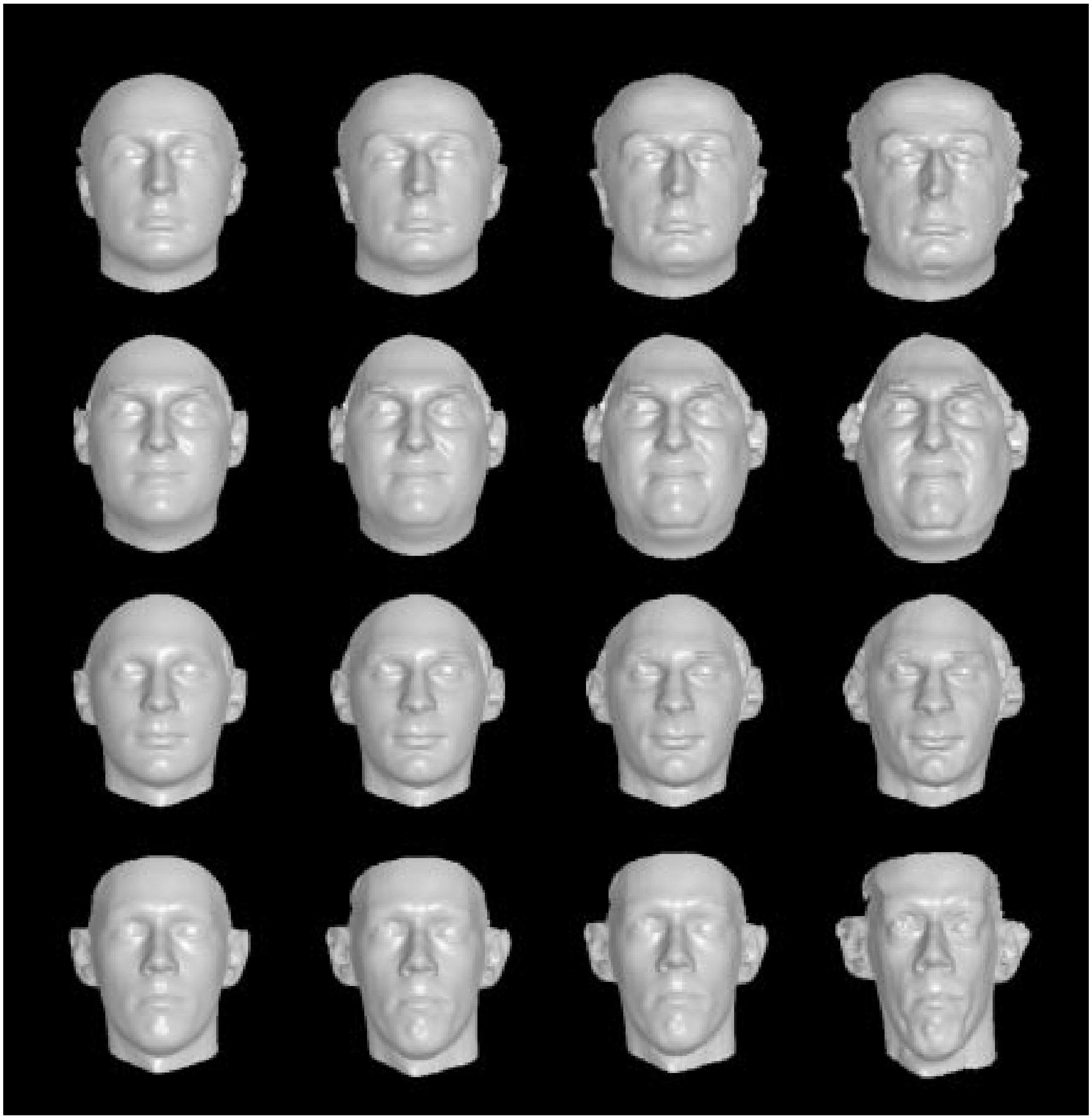

Figure 2: Four caricature levels (columns) of four male heads (rows). Caricature levels from left to right are made at distances of 6.5 (anti-caricature), 10 (approximately veridical), 13.5 and 17 (both caricatures).

.

Faces were presented one at a time on a computer screen, and remained visible until the observer typed in an estimated age. Face order was randomized for each observer.

3.5 Results. Age estimates were compared to the actual ages for each face, and the mean error of the age estimates was computed for each observer in each caricature distortion level. We refer to this measure as the *age error*. A few instances of "typos" by the observers, less than 6 trials across the 10 observers, were eliminated from consideration. These typically contained control characters between the numbers or estimates over 100 years, etc., that made it impossible to determine the age estimate intended by the observer for that trial. The age error data were submitted to a two-factor repeated measures analysis of variance, with caricature level (within subjects) and counterbalance face group (between subjects) as independent variables. No effect of the counterbalance group was found, $F(1,8) = 1.52$, $p > .25$. The caricature level data were remarkably consistent with all 10 observers yielding the same rank order pattern of age estimate errors from least to most caricatured. Not surprisingly, a highly significant main effect of caricature level was found $F(3,27) = 82.17$; $p < .0001$.

The mean age estimate error for each caricature level is displayed in Figure 3. From these data, two points are worth noting. First, the data indicate a baseline error in the accuracy of the age estimates with the three-dimensional representations, such that observers overestimated the age of the veridical faces by about 6 years. Second, perceived face age increased nearly linearly as a function of caricature level. Even considering the 6 year baseline error, for the 17-level caricature distortion group, all 10 observers used ages in the fifties, eight of the ten observers estimated the ages of some faces into be in their sixties, and five observers estimated some faces to be in their seventies and eighties.

It is perhaps worth mentioning that the baseline overestimate of the face age, which is significant in all but the second caricature level condition (see error bars on Figure 3, which are smaller than 6 years in all but the last condition), could have occurred for a number of reasons. The simplest of these is that something about the display of the pure three-dimensional data for faces in this age range makes them look a bit older. Perhaps the image intensities and color information add cues to youthfulness that are important. This is an interesting question for further study.

4 Discussion

4.1 Three-dimensional Caricatures and Facial Age.

Although we can say rather precisely that we manipulated the distinctiveness of the three-dimensional information in faces via a standard caricature algorithm, mapping this description onto more intuitively understandable facial cues is much more difficult. The three-dimensional structure captured by the laser scan data depends on the underlying structure of the skull, on the attached muscle and tissue structures, and finally on the properties of the skin surface. Simple common observation reveals that nearly all of these change with increasing age. More formally, a preliminary list of the kinds of changes that occur to the face with ageing includes: a.) a relaxation in the elasticity of the skin and also its texture, resulting in increased wrinkling; b.) the loss of facial muscle tone; c.) the softening of the cartilage which results in less structural support for the nose and ears making them appear longer; d.) and a decrease in the fatty deposits in the face, resulting in a more bony appearance (Behrents, 1985; Hamra, 1995).

In terms of the skull structure itself, very few longitudinal anthropometric or cephalometric studies have been undertaken on the aged. An excellent summary of the extant literature is reported by Behrents (1985). An important caution in applying this literature to the problem at hand is that cross-sectional studies do not distinguish age-related changes from individual changes, secular trends, or even differences due to possible pathology (e.g., tooth loss, cf., Rogers, 1982). Additionally, the available cross-sectional studies have not used age range cohorts that are consistent enough to compare facial ageing data across the age ranges we are considering (cf., Behrents, 1985).

With that caution in mind, we attempt here to sketch out some of the cues that may have lead to the perceived facial age increases we observed with increased caricaturing. At the most general level, it is worth noting that the three-dimensional caricature algorithm we have applied operates *without* reference to normative data on older faces. This stands in contrast to a recent study by Burt and Perrett (1995), in which faces were aged synthetically by using normative data from older faces. We discuss the study of Burt and Perrett in detail shortly. For present purposes, the primary point we wish to make is that the ageing cues captured in our study are based only on the distinctive aspects of individuals from a relatively small age-range cohort — primarily people in their twenties and thirties.

In this context, of the facial age cues noted above, two seem to be likely candidates for the three-dimensional caricature manipulation we have implemented. First, we think the caricature algorithm does a good job of capturing facial wrinkling as a cue to ageing. According to Behrents (1985), wrinkles are found on the forehead, on the outer corners of the eyes as "crow's feet", between the eyebrows as vertical creases, and as grooves running from the base of the nose past the corners of the mouth to the chin. While wrinkling is indeed a normal phenomenon of facial ageing, the exact placement and shape of wrinkles on a particular face is more or less specific to the particular face (i.e., the exact placement of a smile crease or dimple). Because the three-dimensional caricature algorithm amplifies the contrast of fa-

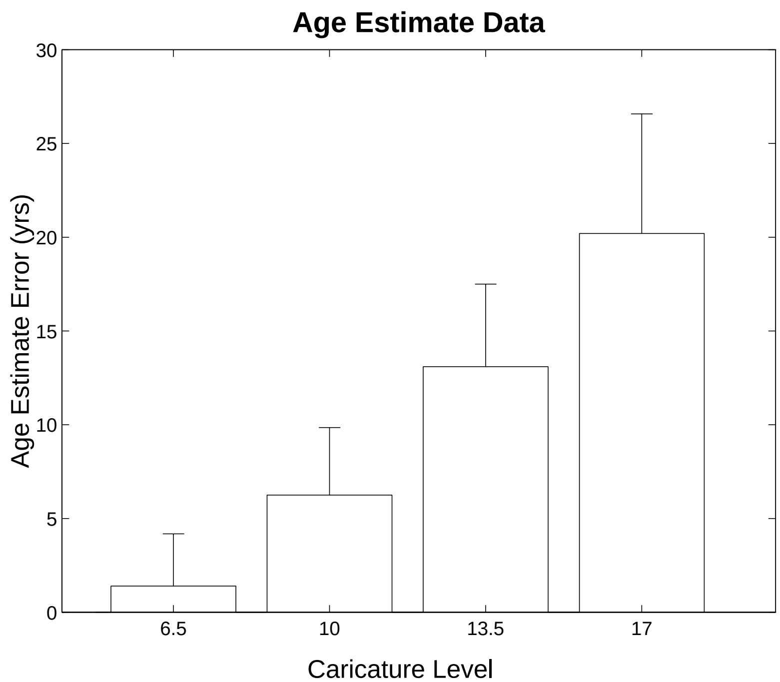

Figure 3: Age estimate errors across the cariacture levels indicate an approximately 6 year over-estimate baseline error, but also a strong and consistent effect of caricature level on the apparent age of the face. Error bars indicate standard errors.

cial creases already present in individual faces, the placement and shape of these amplified creases appear as natural wrinkles appropriate for particular faces. Thus, creases that are barely noticeable in a young face, become more pronounced in the caricatured faces, making these faces appear older. Good examples of these cues can be seen in the fouth male face in Figure 1 and third female face in Figure 2.

A second more global ageing cue that may have been tapped in the caricature algorithm has to do with increasing the prominence of the bony structure of the face (Bartlett, Grossman, & Whittaker, 1972; Behrents, 1985; Hamra, 1995). This is thought to be caused by the loss of fatty tissue between the skull and skin as we age. Indeed, plastic surgeons have designed surgical procedures to mask this ageing effect by repositioning cheek fat to create a fuller appearance around the eyes (Hamra, 1995). Distinctive aspects of the bony structure of individual faces that are visible in these scans (i.e., prominent cheek bones, large or hooked noses) would be further exaggerated in the caricature. The prominent brow of the first male face in Figure 1 and high cheek bones of the second female face of Figure 2 are good examples.

A related phenomenon is the increasing prominence of "jowls" with age. Jowls refer to the flesh under the lower jaw, which when plump or flaccid becomes a noticeable facial feature. In short, these become more accentuated with age due to the loss of skin elasticity and facial muscle tone. The presence of slight jowls is likely to be a distinctive three-dimensional feature of the faces that contain them. With the application of the caricature procedure, they become even more noticeable. Several faces in our data base had small jowls which seemed to have grown with caricaturing. Additionally, while to our knowledge there are no anthropometric studies documenting systematic age-related weight changes that affect the appearance of the face, the caricature algorithm seems to have implemented such changes. Specifically, individuals with relatively thin faces, tended in the caricatures to become even more so — likewise, individuals with slightly chubby faces, seemed also to become more so. This cue seemed to us to be a factor in the age appearance of the face.

Finally, although we believe that this algorithm has captured a number of three-dimensional cues to age, it is clear that these comprise only a subset of the available cues. For example, where skull structure is concerned, there are well-documented, systematic head shape changes with age. These include changes to the angle of the mandible or jaw bone (increased jutting with old age), a reduction in facial height, mainly in the maxilla and mandible, a slight increase in facial width, and a slight increase in facial depth (Bartlett et al. 1992). The present algorithm operated with no reference (explicit or implicit) to normative data on older faces, and hence could not have captured these cues. Interestingly, previous psychological studies have concentrated more on these kinds of normative cues.

4.1.1 In Perspective.

By comparison to the diversity of cues available for determining the age of a face, the perceptual salience of only a few selected cues has been investigated systematically in psychological studies. Studies of the perception of facial age can be categorized according to the kinds of stimuli used and the age range considered. Many studies have made use of line sketches of the outline of the head profile or relatively simple three-dimensional head models (Mark & Todd, 1985; Mark, Todd & Shaw, 1981; Pittenger & Shaw, 1975) and have been concerned with the age changes occuring between infancy and adulthood. These studies have focused almost entirely on the importance of a particular head shape cue, modeled by a geometric transformation called "cardioidal" strain.[6] This cue has been demonstrated to be important for the perception of age between infancy and adulthood.

A smaller number of studies have looked at photographic quality images and have considered age ranges within the span of adulthood (Burt & Perrett, 1995; George & Hole, 1995; Kowner, 1996). Burt and Perrett (1995) and George and Hole (1995) have provided data indicating that human observers are quite accurate at perceiving the age of faces from unmanipulated photographs over the age ranges of 20-60, and 5-70 years, respectively. Interestingly, Kowner (1996) found that face symmetry may be a factor in the perceived attractiveness of older faces — more symmetric older faces were found to be less attractive than less symmetric older faces. This finding did not extend to younger faces.

Directly relevant to the present study, Burt and Perrett (1995) synthetically "aged" faces by applying an algorithm based on standard caricaturing to the two-dimensional shape and colour information in faces. The principle behind their age manipulation was to move individual younger faces, represented as points in a face space, in the direction of the average of older faces — thus, referencing normative data on older faces to make the transformation. Using a data base of frontal views of 147 Caucasian male faces between the ages of 20 and 62 years, Burt and Perrett (1995) divided the faces into 5-year age groups (e.g., 20-24; 25-29, etc.) and obtained separate averages of the two-dimensional shape and colour information, as follows. First, the two-dimensional shape information was defined using 208 manually placed feature points (pixels). Next the locations of these "feature" points or pixels were averaged across the faces in the age group. Average colour information in each age group was obtained by standard two-dimensional image warping (morphing) techniques applied to morph each face to the group average. Average colours were computed by digitally averaging the corresponding pixel colour values.

This process yields separate representations of the two-dimensional shape (i.e., configural) information in a face and the colour information. Again, using a standard caricature approach, Burt and Perrett (1995) attempted to age the faces by using the two-dimensional shape information, the colour information, and both together. An age transformation for shape and colours was estimated by taking the difference between the 50-54 year average and the 25-29 year average and adding it to individual faces. The caricature was applied to 6 faces (two aged 27; two aged 40 and two aged 53). Burt and Perrett (1995) found that the two-dimensional manipulations increased the age estimates on these faces by an average of 4.4 years; the colour manipulation increased the age estimates by an average of 5.7 years; and both the shape and colour manipulation combined increased the age estimates by an average of 8.5 years.

The data of Burt and Perrett (1995) indicate that perceptually salient cues to facial age can be captured in *normative* data on the two-dimensional shape and colour of faces of different ages. The normative aspect of Burt and Perrett's caricatures makes an interesting contrast to the present caricatures, for which the faces were aged by an algorithm that did not reference normative data on older faces. Thus, one might make a distinction between age information that is specific to individuals and age information that can be approximated by general information. The work of

[6] See also Bruce, Burton, Doyle & Dench (1989) for an examination of this cue with minimally-marked laser scan data from views other than the profile.

Burt and Perrett (1995) draws on the more general normative cues, whereas the present study draws on the more individual face-specific aspects of age.

The present result may also be relevant for understanding some well-known findings concerning the attractiveness of averaged faces (Langlois and Roggman, 1990). Langlois and Roggman found that composite faces, created by arithmetically averaging the images of several faces, were judged to be more attractive than almost any single male or female face. Although there is some controversy concerning the interpretation of these results (cf., Alley & Cunningham, 1991; Langlois, Roggman, & Mussleman, 1994; Langlois, Roggman, Mussleman, & Acton, 1991; Perrett, May, & Yoshikawa, 1994; Pittenger, 1991), the findings fit well with older results indicating that human observer ratings of facial attractiveness vary inversely with face recognizability (Light, Hollander, & Kayra-Stuart, 1981).[7] These older results indicate, by implication, that attractive faces may in some ways be "average". In the present study, we defined distinctiveness as the distance to the average face and find that faces close to this average are perceived to be younger than those farther from the average. The important question that arises concerns how the concepts of attractivenes, distinctiveness, and the perception of facial age interrelate. One might speculate that one component of attractiveness may be a youthful appearance, which in the context of a three-dimensional representation would indicate relatively undifferentiated or smooth features. A more definitive answer to these questions, however, would require additional data on the perceived attractiveness and distinctiveness of the caricatures. We are presently in the process of collecting these data.

The rhetorical question that remains is, "As we get older, do we get more distinct?" In one sense, the answer to this question is "yes". As the caricatures illustrate, making the three-dimensional information in faces more distinct or exaggerated captures at least some cues related to the natural process of facial ageing. In a stricter sense, in the present study, we defined our subset of faces to include only narrow range of face ages. Defined with respect to a broader age cohort, it is not entirely clear that older faces would be "more distinct", i.e., in our operational definition, farther from a more age-uniform mean. Although, given the tendency of the placement and shape of wrinkles, etc., to be rather specific to individuals, it may in fact be the case that this result would hold across a more age diverse cohort. A related question concerns the recognisability of these caricatured and aged faces. Prelimary data we have collected on this question indicates that the cariactured faces may indeed be more recognisable than the veridical and anti-caricature faces, though a definitive answer to this question awaits the final results of this study.

In summary, the application of a standard caricature algorithm to a three dimensional representation of faces affected the apparent age of the face. The primary function of a caricature algorithm is to (de-)amplify the "distinctive" information in individual faces. The generic application of a caricaturing procedure requires a commitment to a relatively precise system for representing and quantifying the information in faces. The present exercise indicates that this representation decision has important perceptual consequences. By comparison to other studies (Rhodes et al., 1987; Benson & Perrett, 1991), our results suggest that distinctiveness specified with respect to three-dimensional information in a face may underlie some of the basic information we use to estimate the age of a face. In other words, the present study indicates that there may be an additional perceptual dimension of human facial distinctiveness that is related to the perception of facial age.

[7] Though see also O'Toole, Deffenbacher, Valentin, McKee, Huff, & Abdi, in press; and Shepherd & Ellis, 1973; both of which found the result only for male faces; Light et al. (1981) used only male faces.

References

Adelson E H, Bergen J R 1986 "The extraction of spatiotemporal energy in human and machine vision" In *Proc. IEEE Workshop on Visual Motion* Carlston pp 151-156

Alley T R, Cunningham M R, 1991 "Averaged faces are attractive, but very attractive faces are not average" *Psychological Science* **2** 123-125

Bergen J R, and Hingorani R, 1990 "Hierarchical motion-based frame rate conversion" *Technical Report, David Sarnoff Research Center* Princeton, NJ

Bartlett S P, Grossman R, Whitaker L A, 1992 "Age-realted changes of the craniofacial skeleton : an anthropometric analysis" *Plastic and Reconstructive Surgery* **90** 592-600

Behrents R G, 1985 *Growth in the Aging Craniofacial Skeleton* Monograph 17 Craniofacial Growth Series, Center for Human Growth and Development, The University of Michigan, Ann Arbor Michigan.

Beier T, Neely S, 1992 "Feature-based image metamorphosis" in *1992 SIGGRAPH Proceedings* (Reading, MA: Addison Wesley) pp 35-42

Benson P J, Perrett D I, 1991 "Perception and recognition of photographic quality caricatures : implications for the recognition of natural images" *European Journal of Cognitive Psychology* **3** 105-135

Benson P J, Perrett D I, 1994 "Visual processing of facial distinctiveness" *Perception* **23** 75-93

Bergen J R, Anandan P, Hanna K J, Hingorani R, 1992 "Hierarchical model-based motion estimation" In *Proceedings of the European Conference on Computer Vision* (Santa Margherita Ligure, Italy) pp 237-252

Beymer D, Poggio T, 1996 "Image representations for visual learning" *Science* **272** 1905 - 1909

Beymer D, Shashua A, Poggio T, 1993 "Example-based image analysis and synthesis" *A.I. Memo No.* **431**

Biederman I, 1987 "Recognition by components: A theory of human image understanding" *Psychological Review* **94** 115-147

Brennan S E, "The caricature generator" *Leonardo* **18** 170-178

Bruce V, 1989 *Recognising Faces* (London : Lawerence Erlbaum).

Bruce V, Young A W, 1986 "Understanding face recognition" *British Journal of Psychology* **77** 305-327

Bruce V, Burton M, Doyle T, Dench N, 1989 "Further experiments on the perception of growth in three dimensions" *Perception & Psychophysics* **46** 528-536

Bülthoff H H, Edelman S., 1992 "Psychophysical support for a two-dimensional view interpolation theory of object recognition" *Proc. Natl. Acad. Sci. USA* **89** 60-64

Burt D M, Perrett D I, 1995 "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information" *Proc. R. Soc. Lond. B* **259** 137-143

Cohen J D, McWhinney B, Flatt M, Provost J 1993 "PsyScope : A new graphic interactive environment for designing psychology experiments" *Behavior Research Methods, Instruments & Computers* **25** 257-271

Cootes T F, Taylor C F, Cooper D H, J. Graham J, 1995 "Active shape models : Their training and application" *Computer Vision and Image Understanding* **61** 38-59

Craw I, Cameron P, 1991 "Parameterizing images for recognition and reconstruction" In *Proc. British Machine Vision Conference* Ed Peter Mowforth (London: Springer Verlag) pp 367-370

Duda R O, Hart P E, 1973 *Pattern Classification and Scene Analysis* (New York : John Wiley & Sons)

Enlow D H 1982 *Handbook of Facial Growth* 2nd edition (Philadelphia, PA: W B Saunders Company)

George P A, Hole G J, 1995 "Factors affecting the accuracy of age estimates of unfamiliar faces" *Perception* **24** 1059-1073

Goldstein A G, Chance J, 1980 "Memory for faces and schema theory" *Journal of Psychology* **105** 47-59

Hancock P J B, Burton A M, & Bruce V, 1996 "Face processing : Human perception and principal components analysis" *Memory & Cognition* **24** 26-40

Hamra S T, 1995 "Arcus marginalis release and orbital fat preservation in midface rejuvenation" *Plastic and Reconstructive Surgery* **96** 354-362

Kowner R, 1996 "Facial asymmetry and attractiveness judgment in development" *Journal of Experimental Psychology: Human Perception and Performance* **22** 662-675

Langlois J H, Roggman L A, Mussleman L, Acton S, 1991 "A picture is worth a thousand words: A reply to "On the difficulty of averaging faces" *Psychological Science 2* 354-357

Langlois J H, Roggman L A, Mussleman L, 1994 "What is average and what is not average about attractive faces?" *Psychological Science* **5** 214-220

Lanitis A, Taylor C J, Cootes T F, Ahmad T, 1995, "Automatic Interpretation of Human Faces and Hand Gestures Using Flexible Models" In *Proc. International Workshop on Face and Gesture Recognition*, Ed M Bichsel (Zurich, Switzerland : University of Zurich Press) 98-103

Light L L, Kayra-Stuart F, Hollander S 1979 "Recognition memory for typical and unusal faces" *Journal of Experimental : Human Perception and Performance* **5** 212-228

Logothetis N K, Pauls J, Poggio T, 1995 "Shape recognition in the inferior temporal cortex of monkeys" *Current Biology* **5** 552-563

Mark L S, Pittenger J B, Hines H, Carello C, Shaw R E, Todd J T, 1080 "Wrinkling and head shapes as coordinated sources of age-level information" *Perception & Psychophysics* **27** 245-256

Mark L S, Todd J T, 1983 "The perception of growth in three dimensions" *Perception & Psychophysics* **33** 193-196

Mark L S, Todd J T, 1985 "Describing perceptual information about human growth in terms of geometric invariants" *Perception & Psychophysics* **37** 245-256

Marr D, Poggio T, 1976 "Cooperative computation of stereo disparity" *Science* **194** 283- 287

Mauro R, Kubovy M, 1992 "Caricature and face recognition" *Memory & Cognition* **20** 433-440

Morton J, Johnson M H, 1991 "CONSPEC and CONLERN: A two-process theory of infant face recognition" *Psychological Review 98* 164-181

O'Toole A J, Edelman S, 1996, "Face distinctiveness in recognition across viewpoint change: An analysis of the statistical structure of face spaces" *Proceedings of the International Workshop on Automatic Face and Gesture Recognition* IEEE Computer Society Press.

O'Toole A J, Deffenbacher K A, Valentin D, McKee K, Huff D, Abdi H, in press "The Perception of Face Gender: The Role of Stimulus Structure in Recognition and Classification" *Memory & Cognition*

Perrett D I, Hietanen J K, Oram M W, Benson P J, 1992 "Organisation and functions of cells responsive to faces in the temporal cortex" *Philosophical Transactions of the Royal Society London B* **335** 23-30

Perrett D I, May K A, Yoshikawa S, 1994 "Facial shape and judgements of female attractiveness" *Nature* **368** 239-242

Pittenger J B, Shaw R E, 1975 "Ageing faces as visceral-elastic events: Implications for a theory of nonrigid shape perception" *Journal of Experimental Psychology: Human Perception and Performance* **1** 374-482

Pittenger J B, Shaw R E, 1979 "Perceptual information for the the age level of faces as a higher-order invariant of growth" *Journal of Experimental Psychology: Human Perception and Performance* **5** 478-493

Pittenger J B, 1991 "On the difficulty of averaging faces" *Psychological Science* **2** 351-353

Rhodes G, Brennan S E, & Carey S, 1987 "Identification and ratings of caricatures : Implications for mental representations of faces" *Cognitive Psychology* **19** 473-497

Rhodes G, McLean I G, 1990 "Distinctiveness and expertise effects with homogeneous stimuli: Towards a model of configural coding" *Perception* **19** 773-794

Rhodes G, Tremewan T, 1994 "Understanding face recognition: Caricature effects, inversion, and the homogeneity problem" *Visual Cognition* **1** 275-311

Rhodes G, & Tremewan T, 1996 "Averageness, exaggeration, and facial attractiveness" *Psychological Science* **7** 105-110

Rogers S L 1982 *The Aging Skeleton: Aspects of Human Bone Involution* (Springfield IL: C C Thomas)

Shepherd J W, Ellis H D, 1973 "The effect of attractiveness on recognition memory for faces" *American Journal of Psychology* **86** 627-633

Stevenage S V, 1995 "Can caricatures really produce distinctiveness effects?" *British Journal of Psychology* **86** 127-146

Ullman S, 1981, "Analysis of visual motion by biological and computer systems" *Computer* **14** 57-69

Valentine T, 1991 "A unified account of the effects of distinctiveness, inversion, and race in face recognition" *Quarterly Journal of Experimental Psychology* **43A** 161-204

Vetter T, 1996 "Learning novel views to a single face image" *Proc. of the 2nd Int. Conf. on Automatic Face and Gesture Recognition* (Los Alamitos, CA : IEEE Comp. Soc. Press) pp 22-27

Vetter T, Poggio T, 1996, "Image Synthesis from a Single Example Image" In *Computer Vision – ECCV'96* Eds B Buxton and R Cippola (Cambridge, UK: Springer Verlag), *Lecture Notes in Computer Science* **1065** 652-659

ACKNOWLEGDEMENTS.

A.O. gratefully acknowledges support from NIMH grant 1R29MH5176501A1, the Alexander von Humboldt Foundation, and Texas Instruments Inc. Thanks are due also to Niko Troje for the stimulus creation and processing, and to Ken Deffenbacher, Shimon Edelman, Jennifer Jeffres for helpful comments on this manuscript.