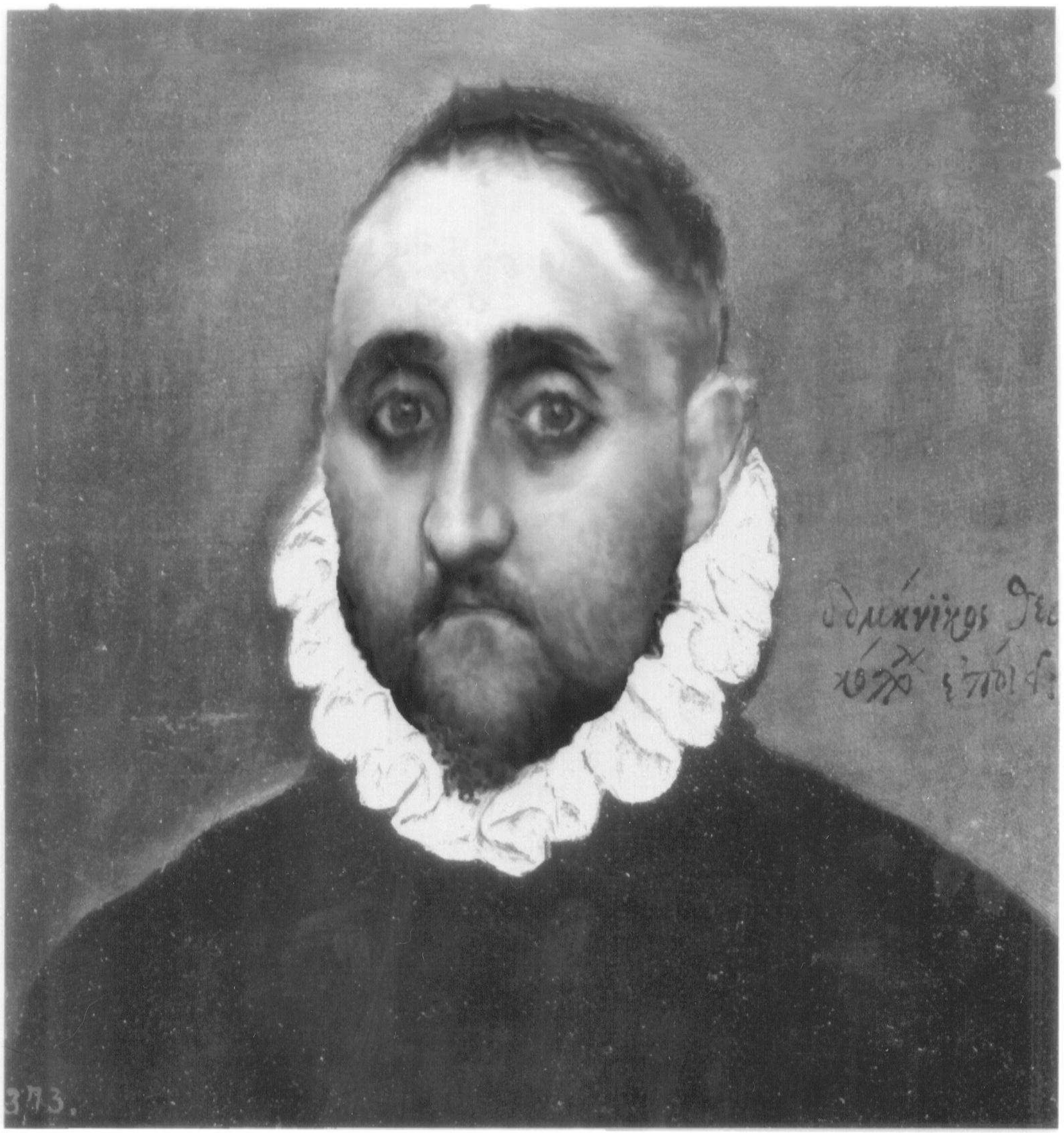
# Computer Graphic Control over Human Face and Head Appearance, Genetic Optimisation of Perceptual Characteristics.

Duncan Andrew Rowland

Doctor of Philosophy

5th January 1998


δομήνικος θεο
κόπ ἐποίει
373.

This thesis also available on the web at:

http://www.mindlab.msu.edu/documents/Thesis_Rowland/title.htm

Declarations:

I, Duncan Andrew Rowland, hereby certify that this thesis, which is approximately 29,000 words in length, has been written by me, that it is the record of work carried out by me and that it had not been submitted in any previous application for a higher degree.

Date……….. Signature of candidate………..

I was admitted as a research student in October 1992 and as a candidate for the degree of PhD in October 1994; the higher study for which this is a record was carried out in the University of St Andrews between 1992 and 1998.

Date……….. Signature of candidate………..

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date……….. Signature of supervisor………..

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulation of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date……….. Signature of candidate………..

This thesis is dedicated to Pokey and the possibilities of the future.

# Acknowledgements:

In no particular order I would like to gratefully acknowledge the help of the following people who have aided me in my studies and in the production of this thesis. My supervisor, Dr David Perrett, provided an enormous amount of encouragement, assistance and freedom - is it has been my joy and privilege to work with him. Keith May helped design the shape transformations. Isono-san and Oda-san at ATR designed the GA interface in experiment two. Stacy and Lisa Kahler proof read. Stuart McMullen and Mathew McEachran assisted in running experiment one. Dr Peter Foldiak and Dr Thomas Vetter conducted a highly constructive viva. Dr Andy Young and Dr Andy Calder wrote many papers and helped my publication record immensely. And finally, Dr John McGaw who never lost interest in shampoo bottles.

I'd also like to take this opportunity to thank: Stacy, for many things, but especially for her love, commitment and seemingly unending ability to accept my irrational behaviour; Pokey, for providing the incentive to get this thing written; and my parents and family whose love and constancy gave me the confidence to grow. I'd especially like to thank all those people who have seen and nurtured a potential in me that I could not see myself. Specifically the following who made my schooling fun: Mr Childs for fuelling curiosity; Don Gee for his tutelage and encouragement; and David Perrett for showing me how potentials can be achieved. I have had a great time in St Andrews, met some remarkable people, and made some very good friends. There are too many of you to thank individually, you know who you are, all the people I've ever got drunk with, all the members of bands I've played with and all the perception lab people, past and present, who made coming to work such a joy. I thank you all.

**Cover Plate:** *The author a la El Greco.* The cover shows a photograph of the author's face transformed in shape and colour using techniques described in the facial transformation section of this thesis. Two averages were created, one from photographs of male Caucasians, and one from the photographs of portraits painted by El Greco. The differences in shape and colour between these two prototypes were then 'added' to the face to induce a look as if painted by El Greco. Placing the resulting facial image into a background from an El Greco painting created the final image[1].

---

[1] The background painting is 'An Unknown Man' *painted 1584-94 (Cossio); 1594-1600 (Mayer), The Prado Museum, Madrid* and was scanned from El Greco by Leo Bronstein, 1966, Thames and Hudson, ISBN 0-500-08046-1).

# Contents

| Page | Section | |
|---|---|---|
| 1 | 1. | Abstract. |
| 2 | 2. | Techniques for the Parameterisation and Manipulation of Faces. |
| 2 | 2.1 | 2-D Introduction. |
| 2 | 2.2 | Global Facial Parameters. |
| 5 | 2.3 | Image Warping Algorithms. |
| 7 | 2.4 | Choice of Warping Algorithm. |
| 8 | 2.5 | Warping Triangles with Barycentric Co-ordinates. |
| 8 | 2.6 | Sampling, Aliasing and Pixel Interpolation. |
| 11 | 2.7 | Delineation. |
| 12 | 2.8 | Morphing: An example Use of Warping. |
| 14 | 2.9 | Real-time Image Morphing. |
| 16 | 2.10 | Forming Facial Prototypes. |
| 18 | 2.11 | Predictive Colour Addition. |
| 20 | 2.12 | Caricaturing. |
| 22 | 2.13 | Transformation Based on Prototypes. |
| 24 | 2.14 | Discussion. |
| 26 | 2.15 | Conclusion. |
| 26 | 2.16 | 3-D Introduction. |
| 28 | 2.17 | 3-D Representations of Faces. |
| 30 | 2.18 | 3-D Prototypes. |
| 31 | 2.19 | 3-D Caricatures. |
| 32 | 2.20 | 3-D Transformations. |
| 33 | 2.21 | Discussion. |
| 35 | 2.22 | Conclusion. |
| 36-53 | | Plates 1-1 to 1-18 |
| 54 | 3 | Evolution of Bottles Design Based on Consumer Preferences. |
| 54 | 3.1 | Introduction. |
| 55 | 3.2 | A Brief History of Beauty. |
| 57 | 3.3 | Introduction: The Genetic Algorithm. |
| 60 | 3.4 | Aims. |

61 3.5 Genetic Evolution of Bottles Experiment.
61 3.5.1 Preparation of Stimuli.
66 3.5.2 Breeding Algorithm.
68 3.6 Experiment 1.1 - Test for Subject Agreement
68 3.6.1 Subjects.
68 3.6.2 Procedure.
68 3.6.3 Evaluation of the GA and Design of Stimuli Set.
68 3.6.4 Results.
69 3.6.5 Discussion.
69 3.7 Experiment 1.2 - Test for Successful Evolution.
69 3.7.1 Procedure.
70 3.7.2 Results.
70 3.7.3 Discussion.
71 3.8 Experiment 1.3 - Determination of Shape Evolution.
72 3.8.1 Method.
72 3.8.2 Results.
72 3.9 Results: Evolution of Colour.
73 3.10 Experiment 1: Discussion.
74 4 Optimising for Consumer Perception of Specific Marketing Concepts.
74 4.1 Introduction.
74 4.2 Computational Methods.
74 4.2.1 Colour Parameters.
75 4.2.2 Genetic Algorithm: Fitness value.
75 4.2.3 Modal Bottles.
76 4.3 Experiment 2.1.
76 4.3.1 Procedure
76 4.3.2 Generation 0.
77 4.3.3 Results: Consistency of Ratings
77 4.3.4 Independence of Ratings
78 4.3.5 Selection of Concepts to Test Evolution.
79 4.3.6 Generation 1.
79 4.3.7 Generation 2.
80 4.3.8 Generation 3 and Modal Bottles.

81 4.3.9 Results: Analysis of Genes (Colour).
83 4.4 Visualisation of Consumer Preferences.
84 4.5 Results: Generation 3 and Modal Bottles.
84 4.6 Viewing Evolutionary Trends.
87 4.7 Discussion and Conclusions.
90-91 Plates 2-1 and 2-2.
92 5 Evolution of Attractive Faces.
92 5.1 Introduction.
92 5.2 Creating a ‘Face Space’.
93 5.3 Principal Component Analysis.
96 5.4 Mapping PCA onto GA.
96 5.5 Psychological Background.
98 5.6 Hypothesis.
98 5.7 Method.
99 5.8 Stimuli.
100 5.9 Procedure.
100 5.10 Test.
101 5.11 Results.
101 5.12 Analysis of Coefficients.
102 5.13 Discussion.
103 6 Thesis Summary.
106-109 Plates 3-1 to 3-4
110 References.

# 1. **Abstract**

The aims of this thesis are two-fold. The first is to develop computer graphics that allow quantitative manipulation of complex visual stimuli. The second is to show that such techniques have utility in the domain of perceptual psychology. There are three main sections to this thesis. The first section creates methods for performing transformations of facial appearance along particular perceptual dimensions. This work begins with 2-D image manipulations and then extends the general principles to 3-D. Effectiveness of the techniques is illustrated with plates showing transformation in age, gender and identity. The second section uses Genetic Algorithms to control the appearance of 3-D computer graphics objects and investigates methods of evolving objects that embody various consumer concepts. Computer graphic models of shampoo bottles are successfully evolved to satisfy a selection of aesthetic and perceptual characteristics. The final section returns to facial stimuli and extends the Genetic Algorithm approach to investigate aesthetic preference for 3-D facial surfaces. The study shows that individual human subjects can evolve facial surfaces based upon their own attractiveness preferences. The faces evolved are non-average and there is consistency between subjects about preferred characteristics. The three parts of this thesis have different theoretical backgrounds and literature relevant to each topic is therefore reviewed at the start of each section.

# 2. **Techniques for the Parameterisation and Manipulation of Faces.**

## 2.1 2-D Introduction.

This section of the thesis is concerned with the development and implementation of techniques for manipulating facial appearance. A review of existing techniques for manipulating 2-D images of faces is given. These include warping and morphing and how these can be used in the formation of a facial prototype (an average appearance of a group of faces, Galton 1879; Benson and Perrett 1994); a caricature exaggeration of the distinctive shape of individual faces (Brennan 1985; Benson and Perrett 1991a,b) and transformations of face properties (Rowland and Perrett 1995). Finally, it will be shown how each technique can be extended for use with 3-D range data and its associated colour information.

## 2.2 Global Facial Parameters.

A 2-D cardiodal strain transformation can be used to model the changes due to bone growth between the shape of an infant's head and the shape of an adult's head (Shaw 1974). This transformation has successfully been applied to individual head shapes in order to age them without affecting identity (Pittenger and Shaw 1975). It should be highlighted that the cardioid strain mimics bone growth (specifically that of the skull) and an attempt to employ it to alter perceived age outwith the range of ages during which bone growth occurs produces a meaningless transformation. Unsurprisingly, in studies where attempts have been made to use a cardioid strain to manipulate age outside the applicable range, it has been found to be an ambiguous cue Bruce *et al* (1989).

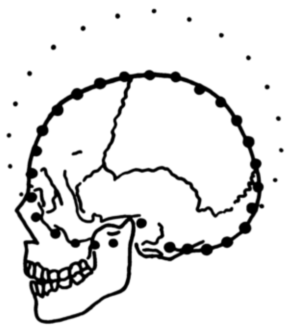


**Figure 2.1.** The figure on the left (taken from Shaw *et al* 1974 and © 1974 Cornell University) shows a regular cardioidal shape (small dots) lying above a human skull. An appropriate transformation maps this shape onto the shape of the skull's profile (large dots).

The cardiodal strain transform is an example of a global parameter transformation, where a ***single*** value can be used to control a perceived attribute (in this case, age). The equations for the cardioid strain are best expressed in a polar co-ordinates system (with the locational origin at the top orb of the ear hole as shown in Fig. 1.1, and horizontal being defined as the Frankfurt horizontal from Pittenger *et al*, 1979). For each point $r, \theta$ (with $r$ representing the distance from the locational origin, and $\theta$ representing angular displacement from rotational origin) the cardioid strain can be expressed as $r' = r(1-\lambda\cos\theta)$ and $\theta' = \theta$. The value of $\lambda$ is the amount of transform to be performed and so is related to time ($\lambda=0$ will produce no change, $\lambda>0$ will predict future growth and $\lambda<0$ will predict previous appearance). This work has been successfully extended to 3-D (notably by Mark and Todd 1983). By the addition of a rotation in a third dimension $(r, \theta, \varphi)$ and representation of the co-ordinate space as a spherical polar space, the equations become $r' = r(1-\lambda\cos\theta\cos\varphi)$, $\theta' = \theta$, $\varphi' = \varphi$.

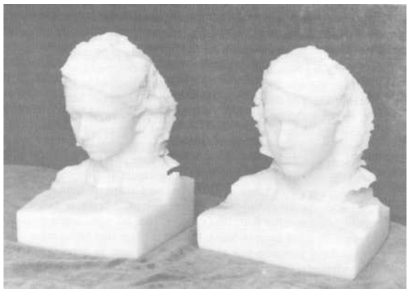


**Figure 2.2.** The figure shows two head models created using scanned data on a computer (Mark and Todd 1983). The bust on the right shows the original shape of a female Caucasian with a mean rated age of 14.5 years (S.D. 1.5). The bust on the left shows the same bust with the 3-D cardioid strain transform applied ($\lambda$=-0.2). In the study this later bust was rated as having a mean rated age of 6.3 years (S.D. 0.9)

It has been stated that global parameters of the type used to specify age may be used by observers in their perception of faces; '… [there is] compelling evidence that perceivers may be sensitive to this kind of global parameter variation between faces' (Bruce 1988). Bruce suggests that evidence for this comes from two main findings. First, the demonstration that the dimensions that account for the majority of the perceived variations between unfamiliar faces are hairstyle, face shape and age (Shepherd 1981). Second, in habituation studies involving babies as subjects, cross category (dimension) sensitivity was found to be higher that within category sensitivity (Fagan and Singer 1979). This suggests that it may be possible to derive transformations that selectively manipulate faces along specific dimensions. Just as growth has been modelled using the cardioidal strain function it may be possible to generate models for other facial parameters (such as age, sex and identity). This would entail the collection of information regarding the facial trait and how, on average, different levels of the trait is manifest in facial appearance. It should then be possible to derive transformations that manipulate the levels of a specific facial trait within a face. These would take the form of global parameter transforms where a single value ($\lambda$) could be used to control the level of the facial characteristic.

The next section of the thesis will describe how techniques were developed to enable the quantification of global parameter transform information, and how this can be used to shift faces along specific global dimensions. Two major issues must be considered when considering parameterising facial information. First, the visual features of a facial image must be located. This can be solved by finding the correspondence between a specific face and an ideal face and is termed the correspondence problem. In the techniques documented in this thesis, correspondence is solved manually by an operator using a mouse (see section 2.7). However, several automated solutions to this problem have been suggested and successfully implemented (Vetter *et al* 1995 and Lanatis *et al* 1995). Second, a method of representation is required to enable the manipulation of the facial information. Morphing provides a method of interpolating between two images (Wolberg1990). There now follows a review of general 2-D image processing techniques (warping and morphing), it will be shown how these have been used to manipulate facial images. These general image manipulation techniques are extended to allow the following tasks to be performed: (i) averaging images of different faces to form facial 'prototypes', (ii) automated caricature exaggeration of the way an individual face differs in shape and colour from a prototype and (iii) transformations manipulating perceived facial qualities while maintaining information about identity (e.g. modifying the apparent sex or age of a face). The techniques developed will then be extended and applied to 3-D head models.

## 2.3 Image Warping Algorithms.

There are several major types of image warping algorithm. Each shall be briefly mentioned and then justification will be provided for the technique selected for the facial transforms implemented here.

The simplest warping technique is a triangle mesh. Here a set of control points defines an original co-ordinate system. These control points are used as vertices for a triangulation of the original space. By moving the position of these control points, whilst keeping the triangulation constant (same combinatorial structure), a well defined mapping can be generated which warps the source image into the shape specified by the target configuration of the control points (Fig 1.3).

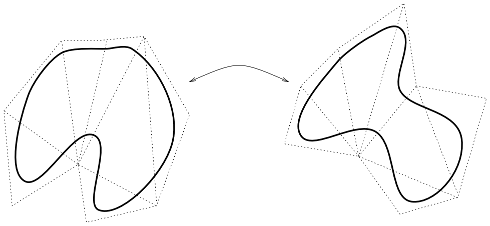
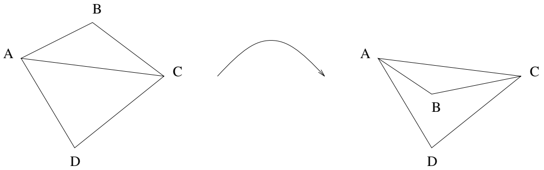


**Figure 2.3.** The figure shows a triangulation and example image (blob). By changing the location of the vertices of the triangle the image contained within each triangle can be deformed.

Delaunay triangulation (Ruppert 1995) is most often used to connect the control points. Delaunay triangulation forces the resultant triangles to have the following characteristic: if you draw a circle passing through each of the three control points making up a triangle, then no other control points will fall within the circle. This restriction causes tessellations to be generated that allow a large movement of the controlling points without the inversion of triangles. Inverted triangles produce incorrect warps and are generally treated as errors. Similarly, the restriction also increases the likelihood that triangles will be generated that are close to being equilateral and this again reduces the chance of a triangle inverting during the warp.

**Figure 2.4**. The figure shows a triangle inverting. Information from ACB is in conflict with information from ACD.



A Barycentric co-ordinates system is a local co-ordinate system defined with respect to known locations within a global co-ordinate system (Birkhoff and Maclane, 1960). Barycentric co-ordinates provide a simple way of mapping the pixels from a source triangle onto the target triangle. The exact equations are given in section 2.5, but essentially every point within each triangle has a 1:1 mapping with a point in the same warped triangle. This correspondence can be used to map the pixel information from a source triangle to the shape of a new triangle. The major ‘disadvantage’ with Barycentric co-ordinate mapping is that it produces discontinuities in the warp (along the edges of the triangles).

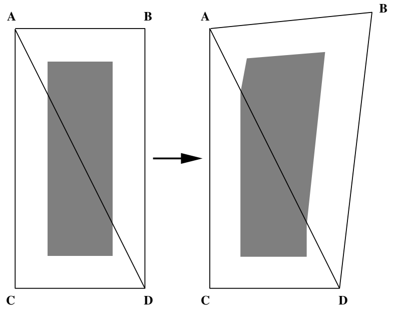


**Figure 2.5.** The figure shows the discontinuities that can introduced due to warping with Barycentric co-ordinates (along A-D). For some purposes these discontinuities can actually be an advantage as shall be discussed later.

To perform a full image warp, each triangle from the triangulation structure is warped into its new shape and the resulting triangles recombined to produce the final image (see fig. 2.8c and Plate 1-1).

A second class of warping *algorithms* is generally termed ‘free-form’ (so called because they use free-form co-ordinate systems). Essentially this means that an initial co-ordinate system is specified using control points. These points are joined up using curves (Splines, Wolberg 1990; Beziers, Sederberg and Parry1986; and others can all be used) to create a co-ordinate space. To warp an image the control points are moved and the new co-ordinate space calculated. The actual warping is accomplished by mapping the new space back onto the initial one. The main disadvantage of this technique is the somewhat unpredictable nature of splines in general. When joining up points with a straight line, it is obvious through which visual features the line will pass. A curve, however, is specified not only by the two points it is joining, but also by the surrounding control points. It is thus more difficult to gauge exactly where the curve will travel.

Another general technique which has been used is Field Based Warping (Beier and Neely, 1992). Line segments are used to specify features on the source image and target shapes. Each change in line (from source to target) thus specifies a global transform for the entire image. The combination of all these transforms into a meaningful warp is done by calculating “influence fields” around each line segment (i.e. points closer to a line are affected more by the displacement of the line). One problem with this approach is that since each line could affect the entire image, many lines must be used to keep the effect of a single line local.

A final noteworthy technique, which is only now becoming possible, is example-based synthesis. This is an automatic technique in that the algorithm automatically solves the correspondence between two source images. Correspondence is solved on a pixel by pixel basis and is calculated using an optical flow algorithm (Bergen *et al.* 1992). An application of this is shown by Beymer, Shahua and Poggio (1993) who teach a neural network a number of different face images (aligned using the above correspondence methodology) along with associated parameters. By giving the network a novel value for a parameter, the network can be used to produce a realistic combination of the input images. For example, the network can be taught face ***A*** with a given parameter set to 1, the network can then be taught face ***B*** with the same parameter set to 0. After training, if the given parameter is set to 0.5 then a face half way between face ***A*** and ***B*** can be produced.

### 2.4 Choice of Warping Algorithm.

The triangle mesh warping technique offers two major advantages over the other warping techniques:

1) It can be performed rapidly; indeed the hardware texture mapping facilities on many of today’s graphics computers allow the warping to occur in ‘real-time’ (see section 2.9).

2) The discontinuities in the warp can be a veritable advantage. The effect of a control point moving is localised to the triangles for which it is a vertex. So if the feature lines from the face (e.g. the perimeter of the lips and eyebrows and the jaw line) are forced

to be edges of triangles in the tessellation, this avoids the blurring of feature boundaries which could occur using the other techniques.

### 2.5 Warping Triangles with Barycentric Co-ordinates.

Each location (P) within a triangle can be specified using vectors ($V_1$ and $V_2$) defined by two different sides of its enclosing triangle (ABC).

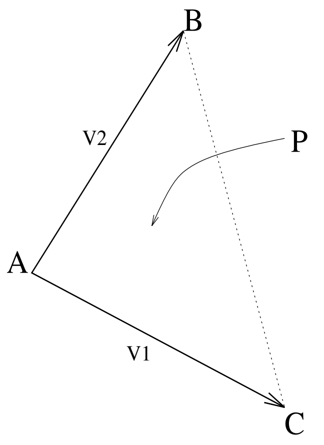


**Figure 2.6.** A point P within triangle ABC can be described using vectors $V_1$and V2

$$V_1 = \overrightarrow{AB} \text{ and } V_2=\overrightarrow{AC}$$

$$P=\lambda_1 V_1 + \lambda_2 V_2$$

To warp an image, it is thus simply the case of moving the defining feature points into the desired shape (making a target set of feature points). Next, for each pixel in the generated image:

1) calculate values for $\lambda_1$ and $\lambda_2$ using the enclosing triangle from the ***target*** set of feature points;

2) look up the colour value from the ***source*** image (using the ***source*** set of feature points) for the same triangle using the values for $\lambda_1$ and $\lambda_2$ calculated in stage 1.

### 2.6 Sampling, Aliasing and Pixel Interpolation.

It is often useful to have labels for the various level of abstraction when considering visual images stored within a computer. The physical universe (P) is the universe of experience and can be defined as being what most people consider reality. Many aspects of this universe can be described using the language of mathematics. This creates a universe (M) of mappings from 'reality' to mathematical construction. A further level of abstraction can be thought of as a universe of representation (R). This universe has symbolic representations of the mathematical universe. Finally, the finite descriptions of this representational universe can be mapped onto an implementation universe (I). This universe contains the specific data structures which will store the information within a computer (Gomes and Velho, 1995b).

A bitmap is mechanism for storing (I) a point sampled representation (R) of a real visual scene (P). From sampling theory, Shannon's theorem intuitively states that, 'if a function has finite energy, and does not have arbitrarily high frequencies, then it is possible to obtain an exact point sampling representation by sufficiently increasing the sampling rate.' (Gomes *et al* 1995a). This means that (assuming the sampling rate is high enough) the mathematical description of the physical scene (M) can be exactly reconstructed from the bitmap. Once the bitmap has been reverted into the mathematical representation, it can be transformed (warped) and then converted back into a bitmap to produce a successful deformation. However, if the original sampling rate is too low, or the mapping from the bitmap (I) into the mathematical representation (M) is not done accurately enough, then aliasing can occur. This reveals itself as spurious high frequencies that can clearly be seen on Plate 1-16.

To avoid aliasing artefacts it is necessary to perform pixel interpolation. This can be performed in several ways, but in general, the task is to create a continuous colour surface from the existing discrete pixel data (i.e. to reconstruct the mathematical description from the point-sampled representation). Colour values can then be read from the surface for locations that are between pixels (to enable resampling). The method chosen for the facial transformations was the simplest, for purely pragmatic reasons, and involves a simple linear interpolation based on the surrounding four pixels. Figure 2.7 shows how this interpolation is performed.

**Figure 2.7.**

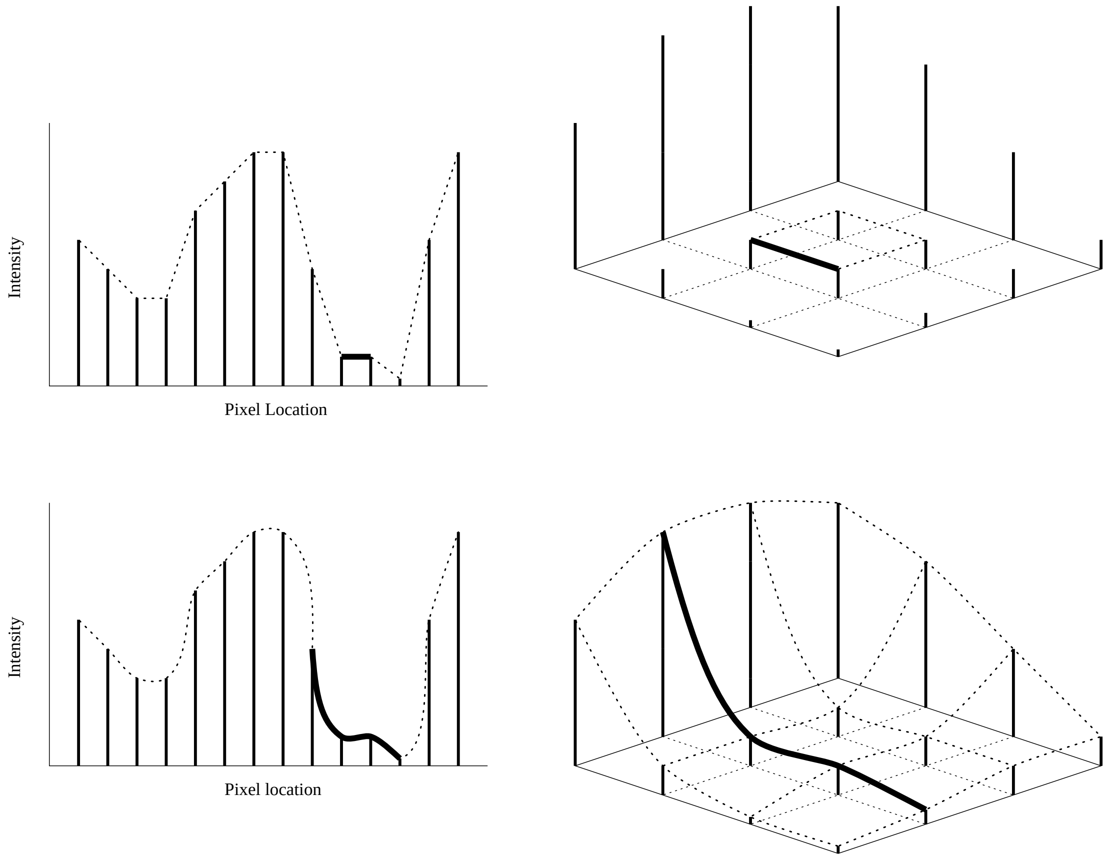


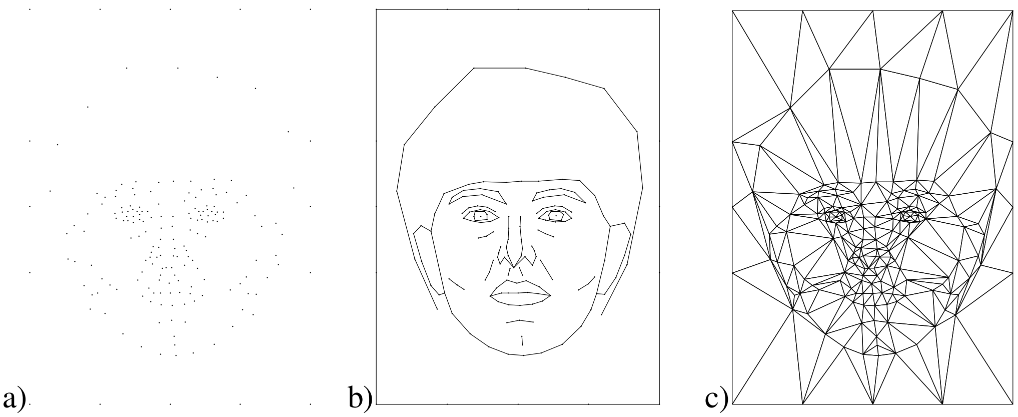
**Figure 2.7.** The figures on the left show a section of a scan line with the height of the lines representing RGB gun intensity at that pixel (i.e. brightness of red, green or blue). Real images are 2-D, however, and so the figures on the right attempts to show a small section of an image (4x4 pixels) again with the height of the lines representing an intensity. With the simple linear interpolation used (top pair of figures) an intensity for a spatial region within the centre square can be calculated from the intensity of the surrounding four pixels. For the second (lower) half of the figure, spline interpolation is used and shows how the surrounding pixels can effect the interpolation within the centre square. This second (Bezier) interpolation is favourable since it produces a result that contains fewer regular aliasing artefacts (i.e. produces a mapping from the bitmap to the mathematical domain that is more accurate). However, experience has shown that for the colour facial work, simple linear interpolation is virtually indistinguishable from the more accurate interpolations. This is not the case for the 3-

D transformations where overlapping discontinuities can produce conspicuous compounded errors.

## 2.7 Delineation.

Each of the techniques to be described later in this thesis relies upon the accurate delineation of 174 (179 with the mouth open) locations. These points are allocated to the 2-D images to demarcate items such as the irises, nose, lips and so on). The set of points closely resembles that defined by Brennan (1992, 1995). The points making up this delineation are termed feature or template points. Currently, these are positioned manually by an operator using a mouse; however some success is being achieved in automating this process. Vetter *et al* (1995) has succeeded in using optical flow algorithms, suggested by Bergen *et al* (1990), to solve this matching task. Lanatis *et al* (1995) uses a different approach, for each feature point a luminance profile for a normal to the feature (a vector perpendicular to the feature line) is matched to an ideal luminance profile for that feature (calculated from many sample faces). At the time of originally writing this section none of these solutions were accurate or robust enough for the manipulations specified in this thesis. However, with the speed of development in this cutting edge computing technology this may no longer be the case (Vetter 1997a,b).

**Figure 2.8** a) The delineation information. The facial image from Plate 1-1 has had 174 standard set of feature points manually positioned to delineate the major facial landmarks (following conventions previously defined, notably by Brennan 1982,1985). Further feature points are added around the perimeter of the image in order to form a limit for the warping, and around the eye to provide greater definition b) the same delineation data connected with line segments; c) a triangulation constrained by the line segments The feature points are used to tessellate the image surface into a set of triangles. A constrained Delaunay triangulation algorithm has been used to encourage maximum area triangles with the strict specification that all the feature lines are included as triangle edges[2].



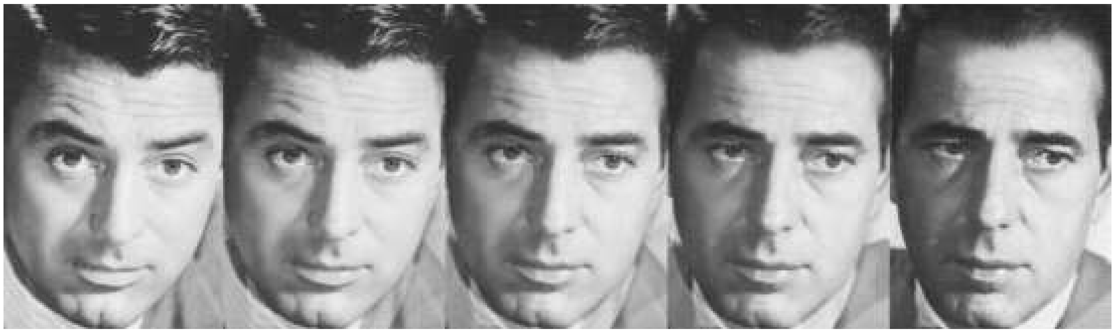
## 2.8 Morphing: An example Use of Warping.

A ‘morph’ is a continuum of images, which changes smoothly from one source image to another. The effect is achieved by a combination of the processes of warping and blending (see Fig. 2.9). This short section is included to illustrate an example use of warping, but mainly to draw attention to the technique of morphing so that it can be seen as fundamentally different to the transforms about to be described.

---

[2] Figure created with software from Shewchuck 1996.

**Figure 2.9**



(‘Grant to Bogey’. From a stimulus set created for Calder *et al* 1996)

**Figure 2.9.** To calculate an image which is n% along the morph between image A (Carry Grant) and image B (Humphry Bogary), where a value of n=0 would yield image A and n=100 would yield image B, a new set of feature point is created a certain (n) percentage between the two source sets of feature points. Image A and image B are then warped to this same intermediary shape and blended together with the weighting specified by n (e.g. n=70 would be a blend with 70% of the warped image A and 30% of the warped image B). This requires both images A and B to be known beforehand. The transformational techniques described later require only one source image and predictively change that image rather than simply combining it with another.

The technique of morphing has proved useful for providing stimuli for psychological studies. It provides a method of creating a stimulus set with quantifiable image parameters (e.g. half way between happy and sad). Examples of studies in which morphing has been used to create stimuli for testing are those conducted by Sprengelmeyer *et al* (1996), where recognition of the expression of disgust is selectively impaired in subjects with Huntington’s Disease; Calder *et al* (1996, 1997) where subjects with bilateral amygdala damage show a differential impairment in the recognition of fear; and Morris *et al* (1996) where neural responses in the human amygdala are measured using a PET scanner whilst subjects view images from a morph. In the latter study, the activity on the amygdala is shown to correlate with the amount (intensity) of expression present in the morph (correlated positively in the case of fear, and negatively in the case of happiness).

## 2.9 Real-time Image Morphing.

A specific problem with experimental design when using a ‘continuum’ of morphed images is the selection of which images from the morph to use as stimuli. There are theoretically an infinite number of images interpolating between the two source (or ‘end stop’) images. One way to overcome this selection problem is to provide a continuous morph, calculating the images whilst the experiment is running. The above mentioned morphing algorithm has been successfully implemented to run in real-time on Silicon Graphics Impact/O2 machines (although the algorithms will work with any system that supports hardware texture mapping and alpha blending). Texture mapping is the process of projecting 2-D bitmap onto a 3-D polygon. This is similar to 2-D warping except that it is performed in 3-D so that additional transformations, such as perspective correction, can be included in the projection. Alpha blending can be thought of as an additional plane in the screen memory. One normally thinks of pixels (**pic**ture **el**ements) as consisting of a triple value with red, green and blue components. An additional component, known as the alpha channel, is similar to the original three components except that it specifies transparency (or rather how the value of two overlaid pixels are to be combined). Thus two images can be blended together in various amounts using alpha blending.

In the hardware morphing algorithm specified here, the original images are first stretched to be 512x512 pixels size (the Impact hardware requiring texture maps of pixel size u x v where u and v = $2^S$, where S = an integer). Feature point data are also stretched by the same factor to maintain registration. By convention, texture maps are mathematically specified to be a unit square (with sides described by vectors U and V) and so the scaling of the image and feature position data is in accordance with this definition. The Delaunay refinement algorithm (Ruppert 1995) is used as implemented by Shewchuck in his public domain ‘Triangle’ package[3]. This creates a triangular tessellation from the feature point data that will be used to project the texture into the desired shape. To do this, an original (unstretched) version of the feature point data is used to map the texture map triangles on screen in the correct dimensions. The texture map projection actually occurs in three dimensions but by projecting the image on a flat surface the resulting image appears 2-D. This is performed under orthographic projection to prevent perspective distortion. Both images (A and B) are warped to the same shape and blended using the hardware alpha blending functions.

---

[3] http://www.cs.cmu.edu/~quake/triangle.html

**Figure 2.10**

**50% along morph from of A to B**

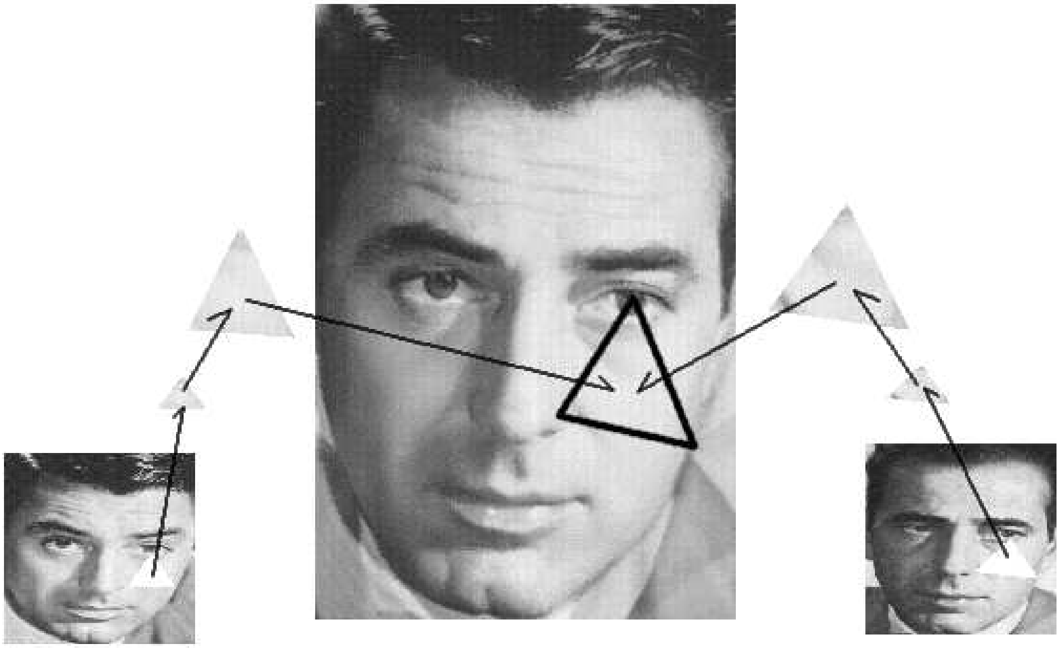


**Figure 2.10.** The figure shows the morphing process for a specific triangle from the triangulation. The pixel information is read from the two (unit square) texture maps and stretched so that the resulting triangle has the correct dimensions. The two triangles are then written to screen on top of one another and combined using alpha blending. In this example the correct dimensions for the triangle are those half way between the delineated face shapes of Carey Grant and Humphrey Bogart, similarly, the triangles are blended with the proportions 50:50. The warping and blending can all take place in hardware and so the whole center image (composed of many triangles) can be generated within a single monitor refresh.

This technique is now being used in several psychological experiments. These include providing subjects with morphs derived from the Ekman photos of standard facial affect (Ekman 1976, Murray 1997). Subjects manipulate a mouse left and right, which causes the morph on the screen to contain more or less of a given facial expression. This allows subjects to choose a level of expression where they are first able to detect the emotion. Similarly, subjects have been given morphs of Japanese to Caucasian faces (Penton-Voak 1997). Here subjects manipulate race and are required to select the point where they first see the face become a specified race. These are being tested, in a similar way to the expression stimuli, and also in cross-cultural studies. Studies of facial attractiveness are being carried out where subjects' are required to select from a continuum the face they perceive as being the most attractive. In these studies the ends of the continuum are masculine verses feminine and own facial traits verses opposite facial traits (Perrett *et al* 1998).

Other possibilities for this technique include allowing subjects to control multiple dimensions simultaneously. For example, using a mouse left/right could control a shape change whilst up/down could control a colour change. This is not possible with a movie (e.g. quicktime, mpeg) since only uni-dimensional control is afforded over which image in a sequence is currently being displayed. It should not be forgotten that the produced faces need not be displayed dynamically; in situations where subjects are selecting from static images, new stimuli could be derived 'on-line'. For example, in studies where the choice of stimuli is governed by the subject's previous responses it may be impracticable to pre-compute all the required images. Real time image manipulation can be used here to calculate only those images actually required at the time of testing.

### 2.10 Forming Facial Prototypes.

In order to define global parameters for facial transformations, one first needs to start by collecting data about how different groups of faces vary. A prototype can be defined as being a representation of the consistencies across a collection of faces. For example, a prototypical male Caucasian face (Plate 1-3, left) contains all the colour and shape information that is consistent across Caucasian faces and can be generated by calculating a mean face from a set of Caucasian faces. The technique for creating a facial prototype from multiple images was pioneered by Galton over a hundred years ago. Galton (1878, 1879) combined faces photographically using multiple exposure of film. He manipulated the orientation and size of component photographs so that the centres of the left and right eyes were constant across the set of faces. Galton's technique has been replicated digitally (Langlois and Roggman 1990, Burson 1992). These digital versions, however, align component images on only two or sometimes three image points (the eyes and mouth centers). Benson and Perrett (1993) improve on this by warping the component images (to average shape) so that the images are aligned not just on the eye centres but on each of 174 (179 if the mouth is open) feature landmarks.

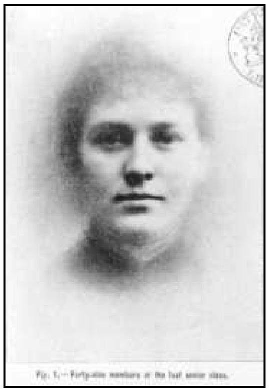


**Figure 2.11.** The figure is from Stoddard (1886), a contemporary of Galton, and was created by compositing the photographs of 49 female undergraduates of Smith College. The constituent photographs were chosen so that the features of the faces would be in similar locations and so could be combined successfully.

In the derivation of prototypes presented here images have been obtained from a variety of sources. Some collections of faces have been deliberately captured under standard lighting with adornments such as jewellery and make-up removed, whilst other collections have been gleaned from magazines and various other less constrained sources.

To derive the prototypical shape for a group of faces the delineation data for each face are first "normalised", making the faces nominally of the same size and orientation. The left and right pupil centres provide convenient landmark points for this process. The first step is to calculate the average left and right eye positions from the group of faces that the prototype is to be derived for. The next step is to apply a uniform translation, scaling and rotation to the positions of all the feature points to normalise each face such that the left eye is located in the average left eye position and the right eye is located in the average right eye position. Strictly speaking, calculating an average should be done *after* normalising rotation (i.e. making the eyes level) but in practice the images used here had more or less horizontal eyes anyway. This process maintains all the spatial relationships between the features within each face but standardises face size and alignment. The average position for each remaining feature point (after alignment) can now be calculated. The resulting data constitutes the mean face shape for the given population.

The prototypical colour information can be obtained by warping each face in the set to the population mean shape. These reshaped images are then blended with equal weighting. Since each reshaped face has been aligned for all features to coincide across the face set, the resultant blend contains the population mean colouration as well as the population mean face shape (see Plates 1-1, 1-3 and 1-6).

As Galton noted, the features of the component faces that are consistent across the group can be visualised in the blend or prototype. Prototypes formed by this process have been used to study the facial characteristics related to age (Burt and Perrett 1994), gender (Brown and Perrett 1995), beauty (Langlois and Roggman 1990; Perrett *et al* 1994) and artistic style (Perrett *et al*. unpublished studies).

### 2.11 Predictive Colour Addition.

A prototype contains the average shape and colour information from the collection of faces from which it was derived. If a specific face (which is similar to the faces from which the prototype was derived) is missing some information, then it is possible to augment the existing information with the information from the prototype. This can ‘fill in the holes’ of the face with relevant average facial information. Since any information added is average, this process should not change the identity of the face. In the example shown on Plate 1-2, a black-and-white (grey scale) image of a face (middle left) is augmented with prototypical colour information (bottom right). This creates an image with plausible colouration. RGB pixels value are used to specify colour throughout this section of this thesis since they closely mimic how colour is represented within a computer. However, RGB colour space is defined by red, green and blue axes and so is not useful if we need to dissociate colour and luminance. HLS (hue, lightness and saturation, Foley *et al* 1990) is an alternative colour space with the three axes representing colour as follows[4]. Lightness is roughly equivalent to luminance with the lowest lightness being black and the highest lightness being white (with all the levels of grey in between). Hue refers to colour and is defined to be an angle (since hue colour space is circular with 0° being red, 60° being yellow, 120° green, followed by cyan, blue, magenta and finally at 360° back to red - this forms a colour hexagon). Saturation refers to the amount of the hue that is present. The resulting space is termed a 'double hexcone' and is the shape of two ice-cream cones abutted (at the open end). It would be useful if this were a perceptually uniform colour space, with unit changes along each axis resulting in unit changes in perceived colour. Unfortunately, it is extremely difficult to create a perceptually uniform colour space. Individual difference (mild and severe colour blindness), room lighting, recent exposure to different lighting conditions, variation in monitor luminance both across monitor and within a single screen, all go towards making a perceptually uniform colour space impracticable. This is one reason why in the HLS colour space L refers to lightness and not luminance (since changes in hue or saturation may create quantitative changes in luminance). Although lightness is not identical to luminance, it is a close approximation and HLS space provides a convenient co-ordinate system to separate colour and luminance information.

To augment a grey scale image of a face with the prototypical colour information, a relevant prototype is first warped to the shape of an example face. The hue and saturation information is then taken from the prototype and combined with the lightness information from the example face. Plate 1-2 shows this whole process and illustrates its success.

---

[4] It should be noted that section 2.11 is the only part of this thesis where the HLS colour space is used. For example it is not useful to calculate average colour values in HLS space due to the circular nature of the hue.

## 2.12 Caricaturing.

Once a prototype has been formed for a collection of faces, it is possible to generate caricatures by accentuating the difference between an individual face and a relevant prototype. After normalising the feature location data from the prototype to the eye positions of an example face, all feature points on the example face can be shifted away from their counterparts on the prototypical face by a given percentage. This percentage is the amount of caricature and can be thought of as extrapolating the shape difference from a morph between a prototype and the example face. If the percentage is -100% then the product of the manipulation will be the example face in the shape of the prototype. If the percentage is -50% then the result will have the shape halfway along the morph between the prototype and the example face. If the percentage is 0% then the example face is returned, if it is +50% then a caricature of the original face is the result. More formally, each feature point is a tuple (x,y) containing positional information. Let $F(i,n)$ contain the (x,y) feature point data of the $n^{th}$ feature point for face $i$. There are 179 feature points, so 1<=n<=179). Let $F(e,n)$ yield the $n^{th}$ feature point data from an example face ($e$), and similarly $F(p,n)$ yield the $n^{th}$ feature point data from the prototype ($p$). To calculate the locations of the feature points in the caricature $F(c,n)$ with a percentage caricature of $\lambda$ (with $\lambda$ =0 returning the original shape and $\lambda$ =100 returning a shape with the differences between the example face and the prototype doubled):

Eq. 1: $$F(c,n) = F(e,n) + (\lambda / 100)\ (F(e,n)-F(p,n))$$

For each feature point n, where 1<=n<=179.

The result of a caricature can be seen in Plate 1-4. Here Hugh Grant has been caricatured away from a prototype made from a collection of male actors. This prototype actually had the mouth very nearly closed and so one can see Hugh's toothy grin is disproportionately exaggerated. This highlights one of the major problems in creating caricatures; the selection of an appropriate prototype from which to caricature away.

Human recognition of face images is improved by caricaturing facial shape, though the improvement has so far only been apparent with highly familiar (famous) faces (Rhodes *et al* 1987 shows the caricature effect with line drawing of faces; Benson and Perrett 1991b, 1993 use warping techniques, similar to those described here, to create photographic quality caricatures). The basic caricature effect has been seen in several instances but always with highly familiar stimuli (e.g. birds for ornithologists, Rhodes and McLean 1990 and modern car shapes, Dodd and Perrett, unpublished studies). The caricature technique is useful for psychological studies because it allows the systematic and quantitative manipulation of facial traits. For example, the degree of fear in a face can be modified in steps from a neutral to normal expression and beyond to a caricature exaggeration of an expression that is more easily recognised (Calder *et al* 1997). This quantitative manipulation of expression allowed correlation with the degree of neural activity in different brain regions that result from viewing fearful expressions (Morris *et al* 1996).

Caricaturing intensity and colour information can be performed in an analogous way to that for manipulating shape. To caricature colour the difference in colour between corresponding pixels in the original image and a relevant prototype is exaggerated. To do this the prototype is first warped to the shape of the original example image. Then the difference in red, green and blue values of corresponding pixels is calculated and a percentage of this difference added to each of the original pixels. More formally, (red, green, blue) colour information can be represented as a triple (R,G,B). Let the $C(i,k)$ yield the RGB colour triple from face $i$ at pixel position $k$. The colour information for the example face ($e$) at pixel position $k$ can be defined as $C(e,k)$. Similarly, the colour information from the prototype face ($p$) at pixel position $k$ can be represented by $C(p,k)$. Images are deemed to have the same fixed horizontal (x) and vertical (y) dimensions, so k represents each pixel position where $0<=k<=((x*y)-1)$. After the prototype and example face have been warped to the caricatured shape a $\lambda$% colour caricature $C(c,k)$ is calculated for each pixel position as follows.

Eq. 2: $$C(c,k) = C(e,k) + (\lambda / 100)(C(e,k)-C(p,k))$$

For each point k, where $0<=k<=((x*y)-1)$.

As an example of colour caricaturing, Plate 1-6 shows an ‘older’ prototype face (made from males aged between 40-50 years old), being colour caricatured away from a ‘younger’ prototype face (made from males aged between 20-30 years old). The final image shows the result of just the colour caricature (100%) and it is readily seen that the colour caricature is older than the original older prototype. It should be noted that no change in shape (i.e. warp) accounts for the increase in age, since all faces were warped to the same shape (that of the older prototype) before caricaturing commenced. However, it is a slightly misleading to say that no shape changes have occurred because 3-D shape represented as intensity information could be affected due to the colour caricaturing.

Caricaturing colour information has been found to facilitate human recognition of famous faces in presentation time paradigms where the face images are presented briefly (Lee and Perrett 1997). Caricaturing intensity information can also improve performance of artificial face recognition systems (e.g. Craw *et al* 1996).

### 2.13 Transformation Based on Prototypes.

Information typical to a class of face (e.g. young, female) is maintained within the class prototype, but any information relating to specific identity is lost (since the prototype has many individuals constituting its make up). The difference between two prototypes can thus be used to manipulate apparent class membership for faces while maintaining identity characteristics. For example, one can change the apparent gender of a female facial image by ‘adding’ to an original face the difference in shape and colour between male and female prototypes (Rowland and Perrett 1995).

To perform the transform, the shape information for male and female prototypes must first be normalised. Currently this is done with respect to the eye positions of the source face, where each prototype is rotated, scaled and translated so that the left and right eye points are co-incident with the left and right eye point of the source face. There is some debate about whether this is the correct normalisation procedure or not: the salient concerns shall be explored in the discussion in section 2.14.

After normalisation, the shape transformation is performed as follows. For each feature point a difference in position is calculated between the two prototypes. This specifies a vector for each point that can be added or subtracted, in whole or in part, to its corresponding feature point from the example face to be transformed. This defines a new set of feature points, forming a shape that is the result of the shape transformation. The source face is then warped into this new shape and the shape transform is complete.

For instance, an example face (***e***) can be changed in shape to produce a gender altered face (***g***) by an amount $\lambda$ using two prototypes (male and female) $p_1$ and $p_2$. Again, letting $F(i,n)$ contain the (x,y) feature point data of the $n^{th}$ feature point for face ***i***, Let $F(e,n)$ yield the $n^{th}$ feature point data from an example face (***e***). Let $F(p_1,n)$ yield the $n^{th}$ feature point data from the first prototype ($p_1$) and similarly let $F(p_2,n)$ yield the $n^{th}$ feature point data from the second prototype ($p_2$). To calculate the locations of the feature points in the gender altered face (***r***) with a percentage transform of $\lambda$ (where $\lambda = 0$ returns the original shape and $\lambda = 100$ returning a shape with 100% of the shape differences between the two prototypes 'added' to the example face):

Eq. 3: $$F(r,n)=F(e,n) + (\lambda / 100)(F(p_1,n)-F(p_2,n)$$

For each feature point n, where 1<=n<=179.

The shape transformation is limited in impact because a considerable amount of (gender) information captured in the prototypes is missing (specifically the colour differences between male and female faces). Similar transformations can be performed in a colour space. The top rights of Plates 1-7 and 1-8 show the impact of a shape transformation while maintaining the original colour of the example face.

To perform the transformations in colour, as well as shape, each prototype is warped into the shape of the example face. The RGB colour difference between prototypes at each pixel location is then calculated and subsequently added to the corresponding value from the original (source) image. More formally, representing colour as an (R,G,B) triple, let the $C(i,k)$ yield the RGB colour triple from face ***i*** at pixel position ***k***.

The colour information for the first example face ***e*** at pixel position ***k*** can be represented by ***C(e,k).*** The colour information for the first prototype $\boldsymbol{p_1}$ (at pixel position ***k***) can be represented by $\boldsymbol{C(p_1,k)}$, and similarly, the colour information for the second prototype $\boldsymbol{p_2}$ can be represented by $\boldsymbol{C(p_2,k)}$, Images are deemed to have the same fixed horizontal (x) and vertical (y) dimensions, so **k** represents each pixel position where 0<=k<=((x*y)-1). After the prototype and example face have been warped to the transformed shape, a $\lambda$% colour transformed face (***r***) is calculated for each pixel position as follows.

Eq. 4: $$F(r,k)=F(e,k)+((\lambda/100)(F(p1,k)-F(p2,k))$$

For each point k, where 0<=k<=((x*y)-1).

Combining the shape and colour transformations (see Plates 1-7 and 1-8) is more effective than either shape or colour manipulations alone. To combine the two transformations, the prototypes are warped to the shape of the image resultant from the shape transformation and the colour difference applied.

Comparing the above equations for transforming faces ($\mathrm{Eq}^{s}$ 3,4) with those for caricaturing faces ($\mathrm{Eq}^{s}$ 1,2), it is now clear that a caricature is merely a special case of the more general transform (where from $\mathrm{Eq}^{s}$ 3,4 the two prototypes $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are replaced in $\mathrm{Eq}^{s}$ 1,2 by the prototype ***p*** and the example face ***e*** respectively).

The transformational techniques have enabled a psychological investigation of visual cues to age (Burt and Perrett, 1995). This study shows that the prototyping process extracts information relevant to perception of age and that this can be applied in the transformational process to increase or decrease perceived age in a realistic manner. This work has successfully validated the transformational techniques documented here and estimates that around 80% of ageing cues are captured by the manipulations.

## 2.14 Discussion.

It has been previously stated in this thesis that there is some debate as to the correct way to normalise prototypes to faces to perform the various transforms. Until recently, both prototypes were normalised exactly on the left and right eye positions of the original face (Perrett *et al* 1995). The problem with the current system arises if there are consistent differences in eye separation relative to the overall size of head. For example, if male and female faces on average have the same head size, but males have greater eye separation, then in a transformation this will cause a change in head size (since eye position is kept constant). It is proposed here that a better method is to normalise on a single point and ensure that rotation is kept constant (i.e. that the eyes are level). This can be achieved by first translating each prototype such that the point mid-way between the left and right eyes is co-incident with the same point on the source face. Rotating each prototype about this same point until the eyes are horizontal (or in alignment with the orientation of the eyes from the face to be transformed) will provide a better normalisation system. The system is better because changes in size are preserved since no scaling of the prototypes is performed. Thus, when a transformation is applied using the new normalisation scheme, changes in eye separation are similarly preserved and no head size artefacts are created. In practice, the proposed normalisation system should only be performed when it is known that the groups of images constituting the prototypes were collected under the same physical conditions. In particular, to be able to use eye disparity meaningfully, the images must all be collected with the subjects a fixed distance away from the camera, and with the camera having a fixed focal length.

The general effectiveness of the transformations is self-evident, but quantification of the manipulations in terms of effect on perception has only been conducted on the age manipulations (Burt and Perrett 1995). This study found that 80% of information relating to age was maintained within a prototype (with roughly half of this coming from the shape information and the remainder coming from the colour). It is surmised here that this missing information is texture. All the bumps and undulations from pores, wrinkles and so on are lost due to the blending (averaging) process used when creating a prototype. Current research is aimed at defining a texture space (rather like a colour space) with which it may be possible to create prototypical texture maps that will capture this missing information (Heeger and Bergen 1995).

It should be noted that these 2-D transformations might, in fact, alter 3-D information present in the images. For example, in the case of a caricature, if a face has unusually hollow cheeks then the shadowing information caused by the indentations will be caricatured along with the rest of the colour information. This could cause the depressions to look even deeper. One possibility is to capture the images to be used for the creation of prototypes under flat lighting conditions.

### 2.15 Conclusion.

This concludes the section describing the 2-D facial transformations that have been developed as part of this thesis. Techniques have been described which allow the consistencies from a group of faces to be extracted in the form of a prototype. These prototypes have then been used to increase the levels of specific facial traits by caricaturing and also to transform (or map) faces of one generic type (e.g. young male) onto a similar generic type (e.g. old female) whilst maintaining identity information. The techniques are already being used to enable new lines of research to be conducted and it is likely that the general methodologies can be applied to other domains. Already research is underway to prototype and caricature dynamic phenomena rather than static images. Pollock (1997) has captured different styles of tennis serves and is applying the techniques of prototyping and caricaturing as a diagnostic tool for players. Motion capture leads the way to other interesting extensions to these techniques. For example, it may be possible to create prototypical gaits for different classes of acted character. An actor could then perform a scene walking normally and then his gait altered later to transform it into that of a person with a limp, or to that of a crow.

### 2.16 3-D Introduction.

It is easy to assume that what is true in the planar space of 2-D is similarly true in the more natural world of 3-D. Ask a friend to name famous painters and they will no doubt be able to recall many. On the other hand, ask them to name famous sculptors and you will be lucky to hear more than half a dozen names. This popularity of 2-D imagery is partly due to the flexibility of the medium. It is possible to depict in 2-D many scenes that could not exist as a 'solid' object that must obey the laws of physical structure (see Figure 2.12).

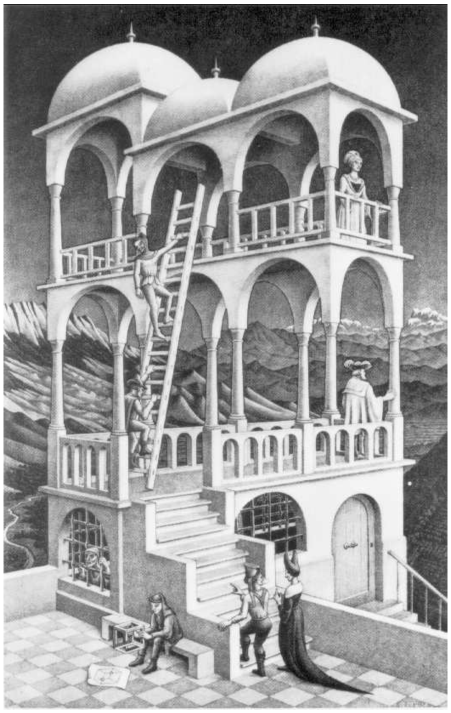


**Figure 2.12.** Paradoxes such as Escher's perpetual staircases and incompatibly structured buildings, or the Penrose triangle exist easily in two dimensions, yet recreating them as physical objects in 3-D is unrealistic. The image on the left is entitled 'Belvedere' and is by Escher (1958). The boy sitting bottom left is shown puzzling over a piece of paper illustrating one of the simplest examples of this type of paradoxical structure, a Necker cube. So it is suggested here that it is not obvious that because the transforms described above work in 2-D it is necessarily the case that they will work in 3-D.

The challenge in this section is to take the same basic methods for quantifying and manipulating images of faces (prototype, caricature and transform) and to discover a method of implementing the same general techniques with 3-D data sets. The adaptation to 3-D will also allow assessment of how well techniques work from different 2-D perspective views. Thus far the graphic processes have been applied solely to frontal images of faces, they may work less well in profile. The pioneering work of Parke in the seventies and early eighties shows the possibility of creating parameterised 3-D facial models. This seminal work creates a parameterised model based on the underlying structure of facial anatomy but includes ad hoc parameters created from observed surface properties. It is suggested by Parke (1982) that an approach combining what is known about an object and what is seen of an object is most likely to create a complete parameter set (i.e. a set of controlling variables capable of producing all possible faces). The work in this thesis takes a slightly different approach and shows that to successfully transform visual appearance all that needs to be known is how visual appearance changes on average (i.e. ignoring the underlying mechanics).

## 2.17 3-D Representations of Faces.

The 3-D data from the face is collected using a Cyberware$^{tm}$ scanner. This device controls a camera that orbits around a stationary head. A thin strip of pixels is recorded every 1/512$^{th}$ of a rotation and an image is produced by concatenating these horizontally. This image is like the ‘photographs’ taken by a photofinish machine at a race course except instead of time being along the horizontal axis, it is angular displacement (the fact that it takes some 20 seconds to perform a 360$^{o}$ scan makes it even more like a photofinish machine! ). This enables the collection of the colour map data. The vertical resolution of the camera is similarly 512 pixels and so this yields a 512 square image that is a cylindrical projection of the head surface onto a surrounding cylinder. The 512 square image is recorded with a colour resolution of 24 bits (8 bits, red, green and blue). To collect the associated 3-D depth data a laser is projected to produce a vertical strip on the head. As the camera is moved, the laser is identically moved around but is positioned at a slight angular offset. The camera thus sees the laser line as tracing out the shape of profiles on the head. This data is concatenated and stored in the same way as the colour data and a 512x512x16bit-depth map data set is produced. Software was written to smooth spikes in the data that occur when the laser passes over a black object and thus no light is reflected. A simple linear interpolation of the surrounding four data points was sufficient for this task. Other software was written to enable the depth map to be viewed in the same way as the colour data (i.e. as a 512x512 image). The two images (colour and depth) are in perfect alignment with pixels in the colour map being coincident with their associated height in the depth map.

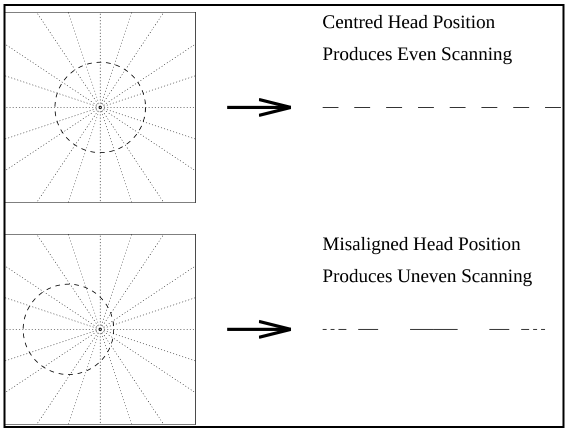


**Figure 2.13.** During the scan, in order to minimise distortions, the centre of the head needs to be positioned at the origin of the camera's rotation. Figure 2.13 shows a top view with circles representing head position and the dotted lines emanating from the centre representing laser scans terminating at the origin of the cameras rotation. On the right of the figure are dotted lines representing a section of the scan from the surface of the dotted circle.

From figure 2.13 one can see that as the surface moves away from the centre of rotation less information is recorded about the surface. Utilising this fact to increase accuracy in the data collection process was considered. Reasoning that by shifting the head back a few inches a greater resolution could be obtained for the face at the expense of the resolution being less for the back of the head (something which was not of primary concern). Unfortunately, since the scanner cannot collect range data past the origin of the camera's rotation this was impossible because data was missed from the neck and so proper chin shapes could not be estimated. Similarly, exact head position is impossible to achieve without restraints that would interfere with the collection of shape. Therefore it was decided to position the head roughly in the center of the scanner's range.

Both the 512x512 colour and depth maps can be delineated in the same way as normal facial images (see Plates 1-11 to 1-13). It is easier to locate most features on the colour map, but for regions defined by a sharp change in surface curvature rather than colouration it may be easier to delineate the depth map. In automated delineation strategies both depth and colour maps could be used to calculate the positions of the feature points.

The head has three degrees of freedom for translation and three degrees of freedom for rotation. As has been said, head position and posture may not be precisely aligned with the scanner centre. Moreover, posture may vary from person to person. Normalisation can be performed on the 3-D data for each head using three facial landmarks (e.g. centre of the two eyes and the centre of the mouth). This is equivalent to translating the 3-D co-ordinate set (so that the points midway between the eyes is aligned with an ideal), rotating about this point so that the axis between the eyes is in alignment with the ideal (i.e. horizontal) and finally rotating about this axis so that the plane defined by the two eye and mouth points is the same as in the ideal. The problem with this method is that resampling is required after the transformation of the data because of the way sample rate alters with distance from the centre of camera rotation. Resampling would involve mapping the transformed data back onto a cylinder (similar to a Lambertian map projection) or alternatively a sphere could be used. In either case the colour maps would have to be recalculated in order to preserve their correspondence with the depth maps. For the 3-D work described here and during the development of the algorithms it was decided not to normalise the data sets in the above way. This approach was considered sensible because of the lessons learned from the 2-D section of this thesis, where it was noted that normalisation on eye position could induce distortions (where changes in eye separation produced in the transform changes in head size).

For the colour facial work, a linear interpolation was found to be almost indistinguishable from the spline interpolation. For the 3-D work, this is not the case and more accurate interpolation is desirable since slight edges on otherwise smooth surfaces can be noticeable when rendered. Moreover, the triangle warping technique is not best suited to warping depth information because of the discontinuities in the warp at triangle edges, which can cause faint ridges around triangles. However, because it is imperative to keep the depth map in alignment with the colour map, it was decided to keep the same warping technique for both colour and depth.

### 2.18 3-D Prototypes.

3-D Prototypes are formed in part using procedures equivalent to those described above for 2-D face images. First, the colour maps for each face in a set need to be delineated. The average shape of the 2-D delineation data is then calculated and each original image colour map warped into this average shape. The aligned colour maps are then blended together to yield the average colour map (Plate 1-14, middle, right).

Since the depth-maps are 512x512 arrays, exactly the same as the colour maps, they can be warped using exactly the same transform. The depth maps for each face in the set are aligned to the average shape of the 2-D delineation data and blended in the same way to yield the average depth map (Plate 1-14, middle, left). The average depth information can be reconstructed as a 3-D object with depth converted back to the distance from the centre of the head (or centre of the scanner). Finally, the average colour information can be texture rendered onto the surface of the 3-D virtual head to synthesise a 3-D prototype (Plate 1-14, bottom). See Plate 1-15 for examples of 3-D shape prototypes.

### 2.19 3-D Caricatures.

Likewise, caricaturing can be performed in a similar way to its 2-D counterpart. The initial steps follow exactly the procedures described above for exaggerating the shape information of the 2-D colour representations. The first stage starts by calculating the difference in shape between the delineation structure of the target face and that of the chosen prototype. Here, shape refers to the x-y position on the cylindrical projection of the face. Shape differences are then amplified by the desired percentage. The depth and colour maps for the target face and depth map of the prototype are then warped into the caricatured x-y shape. The depth map for the target face is now compared to the depth map of the prototype (in the new caricatured x-y shape). Differences in depth at each pixel in the 512 by 512 array are exaggerated by the same percentage. This yields a 3-D caricatured depth map that has deviations from the prototype caricatured in x,y and z (Plate 1-17). Since the colour map information from the re-mapped target face (warped into the caricatured x-y shape) lies in register with the new 3-D caricatured depth map this colour information can be texture-mapped onto the 3-D caricatured shape to create a texture-rendered 3-D caricature.

These procedures caricature only the 3-D shape information. Defining the difference in red, green and blue intensity values between corresponding pixels between the original and prototype and adding a percentage of this colour difference back to the original image can of course, caricature the colour information as well. This can be done only when the x-y shapes of the original and prototype lie in register. If colour differences are to be calculated after shape caricaturing, then the colour map of the prototype needs to be warped to the caricatured x-y shape of the individual. Texture rendering employing both the 3-D caricatured shape and the enhanced colour map yields a caricature that exaggerates all information available from the scan data.

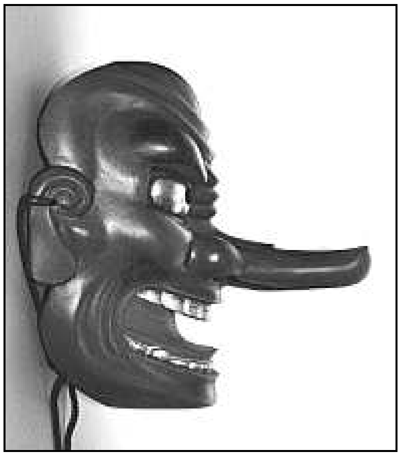


**Figure 2.14.** The image of the left is of a 'Tengu' (Chinese/Japanese Demon). The enormous nose has lead some to suggest that the demon is a representation of the face of a race other than oriental which has been exaggerated (caricatured) against the average (prototypical) oriental features.

### 2.20 3-D Transformations.

As noted above, prototypes can also be used as a basis for changing the apparent characteristics of a face while maintaining identity information (Rowland and Perrett 1995). To perform these operations in 3-D, a two-stage process is needed. The initial operations are equivalent to those described for 2-D. First, the difference in 2-D delineation shape for the two prototypes is calculated. A percentage of this is then added or subtracted from the shape of the delineation for the target face. To modify the apparent gender of a face, for example to induce a more masculine appearance to a female face, the difference between male and female prototypes is added to the original shape of the target face. To increase the femininity of the target female face a percentage of the difference between male and female prototypes is subtracted. Once the new x-y shape of the target face is calculated the colour and depth maps for the original target face and the prototypes are all warped to this shape.

This establishes the desired shape transformation in only x and y. To calculate the shape transformation in the depth plane (z), the difference in depth maps between the male and female prototypes (aligned to the new x-y shape) is found. A percentage of this difference is then added or subtracted from the corresponding pixels in the depth map of the target face (aligned to new x-y shape). This establishes a 'sex transform' transformation in 3-D shape for the target face, producing a head shape with enhanced masculine or feminine structure but maintaining the identity or family characteristics of the individual.

For example, the models in Plate 1-18 were created following this four-stage process:

1) The differences between the average delineation data from the two prototypes are added to the delineation data from the example face. This creates a target shape which all three colour maps (the example face and the two prototypes) are warped into; similarly all three depth maps are also warped into this target shape.

2) The differences between the warped male and female prototypes colour maps are added to the colour map of the example face.

3) The difference in depth between the warped male and female prototype depth maps is added to the depth map of the example face. This is performed in an identical way to the colour addition in step 2.

4) The resulting depth map is then projected back around the cylinder to create the final gender transformed head shape. Finally, this is texture mapped with the resultant colour information from step 2.

### 2.21 Discussion

It is interesting to discuss if and how the documented transforms will fail. For example, a single facial dimension can be envisaged in 2-D (consider gender). If the distribution of faces around the male prototype is larger than the distribution of faces around the female prototype then it is possible for the result of a gender transform to be 'not a face' (Figure 2.15).

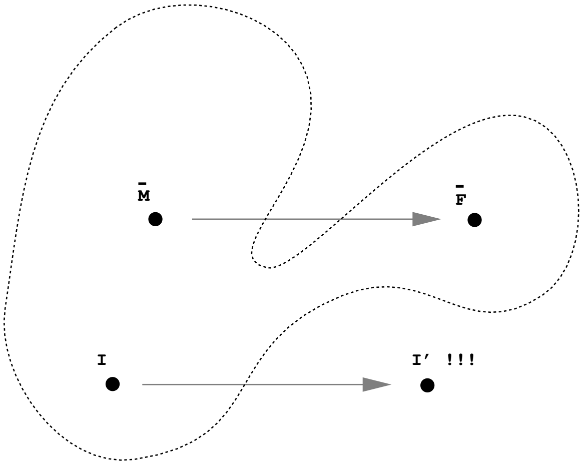


**Figure 2.15.** The vector different between male and female prototypes is shown (M→F). When this vector is added to an individual face (I) the resulting face (I') falls outside of face space and since this cannot be considered a face the transform has failed.

Similarly, a morph between the two face prototypes M to F would first exit and then re-enter the face space. One can avoid this last possibility by assuming that face space is convex, however it will always be possible to provide projections of faces out of face space. For example, feminising an already very female face may create an impossible warp with overlapping triangles. Another interpretation is that linear interpolations between faces are incorrect. For example, if a morph is created between two posed expressions of the same individual (say neutral and happy) then the resulting animations may not be a realistic smile. In actual expressions, eyebrows may raise first followed by a widening of the mouth. In the simulated morphed expression these two events would occur simultaneously. This is perhaps muddying the water slightly, because of the temporal component to an expression, but hopefully the point is made - that however effective they appear, the linear transformations may not be realistic.

A further point worth of note is the representation of the 3-D head data. Since the 3-D data is represented in a cylindrical co-ordinate space the various transforms will have a different effect than if they were performed in a co-ordinate space with 3 orthogonal vectors defining axis (e.g. X,Y and Z). For example, consider a face whose eyes are wider apart than a prototypes. In the latter (Euclidean) co-ordinate system, a caricature would increase the distance between the eyes, pushing them further apart. In the cylindrical co-ordinate space a caricature will push the eyes further around the head, until they finally meet at the back! It is debatable which of these two outcomes is the most appropriate and future psychological studies are aimed at addressing these queries.

### 2.22 Conclusion.

The purpose of this first section was to design and implement computer graphic techniques (the validity and utility of which has been demonstrated in various collaborations and is not the topic of this thesis). The success of the manipulations presented in this first section is apparent, evidence for this comes not only from the many commissions have been received from the entertainment's industry[5], but also from scientific study where the 2-D manipulations have been successfully calibrated and used to control stimulus generation. The effect of the 3-D manipulations is currently under investigation, but it is patent from viewing the effectiveness of the transformations (Plate 1-18) that the technique is as immediately successful as its 2-D counterpart and could be used for an equivalent wealth of psychological studies.

---

[5] Examples include, Television: BBC1, *Face Value*, 1997; Channel 4, *Gay Time TV*, 1996; Channel 4, *Without Walls*, 1994; Newspaper and Magazine: *The Sunday Times*, 5th May, 1996; *GEO* magazine 1996.

**Plate 1-1: Delineation of features on a 2-D face image.** The top row from left to right shows 1) An original face from an example face set (60 female faces ages 20 to 30). 2) This face on which 174 feature points have been manually defined. 3) The same face warped into the average shape for that group (feature points shown) and 4) The same image in colour with the feature points removed. In the bottom left hand corner of the plate and average face shape (dark lines) is shown surrounded by four delineated face shapes taken from the set of constituting faces. The image on the right is the result of the superposition of 60 female faces from the example face set all warped to the same average shape.

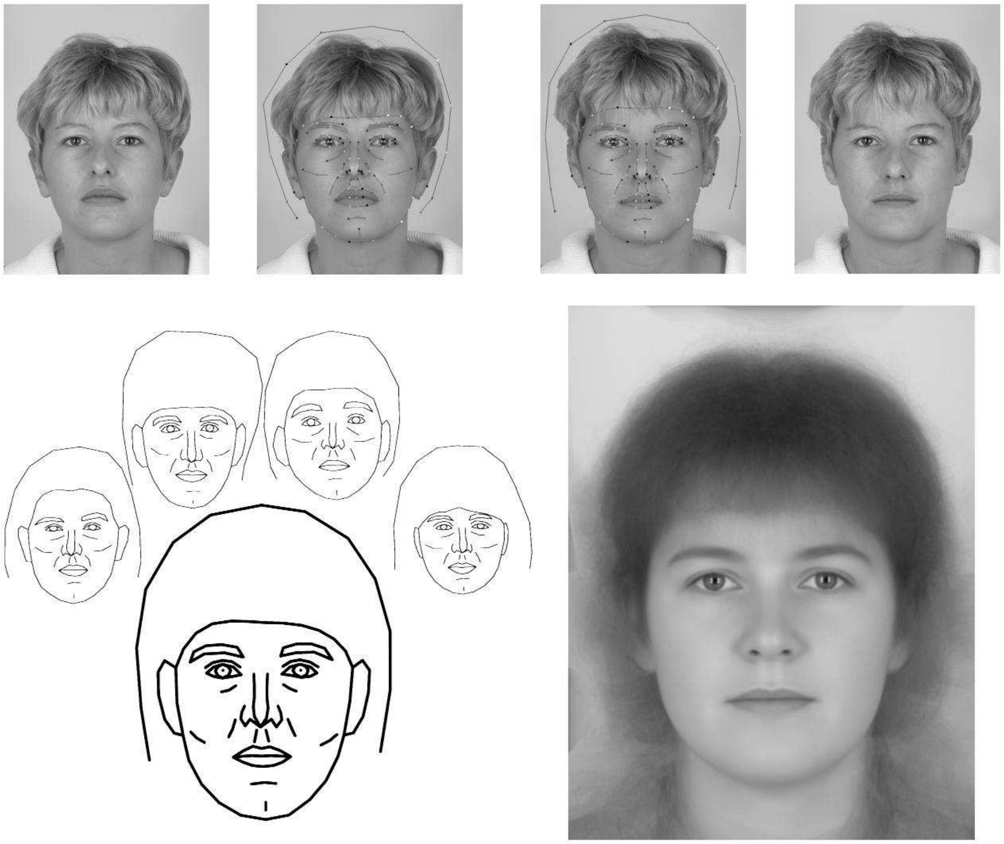


**Plate 1-2: Predictive Colour addition.** The left column, from the top down shows an original colour photograph followed by a lightness only (grey scale version) and finally a reconstructed of colour with the lightness (mid-left) combined with the hue and saturation from the prototype (bottom-right). The right column shows a female prototype. The top image is a female prototype created from the same set of images as the original image top right (although, for this plate only, this original image was omitted from the generation of the prototype). The prototype is warped to shape match the example face on the left. The bottom image shows the hue and saturation from the shape-matched prototype (with a constant lightness, i.e. mid-grey).

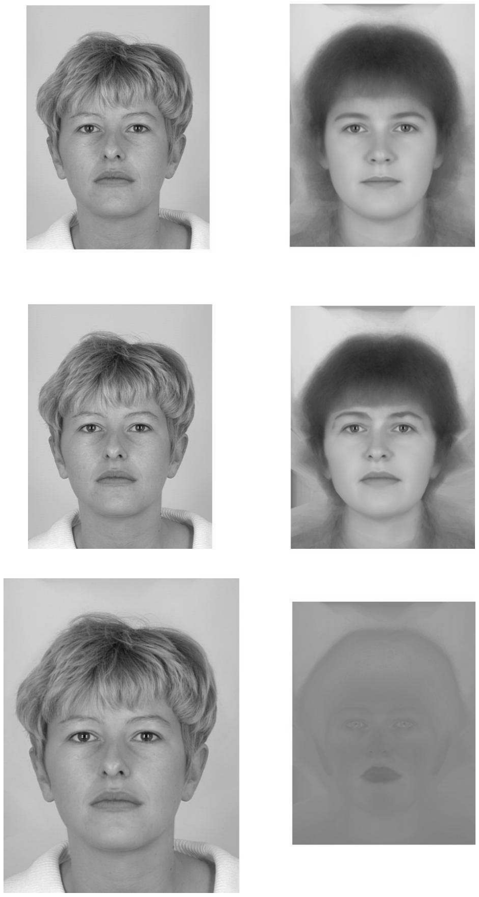


**Plate 1-3: Magazine prototypes.** Prototypes derived from images taken from various magazines. The male prototype was made from 43 Caucasian faces and the female prototype was made from 61 Caucasian faces.

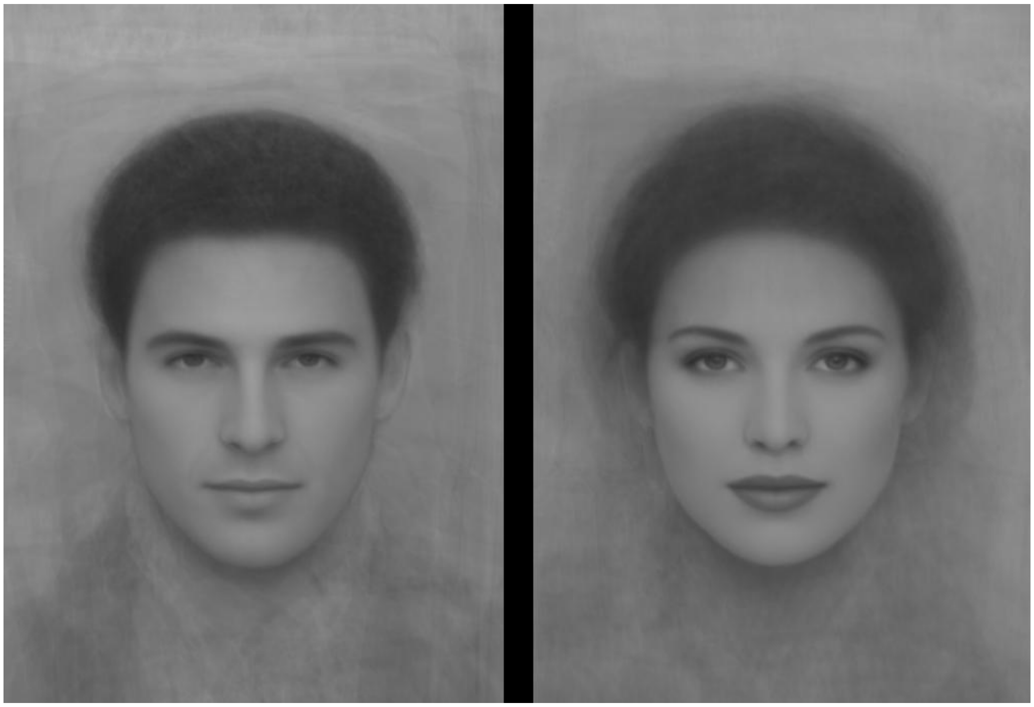


**Plate 1-4: Divine Smile (sic).** An original image of Hugh Grant on the left. On the right is a 50% shape caricature (away from the male prototype from plate 1-3). One can clearly see how Hugh's toothy grin is disproportionately exaggerated since the prototype has a closed mouth.

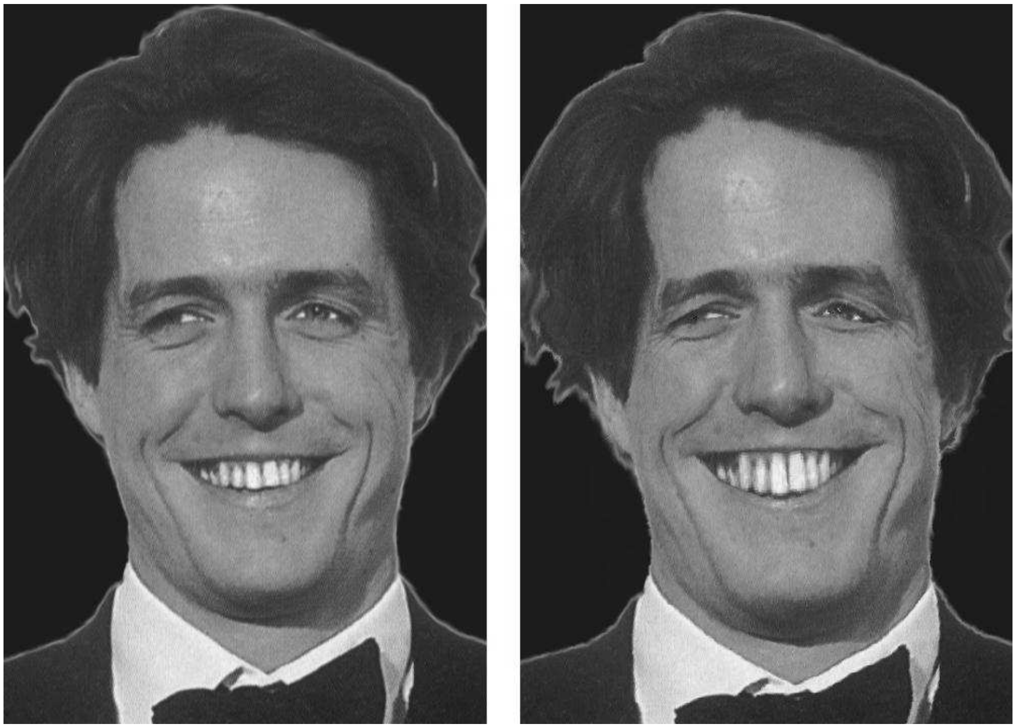


**Plate 1-5: Colour and Shape differences.** The top pair of images are the prototypes from Plate 1-3. The middle pair of images shows the differences in shape between the two prototypes. On the left is the shape information from the female prototype. The lines extending away from this outline show how each point must move to form the male prototype. Similarly, on the right the outline shows the male prototype with lines showing how each point must move to form the female prototype. The movement lines are the same in each image with the male or female prototypes being formed by joining up opposite ends of the lines. These movement lines represent the vector that would be added in order to manipulate the perceived gender of an individual's face shape. The lower pair of images shows the difference in colour between the male and female magazine prototypes. Each prototype has been warped to the same 'androgynous' shape (i.e. half way along the vector between male and female prototype shapes). The colour difference between the image is then found and then added to a mid-grey box. The image on the left shows a grey box with the colour differences between the male and female prototypes subtracted. The image on the right shows a grey box with the colour differences between the male and female prototypes added. In other words, the image on the left is a photographic negative of the image on the right. These colour differences represent the colour changes that would be induced in order to manipulate the perceived gender of an individual's facial colouration.

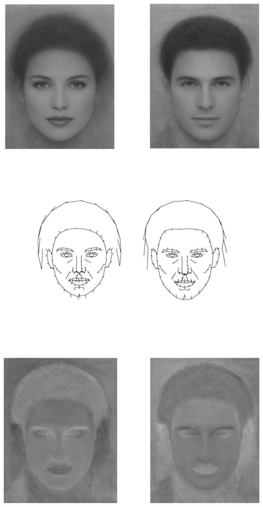


**Plate 1-6: Colour Caricaturing.** From left to right the top set of images show an original younger prototype (made using 36 faces between the ages of 20 and 30, followed by the shapes of the younger male and older male prototype (made using 36 faces between the ages of 50 and 60). Using the shape difference, the young prototype is warped into the shape of the older (bottom left). On the bottom row, colour differences between the younger prototype (warped) and the older prototype (untouched and in the middle) are enhanced to produce the colour caricature on the right. This enhances the colour differences between the younger and older prototype and produces an even older looking prototype. NB. There are no shape differences between the images in the bottom row.

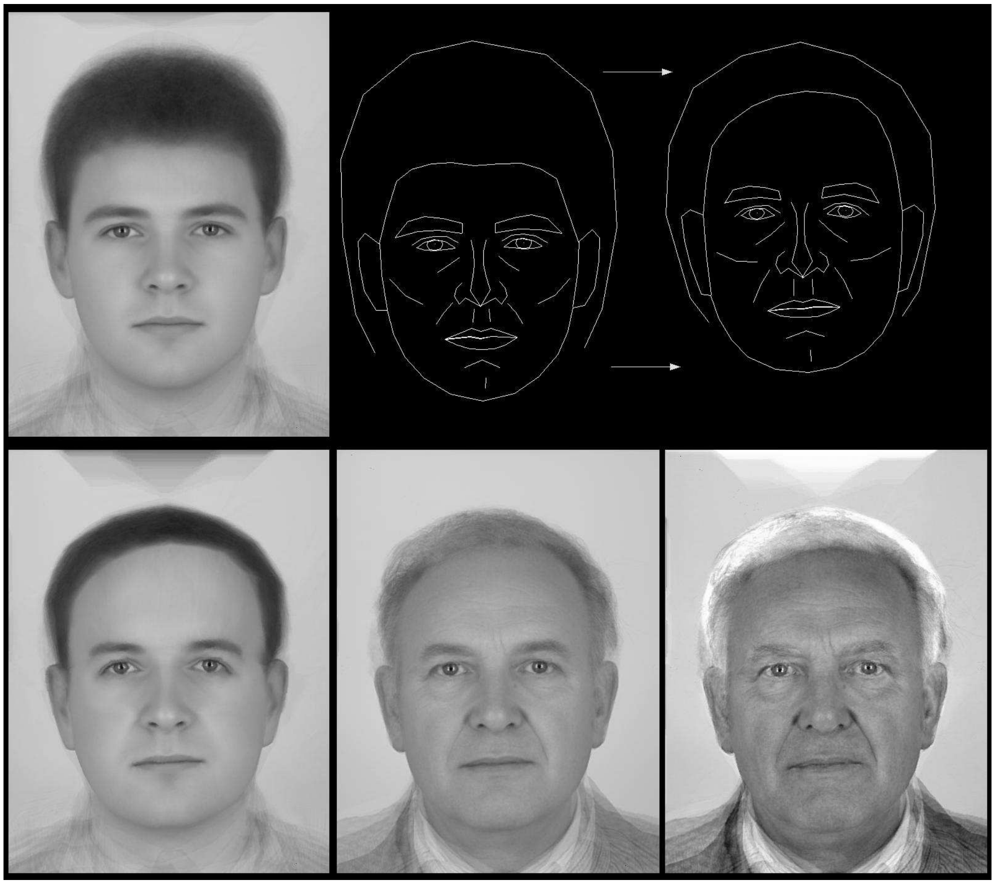


**Plate 1-7: Ageing Transform.** The younger and older prototypes from Plate 1-6 have been used to age the face of an individual (top left). The top right image shows the original image warped into the shape created by adding the difference in feature positions between the younger and older prototypes (from Plate 1-6) to the shape of the original face. The image bottom left shows the original image with the colour difference between the younger and older prototypes 'added'. The bottom right images show a combination of the shape and colour transformation.

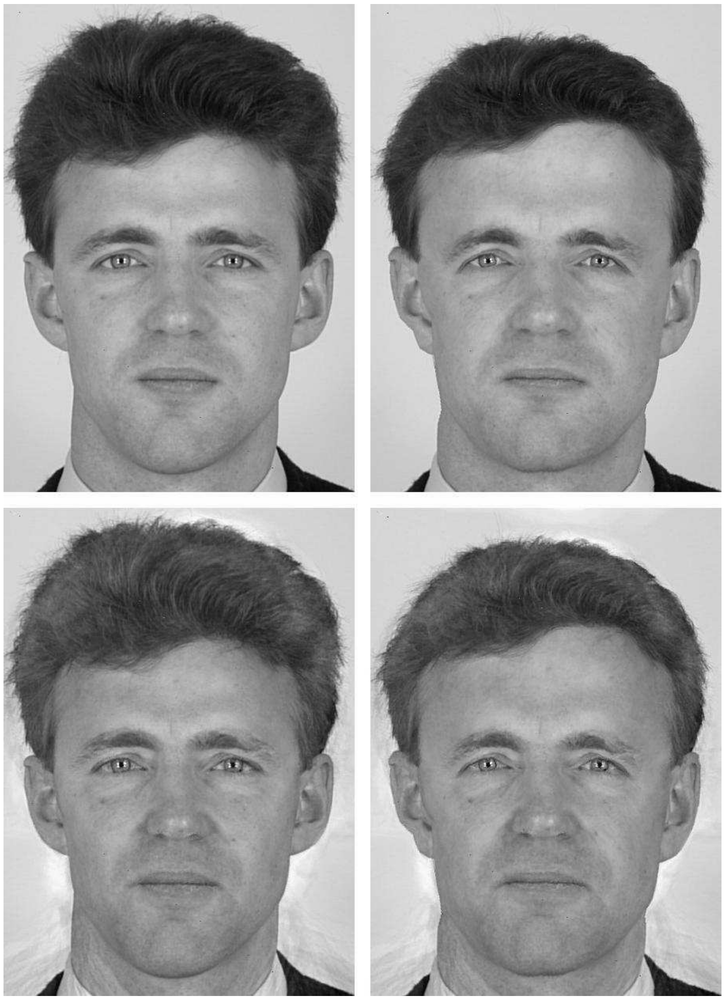


**Plate 1-8: Sex Transform.** The original male face (top left) has had the shape difference between the male and female 'magazine' prototypes, from Plate 1-3, added (top right). The image below the original face has had the colour differences added. And the bottom right image has 100% of both the shape and colour differences between the prototypes.

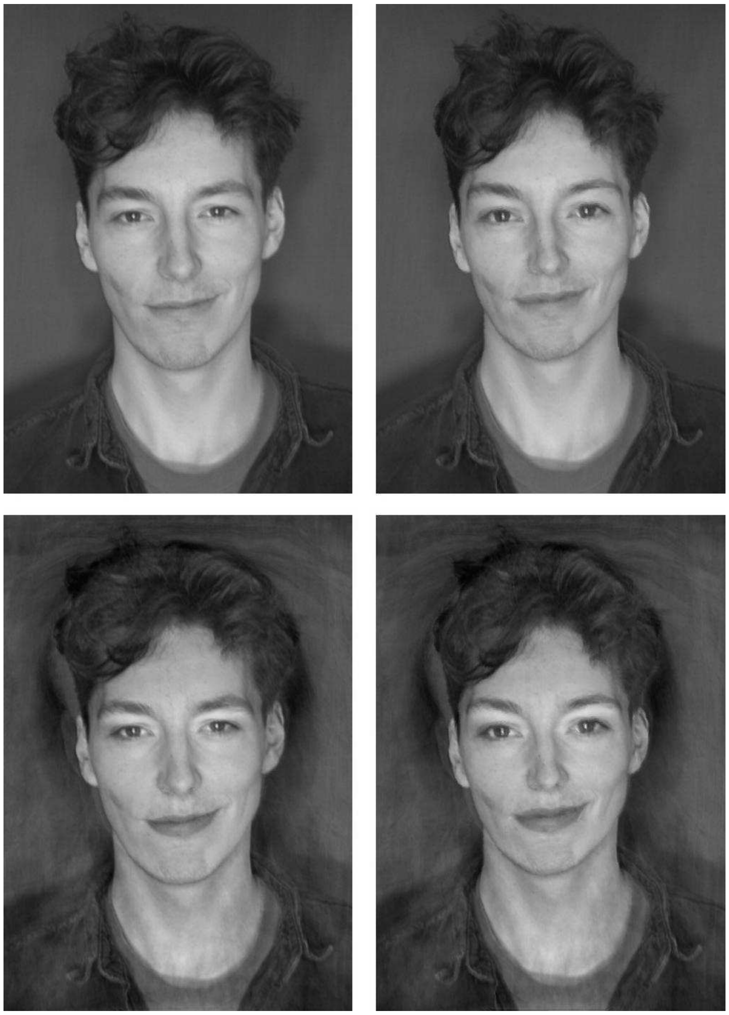


**Plate 1-9: Sex Transform: Different Directions and Degrees.** The bottom row of images shows different degrees of the sex transform. The original image is second on the left. The subsequent three images on the right of the original have had 50%, 100% and 150% of the difference, between the male and female prototypes from plate 1-5, added respectively. The first image on the left has had its masculinity enhanced by 50%. The top row shows an original female face, second from the left. Exactly the same transformation is applied to the face as to the face below except in the opposite direction. This has the effect of feminising the original face (leftmost image) and for the three images on the right hand side of the original, transforming the face 50%, 100% and 150% towards masculine.

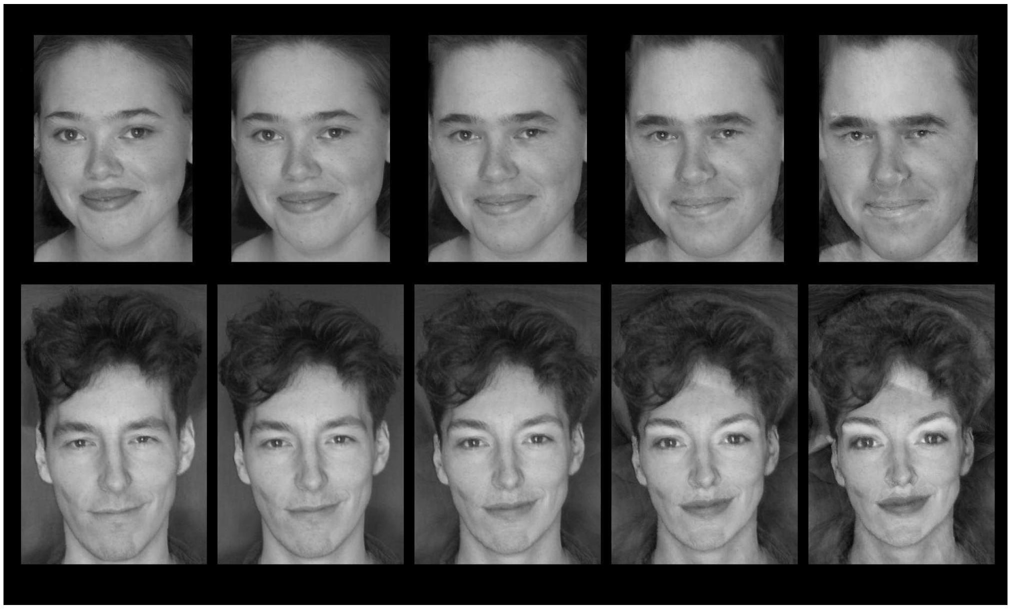


**Plate 1-10: Stoney Bloke.** The original image of Sharon Stone is top right. On the left (from right to left) are progressively increasing feminised versions, whilst below, from left to right are increasingly masculinised versions.

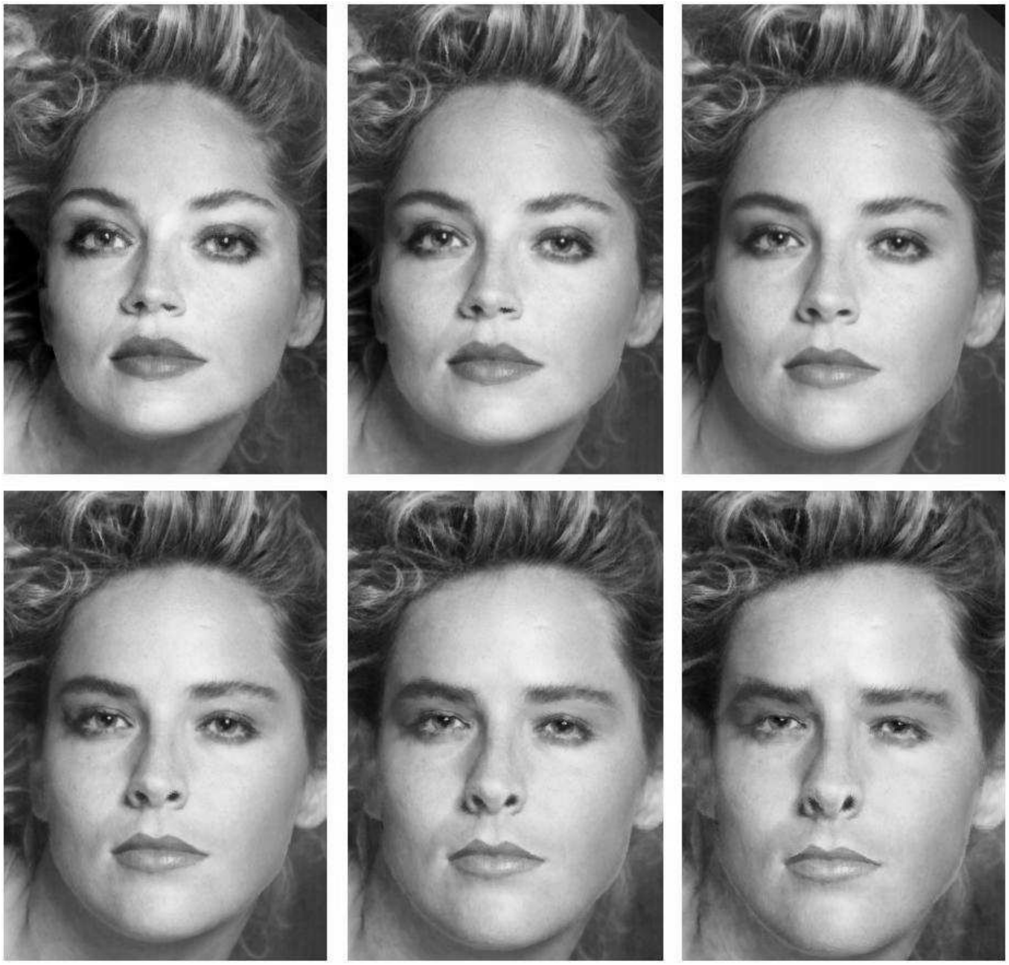


**Plate 1-11: Cyber-Scan Delineation - colour.** The colour map from a Cyber-scan of a human face is shown delineated in a similar way to a normal camera photograph (e.g. from plate 1-1).

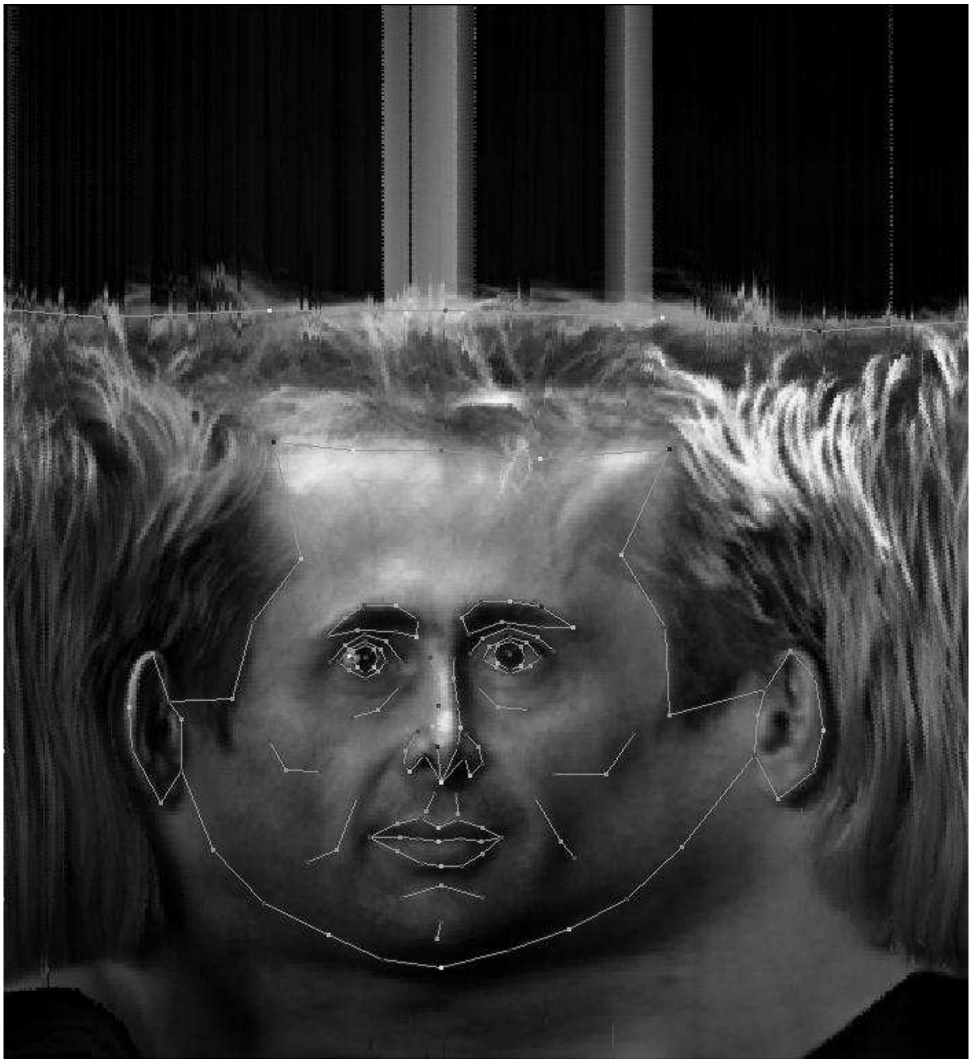


**Plate 1-12: Cyber-Scan Delineation - depth.** Shows a flattened depth map, where the distance from the center of the cylinder has been converted into a height and the cylinder 'unwrapped'. The image is textured with the delineation data from plate 1-11.

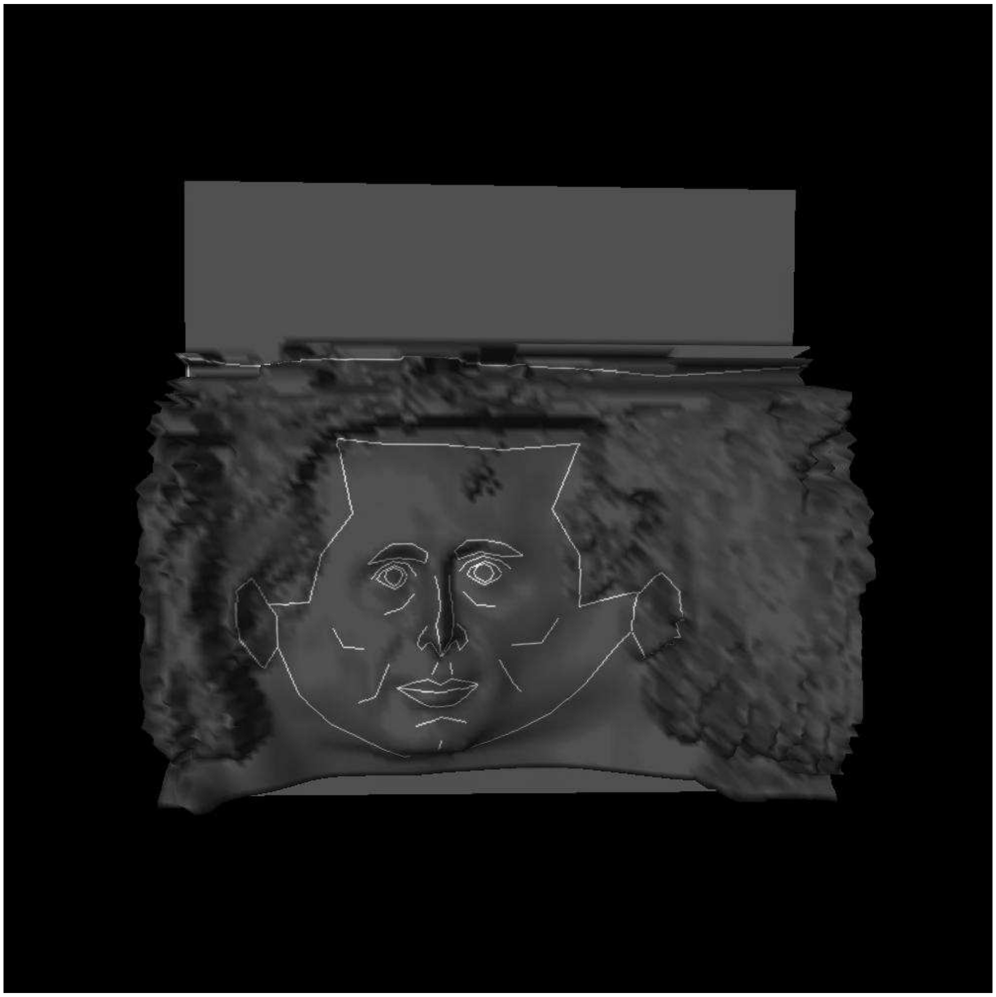


**Plate 1-13: Cyber-Scan Delineation - head.** Shows the flattened depth map from Plate 1-12 wrapped back into a cylinder to give the delineated 3-D shape of the original head.

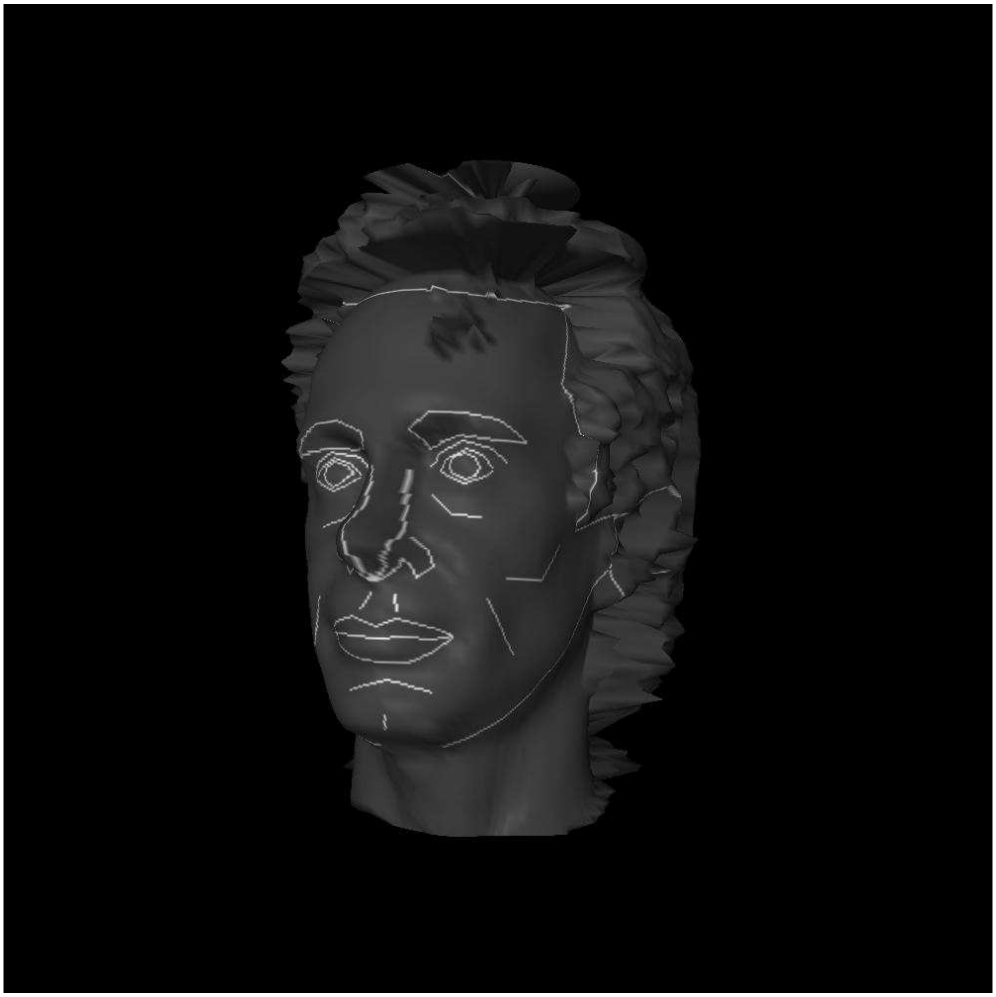


**Plate 1-14. Forming a 3-D facial prototype.** Upper left: the depth maps for two faces from a set of 20 male Caucasian faces. Upper right: the corresponding colour maps for the same faces. Mid-right: the average colour map formed from the 20 component colour maps. Each original colour map was delineated and the average 2-D shape of the delineation data defined. The original colour maps were then aligned by warping them to the average 2-D delineation shape. Finally the aligned colour maps were blended together. Mid-left: the average depth map formed by warping the 20 component depth maps into the average 2-D delineation shape and then averaging the depth values across the aligned maps at each pixel. Bottom: 3-D prototype for the set combining average depth and colour maps. The depth map is now expressed as radial distance from an axis through the centre of the head and the colour has been texture mapped on to the surface.

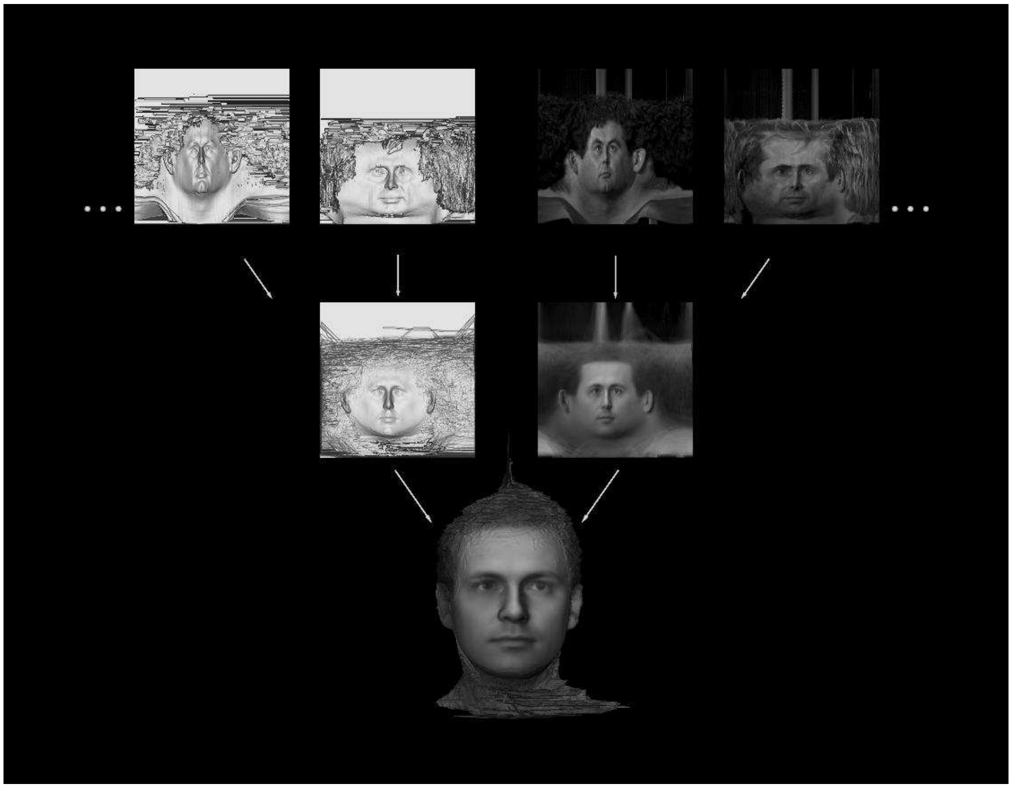


**Plate 1-15: 3-D Prototypes.** Four prototypes are shown. Top left is a Japanese male prototype created from a sample of 18 head scans. Top right is a Japanese female prototype created from 15 scans. Bottom left is Caucasian male made from 24 scans. Bottom right is Caucasian female made from 10 head scans.

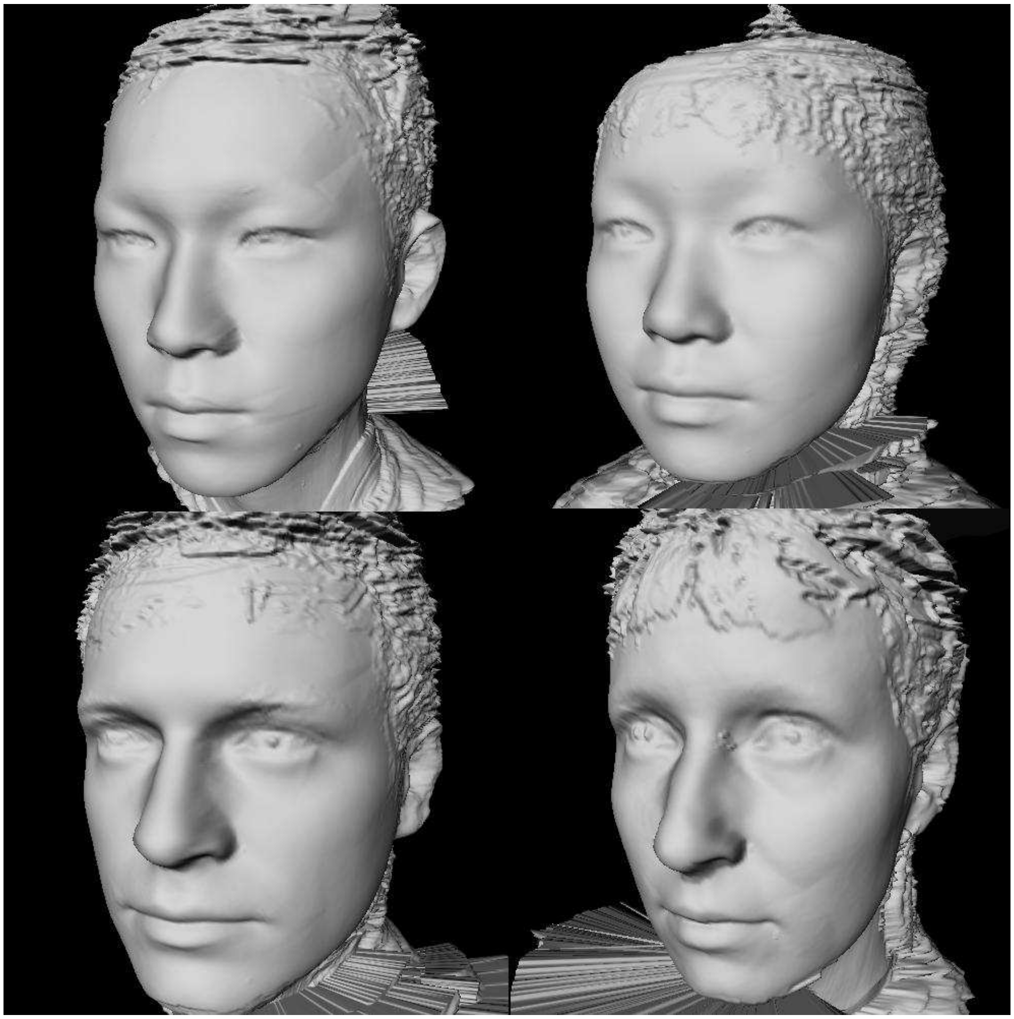


**Plate 1-16: Aliasing.** The image on the left shows a female Japanese prototype after it has been enlarged using the warping algorithm. Aliasing lines are easy to see. Linear interpolation is used for the same transformation for the image on the right.

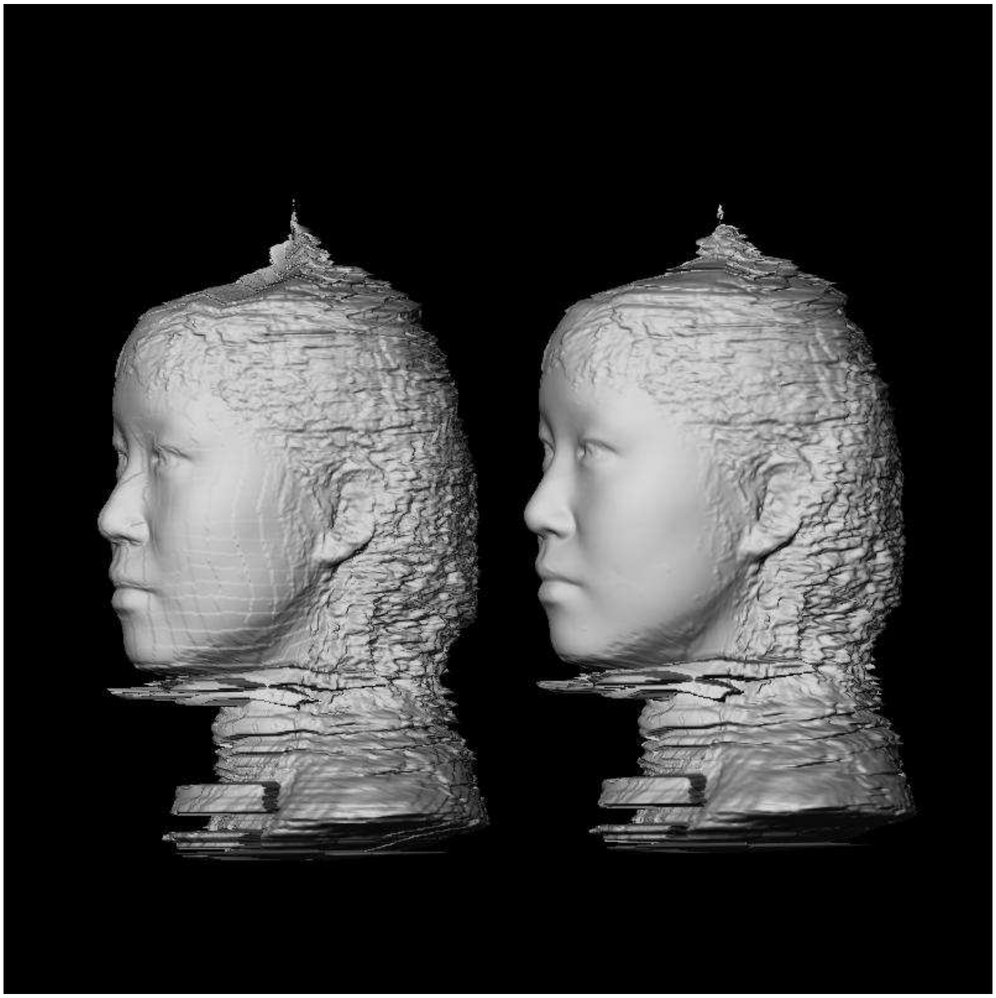


**Plate 1-17: 3-D Caricatures.** The center image shows the original 3-D shape scan. On the left is a caricature made by exaggerating the differences between the original shape and the shape of the male Caucasian prototype by 50%. On the right is a caricature made by exaggerating the differences between the original shape and the shape of the male Japanese prototype by 50%. It can be seen here how typical European-Caucasian facial traits compare with typical Oriental-Japanese. The nose, for example, is exaggerated in length much more when caricatured against the Japanese prototype. This is because the Japanese prototype has a smaller nose that the Caucasian prototypes nose (which is more similar to the head of the example head).

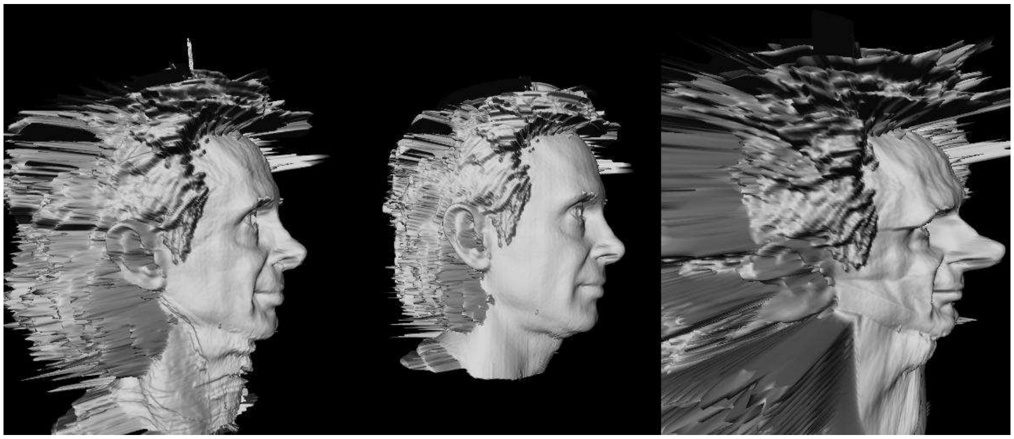
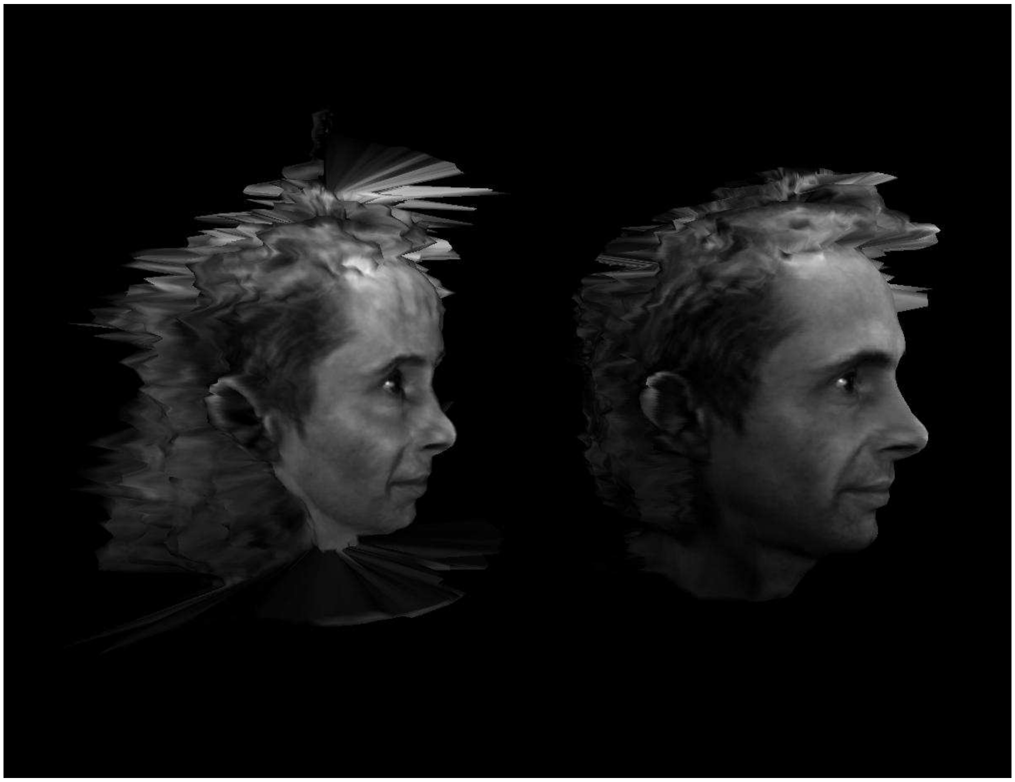


**Plate 1-18: 3-D Gender transformation.** The image on the right shows the original Cyber-scan (with the colour information texture mapped back onto the 3-D shape). Taking the differences between the male and female Caucasian prototypes (from plate 1-15) and adding this information to the face on the left formed the image on the left.



# 3. **Evolution of Bottle Design Based on Consumer Preferences.**

## 3.1 Introduction.

By packaging a product in a certain way it is hoped that the consumer will associate a certain set of ideals and concepts with that product. An obvious example of where packaging has been used to promote a product that otherwise would have been no different from other products in the market place is perfume where the cost of the containing bottle far outweighs the cost of the liquid it contains. The ‘Pammy’ Virgin Cola bottle was simply virgin cola in a bottle designed to have the same curvature as Pamela Anderson. Products like these are bought more from the concepts that the consumers relate with them than the products themselves. It is obviously important, therefore, to design packaging that conveys the required ‘message’. Presently, a great deal of time and effort is employed by costly design agencies, which use a combination of experience and guesswork to create alternative solutions. It is common for consumer research to be conducted to evaluate these alternatives before the product is finally launched. In general, this section of the thesis examines whether ‘consumer research’ can be brought into the marketing process at an earlier stage and, specifically, whether consumer preferences can be used to design the appearance of a product.

The use of computer graphics to create simulated packages is becoming widespread. It is more cost effective than producing physical mock-ups of designs for two reasons. First, the actual production costs are lower since no physical object is to be created, and second, the virtual mock-ups can be produced much more rapidly that their physical counter-parts. The Computer Aided Design (CAD) systems currently available for producing computer graphic prototypes require an expert operator and are generally only used to take designers’ ideas and form *3-D* representations for better visualisation. If it is possible to tap into people’s reactions to various designs and concepts then it may be possible to give control of a CAD system over to a consumer and to let them design their own packaging. This could be achieved by the placement of some interface between the mechanics of a CAD system and some measure of consumer preference. For example, by displaying a selection of packages and asking a consumer which ones best convey a certain concept, it may be possible to generate packages which more strongly display the desired characteristics. One technique that could be applied as an interface is a genetic algorithm (GA). By specifying packages as gene strings and applying consumer choices as a selection pressure it may be possible to evolve an highly suitable package design. Genetic algorithms are appropriate for this style of problem since they allow for the optimisation of many different parameters simultaneously. The prototyping techniques described in the third section of this thesis are not appropriate. The prototyping technique allows the definition of meaningful locations within a multidimensional stimulus space. These are used to create vectors which map instances of one class of stimuli onto the domain of another (e.g. a prototypical young male face and a prototypical old male face can create a vector to age an individual's face). One could imagine that similar techniques could be applied to packaging, perhaps forming a prototypical perfume box. However, for the prototyping technique to work well, the stimulus set must be homogeneous (meaning that all members of the set must share a common set of features that can be delineated). Packages do not form homogeneous sets in this sense and so new technique have been developed to enable the studies presented here to be performed.

Industrial organisations have expressed an interest in this work and have suggested various concepts and objects for which suitable examples could be evolved. Shampoo bottles were selected since it was thought that these would provide a relatively tractable application for computer modelling. Thus, the aim of this study is two fold:

1) Discover whether particular concepts are meaningful (in the sense that subjects will agree on their level of presence in a specific object).

2) Find what the physical manifestations of these concepts are through the development of methods to evolve shape and colour to match perceptual preference.

## 3.2 A Brief History of Beauty.

Are aesthetic judgements universal? Do we intuitively find certain forms and arrangements more pleasing to the eye? Socrates suggests that certain figures, "are beautiful not, like most things, in a relative sense, they are always beautiful in their very nature, and they carry pleasures peculiar to themselves" (Philebus 1757). Antiquity provides a ratio known as the golden section (and approximately equal to 0.618)[6] as a key to beauty. The Golden Section was a great influence on Medieval and Renaissance architecture. Whether this was because of its intrinsic beauty or because of ease of construction is debatable (for example, an arch of 'golden' proportions can be marked out using nothing more than a rope fixed at one end - Coldstream 1991). Regardless, by the 16th century the golden ratio was in widespread usage in all manner of designs and is still a common ratio used in construction (such as most fruit juice cartons!). Discussions on the aesthetic properties of the 'golden ratio' have continued to the present day (Green 1995). Fechner (1876) was the first to apply scientific method to investigate subject preferences for the ratio. Fechner presented subjects with ten white rectangles on a black table. The rectangles ranged in proportion from 1:1 (square) to 2.5:1 and all had the same area. A clear preference for the rectangle with the proportion closest to the golden ratio was shown. Fechner also studied subject preference for ellipses (although Witmer did not publish the results until 1894). A preference (42% of subjects) was shown for the ellipse having its major to minor axis in a ratio of 1.5:1, whilst only 16.7% of subjects preferred the elipse with its axis in the 'golden' ratio (Witmer 1894). Subject preferences for McManus (1980) investigated rectangles with various ratios of width to height more recently. A bi-modal preference was found with one (less significant) preference being based on the square, and the other upon a ratio close to that of the golden section. The consistancy of subject preference for shape is mirrored in literature concerned with subject preference for colour. Granger (1955), Guildford and Smith (1969) and Helson and Lansford (1970) conduct colour preference studies where colours are presented to subjects on 'colour chips'. All three studies place hues in the following (decreasing) order of popularity: blue, green, purple, red and yellow. More recently, McManus *et al* (1981) attempted a replication of these studies and found that the majority of subjects preferred blue hues and least liked yellow hues (no consistent pattern emerged for the intermediary hues). It should be noted here that the previous studies of colour preference were all carried out using abstracted colour presentations. That is, the colours were shown on colour chips, and thus without a meaningful shape. It is interesting to postulate then that when a colour is presented on a form, preference for the colour could be modulated (indeed the same could be true for colour modulating the preference for a shape).

---

[6] The Egyptians are known to have had a value for the golden section within 0.5% of 0.618 (Ghyka 1946/77).

On a level slightly more relevant to the main tops in the thesis, faces offer a highly familiar, complex stimuli set which have previously been used in studies on preference (attractiveness). Perrett, May & Yoshikawa (1994) ask subjects to rate the attractiveness of prototype faces which differ only in shape (one shape derived from faces rated highly for attractiveness, and one shape derived from faces rated less-highly for attractiveness). Japanese and Caucasian observers both preferred the same (more attractive) face shape. This study shows that subjects not only agree to a high degree about what is attractive, but that this agreement is consistent across different cultural backgrounds.

The studies presented in this short review of aesthetics support the notion that 'beauty' is not 'in the eye of the beholder' (as commonly believed) with subjects actually showing a large degree of consistency in their preferences for shape and colour. Since people can agree on aesthetic values in simple and complex stimuli then it might be possible to optimise a given stimulus with respect to these aesthetic qualities. While the initial aims of this work are to design 'simple' bottle shapes, in principle the techniques can be extended to faces (see section 5).

### 3.3 Introduction: The Genetic Algorithm.

Since their conception in the mid 60's (see Holland 1992), evolutionary algorithms have provided many novel and interesting solutions to problems in a wide variety of domains. From the design of turbine blades and simple autonomous robots to the optimal geographical placement of a series of gas filling stations, the fundamental techniques involved in all these examples are identical and are based on nature's own design strategy (Goldberg 1989). In a very simplistic sense this is survival of the fittest. Many different entities are created and vie for the chance to reproduce and so pass on their genetic information. Thus more successful genetic information is created whilst less successful solutions are discarded. The most popular and successful type of evolutionary algorithm is currently a 'Genetic Algorithm' (GA).

There are 3 main areas of interest when developing a GA:

1) How to map entities onto gene strings (encoding);
2) How to calculate the successfulness of a solution (fitness);
3) How to allow the combination of solutions (breeding).

Encoding of problem space can be achieved in many different ways, but in essence most GA's use a simple 'alphabet' from which 'words' can be made up which describe an entity. Roughly speaking, in human DNA this alphabet consists four different nucleotides: adenine, thymine, cytosine, and guanine (A,T,C,G). These are put together to make a long word (our gene string, genetic code or genotype) which fully describes our individual physical make-up. An individual's genotype is the product of millions of years of evolution, and is created as a combination of the parents' code. Which code is 'good' and which code is not is defined by a 'fitness function'. In natural evolution, the fitness of an entity can only be gauged with hindsight and relates to the successfulness at becoming an ancestor. In a GA the fitness function is used to evaluate how well an item fulfils a specified selection criteria. This is then used to define how many offspring the entity will produce.

In these studies investigations are being conducted about how various attributes of bottle design relate to perceptual judgements about that bottle. In this case, the fitness function is related to what percentage of people sampled have a preference for an individual stimulus bottle and is used to preferentially enhance the appearance of the genetic code for these likable traits in future bottle generations.

Finally, what does 'breeding' mean? How can bottles possibly breed!? Once a problem domain (shampoo bottles) has been encoded into a simple alphabet (in our studies we used just two 'letters' 0 and 1) an initial 'random' population can be created by creating an initial gene pool (random strings of a specified length of 0's and 1's) and generating the bottles which these strings specify. For two bottles to breed, one simply takes a portion of the genetic string from one bottle and the rest of the genetic string from the other bottle and combines the two (a percentage of the letters are changed randomly to simulate mutation) and the bottle generated which the resultant string specifies. Mutation is important since it helps stop stagnation. If there is not enough diversity in a population it may not be possible to create any improvements. By randomly mutating a percentage of the letters, new areas of the problem space can be explored and the population is less likely to get stuck at a local maximum.

The application of evolutionary algorithms to optimise aesthetic values is still in its infancy and there is limited literature on the subject.

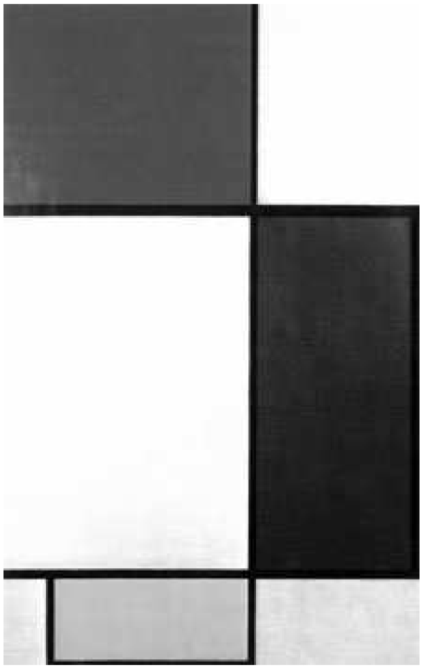


**Figure 3.1: Mondrian (1928) - "Composition with Red, Yellow and Blue"**

In a study of Mondrian's art, McManus *et al* (1993), tested whether an evolutionary technique could be used to allow subjects to create their own Mondrian style paintings. A first study demonstrated that subjects preferred Mondrian's paintings to those created by the computer (showing that consistent preferences did indeed exist). A second study used a hill climbing approach, whereby a small random modification to the images was made (a line moved up or down, left or right) and the painting rejudged. If the alteration improved the quality of the solution the alteration was kept and a further modification made; otherwise the alteration was discarded and a different one tried. The idea being that subjects would select pictures that they preferred (from a generation of 25 pictures) and evolve their own optimal piece. It was expected that the final composition would contain the most pleasing arrangement for a specific subject. Neither consistency between the final paintings was found, nor was there directional preference towards Mondrian's sense of proportions. In this case it was not possible to evolve preferred shapes.

There are several possible reasons for the failure of the McManus *et al* study. Perhaps the algorithm used was not suitable to the problem. A better solution may have been to use a GA. Alternatively it may simply be the case that the reason Mondrian's sense of proportions were so popular in the first study was because the subjects had been previously exposed to them. Indeed current priming studies show that priming can occur months, even perhaps years before a recognition event and still have an effect (Bruce 1994). Bruce's studies have shown a mere exposure effect where priming can occur from the posters advertising for subjects to take part in the experiment. As has been previously noted, advertising plays an important role in determining what concepts consumers associate with what product. For this reason it is not possible to use existing shampoo bottle designs in an experiment to find out what concepts are associated with which design. An entirely novel stimuli set needed to be created. One that would mimic the variability inherent in the current range of shampoo bottles, but one that would not be so reminiscent to instantly remind one of a certain brand.

A study conducted by Johnston and Franklin (1994) allowed subjects to evolve attractive face shape using a GA. The technique used is rather crude and involves 'cutting and pasting' a variety of facial components together (in a configuration specified by a gene string). The computer first generated a small population of 'random' faces. Each of the populations of faces was rated by a subject for beauty and the fittest faces were bred together. The process was reiterated for several generations until the change in fitness from one generation to another was negligible. The faces evolved by subjects using this methodology were consistently different in shape from average. This supports the findings of Perrett *et al* (1994) that the concept of facial beauty is not 'in the eye of the beholder' but rather involves shape attributes about which subjects agree.

## 3.4 Aims.

There are two aims of the study:

1) The assessment of whether subjects agree on perceptual attributions.
2) The evolution of bottles that subjects agree displays a specific perceptual attribute.

## 3.5 Genetic Evolution of Bottles Experiment.

### 3.5.1 Preparation of Stimuli.

The bottles were created using specifically designed software. Examination of the shape of shampoo bottles currently on the market indicated that most were rotationally symmetric. Shampoo bottle shapes could thus be created by the rotation of a curve about a central axis (see Figure 3.2). A method of generating the curve was required which would be flexible enough to allow the creation of the majority of bottle shapes presently available whilst making the generation of unrealistic bottles unlikely. There are many types of curve that could have been used but most fall into the category of parametric cubics. Parametric equations take the form of a function that takes a parameter and returns a value. For a location (**t**) along the curve (with **t=0** at the start of the curve, and **t=1** at the end of the curve) the **X** value is given by a function $\mathbf{F_x(t)}$ and the **Y** value by a function $\mathbf{F_y(t)}$. To draw the curve a set of successive values of **t** (usually evenly spaced) going from 0 to 1 are used to calculate successive values of **X** and **Y**. These successive **X,Y** co-ordinates are joined together using straight lines. If the sampling rate is high enough (that is, there are enough values of **t**) then the result is a curve of smooth appearance. The shape of a curve is generally specified by a set of control points. In *2-D* these are (**X,Y**) co-ordinates whose values are used to generate the functions $\mathbf{F_x}$ and $\mathbf{F_y}$. **X** values are used to generate $\mathbf{F_x}$ and **Y** values are used to generate $\mathbf{F_y}$ (this scheme easily extends to higher dimensions). With long curves (that is, curves with many control points) there is little point in considering the location of control points at the start and end of the curve when calculation the position of the curve in the middle. So long curves are generally considered to be made out of segments **Q(t)** with a local subset of control points specifying the curve segments shape. Continuity of curves is an important concept since it describes how curve segments join together. There are two types of continuity *geometric continuity* and *parametric continuity*. If two curve segments join together then the resulting curve is said to have $G^0$ *geometric continuity* (note, this can include shape corners, but not gaps). If when the two curve

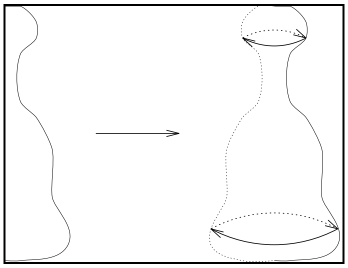


**Figure 3.2 shows extrusion of a curve to produce a bottle shape.**

segments join together they have the same direction then the curve is said to have $G^1$ geometric continuity. In the case of $G^1$ it is important to note that the magnitudes of the curve segments at that join point need not be equal (e.g. if the curve was used to position a camera at time proportional to t then there may be a jolt at the intersection of the two curve segments). If when the two curve segments join together the two tangent vectors at the joint point are equal (i.e. equal in direction and magnitude) then the curve has first-degree continuity in the parameter t, or *parametric continuity*. In this case the curve is said to be $C^1$ *continuous*. Similarly, for higher order derivatives, if the direction and magnitude of $d^n/dt^n[Q(t)]$ are equal (for 0..n) then the curve is said to be $C^n$ *continuous*. For example, with the analogy of the camera path one would probably desire both the velocity *and* acceleration to be continuous to provide a smooth track, this would require $C^2$ continuity.

To allow the variations required in curve shape to produce shampoo bottle cross-sectional halfs, the curves should allow for a variety of connectivity constrained at different regions by various means.

1) The curve must start and terminate at the center top and center bottom of the bottle respectively. This is to allow a solid shape to be generated when the curve is rotated such that there are no holes (or overlap) at the top or bottom.

2) The top of the bottle should be constrained such that a lid section is generated. This may (or may not) have angular sections, but should allow for fluctuations as apparent in existing shampoo bottles.

3) The neck to the shoulder of the bottle should widen and become the body of the bottle that should be allowed a large variation in shape and curvature.

4) The bottom of the bottle should be flat.

With these restrictions in mind, it is possible to judge commonly used parametric cubic curves for their appropriateness to the task of generating the shampoo bottle cross-sectional half curves.

Hermite curves (Foley *et al* 1990) are specified by two end points and starting and finishing vectors. Both start and end points are interpolated in the curve and the direction of the curve at the start and end is totally specified by their associated vectors. To allow for $C^1$ continuity then, the vector at the end of a curve section must be equal (or opposite) to the vector at the start of start of the next curve section. This does not give a high degree of shape flexibility and so makes Hermite curves unsuitable.

Bezier curves (Bezier 1970, 1974) have four control points and interpolate the first and the last. The middle two control points are used to control the shape of the curve and are approximated, but not necessarily interpolated within the curve. The problem with Bezier curves is similar to that for Hermite curves and in general is that the techniques are not flexible enough. For example, for some shampoo bottle shapes it may be important for a specific control point to be interpolated within a curve, yet for another shape it may be required that the same control point is merely used to deflect, and is not itself included in the curve.

Splines were originally springy strips of metal used by draughtsmen to plan surfaces of boats, planes etc. Weights were attached to the spline in order to change in shape and the curve generated would have $C^2$ continuity. This is one degree more continuity than present in either the Hermite or Bezier curves and so splines can produce smoother curves. Natural splines, which closely model the original strips of metal, have the unfortunate property that in order to calculate a point of the curve the position of all n control points must be taken into account. For the purposes of generating shampoo bottles, it is important that local control be acquired over the curve so that a change in the body of the bottle does not change the lid. This is required so that entities that have successfully evolved a pleasing lid do not have to compromise it by improving their body. B-splines (Bartels 1987) consist of curve segments whose shape depends on only a few control points. B-splines have the same continuity as natural splines but they do not necessarily interpolate their control points. Cubic B-splines approximate a set of m+1 control points $P_0, P_1, \ldots P_m$, $m>=3$. This is achieved by the creation of a series of m-2 curve segments $Q_3, Q_4 \ldots Q_m$. These are joined together at knots, so including the start and the end points of the curve as knots there are m-1 knots in total. Different coefficient values at each knot specify how the curve acts at the joint and effect continuity. This method provides the required level of flexibility and control. Uniform B-splines require these knots to be at equal spacings with reference to the parameter t. To allow maximum control, it was decided to use Non-rational Non-uniform Cubic B-splines for the generation of the shampoo bottle shapes. Here the knot values are used to allow for multiple control points to be placed coincidentally. This is used to create sharp corners. For example, the knot sequence to specify the three curves in Figure 3.3 would be: a) 0,1,2,3,4 b) 0,1,1,3,4 c) 0,1,1,1,4

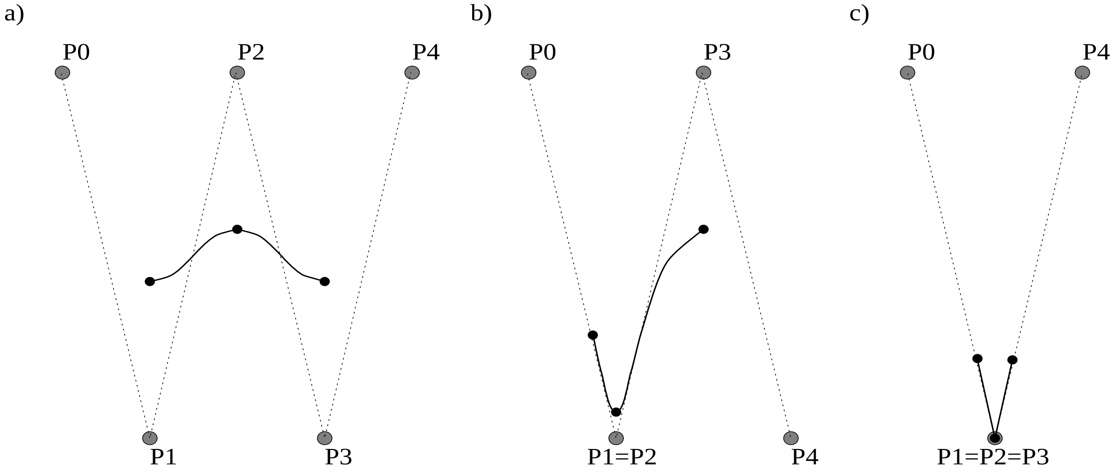


**Figure 3.3: The effect of overlapping control points on non-uniform non-rational cubic B-spline curves (dark line).**

Each shampoo bottle was specified by 11 control points which were used to generate a B-spline curve (see figure 3.3). Each control point has a related multiplicity (or “power”) that can vary at most between one and three. This power controls the minimum distance between the curve and that point (see Figure 3.3): A power of one merely causes a deflection of the curve whilst a multiplicity of three forces the curve to pass through the control point. The locations which control points could take was broad enough to enable the creation of a wide variety of bottle shapes but restrictive enough such that most locations of control points would yield a believable bottle shape (see Figure 3.4). The generated curve is rotated about the vertical axis to form the bottles surface, and colour and sheen values are applied (note sheen was not used in Experiment 1).

**Figure 3.4**

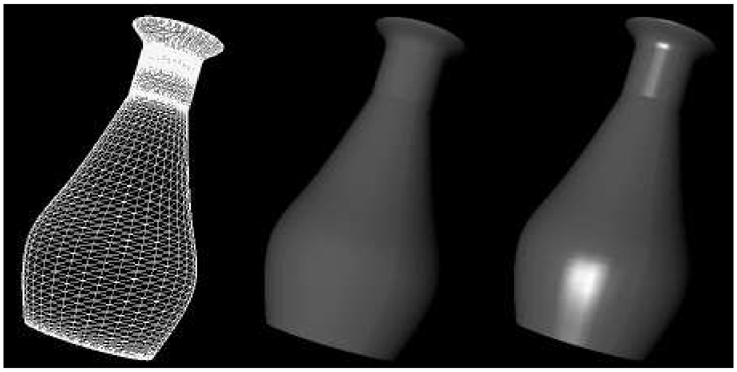


Figure 3.4. On the left is shown the wire frame (quadmesh) version of the bottle. The quadmesh contains 36x***height*** data points. The height varies because there may be overlapping data points. There are 20 sampled data points between each of the 11 control points and so the height is in the order of 200 points. The middle figure shows the quadmesh rendered with diffuse lighting. The final image, on the right, includes specular lighting effects that were included in Experiment 2. The Silicon Graphics *3-D* toolkit (Inventor 1.0) was used to render the bottles and hardware routines used to calculate the effects of the values for diffuse light colour/amount and the specularity (Werneck 1995). A record with the following fields and types specifies each bottle:

| | |
|---|---|
| *Int* | $POWER_1$ |
| *Real* | $X_1$, $Y_1$ |
| *Int* | $POWER_2$ |
| *Real* | $X_2$, $Y_2$ |
| " | " " |
| " | " " |
| *Int* | $POWER_{11}$ |
| *Real* | $X_{11}$, $Y_{11}$ |
| *Real* | R,G,B |
| *Real* | Sheen |

(see Figure 3.5)

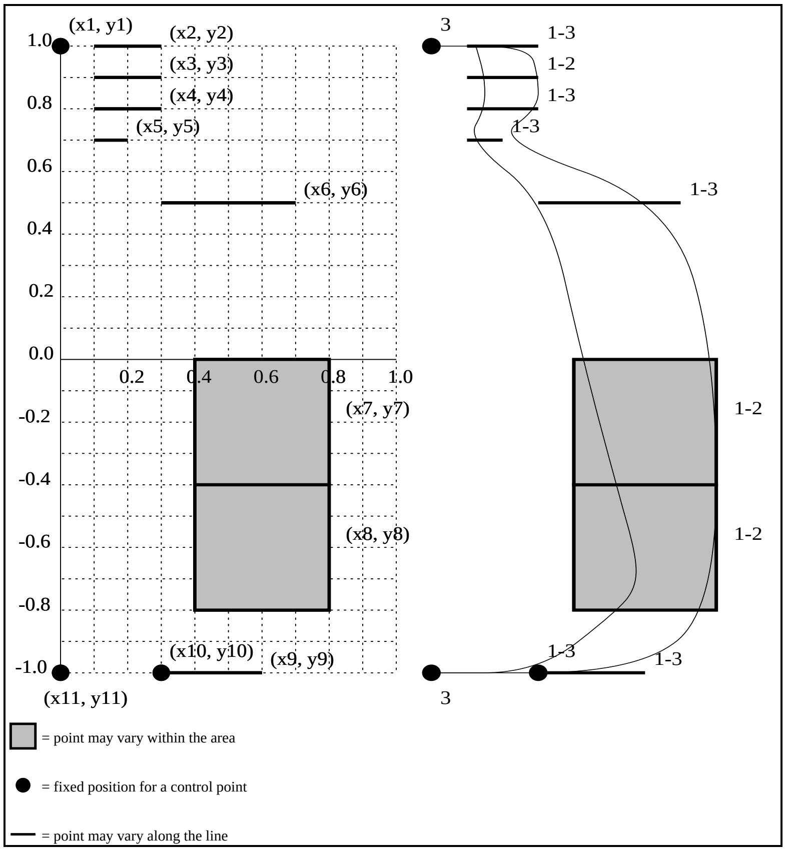

(x1, y1)
(x2, y2)
(x3, y3)
(x4, y4)
(x5, y5)
(x6, y6)
(x7, y7)
(x8, y8)
(x9, y9)
(x10, y10)
(x11, y11)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
0.2
0.4
0.6
0.8
1.0
3
1-3
1-2
1-3
1-3
1-3
1-2
1-2
1-3
1-3
3
▨ = point may vary within the area
● = fixed position for a control point
— = point may vary along the line

**Figure 3.5.** The potential position of control points and their powers is shown. Two possible outlines of bottle stimuli are illustrated on the right.

An initial population of 90 bottles was created from randomly computed values for each of the bottle record fields. The stimuli were printed out with a black background at 150 dots per inch, 18 per A4 sheet size, using a Mitsubishi sublimation printer (S3410/30). Four colour ink roll was used (Magenta, Cyan, Yellow and Black) and the stimuli were cut out individually to give 90 cards (4.5cm × 2.5cm each).

### 3.5.2 Breeding Algorithm.

For the successive generations the selection rate for each bottle in the previous generation was used to calculate the current generation. A bottle was selected for mating with a probability related to how many times it had been rated as good (i.e. the more highly a bottle was rated, the more likely it was to be mated).

A specific bottle was selected for breeding by the ratio of the cube of its ratings to the sum of the cubes of the ratings of each of the bottles in the whole population.

Ratings were cubed in order to give more weight to highly rated bottles and to hasten the evolution. Higher exponents proved to limit the breeding pool too much, and lower valued exponents proved not to give enough weighting to the best bottles

Two different bottles are selected and mated as follows: Integers are ‘mated’ by simply selecting the value from either the first bottle or the second (with equal probability); Real numbers are ‘mated’ by a rather longer process (Syswerda, 1989). Real values are first scaled to be within the range zero to the maximum gene value ($2^{24}$, since bits are either 0 or 1 and there are 24 per gene). Values for corresponding fields from two bottles that are mating are then converted into binary and then for each bit position (0..23) a bit value selected at random from the corresponding bit position from either of the two genes. For example, consider point 2 (x2 and y2 on Figure 3.5), point 2 may appear anywhere along the shown line, but the y value is fixed. The x and y values are first converted into a 24 bit binary word using standard binary notation. Floating point representations are not used because of the discontinuities in the representation.

The resulting y value is the same as the parents (since it was the same for both bottles):

Parent 1 y2 = **111**110 **11**1111 **0**00000 **011**010
Parent 2 y2 = 111**110** 11**1111** 0**00000** 011**010**
Offspring y2 = **111110 111111 000000 011010**

The resulting x value will be a product of uniform crossover:

Parent 1 x2 = **11**0000 **00**1**000** 1**100**11 1**0**1**111**
Parent 2 x2 = 00**1111** 11**0**100 **1**110**00** **0**1**0**010
Offspring x2 = **111111 000000 110000 000111**

This mating procedure is performed for all the real valued field of the gene string.

A mutation rate was set to 0.1% per bit in keeping with most current GA research (Bäck and Schwefel 1993). This means that on average every $1000^{th}$ bit it flipped from 1 to 0 (or vice-versa).

## 3.6 Experiment 1.1

### 3.6.1 Subjects.

40 subjects who had not taken part in the similar experiments were used in each population group. In total 320 subjects (approximately equally distributed in terms of gender) took part in the following experiment.

### 3.6.2 Procedure

The experimenter first shuffled the 90 cards from the current population and handed them face up to the subject a pack. Each subject was asked to sort the cards into two piles according to whether they liked or did not like the bottle depicted. After the subject had completed the task, the bottles placed in the 'like' pile were recorded on a score sheet. This method was applied for two sets of forty subjects for the initial 'random' population and then for one set of forty subjects for the subsequent populations.

### 3.6.3 Evaluation of the GA and Design of Stimuli Set.

In order to discover whether the stimulus type proposed for the evolution of bottles was suitable (i.e. that subjects could agree on particular perceptual traits when selecting bottles varying is shape and colour) the initial 'random' population of bottles was rated by two different groups.

### 3.6.4 Results.

A 'summed-like' score was calculated for each bottle by counting how many times it was placed in the 'like' pile. The total for each bottle from the first forty subjects was compared with the total from the second forty subjects for the same bottle. A Spearman’s r correlation was calculated across all 90 bottles and it was found that there was a significant correlation ($r = 0.719$, $n=90$, $p< 0.001$).

### 3.6.5 Discussion.

The significant correlation between the two groups supports the hypothesis that subjects agree on which bottles are liked. It also successfully demonstrates that the set of bottle stimuli is appropriate for ratings to be made upon them.

## 3.7 Experiment 1.2 - Test for Successful Evolution

Five generations of bottles were generated with 40 subject rating each new population and the ‘like’ data used as a fitness value to produce the next generation (as described in Breading Algorithm above). In order to discover if the GA was successfully evolving better bottles, the best group of bottles from each generation was compared to each other. If better bottles were being evolved then there would be a progression with bottles from each successive generation being rated more preferentially with the best/most popular bottles appearing in the last generation.

### 3.7.1 Procedure.

The ten most popular bottles (i.e. the ones that were selected most often for the ‘like’ pile) were extracted from each generation and pooled together to give a set of 50 stimuli. As in experiment 1.1, these stimuli were rated by 40 subjects.

### 3.7.2 Results.

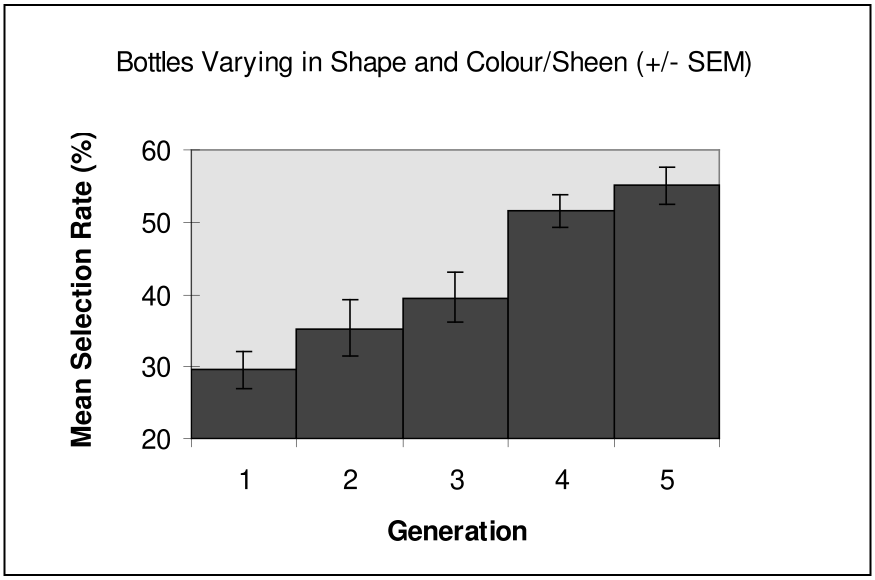


**Figure 3.6.** Selection rates are shown for bottles in the final test population (that is the population made up from the best bottles from the previous generations). The graph clearly shows that more evolved bottles were preferred over less evolved bottles.

A one-way analysis of variance was performed on the selection rate for each bottle from each generation. This indicated that the generation from which bottles originated had a significant impact on rating [$F_{4,9}$= 13.3, $p< 0.05$] and as is shown in Figure 3.6 there is an increase in bottle selection rate with generation number.

### 3.7.3 Discussion

The results show that the genetic algorithm successfully created bottles that subjects more often liked

**Figure 3.7**

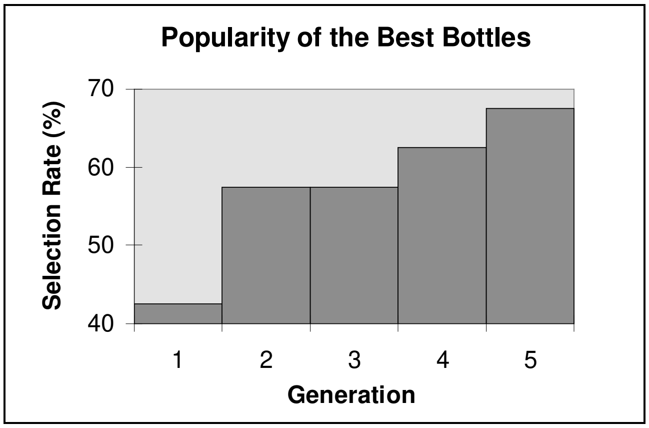


Figure 3.7 shows the selection rate of the most popular bottle from each generation (selected during the final test population). A clear preference for more evolved bottles is shown.

During generations one to five subjects could have selected the bottles that they disliked least and were in fact evolving towards average (i.e. bottles that did not have anything wrong with them) and not towards any specific better type of bottle. Figure 3.7 shows that this is not the case. 67% of the subjects chose the top bottle which came from the fifth generation whilst only 43% for the top bottle from the first generation. This illustrates that bottles were getting better in a direction that was not average (since an average bottle would have appeared in all generations).

### 3.8 Experiment 1.3 - Determination of Shape Evolution.

The previous experiment successfully showed that the bottles evolved over the five generations increased in popularity. The first two experiments (2.1 and 2.2) show that subjects agree on bottle preference, although it could be that the subjects were selecting bottles based on trivial decisions and not taking the whole bottle appearance into account. For example, subjects may have been basing their decision purely on the colour of the bottles, or indeed purely on their shape. To determine if this was the case or not, these factors must be analysed independently of each other. The evolution of the colour genes can be assessed with a statistical analysis of the RGB values, but since shape factors are less easy to qualify a further supplementary study (Experiment 1.3) was conducted.

### 3.8.1 Method.

The role of shape was examined by creating stimuli with the shape from the 10 best bottles from each generation. These were re-rendered to have the same surface colour (mid grey). Subjects (40) then sorted the bottles in the standard manner.

### 3.8.2 Results

A one-way analysis of variance was performed on the selection rate for each bottle from each generation. This showed that the generation from which bottles originated had a significant impact on rating [$F_{4,9}= 6.972$, $p< 0.05$]. Similarly, when colour was included, there is also an increase in selection rate as generation increases. Interestingly, the F value is not as high as when colour was included ($F_{4,9}=13.3$). This is perhaps indicating that although the shape is evolving, the colour is also evolving and playing a part in the attractiveness of a specific bottle.

## 3.9 Results: Evolution of Colour.

With no selection pressure applied (i.e. bottles left to breed at random) no evolution would occur and the average population value for each gene would remain at 50% of the maximum value. A significant deviation away from this population mean starting value can then be seen as evidence of evolution.

A one way analysis of variance was performed on the RGB (colour) fields of the bottles from population one and population five. This indicated that there was a significant difference for the mean Red value [$F_{1,179}= 5.87$, $p< 0.05$] between the 90 bottles from the first and the 90 bottles from the fifth generation but not for the mean Green or Blue values ($F_{1,179}= 0.13$, $p> 0.05$) and ($F_{1,179}= 0.50$, $p>0.05$) respectively. This can be more clearly seen in Figure 3.8.

**Figure 3.8**

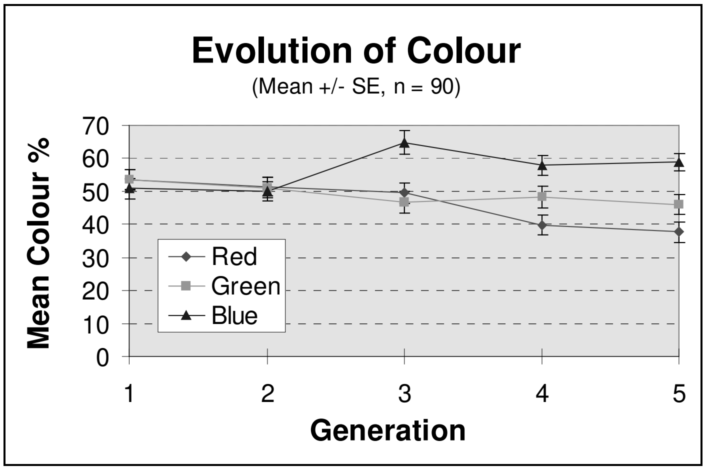


Figure 3.8 shows how the mean values of the red, green and blue colour genes all start at value of 50% (of the maximum colour value) and then evolve to different amounts after five generations.

### 3.10 Experiment 1: Discussion.

The results presented in experiment 1.3 indicate that both the shape and colour of the bottles are playing a role in their evolution. Over two thirds (67%) of the subjects chose the most popular bottle from the fifth generation (which, coincidentally, was also the most selected bottle from all populations) as opposed to under half (47%) for the most popular bottle from the first generation (Figure 3.7). Thus, it is clear that subject responses were allowing the evolution in shape to produce more popular designs (if no increase in mean ratings had occurred then it would only have been fair to deduce that subjects were evolving bottles according to colour). Counterwise, is it possible that subjects made their selections purely on shape attributes? This is unlikely as an analysis of RGB values shows that colour is also evolving. Figure 3.8 clearly shows the divergent evolution of the different colour genes (the mean R, G and B values initially group around the mid value, but by generation five the mean values for both the Red and Blue genes are significantly changed). All this suggests that both colour and shape are successfully evolving in harmony to produce appealingly shaped bottles with appropriate colouration.

# 4. **Optimising for Consumer Perception of Specific Marketing Concepts.**

## 4.1 Introduction.

The first optimisation experiment demonstrated that GAs could be employed to evolve stimuli towards more desirable shapes and colours using subject preferences. An extension of this research would be to find if the same technique could be used to evolve along more specific consumer concepts rather than general aesthetic preferences. The concepts were provided by industrial collaborators and were chosen to be salient to the shampoo bottle market. In order to test whether the evolution occurring was successfully improving the perceived suitability of the bottles to the specified criteria, two concepts were selected and an attempt was made to evolve separate populations for each. One predicts that if bottle concepts can be successfully evolved then:

(a) Bottles from successive generations should achieve higher ratings;

(b) Bottles bred for one quality should obtain higher ratings for that quality than bottles bred for the alternative quality.

## 4.2 Computational Methods.

### 4.2.1 Colour Parameters.

In experiment 2 bottle surfaces were assigned seven colour properties (as opposed to the three local colour properties from experiment 1): RGB values for red, green and blue local (diffuse) colour; RGB values for red, green and blue reflected (or sheen) colour; and an amount of a property S which relates to the overall surface reflectivity (high values produce more localised ‘metallic’ highlights, low values produce more matte diffuse reflection, Wernecke 1995) (see Figure 3.4).

### 4.2.2 Genetic Algorithm: Fitness value.

A fitness value was calculated for each bottle in a similar way to experiment 1, except in Experiment 2 subjects were rating the bottles (from 1 to 5, 5 being the best). The values from all subjects were summed to give a fitness value. All bottles that had less than the average fitness value had their fitness' set to 0. This was performed in order to expedite the evolution by killing off low performers as well as allowing the most popular to procreate. With the remaining values, a specific bottle in the remaining sub-population was selected for breeding by the ratio the cube of its ratings to the sum of the cubes of ratings of each of the bottles in the sub-population. Again, fitness values were cubed in the above expression in order to give more weight to highly rated bottles to hasten the evolution.

### 4.2.3 Modal Bottles.

It cannot be assumed that the distribution of a specific gene's values within a population is uni-modal. For data that had more than one peak in its distribution, mean values may not convey useful information. For this reason, when creating an average bottle for a population, a mode value should be taken rather than a mean. Modal bottles were created in the following way.

Each gene code making up a bottle was in the range [0...1]. These were quantised into the ranges [0 to <0.1], [0.1 to <0.2], ... , [0.9 to 1.0] and the value 0.05, 0.15, ... , 0.95 recorded respectively. This was repeated for each bottle in the population. The modal (most frequent) value was then obtained for each gene parameter constituting the bottle descriptions. The modal bottles for a given generation and categories were generated using these modal gene codes.

## 4.3 Experiment 2.1.

### 4.3.1 Procedure.

A constant pool of 20 female subjects (mean age 30, minimum age 25 and maximum age 35) was used throughout the study (except during the ratings of the modal bottles). In each phase of the study the following method of data collection was performed. An individual subject was given a card (4 by 6 cm) displaying a single shampoo bottle (printed to near-photographic quality on a dye-sublimation printer at 150 dpi). The subject was then given a category on which to rate the stimuli (using a 1-5 Likert scale, 1 = low rating) and their response recorded. This was repeated for the requisite number of bottles and categories.

### 4.3.2 Generation 0.

The first generation consisted of 90 shampoo bottles that were generated from random genetic material. The bottles were presented in a random order and one of the following industrially inspired categories/consumer concepts given to the subject with which to rate it. This was repeated until all 20 subjects had rated all 90 bottles for all five concepts. The five concept were:

> Nourishes the roots,
> Pampers your hair for that special occasion, (*pampers*)
> Leaves your hair shiny and healthy looking, (*shiny*)
> Cares for and protects the hair,
> Is exotically enriched to leave the hair soft.

### 4.3.3 Results: Consistency of Ratings.

Table 1. Kendal’s Coefficient of Concordance Analysis[7] (Howell 1992).

| | W | $Chi^2$ | *p* |
|---|---|---|---|
| Nourishes the roots | 0.14 | 237 | <0.0001 |
| Pampers your hair for that special occasion | 0.14 | 236 | <0.0001 |
| Leaves your hair shiny and healthy looking | 0.17 | 294 | <0.0001 |
| Cares for and protects the hair | 0.14 | 248 | <0.0001 |
| Is exotically enriched to leave the hair soft | 0.13 | 221 | <0.0001 |

From this analysis it can be seen that subjects show a significant level of agreement in their ratings of bottles along each of the five concepts. However, this agreement does not mean that each of the rated dimensions is independent.

### 4.3.4 Independence of Ratings

To analyse the degree to which rated concepts were independent, total ratings were calculated for each bottle by summing all 20 ratings from the individual subjects for each of the 90 bottles. A Spearman’s *r* correlation performed on all possible pairs of comparisons (see Table 2) revealed that each of the consumer concepts were significantly correlated (df= 89, p<0.05 each comparison).

**Table 2.** *r* values relating different concepts.

| | *Care+p* | *Exotic* | *Nourish* | *Pamper* | *Shiny+h* |
|---|---|---|---|---|---|
| Care+p | | | | | |
| Exotic | 0.47 | | | | |
| Nourish | 0.68 | 0.52 | | | |
| Pamp | 0.36 | 0.77 | 0.40 | | |
| Shiny+h | 0.59 | 0.56 | 0.78 | 0.43 | |

---

[7] Kendall's W is 0 for total disagreement between raters and 1 for total agreement between raters.

A factor analysis (PCA, Table 3) performed on the total ratings for the five concepts and 90 bottles, revealed only one factor (as would be suggested from the above correlations). A second factor with an Eigenvalue of 0.92 was suggested but did not quite reach significance.

**Table 3.**

| Factor | Eigenvalue | % of Var. | Cum % |
|---|---|---|---|
| | 3.23 | 64.7 | 64.7 |
| | 0.92 | 18.4 | 83.1 |
| | 0.42 | 8.4 | 91.5 |
| | 0.22 | 4.4 | 95.9 |
| | 0.21 | 4.1 | 100.0 |

### 4.3.5 Selection of Concepts to Test Evolution.

In order to give the genetic algorithm the best chance of successfully breeding bottles optimized for different consumer concepts it seemed sensible to try and breed based on:

1) A concept about which there was most consumer agreement and

2) Two different concepts that were correlated least in consumer ratings.

From the Kendal's concordance analysis the concept 'Leaves your hair shiny and healthy looking' produced the greatest degree of agreement between subjects. From the correlation of ratings (Table 2) it can be seen that the concept which is most different from (correlates least with) the *shiny* concept is the 'Pampers your hair for that special occasion' category.

By looking only at the (top 10) highest rated bottles from these most separable of categories there is a suggestion that the two concepts are divergent (Figure 4.1). The *Shiny* and *Pampers* concepts were therefore selected for construction of generations 1, 2, and 3.

**Figure 4.1**

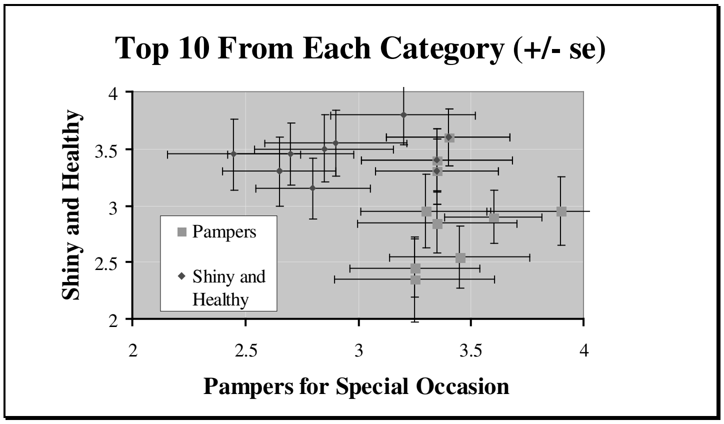


Figure 4.1 shows average ratings for two concepts plotted against each other. It can be seen that the bottles rated highest for *shiny and healthy* and are mostly different to those rated highest for *pampers*.

### 4.3.6 Generation 1.

The consumer concept which showed the most agreement between subjects (***shiny***) and the least correlated with this (***pampers***) were chosen for selective breeding. Bottles were mated with a probability based on their fitness value (as described in the computational methods in section 3.5.2). This created a new population of 90 bottles. This was done twice, once with the fitness values derived from the *pampers* ratings and once with the fitness values derived from the *shiny* ratings. This created a total of 180 bottles (2 sets of 90). These bottles were rated by subjects in the same way as for the generation 0 bottles except that the bottles derived from the *pampers* ratings were rated solely on the *pampers* criteria, and the bottles derived from the *shiny* ratings were rated only on the *shiny* criteria.

### 4.3.7 Generation 2.

Another 180 bottles were generated from the genetic code from generation 1 and their relative fitness value established from the ratings of generation 1. 90 bottles were generated from the generation 1 *shiny* bottles and their *shiny* rating values and 90 were generated from the generation 1 *pampers* bottles and their fitness values. These 180 bottles were intermixed and combined with additional stimuli selected as follows: the top 20% (the 18 bottles with the highest fitness values) from generation 1 from the *pampers* group; the top 20% of the bottles from generation 1 from the *shiny* group;

the top 20% of the bottles from generation 0 rated on the *pampers* criterion and the top 20% of the bottles from generation 0 rated on the *shiny* criterion. (NB. five of the best *shiny* bottles from generation 0 were also the best *pampers* bottles from generation 0 and these were only included once.) This gives a total number of bottles for generation 2 of 180+18+18+18+13=247. In order to see if evolution was occurring along the required dimensions, all 247 bottles were rated by subjects on both the *shiny* and the *pampers* concepts.

### 4.3.8 Generation 3 and Modal Bottles.

A final *shiny* and *pampers* population were created to discover whether the two populations as separate wholes evolved diversely (generation 3). To investigate whether the *pampers* bottle selective breeding had created a bottle better suited to the *pampers* category (and similarly whether the *shiny* evolution was creating a bottle better suited to the *shiny* category) modal bottles were created for each concept for generation 3. These were used as stimuli for a new group of subjects who had not previously been exposed to any of the bottle stimuli. 30 subjects (15 male, 15 female, age 20-40 years) resident in St. Andrews were tested. Test cards with the two modal bottles were generated with either the question "Which bottle do you think contains shampoo which would leave your hair shiny and healthy looking?" or the question "Which bottle do you think contains shampoo which would pamper your hair for that special occasion?". The two modal bottles were printed above the question. Two cards were then given to each of 30 subjects, one card asking the *shiny* question and the other card asking the *pampers* question. Side of presentation was randomized (either *pampers* modal on left and *shiny* modal on right or vica-versa). Similarly, order of question was randomized (pampers question first followed by shiny question or shiny question first followed by pampers question). Responses were recorded as correct when the subject replied that, out of the two, the bottle bred for the questioned criteria was most suitable to the questioned asked.

### 4.3.9 Results: Analysis of Genes (Colour).

A three-way analysis of variance was performed on all the colour parameters for the 90 bottles from generations 1, 2 and 3 from the *pampers* and *shiny* bottle sets. Generation 0 items was not entered into the analysis because the 90 items of generation 0 belonged to both *pampers* and *shiny* categories and because each gene set would have a flat distribution (since genes for the initial population were generated at random).

Structure of ANOVA:

| | |
|---|---|
| 1) Concept | 2 levels FIXED (Pampers or Shiny) |
| 2) Generation Number | 3 levels FIXED (Generation 1,2,3) |
| 3) Colour | 7 levels FIXED (RGB local, RGB reflected and S) |
| 4) Bottle Number | 90 levels FIXED nested within Generation Number |

There was a main effect of colour, ($F_{6,1602}$=77.0, p<0.0001) indicating that the intensity levels of RGB were not equal. This suggests evolution in colour was occurring (since the average bottle colour was not mid-grey as it had started in generation 0). Other main effects were non-significant.

More interestingly, colour levels interacted with concept ($F_{6,1602}$=34.9, p<0.0001). This means that the concepts of *pampers* and of *shininess* were portrayed best by bottles with different surface colours.

There was a significant interaction between generation and colour ($F_{12,1602}$=6.7, p<0.0001) indicating that the colour of the bottles changed with the evolution. Bottle colours start out, on average, as mid-grey, but colours move away from grey with successive generations. The overall change in colour with evolution could reflect a general subject preference for particular colours (e.g. generation 3 bottles have higher specular red than generation 1 bottles; this being true both for bottles evolved for the *shiny* and *pampers* concepts).

A significant third order interaction between concept, colour and generation ($F_{12,1602}$=14.7, p<0.0001) establishes that there was a different direction of evolution in colour parameters for the two consumer concepts. For example, by generation 3, low green values are established when bottles are bred for pampers, but this is not the case when bottles are bred to denote the concept *shiny and healthy*. In essence the 3rd order interaction indicates that the genetic algorithm can be used to evolve bottle qualities that convey different consumer qualities.

The evolution on the colour genes values can be seen in Figures 4.2 and 4.3. For generation 0 each gene value varies randomly between 0 and 1 and thus has a mean value of around 0.5. As the evolution progresses these mean values change in different ways for the two consumer concepts.

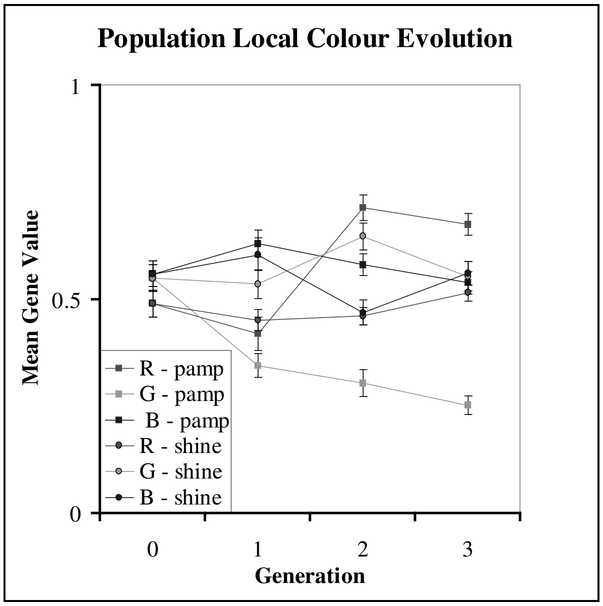
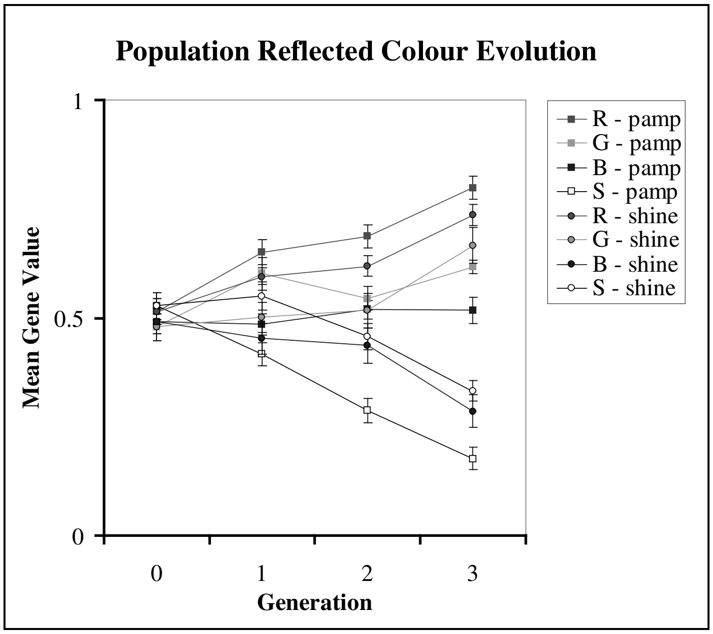


**Figure 4.2.** Mean red green and blue (RGB) values (+/- 1 S.E.) are displayed for different generations.

**Figure 4.3.** Evolution of reflected (specular) colour is shown. Mean red, green and blue (RGB) and shininess (S) values (+/- 1 S.E.) are displayed for different generations.



### 4.4 Visualisation of Consumer Preferences for More Evolved Bottles.

The information from the subject responses for the 180 bottles from generation 2 (and the 67 additional bottles from the best of the previous generations 0 and 1) contains not only *pampers* bottles rated on the *pamper* criterion and the *shiny* bottles rated on the *shiny* criterion, but also *shiny* bottles rated on the *pampers* concept and *pampers* bottles rated on the *shiny* concept. Examination of the ratings for the generation 2 set allows a direct test of whether more evolved bottles attain appropriately higher ratings.

Figure 4.4 illustrates the mean ratings for both consumer concepts. The Figure shows that more evolved bottles do indeed achieve the higher ratings. It is interesting to note, however, that whilst the bottles bred for *pampers* achieve a higher score when rated for *pampers* than the bottles bread for *shiny*, the converse is not true. Bottles bred for *shiny* do not achieve a higher score when rated for *shiny* than the *pampers* bottles.

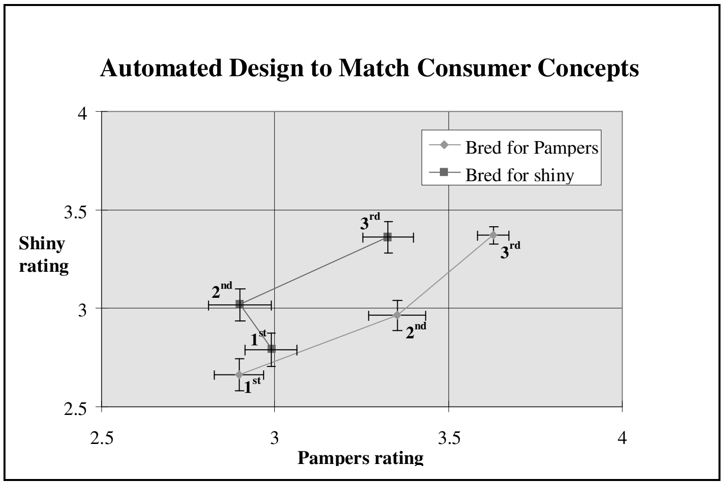


**Figure 4.4.** Mean (+/- 1 SE) of ratings for *pampers* and *shiny* for the top 20% of bottles of each generation.

### 4.5 Results: Generation 3 and modal bottles.

A binomial test showed that when asked to choose the bottle most appropriate for the concept *pampers*, subjects chose the bottle bred for this quality a significantly greater number of times than bottle bred to denote *shiny* (22/30 $p< 0.05$). Similarly, when subjects were asked to choose the bottle most appropriate for *shiny* subjects chose the bottle bred for this quality over the bottle bred for *pampers* (21/30, $p<0.05$).

### 4.6 Viewing Evolutionary Trends

Now that the general success of the procedure has been established, it is interesting to examine how the genetic algorithm arrives at solutions and whether subsequent generations would continue to improve the bottles. In order to view how the distribution of gene values within a population is changing, one can draw a *3-D*-histogram. A specific gene is selected (local red for example) and then for each population of 90 bottles a histogram can be drawn by quantising the values the genes represent. A histogram was generated for each generation and these plotted together to show a *3-D*-histogram with peaks representing commonly occurring values in the generation for the selected gene. Figure 4.5 shows that the frequency of bottles with low shininess values steadily increases across generation.

Genes for red colour by contrast do not show a uni-modal peak in the most evolved generation (see Figures 4.6 and 4.7). The mean local red value for the pampers population is high in generation 3 (around 60%), however, Figure 4.6 displays three peaks and the modal red value is substantially lower (around 40%).

The existence of multiple peaks indicates that there may be several colour values that are good at displaying a given concept.

**Figure 4.5**

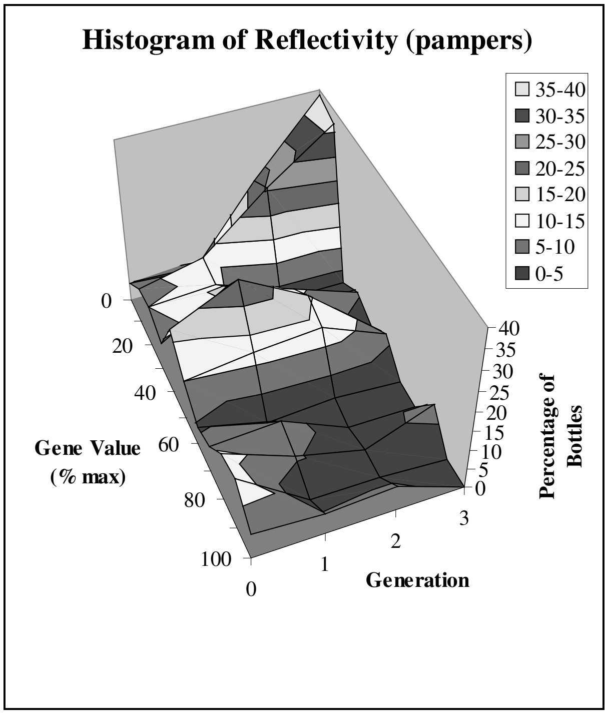


Figure 4.5 shows a histogram plane generated from the reflectivity gene from the pampers generation. Generation number and Gene value form the plane and the height of the plane is the percentage of bottles with gene values in the specific range in a specific generation. Gene Values were put into bins of size 10%. By generation 3 it is clear that low reflectivity values are much preferred in the population as the incidence of higher values is very low.

**Figure 4.6**

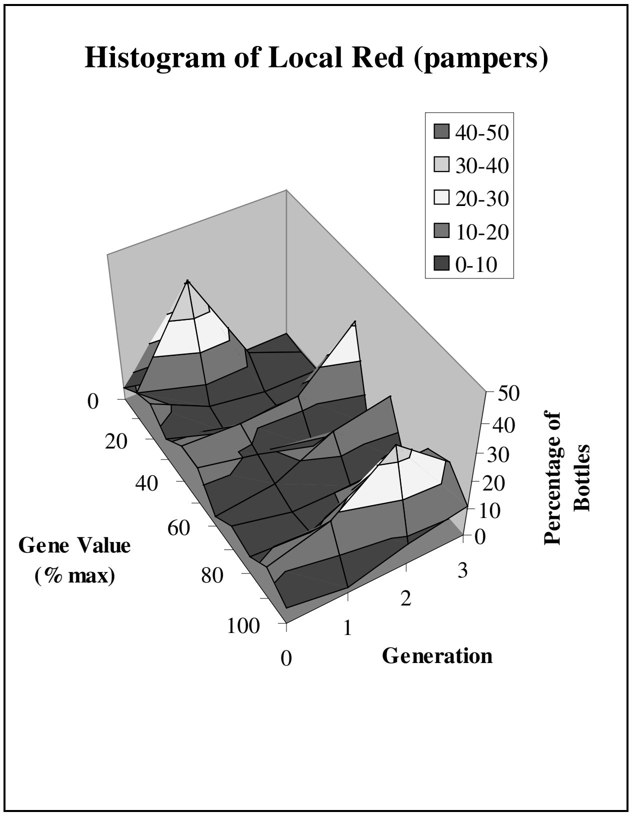
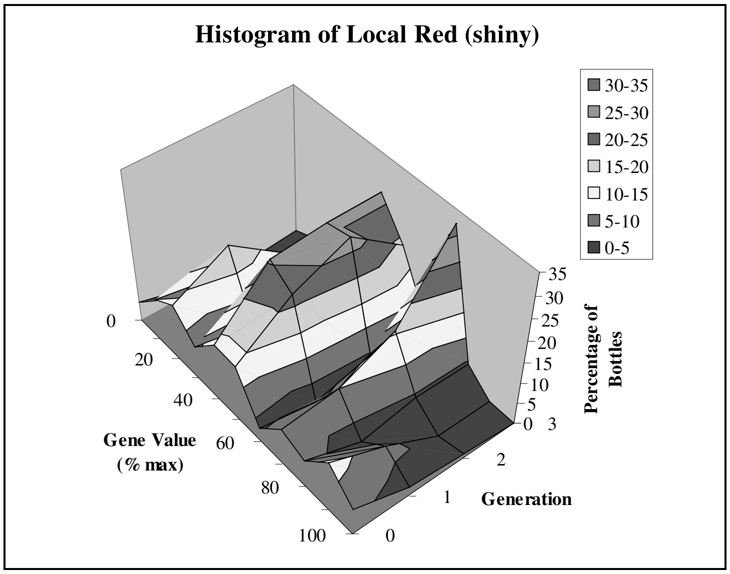


Figure 4.6 shows a histogram plane generated from the local red gene from the *pampers* generation (bin size was again 10%). Several changes in the population are evident here, which suggest evolution is progressing correctly and has not been biased by the initial populations' distribution. Bins from the population that initially have many members are reduced to having few and bins that initially have few members increase to having many.

**Figure 4.7**

Figure 4.7 shows a histogram plane generated from the local red gene from the shiny generation (bin size was again 10%). The data from generation 0 is the same for both the pampers and the shiny evolution and so is the same on both this figure and figure 2.14. It is interesting to note the how from the same starting point such a divergent evolution has occurred.



Future work could isolate these multiple optimal solutions and relate them to subject characteristics. Such work would be based on implementing a genetic algorithm process that interactively optimises a product for a given subject and then compares solutions reached. Niche optimisation GAs are currently being investigated which can perform this task (Goldberg 1987, Mahfoud 1995)

## 4.7 Discussion and Conclusions.

The main aim of this whole section was to discover whether subject preferences could be used to evolve a bottle with regard to shape and colour. Bottles were successfully evolved and in the second and thirds stages further experimental evaluations conducted to find whether the evolved bottles were more aesthetically pleasing. These evaluations successfully showed an increase in preference with generation number. So it can be confidently stated that stimulus optimisation with reference to how likeable the stimuli were was indeed taking place.

The results of the colour evolution are similar to the findings of the various previously mentioned studies (McManus *et al* 1981). Colours previously found to be unpopular (yellow) were first to disappear (see plate 2-1). As can be seen in Figure 3.8 the red value is declining with each successive generation whilst the green value has remained practically constant. Blue, by contrast, has significantly increased. Perhaps it could be argued from the literature that these colour preference are simply indicating popular colours, and not colour which is specifically appropriate for aesthetically pleasing shampoo bottles. To address this question Experiment 2 was conducted to discover whether different directions of evolution could be achieved for different concepts.

It is possible that the initial set of bottles could bias the development of future generations. However, this is unlikely due to the large (90) population size, but repetition of the initial experiment would provide validation of this. Similarly, five generations is precious few in which a GA could arrive anywhere near an optimal solution. Although, as stated, the main aim of Experiment 1 was to discover whether optimisation of a given stimulus could occur at all (and was not to find the perfect bottle). Evidence of evolution was successfully shown, thus validating the GA technique and computer graphic software.

The second experiment attempted to find whether the genetic algorithm technique could be driven by consumer preferences to subtle market/industrial concepts. At the outset of the study, it was not clear if the five chosen concepts had any independence or perceptual validity. Consumers (and ‘ad-men’) may understand what is meant by each concept, but there was no reason to believe that there would be a consensual view as to what the physical manifestations of the concept would be. Indeed, even after the original ratings for population 0, there was no firm evidence for separability of concepts (since ratings for five the concepts were inter-correlated). Despite this lack of evidence, definition of distinct properties denoting the two market concepts was established relatively quickly (within three generations). Further generations could well have produced stimuli attaining even higher ratings for each concept, but proof of principle has been established.

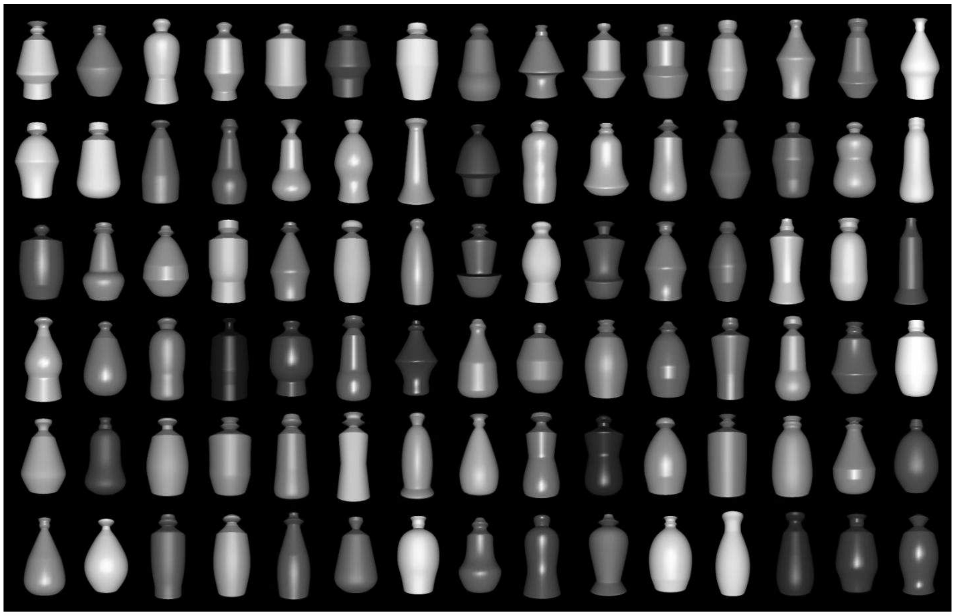
Possible extensions to these investigations are many. One possibility is the creation of multiple good solutions for a given concept. The ‘Viewing evolutionary trends’ section (4.6 above) includes figures showing the distribution of colour values. These figures suggest that there is more than one set of colour values for bottles rated highly on a given concept. The present study has focused on the mean and modal properties established by the genetic algorithm technique targeting the existence of a single best solution. Further work could establish for a given concept the existence and nature of several types of product characteristics that match a given concept and how these alternative solutions relate to subject characteristics. For example, young female subjects may have a different perception of what properties denote a healthy and shiny shampoo bottle from middle aged male subjects. The techniques developed in this thesis could be extended and used to create bottles defined by the preferences of individuals within a group. Hence bottles evolved for subgroups of the population and individuals could be compared and contrasted. The present study indicates some generality of the results with the qualities evolved to match consumer concepts with one panel of subjects, generalise to a wider populous (since subjects in the modal bottle experiment had nothing to do with the bottle genesis).

Perhaps of more industrial salience is the question of generalisation. Will the results of this study vary depending upon the type of computer-generated stimuli and the method of presentation? If the same results are obtained using ‘real world’ models (instead of computer generated ‘virtual’ bottles) then the implications for industrial application are enormous. For example, in the more applied setting real bottles could be created and filled with identical shampoo liquid. Consumers could then try the product for a week and give their opinions of the substance. Such an extension, unfortunately, is beyond the boundaries of this thesis.

GA's are often maligned as not providing meaningful solutions to problems. Often there is no way of calculating whether the result of a GA search is optimal and no knowledge of the solution landscape is produced as a result of the search. However, within the scope of this thesis they have proved useful tools for the examination of perceptual preference. Further work could compare GA performance with other search strategies and could perhaps be used to discover an absolute value of fitness for the results of the evolutions described in this section.

Obviously, shampoo bottles are not the only stimulus group to which the GA technique could be successfully applied. The following section considers the mapping of *3-D* facial surfaces onto gene strings and their implementation within a GA paradigm experiment. The section also attempts to extend the technique so as to evolve stimuli, not for a group, but for the individual.

**Plate 2-1:** From experiment 1.1: To be read left to right, top left to bottom right for each of the two generations. Top figure shows generation one (created 'randomly') after it has been ordered with respect to mean preference ratings. The lower figure shows the same for generation 5.



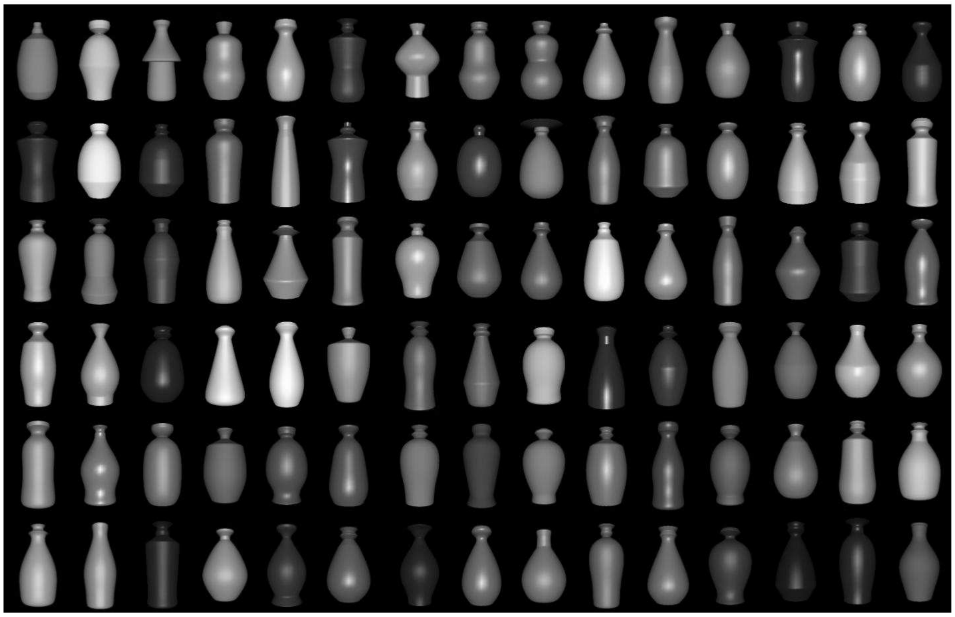
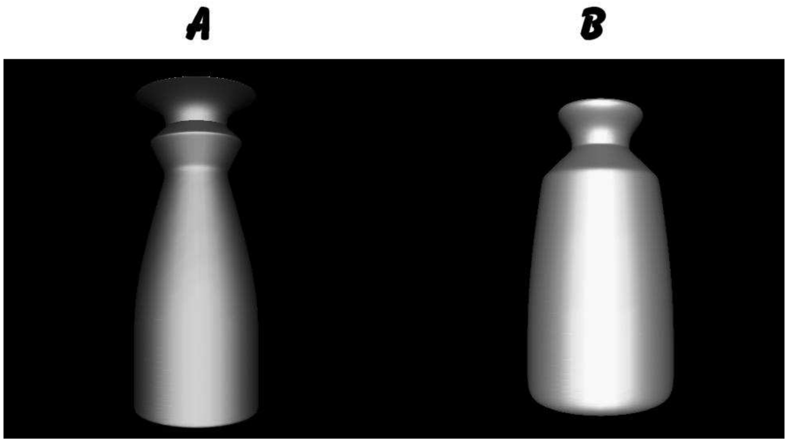
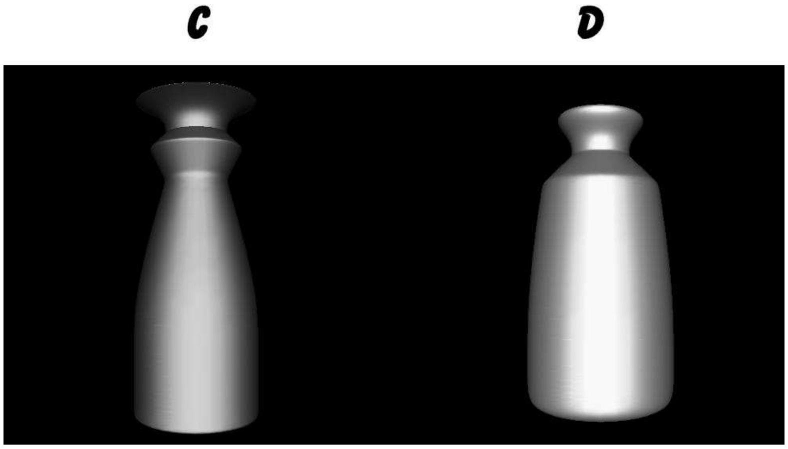


**Plate 2-2:** From experiment 2.3: Modal bottles with questions.



Which bottle do you think contains shampoo which would -
- "Pamper your hair for that special occasion" ?



Which bottle do you think contains shampoo which would -
- "Leave your hair shiny and healthy looking" ?

# 5. Evolution of Attractive Faces.

## 5.1 Introduction.

In the preceding section shampoo bottles were selected as stimuli to test the applicability of GAs for optimisation of stimuli based on subject preference. Faces are a more difficult object class for which to create a stimulus space because they can vary in many different and subtle ways (Caldwell 1992). Similarly, the surface properties of human skin are far more complex than that of plastic. One major problem in using faces as stimuli in psychological studies is that a given set of faces only contains a finite number of exemplars. One method of parameterising faces has been described in the first section of this thesis (that is, using prototyping to create dimensions which faces can be transformed along). An extension of this technique employing a Principal Component Analysis (PCA - Jolliffe 1986) is now described in order to produce encoded faces suitable for representation as genes.

## 5.2 Creating a 'Face Space'

It has been shown how, by delineating a set of features on a collection of faces, prototypes could be formed for various sub-sets. For example, from just the male faces in a collection a prototypical male face could be created, and similarly, from just the female faces a prototypical female face could be formed. These prototypes contain no information other than that consistent between the constituent faces. Thus, the prototypes have none of the idiosyncrasies of identity or individuality but rather represent a conceptual location within a face space. It has also been shown how the difference between two prototypes (e.g. male and female) can be used to create transforms for selectively manipulating the apparent gender of an individual's face. In effect, two conceptual locations within the face space are used to create a vector (or axis) which, for example, defines gender. By moving a face in the direction defined by the axis (i.e. manipulating both shape and colour information) only its gender is altered.

If one considers a set of faces (A..Z) then a set of vectors can be defined between any pair of faces by defining vectors from one particular face (or average) to all others (e.g. A→B, A→C,…A→Z). For example to go from face E→G one could traverse (A→G)-(A→E), and to find the prototype vector from all (say 26) faces [(A→B)+(A→C)+…(A→Z)]/26. These vectors can be used to define a face space by the definition of an origin. It is most convenient to define the origin to be the average of the face set (since the sum of the size of all deformations required to produce all the individual faces from this original will be minimised). Thus, to keep the origin at the conceptual centre of the face space, dimensions can be thought of as being defined as the vectors going from the origin to each of the individual faces (e.g. average→A, average→B,… average→Z). A face space can be defined in this way by the average (prototype) and a set of vectors to exemplars. Such a space could be used for GA evolution where genes would map on to the prevalence of each component vectors. For the examples space (A..Z) 26 genes could define the space and all faces within it.

For the evolution of faces the above definition of a face space was originally deemed problematic because it is not constructed in a principled or compact way. It is possible, for example, that several faces could vary from average in a relatively similar manner. Hence some dimensions would be redundant and others would carry disproportionately small weight. As is discussed later, this may not in fact be such a concern for the current study; however, it could be more problematic if the set size, and hence dimensionality, were to be extended.

In order to reduce the dimensionality of the face space a principal component analysis was performed.

## 5.3 Principal Component Analysis

Principal component analysis (PCA) can be used as a linear method of data reduction (and is an expansion of the Karhunen-Loéve representation, see Kirby 1990 and Chellappa 1995). Essentially, a PCA examines data and creates an ordered co-ordinate system based upon the distribution of the data in a high dimensional space. The vector describing the first dimension in the new co-ordinate system is the vector that describes the most variance in the data. For example, if the data formed a *3-D* cloud in the shape of a rugby ball then the first dimension would be the long axis of the ball - end to end. The second dimension is perpendicular to the first and is the vector that accounts for the next most variance. Similarly, the third dimension is perpendicular to the first two and is the vector that accounts for the remaining variance (in the case of the 3-D ball example). This process can be similarly applied to higher dimensional spaces where all variance in the original data can be expressed in the derived co-ordinate system. In such a situation no data reduction is likely to occur since there will be the same number of dimensions as there were in the original set (e.g. for a space made from 36 exemplars there are 36 principal component that fully describe the space). Note though that the majority of variance may be accounted for by the first few dimensions (although if the example were a football, and not a rugby ball, all three dimensions would be equally important in terms of variance). Sirovich and Kirby (1987) were first to perform Principal Component Analysis on face images. Their representation utilises the pixel by pixel difference of a set of faces image to define a low dimensional space describing the majority of variance in the set. This is possible because of the relative similarity between images of faces. Subsequently, many have taken this technique and used it as a compact way of representing a set of faces (Kirby and Sirovich 1990, Turk and Pentland 1991). The basic representation has now been extended to included shape information (Craw and Cameron 1991, O'Toole *et al* 1993, Vetter and Troje 1995, Costen *et al* 1996). For example, with a sample set of *2-D* facial images, each face can be delineated and warped to an average shape (as in section one of this thesis). For each of the faces, the distortion[8] applied along with the resulting image can be fed into a PCA as one data point. A multi-dimensional 'face space' would be created with 'Eigenfaces' (see Figure 5.1) representing the axes (the centre of the space would be the average of all the original faces). The representation of faces utilised in this thesis is most similar to these latter representations since both delineated shape and depth information is included in the PCA analysis.

---

[8] In the original work of Kirby and Sirovich, the PCA was performed on images that did not have the facial features in alignment. For face image reconstruction, feature alignment is preferred because otherwise the images produced are blurry.

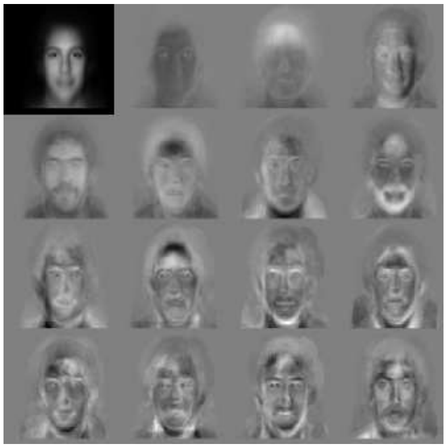


**Figure 5.1** The first 15 Eigenfaces are shown calculated from a source population of 128 images together with the average (top left). Eigenface 1 is on the right of this prototype with Eigenface 15 in the bottom right.

Note: no feature alignment (apart from the eyes) has been performed.

Image credits Moghaddam (1996).

Since the majority of the differences between the faces will be taken into account by the first few dimensions the remaining dimensions can be disregarded. For example, with specific reference to face recognition, the MIT software 'picturebook' uses a modular approach creating 20 principal components for each feature, derived from 128 original images (with separate PCAs being calculated for the head as a whole, the two eyes the nose and the mouth). This system gives a 95% correct recognition rate from a search of 3000 individuals faces (Pentland 1994, Moghaddam 1996). In terms of face reconstruction, this enables all faces in the original set to be recreated using a reduced amount of information (Phillips 1997) - albeit at a loss of some fidelity.

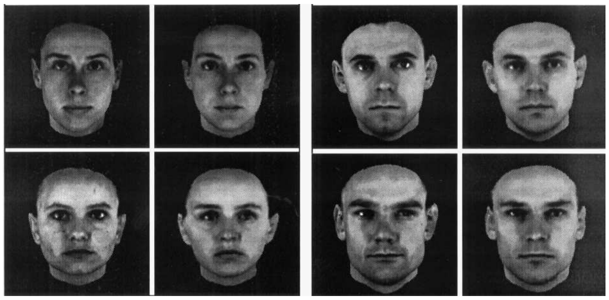


**Figure 5.2** Four pairs of faces are shown, the leftmost of each pair is an original face whilst the rightmost image is constructed using all the principal components from 49 other face images.

Image Credits Vetter and Troje (1995).

Reconstructions of faces using 49 principal components have been made and the majority of faces look similar to the originals (see Figure 5.2) "the fact that projections of a new face into the space spanned by the 49 remaining faces yields a fairly good reconstruction, shows that the dimensionality of the space might not be much larger" Vetter and Troje (1995).

This method of compact representation can be extended to *3-D* by delineating and warping depth maps in an analogous way to the description in section one of this thesis.

## 5.4 Mapping PCA onto GA

The resulting multidimensional 'face space' (created by performing a PCA) can be explored using a GA to discover maxima (or minima) for appropriate fitness functions (Holland 1975, Goldberg 1989). The advantage of using the dimensions created by the PCA rather than the dimensions inherent in the original face set is that redundant information (that is held in the lower order dimensions) can be omitted and the actual gene representation can be compact. For example, for the set of 3000 faces used by Moghaddam *et al* 1996, the first 20 principal components gave a recognition rate of 95%. Thus, a smaller number of genes are needed in order to encode a face. In the actual experiment reported here, all components were included as genes in the GA. This was done because the available face set was limited in size and also simply for completeness since at the end of the evolution low impact component would simply vary randomly affecting the generated face little.

## 5.5 Psychological Background

As previously discussed in this thesis, there is often a great deal of agreement about aesthetic judgements (essentially implying that beauty is not is the eye of the beholder). Shampoo bottles have been successfully 'bred' to enhance attractiveness and those characteristics which subjects found most befitting to ascribed consumer concepts. Whilst not a great deal of published work has previously been conducted on such judgements about shampoo bottles, there is wealth of reporting aesthetic judgements of another homogeneous stimulus type, namely faces.

Much scientific interest in the study of facial attractiveness (or beauty) is based upon the notion that the face, is in some way an advert for evolutionary fitness. One could imagine, for example, that to grow something symmetrically requires development at an equal rate on each side. Thus external factors (such as parasites) are not stopping the normal symmetric growth. So it may be fair to assume that the organism has good immunity to the negative influences in its surroundings. Similarly, Symons (1979) suggests that closeness to averageness maybe the best indicator of  “good genes”. In agreement with this Langlois and Roggman (1990) found subjects prefer average faces to any of the examples making up the average. However, later studies have indicated that whilst average faces are attractive, the notion that they are best is confounded by the averaging of texture which are used to create the faces. For example, if smooth texture (indicating perhaps youth) is preferred to rough, idiosyncratic texture then the averaged faces may be preferred to real faces for that reason alone. Perrett *et al* (1994) showed that by keeping texture constant, a preference is found for non-average face shapes. Perrett *et al* created an image with average skin texture and shape which was average for the whole set of sample faces; and another image ‘high attractive’, with the same average texture, but in a shape created as an average of the most attractive 25% of faces. A clear preference was found for the shape created from the most attractive faces. This implies that attractive faces differ in a consistent way from average faces. This preferred direction of deviation away from the population mean shape persisted when a further image was made by caricaturing the high attractive shape away from the average shape. This super-high attractive shape was rated as being more attractive than either of the two previous facial images. This study confirms that attractive face shapes are not average (although no evidence was given on what direction the deviation from average might take).

Much has been written about the high degree of agreement there is on facial beauty between cultures (Perrett *et al*1994, Cunningham *et al* 1995, Jones 1995). Such claims are not only recent, with Darwin in 1871 commenting that the explorers of the time were reporting they saw the same level of beauty emanate from a variety of indigenous peoples. Of particular interest to the current study is the finding from Perrett *et al* (1994) that both Japanese and Caucasian subjects find the same faces attractive whether the faces presented are Japanese or Caucasian in origin.

This section of the thesis extends the GA paradigm to explore facial surfaces. Each subject will rate face surfaces on attractiveness and evolve there own preferred face shapes.

## 5.6 Hypothesis

1) Subject preference will be a viable method of evolving preferred *3-D* head shapes using a GA;
2) The direction of evolution will be similar across subjects;
3) The direction will be away from the average shape.

## 5.7 Method

Previously, this thesis has shown how *3-D* head models can be expressed as *2-D* depth maps and delineated in an analogous way to *2-D* images of faces. Isono (1996) uses a simpler 14 point delineation strategy and warps boxes (rather than triangles) to normalise for shape (see Figure 5.3). In addition, the box formed by the extreme points crops the data. To normalise in *3-D* three points are required, the outer eye points are used and an additional point (the center of the mouth) is required.

Once the depth map is normalised for shape and cropped a PCA can be performed with the deformation information and the normalised depth map. This produces a number of principal components that can be used as genes in a genetic algorithm. A specific face is generated as a summation of each Eigenface multiplied by a coefficient. These coefficients can vary enormously in value, however their mean value is zero (since their distribution is the result of the PCA). Since the GA implemented is binary coded, the coefficients need to be converted into a bounded integer form. This was achieved by scaling each coefficient in the set by six times the standard deviation of that coefficient (from all the other faces) and by adding a constant so that the middle of the coefficient’s range was aligned with the middle of the binary representation's range. This had the effect of keeping all coefficients with +/- 3 standard deviations. Each of the principal components was by represented by 16 bits and a mutation rate was set to

1:0.001. During the evolution phase of the GA a uniform crossover method was used as it was in the evolution of shampoo bottles.

**Figure 5.3**

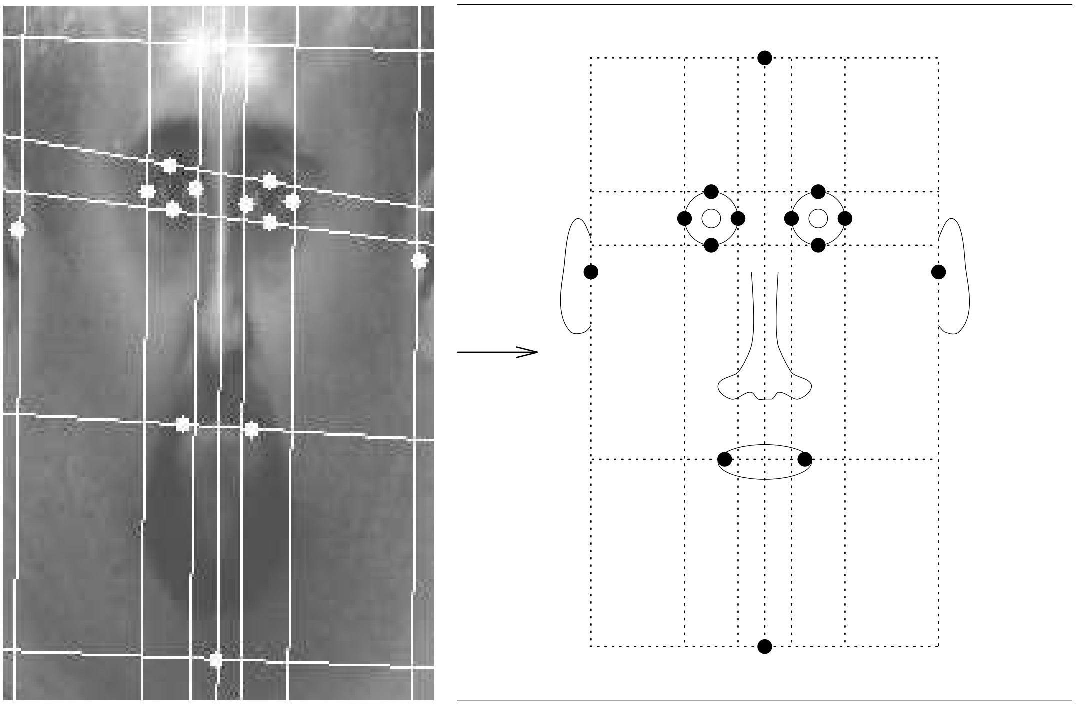


Figure 5.3 shows 14-point delineation after Isono (1996). The image on the left shows the colour map from a cyberscan delineated with 14 points. The image on the right is a schematic to show how the features will be aligned after *2-D* (shape) normalisation.

### 5.8 Stimuli

36 head scans of Japanese females were taken using a Cyberware$^{tm}$ 3030 scanner. These were delineated and a PCA performed from which 35 components were obtained. A head was then represented as a combination of 3D Eigenhead coefficients. Head images were presented as Gouraud (1971) shaded objects with colour maps omitted (hence facial colouration cues to attractiveness were removed to allow the just the evolution of shape - see Plates 3-1 and 3-2)

### 5.9 Procedure

Six subjects, three males and three females, ages 23-25 took part in the study. Each was tested individually in an attempt to evolve each their own preferred head shape. Stimuli were presented on three successive pages (the first and second pages containing fifteen head and the third page containing six heads - see plate 3-2). Subject were free to go backwards and forwards through these pages and were requested to rotate the individual heads, using the mouse, in order to see them from different views (see plate 3-3). Subject were initially presented with the 36 original heads (on the three separate pages) and requested to examine all 36 heads and pick the one they found most attractive. 7 points should be awarded to this head (by clicking on the box below it and entering the digit 7 using the keyboard). The subject was the requested to rate each of the remaining heads for attractiveness. The scale to be used ranged from 1 (being the least attractive) to 7 (being the best and of equal attractiveness to their favourite). After all 36 heads had been rated the subject clicked on an ‘evolve’ button and the GA used the ratings to create a new stimuli set of 36 heads (using a crossover and mutation rate as specified above). The subject rated this new set of 36 heads in the same way and a further set of heads created. This process continued for a total of 6 generations.

### 5.10 Test

The six highest rated heads for the subject from each of the 6 generations were then recalled from each generation and presented together as a final test generation (of 36 heads). The subject was required to rate these heads in the same method as before. If subjects had evolved faces to become more attractive than in this test phase, then one would expect faces from later generations to be given higher ratings than those from earlier generations.

## 5.11 Results

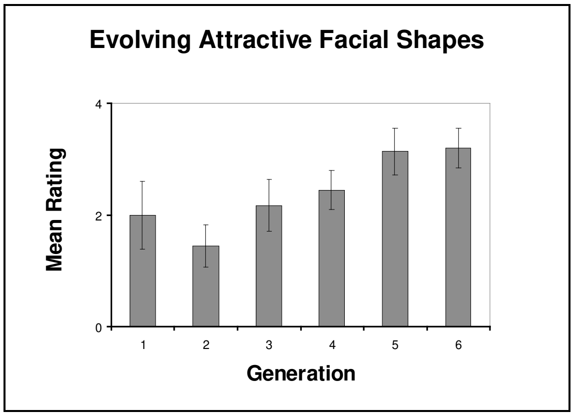


**Figure 5.4.** The figure on the left illustrates data from the final test population and shows the mean rating increasing as the faces become more evolved. The mean was calculated from each of the subjects ratings of their own favourite six faces for each generation from the test phase

Essentially the graph shown in Figure 5.4 suggests that subject have a preference for faces from generation 5 or 6 faces (rather than those from earlier generations - see plate 3-4 for a specific example). One-way analysis of variance of the average ratings for each subject for the top six faces per generation showed a significant main effect of generation number ($F_{5,25} = 5.74$, $p=0.001$). Posthoc comparisons show a progressive improvement of the ratings across generation. The average rating for the best six faces in generations 6 and 5 were higher than those for generations 1, 2 and 3 ($p<0.05$, each comparison, PLSD[9] test). The final two generations (5 and 6) were not rated differently ($p>0.05$) but there was a trend for these generations to receive higher than the ratings for generation 4 ($0.1> p>0.05$). The apparent dip in attractiveness between generation 1 and 2 was not significant ($p=0.18$)

## 5.12 Analysis of Coefficients

If different subjects select attractive faces on the basis of the same facial surface cues then faces in generation 6 should have similar principal component coefficients. Statistical analysis was performed on the coefficients used to create the faces in order to discover whether the more evolved faces were systematically deviating from an average shape. All principal component coefficients were first converted into Z-scores so they could be reasonably compared (an average shape would accord with all coefficients for a specific head being 0). A mean value for each of the 35 coefficients was calculated from each of the best six faces from both generations. A similar set of mean values was calculated for generation 6 and a (2-tailed) Spearman's rank correlation use to compare the two setsof means. This test showed a significant correlation (r=0.558, p<0.001, N=35) indicating that direction and relative size of deviation for preferred genes was similar between the best faces from generation 1 and generation 6. A matched pairs t-test applied to the absolute value of the same coefficients revealed that the deviation away from average (coefficient = zero) was significantly larger in generation 6 (t=4.98, df=35, p<0.001) - see Figure 5.5.

---

[9] See Snedecor and Cochran (1980).

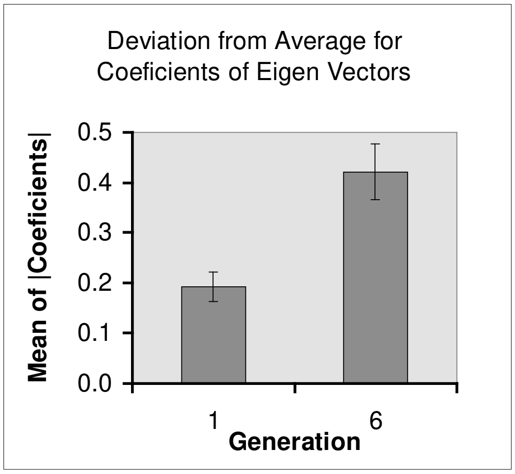


**Figure 5.5.** The figure on the left shows a significant increase in the magnitude of the deviation away from average when comparing the coefficients from generation 1 and generation 6.

### 5.13 Discussion

The use of a PCA is, perhaps, redundant in the last section. With reference to face recognition literature several concerns have been addressed by Kouzani *et al* (1997) "Even in a simplified case [Turk and Pentland 1991], if more known faces are added to the database, then (i) the calculation of the basis becomes computationally expensive, (ii) the compression ratio declines, (iii) the basis must be recalculated, and (iv) all faces within the database must be re-compressed using the new basis". Kouzani *et al* go on to suggest a fractal model of face representation and recognition. In the representation schema a face image is represented by its fractal model (which is a small set of parameters). With this new technique, it may be possible to create a gene mapping for faces which contains less redundant information than the PCA approach and quicker to administer. However, since a genetic algorithm will automatically find values for important genes and let unimportant genes vary, it matters little if some redundant information is encoded along with the vital. Similarly, it is not important for a GA to have dimensions that are orthogonal. A better encoding strategy may be simply to have faces defined as being a weighted summation of all faces in a sample set as has been described earlier.

The main finding from the results of this last study concurs with Perrett (1994) in that the preferred face shapes were significantly different from average. The evidence for the successful evolution of more attractive faces, shown by the analysis of coefficient, reveals that there is a progressive increase in deviation away from average as faces become more evolved. The direction of deviation is the same as the direction selected for in generation 1, showing that not only were faces becoming more evolved, but that the direction for which the faces were being selected remained constant. This shows that it is highly unlikely that the increase in preference is due simply to the elimination of less desirable faces. Similarly, since subjects concur on which directions away from average are most attractive, evolution occurs in broadly the same direction as shown by the matched pairs t-test. Together with the Spearman's rank correlation these tests show that the coefficients from generation 6 were amplifications of the coefficients values from the best faces from population 1. This indicates that the attractive faces in generation 1 deviated from average in particular ways and that the level of these favoured traits was increased in the more evolved faces of generation 6. Thus, the study has successfully shown all of the hypothesised statements as being true. Subject preference is a viable method of driving a GA to evolve preferred *3-D* head shapes, subjects do agree on the direction of evolution and that direction is away from an average shape.

# 6. **Thesis Summary**

This thesis has developed a variety of novel computer graphic processes and has documented psychological studies that have been performed to gauge the performance and illustrate the applicability of the techniques. The first section sought to develop techniques that could be used to parameterise and systematically alter facial appearance. Existing techniques for delineating and prototyping faces were used to find consistencies from groups of faces. Once these group traits were discovered, new technology was implemented which first enabled the inclusion of average information in place of absent information (i.e. the addition of average colour information to successfully ‘colourize’ a black and white photograph) and second, allowed the difference between prototypes to be used to create transforms which selectively alter perceived group membership. Quantification of the perceived effectiveness of these transforms has been performed by others (for ageing see Burt and Perrett 1995 and Plate 1-8) but the success of the gender manipulations and the ageing transforms is immediately apparent (see Plates 1-9,1-10 and 1-11). It was noted that previous automated shape caricaturing techniques are simply a special case of the more general transform techniques described in this thesis. The development shape and colour transforms has enabled others to investigate the role of colour in facial perception and the success of caricaturing colour to enhance recognition accuracy (Lee and Perrett 1997). The transformational techniques have been further extended by their application to the delineation, prototyping, caricaturing and transformation of *3-D*-head models. The success of these extensions is apparent in plates 1-15, 1-17 and 1-18 (although work is underway to conduct psychological evaluations of the impact of the transforms). The *2-D* transforms have been shown to capture 80% of the visual cues, at least in the case of ageing (Burt and Perrett 1997) . It is suggested here that the remaining 20% could be mostly made up of textural cues. For example, since the prototyping process removes all fine lines and wrinkles these obvious traits of ageing are not included in the transformation. Current work is exploring ways to quantify these textural cues and to include their effect in the transforms.

In the second section of this thesis a homogeneous stimulus class was selected and parameterised. A genetic algorithm was implemented to allow search in the stimulus domain with fitness defined by subject preferences. Shampoo bottles were used since they provided ‘easily’ creatable virtual objects. It proved possible to evolve the shape and colour of shampoo bottles using subject preferences so that the end products were progressively more popular. The second study successfully proved the applicability of genetic algorithms to such perceptual optimisation tasks when the concepts being optimised were defined from an industrial setting and were closely related.

The final section of the thesis brings together the technology developed in the first two sections and applies it in a psychological investigation into the optimisation of *3-D* head models. Subject preference for attractiveness was successfully used to evolve the *3-D* shape of faces. This last section is perhaps the most exciting since it uses techniques developed earlier in this thesis to open up new ways of investigating complex stimuli types, and indeed the technology is already being extended by others (Isono *et al* 1997 use the technique to investigate the effect of view on the creation of subjective personality traits of faces). The findings from this last section concur with current psychological literature and indicates: i) a high agreement between subjects on which faces are attractive, and ii) that attractive face shapes are not average. This successfully validates the technique and methodology. Current work on this topic is directed at identifying the preferred direction away from average.

**Plate 3-1:** Fifteen cyberscanned heads are shown with texture maps applied.

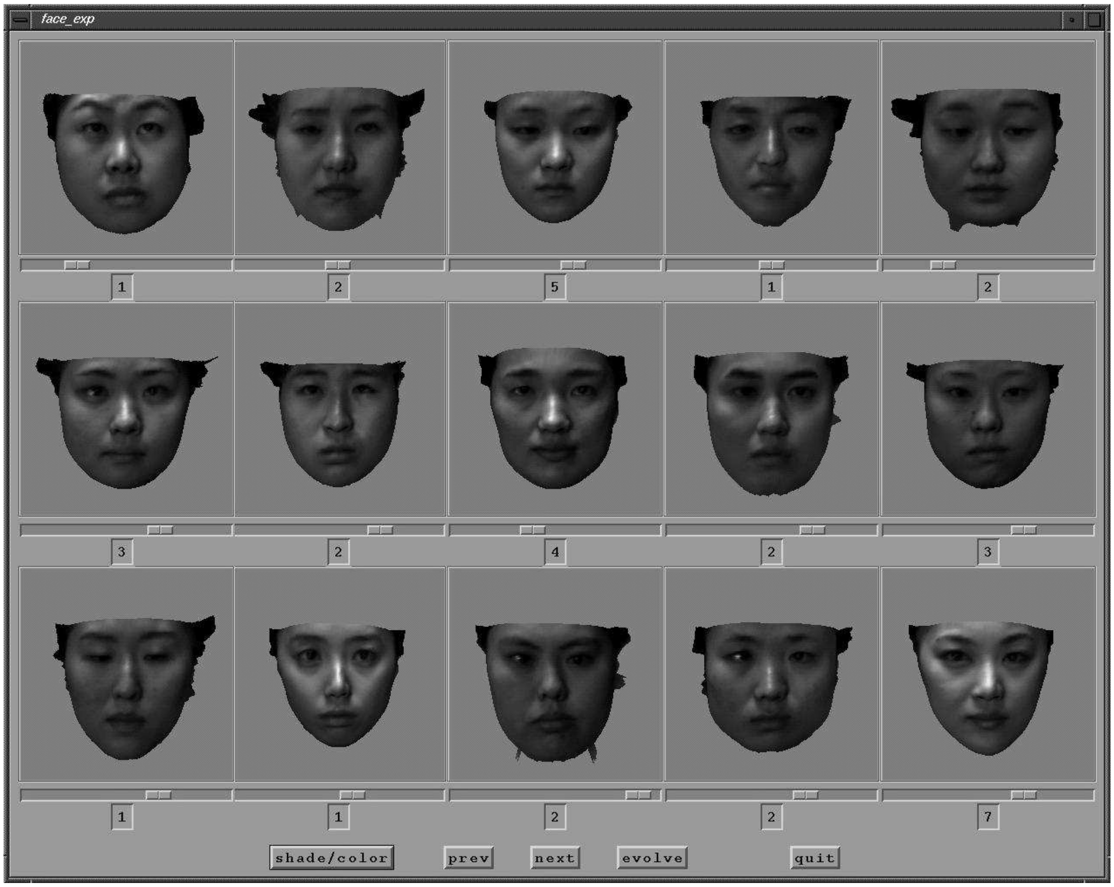

face_exp
1
2
5
1
2
3
2
4
2
3
1
1
2
2
7
shade/color
prev
next
evolve
quit

**Plate 3-2:** The same 15 cyberscanned heads as in plate 3-1 are shown with the texture removed and Gouraud shading applied.

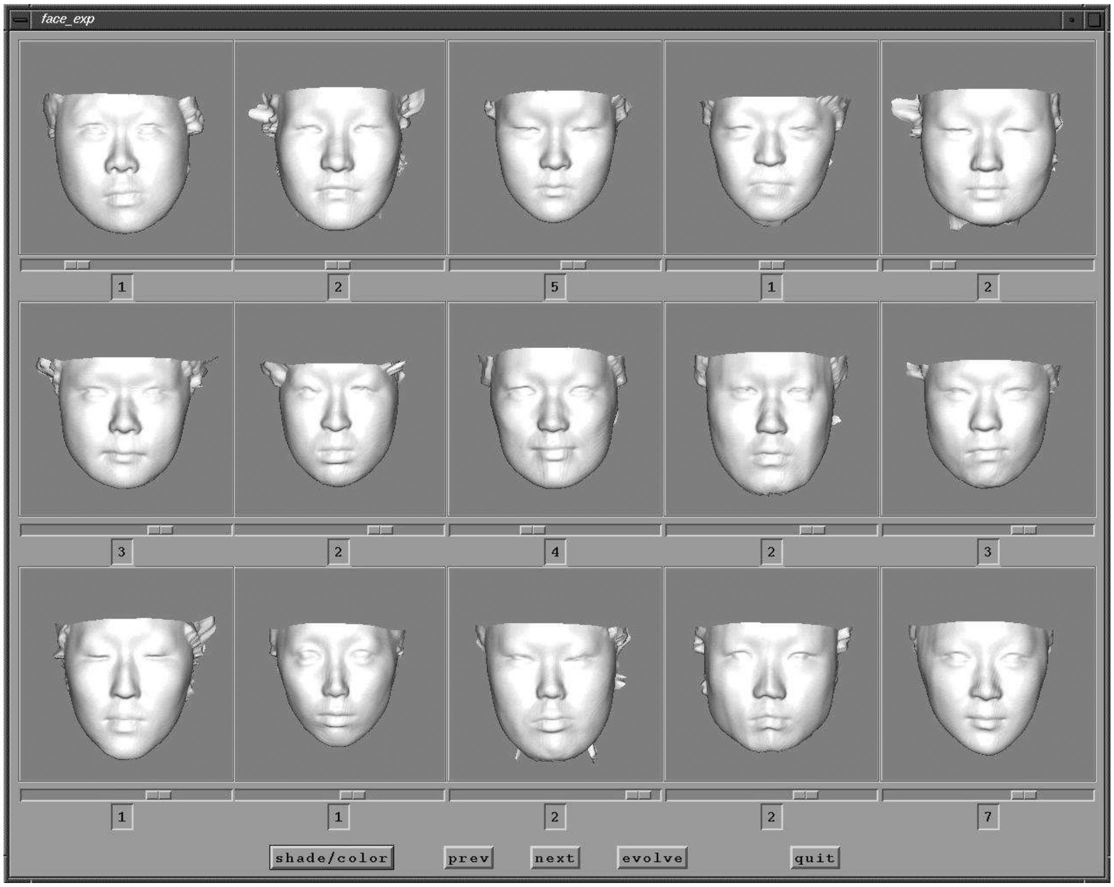

face_exp
1
2
5
1
2
3
2
4
2
3
1
1
2
2
7
shade/color
prev
next
evolve
quit

**Plate 3-3:** The same 15 cyberscanned heads as in plate 3-1 are shown ‘mid experiment’ in various poses and with various ratings having been submitted (the rating is the value in the small box below each face).

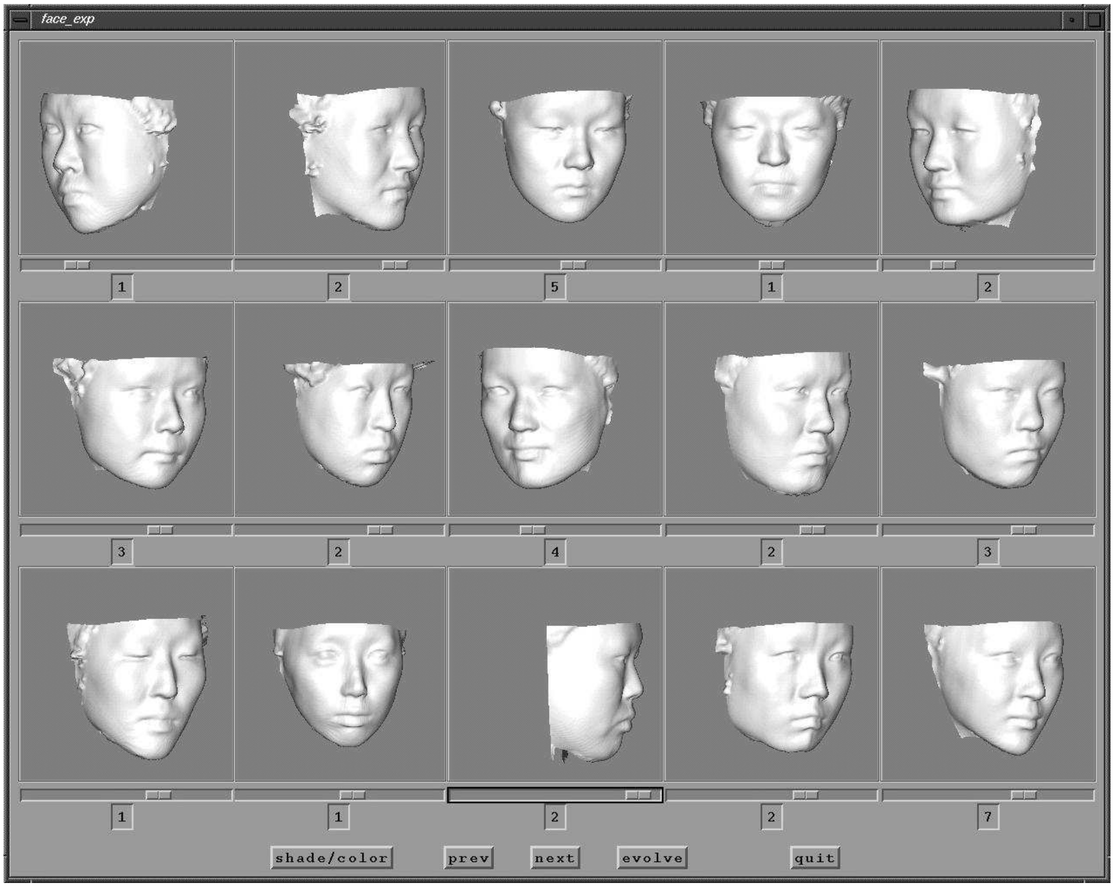
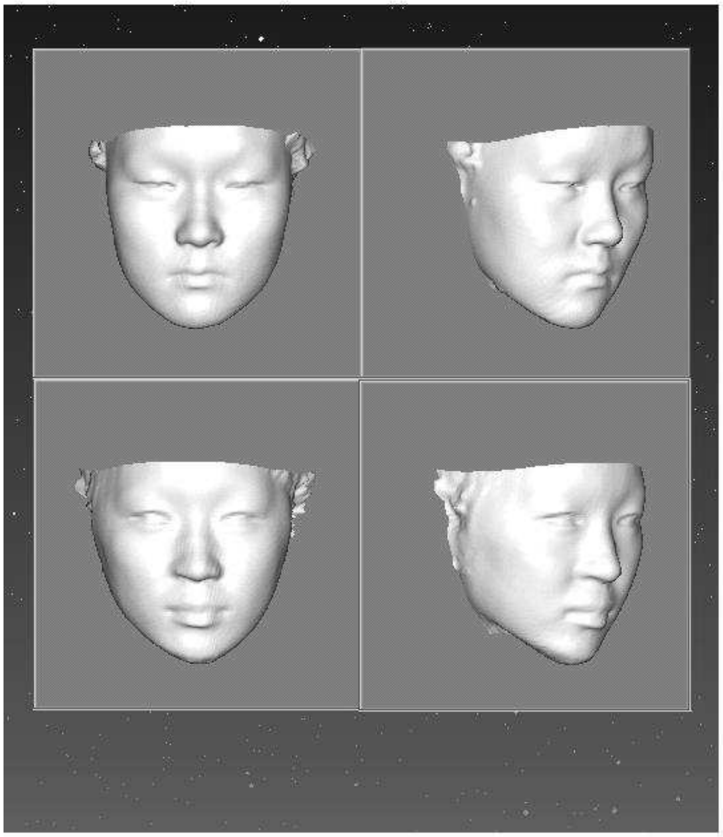

face_exp
1
2
5
1
2
3
2
4
2
3
1
1
2
2
7
shade/color
prev
next
evolve
quit

**Plate 3-4:** On the top of the plate is shown the favourite face from generation 1 for a specific subject (2), below this the same subject's favourite face from generation 6 is illustrated. The subject preferred the more evolved face from generation 6 to the original face from generation 1.

**Favorite Faces for Subject 2**



**The more evolved face (from generation 6) was rated more highly than the most popular face from the original population (generation 1).**

# 7. **References.**

Bartles, R., Beatty, J. and Barsky, B., (1987) “An Introduction to Splines for Use in Computer Graphics and Geometric Modeling”, *Morgan Kaufmann*, Los Altos, CA.

Bäck, T. and Schwefel, H-P. (1993) “An overview of evolutionary algorithms for parameter optimisation”, *Evolutionary Computing*, 1, 1-23.

Beier, T. and Neely, S. (1992) "Feature-based image metamorphosis", *Computer Graphics (SIGGRAPH ’92 Proceedings),* 26, 35-42.

Benson, P.J. and Perrett, D.I., (1991a) "Synthesising continuous-tone caricatures", *Image and Vision Computing*, 9, 123-129.

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

Benson, P.J. and Perrett, D.I., (1993) "Extracting prototypical facial images from exemplars", *Perception*, 22, 257-262.

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

Bergen, J. and Hingorani, R., (1990) "Hierarchical motion-based frame rate conversion." *Technical report*, David Sarnoff Research Center, Princeton, NJ, 08540.

Bergen, J., Anandan, P., Hanna, K. and Hingorani, R. (1992) “Hierarchical model-based motion estimation”, *Proceedings of the European Conference on Computer Vision*, 237-252, Santa Mergherita Ligure, Italy.

Beymer, D., Shashua, A. and Poggio, T. (1993) “Example based image analysis and synthesis”, *M.I.T. Artificial Intelligence Laboratory*, A.I. Memo 1431.

Bezier, P. (1970) “Emploi des Machines a Commande Numerique”, *Masson et Cie*, Paris, 1970. Translated by Forrest, A.R., and Pankhurst, A.F., as Bezier, P., Numerical Control - Mathematics and Applications, *Wiley*, London.

Bezier, P. (1974) “Mathematical and Practical Possibilities of UNISURF”, in Barnjill, R.E. and Risenfeld, R.F., Eds, *Computer Aided Geometric Design*, Academic Press, New York.

Birkhoff, G. and Maclane, S. (1960) “A Survey of Modern Algebra*”, Macmillan Company*, New York.

Brennan S E, 1982, Unpublished Thesis, MIT.

Brennan S E, 1985 "The caricature generator", *Leonardo*, 18, 170-178.

Brown, E. and Perrett, D.I. (1993) "What gives a face its gender?", *Perception*, 22, 829-840.

Bruce, V. (1988) "Recognising Faces", *Lawrence Erlbaum Associates Publishers*, London. ISBN 0-86377-142-4.

Bruce, V., Burton, M., Doyle, T. and Dench, N. (1989) “Further Experiments on the Perception of Growth in Three Dimensions”, *Perception and Psychophysics*, 46, 528-536.

Burnham, M.W., Hanes, R.M. and Bartleson, C.J. (1963) “Colour: A guide to basic facts and Concepts”, *John Wiley*, New York.

Burson, N., (1992) “Faces”, *William Morrow & Co*., New York.

Burt, D.M., Perrett, D.I., (1995) "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information", *Proceedings of the Royal Society of London - B*, 259, 137-143

Calder, A.J., Young, A.W., Karnat, A., Sprengelmeyer, R. Perrett, D.I. and Rowland, D., (1997) "Selective deficits in emotional expression recognition following brain damage.", *International Journal of Psychophysiology*, 25, 69

Calder, A.J., Young, A.W., Rowland, D.A., Perrett, D.I., Hodges, J.R. and Etcoff, N.L. (1996) “Facial Emotion Recognition After Bilateral Amygdala Damage - A Differentially Severe Impairment of Fear”, *Cognitive Neuropsychology,* 13, 699-745.

Caldwell, C. and Johnston, V.S. (1992) "Tracking a Criminal Suspect through ‘Face Space’ with a Genetic Algorithm"*, In Proceedings of International Conference on Genetic Algorithms,* 416-421

Carey, S. (1992) “Becoming a face expert*”, Philosophical transactions of the Royal Society of London*, 335, 95-103.

Chellappa, R., Wilson, C.L. and Sirohey, S. (1995) "Human and Machine Recognition of Faces: A Survey", *Proceedings of the IEEE*, 83, 705-740.

Coldstream, N. (1991) “Medieval Craftsmen: Masons and Sculptors”, *University of Toronto Press*, Toronto.

Costen, N., Craw, I., Robertson, G. and Akamatsu, S. (1996) "Automatic Face Recognition: What Representation ?", *Computer Vision - ECCV '96, Lecture Notes in Computer Science*, Cambridge, UK, Eds. Buxton, B., Cipola, R., 504-512.

Craw, I., Costen, N., Kato, T., Robertson, G., Akamatsu, S. (1996) "Automatic face recognition: combining configuration and texture" (in press)

Craw, I. and Cameron, P. (1991) "Parameterizing Image for Recognition and Reconstruction", *Proceedings of the British Machine Vision Conference*, 367-370.

Cunningham, M.R., Roberts, A.R., Barbee, A.P. and Druen, P.B. (1995) "`Their ideas of Beauty are, on the whole, the same as ours': Consistency and Variability in the Cross-Cultural Perception of Female Attractiveness", *Journal of Personality and Social Psychology*, 68, 261-279.

Darwin, C. (1871). "The descent of man, and selection in relation to sex*", John Murray*, London.

Dodd, J.V. & Perrett, D.I. (in submission) "The effect of caricaturing on learning and recognition of car shapes".

Eysenck, H.J., (1941) "A critical and experimental study of colour preference", *American Journal of Psychology*, 54, 385-394

Fagan, J.F. and Singer, L.T., (1979) "The role of simple feature differences in infants' recognition of faces", *Infant Behaviour and Development*, 2, 39-45

Fechner (1876) "Vorschule der Aesthetik", *Brietkopf and Hartel*, Leipzig.

Foley, J.D., van Dam, A., Feiner, S.K. andHughes, J.F., (1990) "Computer Graphics Principles and Practice, 2nd Edition", *Addison-Wesley*, Reading Mass., 592-595. ISBN 0-201-12110-7

Galton, F.J. (1878) "Composite portraits", *Nature*, 18, 97-100

Galton, F.J. (1879) "Composite portraits, made by combining those of many different persons into a single, resultant figure*", Journal of the Anthropological Institute*, 8, 132-144.

Ghyka, M. (1997) "The Geometry of Art and Life", *Dover*, New York. (Original work published in 1946).

Goldberg, D. E., and Richardson, J., (1987) "Genetic Algorithms with Sharing for Multimodal Function Optimization*," Genetic Algorithms and their*

*Applications: Proceedings of the Second International Conference on Genetic Algorithms*, 41-49.

Goldberg, D.E. (1989) "Genetic Algorithms in Search, Optimisation and Machine Learning", *Addison-Wesley*, New York.

Gomes, J. and Velho, L., (1995b) "Abstraction Paradigms for Computer Graphics", *The Visual Computer*, 11, 227-239.

Gomes, J., Berton, J. Costa, B., Darsa, L., Velho, L. and Wolberg, G. (1995a) "Warping and Morphing of Graphical Objects", *Course Notes*, SIGGRAPH '95, Association of Computing Machines.

Gouraud, H. (1971) "Continuous Shading of Curved Surfaces*", IEEE Transactions on Computers*, 20, 623-628.

Granger, G.W. (1955) "An experimental study of colour preferences*", Journal of General Psychology*, 52, 3-20.

Green, C.D. (1995) "All That Glitters: A Review of Psychological Research on the Aesthetics of the Golden Section", *Perception*, 24, 937-968.

Heeger, D.J. and Bergen, J.R. (1995) "Pyramid-Based Texture Analysis/Synthesis", *Computer Graphics (SIGGRAPH '95 Proceedings),* 229-238.

Helson, H. and Lansford, Y. (1970) "The role of spectral energy of source and background colour in the pleasantness of objects", *Applied Optics*, 9, 1513-1562.

Holland, J.H. (1975) "Adaptation in Natural and Artificial Systems*", The University of Michigan Press.*

Holland, J.H. (1992) "Genetic Algorithms", *Scientific American*, July 1992, 44-50.

Howell, D.C. (1992) "Statistical Methods for Psychology", 3rd ed, Boston, Massachusetts: PWS-KENT, ISBN 0-534-92955-9.

Humphrey, G.K., Goodale, M.A., Jakobson, L.S. and Servos, P. (1994), "The Role of surface information in object recognition: studies of a visual form agnosic normal subjects", *Perception*, 23, 1457-1481.

Isono, K. and Akamatsu, S. (1996) "A Representation for *3-D* faces with better feature correspondence for Image Generation using PCA*", Technical Report of the IEICE*, HIP96-17.

Isono, K., Oda, M. and Akamatsu, S. (1997) "Generating Subjective Impression for the Face using Genetic Algorithm*", Technical Report of the IEICE*, HIP97-4

Johnston, V.S. and Franklin, M. (1993) “Is beauty in the eye of the beholder?” *Ethnology and Socio-biology*, 14, 183-199.

Jolliffe, I.T. (1986) "Principal Component Analysis", *Springer-Verlag*, New York.

Jones, D. (1995) "Sexual Selection, Physical Attractiveness and Facial Neoteny." *Current Anthropology*, 36, 723-748.

Kirby, M. and Sirovich, L. (1990) "Application of the Karhunen-Loéve Procedure for the Characterization of Human Faces," *IEEE Transactions On Pattern Analysis and Machine Intelligence*, 12, 103-108.

Kouzani, A.Z., Sammut, F.He. and K. (1997) "Fractal Face Representation and Recognition*", SMC Conference Proceeding, IEEE International Conference on Systems, Man, and Cybernetics - Computational Cybernetics and Simulation*, Florida, 1609-1613.

Lalo (1908), “L’Esthetique experimentale et contemporaire”, *Alcan*, Paris.

Lanatis, A., Taylor, C. Cootes, T. and Ahmad, T. (1995) "Automatic interpretation of human faces and hand gestures using flexible models." In Bichsel, M., editor, *Proceedings of the International Workshop on Face and Gesture Recognition*, 98-103, Zurich, Switzerland.

Langlois, J.H. and Roggman, L.A. (1990) "Attractive faces are only average" , *Psychological Science,* 1, 115-121.

Lee, K.J. and Perrett,D.I. (1997) "Presentation Time Measure of the Effect of Manipulations in Colour Space on Discrimination of Famous Faces", *Perception*, 26, 733-752.

Mahfoud, S.W., (1995) “PhD thesis in Computer Science”, *Graduate College of the University of Illinois*, Urbana-Chamaign.

Mark, L.S. and Todd, J.T. (1983) “The perception of growth in three dimensions”*,Perception and Psychophysics*, 33, 193-196.

McManus, I.C. (1980) “The aesthetics of simple figures”, *British Journal of Psychology*, 71, 505-524.

McManus, I.C., Jones, A., and Cottrell, J. (1981) “The aesthetics of colour”, *Perception*, 10, 651-666.

McManus, I., Cheema, B., and Stoker, J. (1993) “The Aesthetics of Composition: A Study of Mondrian”, *Empirical Studies of the Arts*, 11, 2, 83-94.

Moghaddam, B., Naster, C. and Pentland, A. (1996) "Bayesian Face Recognition using Deformable Intensity Surfaces*", In Proceedings Computer Vision and Pattern Recognition,* 96, 638-645 (http://www-white.media.mit.edu/vismod/demos/facerec/)

Morris, J.S., Frith, C.D., Perrett, D.I., Rowland, D., Young, A.W., Calder, A.J. and Dolan, R.J. (1996) "A differential neural response in the human amygdala to fearful and happy facial expressions", *Nature*, 383, 812-815.

Murray, L.K. (1997) "Unpublished 1st year PhD report"*, School of Psychology, University of St Andrews*, Fife, Scotland.

O'Toole, A.J., Abdi, H., Deffenbacher, K.A. and Valentine, D. (1993) "Low-Dimensional Representation of Faces in Higher Dimensions of the Face Space", *Journal of the Optical Society of America A*, 10, 405-411.

Osgood, G.E., May, W.H., Miron, M.S. (1975) "Cross-cultural Universals of Affective Meaning", *University of Illinois Press*, Urbana, IL.

Parke, F.I. (1982) "Parameterized Models for Facial Animation", IEEE Computer Graphics and Applications, 2, 61-68.

Pentland, A., Moghaddam, B., Starner, T., (1994) "View-Based and Modular Eigenspaces for Faces Recognition"*, IEEE Conference on Computer Vision and Pattern Recognition.*

Penton-Voak, I. (1997) "Unpublished 1st year PhD report"*, School of Psychology, University of St Andrews*, Fife, Scotland.

Perrett, D.I., Lee, K.J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castle, D. and Akamatsu, S. (1998) "Effects of Sexual Diomorphism on Facial Attractiveness", *Nature, under review.*

Perrett, D.I., May, K.A. and Yoshikawa, S. (1994) "Facial Shape and Judgement of Female Attractiveness", *Nature*, 368, 239-242.

Phillips, J.P., Moon, H., Rauss, P. and Rizva, S.A., (1997) "The FERET September 1996 Database and Evaluation Procedure", *First International Conference on Audio and Video-based Biometric Person Authentication*, Crans-Montana, Switzerland.

Pittenger, J.B. and Shaw, R.E. (1975) "Ageing faces as viscal-elastic events: Implications for a theory of nonrigid shape perception", *Journal of Experimental Psychology: Human Perception and Performance*, 1, 374-382.

Pittenger, J.B., Shaw, R.E. and Mark, L.S. (1979) “Perceptual Information for the Age-Level of Faces as a Higher-Order Invariant of Growth”, *Journal of Experimental Psychology: Human Perception and Performance*, 5, 478-493.

Pollock, F., (1996) personal correspondence, ATR, Kyoto, Japan.

Rhodes, G., Brennan, S.E. and 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

Rowland, D.A. and Perrett, D.I. (1995) "Manipulating facial appearance through shape and colour", *IEEE Computer Graphics and Applications*, 15, 70-76.

Rowland, D.A., Perrett, D.I., Burt, D.M., Lee, K.J. and Akamatsu, S. (1997) "Transforming Facial Images in 2 and 3-D*". Imagina-97, Conference Proceedings*, Monaco.

Ruppert, J. (1995) "A Delaunay Refinement Algorithm for Quality Mesh Generation", *Journal of Algorithms*, 18, 548-585.

Sederberg, T. and Parry, S. (1986) “Free-form deformations of solid geometric models”, *Computer Graphics* (SIGGRAP ’86 Proceedings), 26, 25-34.

Shaw, R.E., McIntyre, M. and Mace, W. (1974) “The role of symmetry in event perception”, In R.B.MacCloed & H.L.Picks (Eds.), *Perception: Essays in honour of James J. Gibson*. Ithaca, Cornell University Press, New York.

Shepherd, J.W., Davis, G.M., and Ellis, H.D. (1981) “Studies of cue Latency”, In G. Davies H. Ellis & J. Shepherd (Eds.), *Perceiving and Remembering Faces*, Academic Press, London.

Shewchuck, J.R. (1996) "Triangle: Engineering a 2-D Quality Mesh Generator and Delaunay Triangulator", *First Workshop and Applied Computational Geometry*, Philadelphia, Pennsylvania, 124-133.

Sirovich, L. and Kirby, M. (1987) "Low-Dimensional Procedure for the Characterisation of Human Faces", *Journal of the Optical Society of America A*, 4, 519-554.

Snedecor, G.W. and Cochran, W.G. (1980) “Statistical Methods”, Ames, Iowa: The Iowa State University Press.

Socrates - Philebus in Burke, E. (1757) “A philosophical enquiry into the origin of our ideas of the sublime and the beautiful”, *R & J Dodley*, London.

Sprengelmeyer, R., Young, A.W., Calder, A.J., Karnat, A., Lange, H., Homberg, V., Perrett, D.I. and Rowland, D.A. (1996) "Loss of Disgust - Perception of Faces and Emotions in Huntingtons Disease", *Brain*, 119, 1647-1665.

Stoddard, J. T. (1886) “Composite Portraiture”, *Science*, 8, 89-91.

Symonds, D. (1979) "The Evolution of Human Sexuality", *Oxford University Press*, Oxford.

Syswerda, G. (1989) "Uniform Crossover in Genetic Algorithms," *Proceedings of the Third International Conference on Genetic Algorithms*, Schaffer, J. (Ed.), *Morgan Kaufmann Publishers*, Los Altos, CA, 2-9.

Turk, M.A., and Pentland, A.P. (1991) "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, 3, 71-86.

Vetter, T. and Troje, N. (1995) "A Separated Linear Shape and Texture Space for Modelling Two-Dimensional Images of Human Faces", *Technical Report No. 15*, Max-Planck Institut für Biologische Kybernetik.

Vetter, T. (1996) “Synthesis of Novel Views from a Single Faces Image*”, Technical Report No. 26*, Max-Planck Institut für Biologische Kybernetik.

Vetter, T. (1997a) “Automated Face Morphing and Image Based Modelling of Faces”, *Imagina 1997*, 131-138.

Vetter, T. (1997b) “A Bootstapping Algorithm for Learning Linear Models of Object Classes”, *Technical Report No. 48*, Max-Planck Institut für Biologische Kybernetik.

Wernecke, J., (1995) “The Inventor Mentor - Programming Object-Oriented *3-D* Graphics with Open Inventor, Release 2*”, Addison-Wesley Publishing Company*, ISBN 0-201-62495-8.

Witmer, L. (1894) “Zur Experimentellen Aesthetik Einfacher Räumlicher Formverhältnisse [On the Experimental Aesthetics of Simple Spatial Relations]”, *Philosophische Studien*, 9, 96-144; 209-263.

Wolberg, G., (1990) "Digital Image Warping", ISBN 0-8186-8944-7.