PHD

Making FACES: The facial animation, construction and editing system

Patel, Manjula

Award date:
1991

Awarding institution:
University of Bath

Making FACES

The Facial Animation, Construction and Editing System

submitted by **Manjula Patel**

for the degree of **Ph.D**

of the **University of Bath**

December 1991

Summary

The human face is a fascinating, but extremely complex object; the research project described is concerned with the computer generation and animation of faces. However, the age old captivation with the face transforms into a major obstacle when creating synthetic faces. The face and head are the most visible attributes of a person. We master the skills of recognising faces and interpreting facial movement at a very early age. As a result, we are likely to notice the smallest deviation from our concept of how a face should appear and behave.

Computer animation in general, is often perceived to be "wooden" and very "rigid"; the aim is therefore to provide facilities for the generation of believable faces and convincing facial movement. The major issues addressed within the project concern the modelling of a large variety of faces and their animation. Computer modelling of arbitrary faces is an area that has received relatively little attention in comparison with the animation of faces. Another problem that has been considered is that of providing the user with adequate and effective control over the modelling and animation of the face. The *Facial Animation, Construction and Editing System* or FACES was conceived as a system for investigating these issues.

A promising approach is to look a little deeper than the surface of the skin. A three-layer anatomical model of the head, which incorporates bone, muscle, skin and surface features has been developed. As well as serving as a foundation which integrates all the facilities available within FACES, the advantage of the model is that it allows differing strategies to be used for modelling and animation.

FACES is an interactive system, which helps with both the generation and animation of faces, while hiding the structural complexities of the face from the user. The software consists of four sub-systems; CONSTRUCT and MODIFY cater for modelling functionality, while ANIMATE allows animation sequences to be generated and RENDER provides for shading and motion evaluation.

Dedication

To

Andrew

Acknowledgements

Thanks are expressed to my supervisor, Dr. Philip J. Willis, for support and advice during this research project. My appreciation also extends to other members of the Computer Graphics Group at the *University of Bath*, both past and present, in particular: Dr. N. Michael Brooke, Dr. Andrew Hunter, Dr. John N. Spackman, Dr. Nigel John, Mr. Robert J. Oddy and Mr. Paul D. Templeton.

To the following, I am grateful for provision of valuable information and extremely useful discussions: Dr. Tony Kitson of the *Scientific Research and Development Centre* of the Home Office; Mr. Richard Neave, Director of Medical Illustration of the *University of Manchester Medical School*; and Dr. Johnathan Musgrave of the *University of Bristol.*

Thanks also to Dr. Keith Waters for permission to use adaptations of his diagrams of the three muscle models, and to Mr. Steve Caplin for permission to use the newspaper cartoon strip called *'Camera Obscura'*.

I am greatly indebted to Mr. Andrew Cook for taking on numerous rôles including those of: proof-reader, constructive critic and photographer extraordinaire; but most of all for moral support, encouragement and large doses of enthusiasm. Gratitude is expressed to Mr. Peter Marshall, Dr. Steve Hickey and Dr. Hilary Roberts for useful discourse during the early stages of the project.

A special thanks to my sister, Hasu, without whose initial support I would probably never have experienced university life. Finally, financial support provided by the *Science and Engineering Research Council* from October 1988 to September 1990 is very gratefully acknowledged.

Preface

The following description is an organisational map to guide the reader through the structure of the thesis. Four major sections constitute the overall form of the thesis.

In chapter 1, contemporary problems in the area of computer generated facial animation are identified and the major goals of the research project established. This chapter also forms a discussion of the general area of face processing and ascertains the nature of typical applications that may benefit from a computer based approach to the generation and animation of faces. It provides a context within which the research undertaken may be considered. In addition, a brief overview of the *Facial Animation, Construction and Editing System* or FACES is provided.

Chapters 2, 3 and 4 cover the groundwork on which FACES has been developed. This groundwork forms the foundation upon which facilities for the user are provided. The anatomy of the head and face are examined in chapter 2 to demonstrate the importance of bone, muscle and skin to the objectives of the system. Development of a three-layer anatomical representation for the head is described in chapter 3, while simulation of facial motion is discussed in chapter 4.

Chapters 5, 6 and 7 describe the rationale behind the facilities provided in FACES. At this stage, emphasis is placed on the user's perspective of the system. Chapter 5 is concerned with the modelling of faces and the functionality available in the CONSTRUCT and MODIFY sub-systems. Facial animation, motion specification and control are the themes of chapter 6, in which the ANIMATE sub-system is studied. Chapter 7 addresses rendering in terms of generating visually realistic images, control over colour and evaluation of motion.

In chapter 8 we investigate issues relevant to the usability of the system and the user interface. Some thoughts are aired with regard to further development of the system in chapter 9, while chapter 10 concludes by drawing together the major achievements of the project.

Contents

1 **Synthetic Faces** 1

1.1 Introduction . . . 2

1.2 Objectives of the Research Project . . . 2

1.2.1 Modelling . . . 3

1.2.2 Animation . . . 3

1.2.3 User Control . . . 3

1.3 Research Context . . . 4

1.3.1 Identification and Communication . . . 4

1.3.2 Range of Face Types . . . 4

1.3.3 Reality as a Philosophy . . . 5

1.4 FACES . . . 5

1.5 Applications . . . 6

1.5.1 Advertising . . . 6

1.5.2 Entertainment . . . 6

1.5.3 Medicine . . . 7

1.5.4 Speech Synthesis . . . 7

1.5.5 Criminology . . . 8

1.5.6 Computer Generated Characters . . . 9

1.6 Research Management . . . 10

1.6.1 Multi-Disciplinary Approach . . . 10

1.6.2 Interviews . . . 10

1.6.3 Hardware and Software 10

1.7 Summary and Conclusions 11

2 Anatomy of the Human Head 12

2.1 Introduction 12

2.2 Skin 13

2.2.1 Structure 13

2.2.2 Movement 13

2.3 Muscles 14

2.3.1 Structure 14

2.3.2 Contraction 14

2.3.3 Movement 15

2.4 Bone 17

2.4.1 Race, Gender and Age 17

2.4.2 Arrangement 18

2.4.3 Movement 19

2.5 Eyes 19

2.5.1 Structure 19

2.5.2 Movement 20

2.6 Summary and Conclusions 20

3 Representation of the Head 22

3.1 Introduction 22

3.2 Facial Modelling 23

3.2.1 Morphology 23

3.2.2 Movement 24

3.3 Representation of Skin 24

3.3.1 Techniques for Extraction of Polygonal Data 26

3.3.2 Definition of Facial Topology 26

3.3.3 Acquisition of Topological Data 27
3.4 Representation of Muscle 28
3.5 Representation of Bone 28
3.6 Extensions to the Mesh Model 30
3.6.1 Neck 30
3.6.2 Eyebrows 30
3.6.3 Eyelids 31
3.6.4 Eyeballs 31
3.7 Eye Focusing and Tracking 32
3.8 Summary and Conclusions 33

4 Dynamics of the Head **35**

4.1 Introduction 36
4.2 Facial Movement in FACES 36
4.3 Kinematic Models 37
4.3.1 Shape Interpolation 38
4.3.2 Parameter Interpolation 39
4.3.3 Rotoscopy 40
4.3.4 Performance Driven Animation 41
4.4 Dynamic Models 41
4.4.1 Tension Nets 42
4.4.2 Simulation of Muscles 43
4.4.3 Abstract Muscle Actions 43
4.5 Modelling Facial Dynamics 44
4.5.1 Elasticity of Skin 45
4.5.2 Muscle Contraction 45
4.5.3 Muscle Interaction 49
4.5.4 Parameters of Muscle Models 49

4.6 Facial Action Coding System 52
4.6.1 Suitability 52
4.6.2 Limitations 53
4.7 Time-Variance of Facial Movement 54
4.7.1 Path of Motion 54
4.7.2 Duration of Motion 56
4.7.3 Amount of Movement 57
4.8 Summary and Conclusions 57

5 Creating Faces **59**
5.1 Introduction 60
5.2 Determinants of Facial Form and Appearance 61
5.2.1 Psychology of Recognition 61
5.2.2 Criminology and Identification 62
5.2.3 Facial Reconstruction and Anthropology 65
5.2.4 Medicine and Anatomy 66
5.2.5 Portraiture, Sculpture and Proportion 67
5.3 Computer Representation of the Face 68
5.3.1 Digitisation 68
5.3.2 Shape Interpolation 69
5.3.3 Local Deformation 70
5.3.4 Kit of Parts 70
5.4 Conformation Modelling in FACES 71
5.5 The CONSTRUCT Sub-System 73
5.5.1 Global Control 74
5.5.2 Regional Control 75
5.5.3 Local Control 78
5.6 The MODIFY Sub-System 81

5.6.1 Global Control 81
5.6.2 Regional Control 84
5.6.3 Local Control 84
5.7 Summary and Conclusions 87

6 Animating Faces **91**

6.1 Introduction 91
6.2 Animation in FACES 92
6.3 The ANIMATE Sub-System 94
6.4 Creation of Animation Sequences 94
6.5 Conformation Animation 96
6.6 Facial Animation 97
6.6.1 Expression and Gesture in Emotion 97
6.6.2 Components of Facial Movement 98
6.6.3 Expression Animation 100
6.6.4 Auxiliary Movement 103
6.6.5 Generic Expressions 107
6.6.6 New Expressions 107
6.6.7 Bodily Requirements 111
6.7 Storage and Retrieval of Sequences 111
6.8 Generation of Frames 113
6.9 Summary and Conclusions 118

7 Rendering Faces **119**

7.1 Introduction 119
7.2 Rendering in FACES 120
7.3 The Rendering Pipeline 121
7.3.1 Viewing 121
7.3.2 Visible Surface Determination 121

7.3.3 Scan Conversion 123
7.3.4 Illumination Model 123
7.3.5 Shading Polygonal Meshes 125
7.4 The RENDER Sub-System 126
7.5 Colouration of Features 127
7.5.1 Appearance 128
7.5.2 Emotion 129
7.6 Examination of Frames 129
7.7 Evaluation of Motion 129
7.8 Summary and Conclusions 136

8 FACES for the User 137

8.1 Introduction 137
8.2 The User's Conceptual Model 138
8.2.1 Control for Modelling 139
8.2.2 Control for Animation 140
8.2.3 Control for Rendering 142
8.3 Design of the User Interface 142
8.3.1 Interaction 142
8.3.2 Feedback 144
8.3.3 Experimentation 145
8.3.4 Practicality 146
8.4 Summary and Conclusions 146

9 Further Development 148

9.1 Introduction 149
9.2 Head Model 149
9.2.1 Surface Patches 149
9.2.2 Additional Features 149

9.2.3 Modelling Deformation 150
9.3 Muscle Dynamics 150
9.4 Modelling Faces 151
9.4.1 Skull Prototypes 151
9.4.2 Growth and Ageing 151
9.4.3 Face Composition 152
9.4.4 3D Sculpting 153
9.5 Animating Faces 154
9.5.1 Skin Wrinkles 154
9.5.2 Smoother Motion 154
9.5.3 Timing for Facial Movement 155
9.5.4 A Parser for Action Units 155
9.6 Rendering Faces 155
9.6.1 Illumination and Shading 155
9.6.2 Skin Tone and Texture 156
9.6.3 Colour for Emotion 156
9.6.4 Hair 157
9.7 Speech Synthesis 157
9.8 User Interface 158
9.9 Summary and Conclusions 158

10 Concluding Remarks **160**
10.1 Introduction 160
10.2 Synthetic Faces 161
10.3 Anatomy of the Head and Face 161
10.4 Representation of the Head 162
10.5 Dynamics of the Head 162
10.6 Creating Faces 163

10.7 Animating Faces 164

10.8 Rendering Faces 166

10.9 User Interface 167

10.10 Further Development 168

10.11 Summary and Conclusions 169

Appendices

Origins and Insertions of Facial Muscles . . . . A

Facial Action Coding System . . . . B

Data-Files for FACES C

Video Sequences . . . . D

Structure of FACES . . . . E

Refereed Paper . . . . F

Figures

3.1 Mesh Model 29

3.2 Construction of Eyeball 31

3.3 Eye Focusing and Tracking Mechanism 32

4.1 Geometry of Linear Muscle Model 46

4.2 Geometry of Sphincter Muscle Model 47

4.3 Geometry of Sheet Muscle Model 48

4.4 The Four Laws of Motion 55

5.1 The CONSTRUCT Sub-System 73

5.2 Generation of New Heads 76

5.3 Global Scaling of Skull 77

5.4 Changes to Relative Proportions of the Head 79

5.5 Modifications to Bone 80

5.6 The MODIFY Sub-System 82

5.7 Changes to Overall Proportions of the Face 83

5.8 Changes to Relative Proportions of the Face 85

5.9 Modifications to Features 86

6.1 Animation in FACES 92

6.2 Interpolation between two well-known Faces 93

6.3 The ANIMATE Sub-System 95

6.4 Specification of a Conformation Animation Sequence 96

6.5 Specification of a Facial Animation Sequence 102

6.6 Wireframe Representation of Generic Expressions 108

6.7 Creation of New Expressions 112

7.1 The RENDER Sub-System 127

7.2 The PLAYBACK Facility 135

8.1 Top-level Menus 144

Colour Plates

3.1 Plastic Mask 25

4.1 Linear, Sphincter and Sheet Muscles 51

5.1 Modelling a Variety of Faces 88

6.1 Eyeball Rolling 104

6.2 Fish-Eyes 105

6.3 Winking 105

6.4 Head Movement 106

6.5 Generic Expressions of Emotion 109

6.6 Facial Animation 115

6.7 Conformation Animation without Facial Movement 116

6.8 Metamorphosis with Facial Animation 117

7.1 Constant Shaded Face 126

7.2 Examples of Skin Tones 130

7.3 Colouration of Features 132

7.4 Blushing 133

7.5 Paleness 133

7.6 The STORYBOARD Facility 134

Tables

4.1 Parameters of the three Muscle models 50

5.1 Characteristics of Facial Features 63

6.1 Facial Actions available in FACES 99

Chapter 1

Synthetic Faces

1.1 Introduction 2

1.2 Objectives of the Research Project 2

1.2.1 Modelling 3

1.2.2 Animation 3

1.2.3 User Control 3

1.3 Research Context 4

1.3.1 Identification and Communication 4

1.3.2 Range of Face Types 4

1.3.3 Reality as a Philosophy 5

1.4 FACES 5

1.5 Applications 6

1.5.1 Advertising 6

1.5.2 Entertainment 6

1.5.3 Medicine 7

1.5.4 Speech Synthesis 7

1.5.5 Criminology 8

1.5.6 Computer Generated Characters 9

1.6 Research Management 10

1.6.1 Multi-Disciplinary Approach 10

1.6.2 Interviews 10

1.6.3 Hardware and Software 10

1.7 Summary and Conclusions 11

1.1 Introduction

For millenia mankind has been intrigued with the human face. As a result of this fascination, an intense desire has developed to reproduce the facial form in a variety of media including painting, sculpture and photography. In recent decades, this objective has been pursued through the use of computer technology, in particular since Frederike Parke's pioneering work in the early 1970s [115]. Today, the same interest is the motivation behind the research project described in this thesis. More specifically, the project is concerned with the computer generation and animation of faces.

Within the field of computer generated facial animation, there are several important issues which need to be addressed. In section 1.2 we shall examine these outstanding problem areas to establish the overall objectives of the project. Research was undertaken against a backdrop of aspects concerning the face; these are described in section 1.3. An overview of the software system resulting from the project is presented in section 1.4.

The face is used in many applications for as many different reasons, a few selected applications are described in section 1.5. A major advantage of a computer based approach is that it provides the potential for adaptation to different applications for a large variety of uses. Management of the research project is considered in section 1.6.

1.2 Objectives of the Research Project

In a similar manner to other areas of computer animation, facial animation falls into the two major categories of modelling and animation. The reason why representation and animation of synthetic faces is considered to be special and a major challenge is that our criterion for judging success is extremely stringent.

We are very familiar with faces and facial movement, mastering the skills to interpret faces at a very early age. We therefore have a well developed sense for distinguishing which expressions are natural for a face. The result is that we are likely to notice the smallest deviation from our concept of how a face should appear and behave. In fact, researchers have found that viewers are unsympathetic to flaws, no matter how minor, in a model that claims to be realistic, while readily accepting the inadequacies of a model that is obviously a caricature [118, 176].

Synthesis of the face requires an interactive ability to create face models and to generate and control simulated expressions on such models. At present three major areas of research, each with its own specific problems, can be identified. These areas concern modelling, animation and user control, which form the major objectives of the research project in terms of:

- Investigation of methods for composing and creating synthetic faces.
- Animation of any face constructed within the system in terms of facial movement and non-verbal communication.
- Examination of the most appropriate techniques for giving the user control over both modelling and animation facilities.

1.2.1 Modelling

There is a requirement to be able to produce a wide range of faces. This issue is part of a more general problem; that of modelling and modifying irregular 3D structures. With regard to the face there are two sub-aspects which need to be dealt with. First of all, faces are irregular structures which cannot be as easily represented as regular geometric shapes. Secondly, faces vary from person to person. This latter aspect involves the investigation of *conformation* [118], which is concerned with modelling the form of different faces.

1.2.2 Animation

Generation of natural facial movement is an issue which is particularly important for communication. It is necessary to make face models and their movement more naturalistic and realistic, in the sense of improving accuracy in order to enhance comprehension. This involves investigation of the motion and behaviour of faces and comprises three further issues involving the generation of: facial expressions which are readily comprehensible; believable head movements and gestures; and speech-synchronised sequences.

Speech synthesis has been excluded as an initial goal since it is a major area of research in its own right [15, 17, 53, 68, 123, 156]. However, note that facial animation does include some lip movement.

1.2.3 User Control

It is necessary to provide the user with a flexible, efficient and intuitive manner in which to create faces and to control the subtle movements that make up both facial expressions and lip movement for speech. This last area, regarding the provision of adequate control and support for users is becoming increasingly important. Users nowadays expect high-level systems such that they can specify **which** actions they would like to take place without being concerned about the esoteric details of **how** those actions actually occur.

1.3 Research Context

Several aspects with regard to the face were borne in mind during the development of the research project. These aspects concern: the functions of the face; the range of facial types; and the use of reality as a frame of reference.

1.3.1 Identification and Communication

The face serves two primary functions, those of identification and communication. Infinite variability of structure enables the face to be a unique form for each and every person. Such individuality provides the only attribute that guarantees visual identification. Other characteristics such as gait, voice and mannerisms are sometimes used; however, it is not until we see the face that we are certain of really knowing who a person is.

The face is also of great importance for communication in terms of both speech and facial expression. The significance of these modes of verbal and non-verbal communication as individual factors should not be underestimated. When combined however, they become an even more powerful and versatile means of conveying information.

1.3.2 Range of Face Types

The appearance of the face can range from *stylistic* to *realistic* through to *specific*, according to the fidelity that an application requires in representing both the structure and movement of the face. These three types of facial structure should not be viewed as discrete classes, but rather as particular forms in a continuous range from which the preciseness of a model can be selected.

An example of the stylistic approach is caricature animation such as that used in *Fluck and Law's* highly popular television series called *'Spitting Images'*. Here exaggeration is used as a medium for communicating ideas and messages in order to entertain. The realistic category is distinguished by its requirements for an accurate portrayal of the human face and the plausibility of the motion. Such faces may be very useful for studies undertaken by psychologists for example. Models of specific faces are largely used in the field of medicine which requires a precise representation of particular faces in order to undertake pre-operative investigations.

1.3.3 Reality as a Philosophy

The applications mentioned later in section 1.5 provide an indication of the most appropriate method of meeting the goals of the project; each uses reality as a basis or frame of reference. Since the requirements of applications differ, it is necessary to make the system as flexible as possible while its potential is under investigation. Although there is a need to simulate components of the face, the complexity of the model will depend on the particular application.

Another reason for using reality as a basis is that one of the commonest criticisms of computer generated animation is that it lacks subtlety of movement. Computer animated characters tend not to be as 'alive' as those drawn by hand; they are generally perceived as being 'wooden' and very 'rigid'. The closer a model corresponds to the 'real thing' the greater the opportunity for creating naturalistic, realistic and accurate facial expressions which are believable. Further, by basing a face model on the constraints of real faces, much of the structural complexity can be hidden within the model.

The aim however, is not to allow reality to constrain the functionality in the system. Wherever appropriate, the goal is to provide the user with facilities to take reality further and caricature it rather than merely to imitate it. Herein we encounter the familiar conflict between simulation and animation. Simulation involves exact modelling, with all the constraints that it entails, while animation is the use of a medium for communicating ideas, which often requires artistic license to improve comprehension by focusing on relevant information only. The view that underlies the research project is that not only do both simulation and animation have their place in facial animation, but that the facilities provided for the user can be enhanced by using reality as a frame of reference, while at the same time allowing the user flexibility to override constraints when necessary.

1.4 FACES

FACES is an acronym for the *Facial Animation, Construction and Editing System.* The project aims to provide a software emulation of the human face in 3D. This involves the interactive construction and modification of a head model and its subsequent animation.

The hierarchical nature of FACES is illustrated in Appendix E; it consists of four sub-systems, named: CONSTRUCT, MODIFY, ANIMATE and RENDER. The CONSTRUCT and MODIFY sub-systems enable changes to be made to the structure of the head, at both global and local levels, enabling distinct faces to be created. The ANIMATE sub-system caters for motion specification and control, permiting faces to be animated. The remaining part of the system, RENDER, facilitates the generation of shaded images. This includes control over colour which is important in determining the appearance of the face. Facilities for examining frames and for near real time playback of sequences also exist in the RENDER sub-system.

1.5 Applications

Below are identified several applications which could benefit from a computer based system for modelling and animating the face. Note that the applications cover a large range of disciplines demonstrating the versatility and ubiquitous nature of the face.

1.5.1 Advertising

Frequent use is made of characters and in particular facial animation in advertising to enable unambiguous messages to be broadcast. Expressions and gestures are both used to bring abstract and often inanimate forms to life. Examples include: models of animals such as tortoises, penguins and bears to advertise electricity, based on the original short animated film *'Creature Comforts'* produced by Nick Park of *Aardman Animations*; the *Big Dom* advert for detergent, in which a bottle of detergent is portrayed as a man; and the credit card promoting itself as a 'flexible' friend through the use of facial expressions and body gestures.

The manual method of generating such sequences is to create physical models and then to painstakingly record their movement using stop-frame techniques. The time, expense and effort expended is therefore enormous. A computer based approach would not only provide speed, but also flexibility and cater for experimentation.

1.5.2 Entertainment

Entertainment, like advertising, also relies heavily on effective character and facial animation. Both stylistic and the more realistic characters are in common use.

Cartoon features from Walt Disney, Warner Brothers and others have put smiles on our faces for many years. Such animators are very skilled in the art of telling stories which are funny and endearing. They have mastered the art of creating empathy and human emotion towards what are intrinsically mere drawings. We find cartoons entertaining because the characters are in fact caricatures of human emotions and personalities. These aspects are communicated through both facial expression and bodily gestures.

An ex-Disney animator, Lasseter, has demonstrated the potential of 3D computer graphics for character animation. This has been exemplified by a number of short films, namely: *'Luxo-Junior'*, *'Red's Dream'*, *'Tin-Toy'* and *'Knick-Knack'*. These productions have shown that many of the techniques and principles used in conventional 2D character animation are equally applicable in 3D computer animation [87]. In terms of facial animation the most important of the principles are *squash*, *stretch* and *exaggeration*.

1.5.3 Medicine

Previews of corrective plastic surgery and dental treatment on specific faces are of great interest to both practitioners and patients alike [34, 192]. Such applications demand precise models of particular individuals based on the bone and soft tissue of the head. These techniques further require a means of interacting with the model and quantification of any changes that are made.

A computerised system, which incorporates an anatomically complete model of the head and face, would provide surgeons with the capability to plan, and even rehearse, complex operations without the need to undertake costly and potentially dangerous exploratory surgery.

1.5.4 Speech Synthesis

In every-day life speech is probably the most important form of communication in common use. However, facial expressions have an important complementary rôle to play in emphasising the tone of the audible signal in order to provide a clearer message [16].

There is a distinction to be made between producing synthetic sound and that of creating images of lip movements synchronised to a prerecorded sound track. Both of these processes require investigation and analysis of visible and acoustical speech signals [15]. With regard to facial animation the focus has so far been on producing speech-synchronised animated sequences.

A face model incorporating accurate speech synthesis could prove to be useful for the deaf and hard-of-hearing. Pearson describes methods of extracting features from images of the face and hands to enable deaf people to converse with one another over a telephone network using sign language [124].

Accurate speech synthesis is however, proving to be extremely difficult [15, 17, 53, 68, 123, 156]. Most realistic talking sequences are still produced by tracing individual frames from live-action footage [115, 131, 175]. This is an extremely tedious, expensive and time-consuming process. A promising approach is to simulate the physical characteristics of the voice-producing parts of the head [90].

1.5.5 Criminology

Recognition and identification of faces is an important aspect of human psychology, particularly in the field of criminology. A considerable amount of research has been undertaken by psychologists with respect to the face. Particular issues addressed include how faces are represented in memory and which facial features are used most often in the identification process. On the one hand answers to such queries can provide important guidelines for the construction of 3D computer models of realistic heads and faces. Conversely, a system for modelling faces and their motion could prove to be of great use in studying the recognition process itself [18].

Reconstruction of realistic faces from skeletal remains is of immense interest in forensic medicine as well as in archeology. Victims are sometimes found with no clue to their personal identification [141]. When other procedures such as dental checks cannot be used, facial reconstruction can be employed to assist in the identification of the victim. The process is a difficult one and consequently, there is only one person in the United Kingdom and very few in the world who have the required expertise [54, 141]. Facial reconstruction was first used by His in the 17th Century [85] to reconstruct the face of Bach, and more recently by a French anthropologist to create Mozart's face from his skull [8, 65].

Neave, who is the leading proponent of the technique in this country, has employed facial reconstruction in both archeology and to aid in the identification of unknown persons. With considerable success, he has reconstructed the faces of: Philip II of Macedon, King Midas, three Egyptian mummies and Lindow Man [106, 109, 141] as well as those of a Finnish nurse, Sabbir Kussam Kilu and Karen Price [107, 141]. However, reconstruction artists such as Neave, have always maintained that the method is not 100 per cent accurate; the technique provides a likeness of the original person, but not a portrait [54, 107, 141]. In fact there has been considerable concern about the accuracy and reproducibility of results by different people and the amount of subjectivity on the part of the artist that can creep into the model [108, 170].

A computer based approach to the process of facial reconstruction would provide two advantages over a physical model. Firstly, it would reduce the amount of subjectivity that would otherwise be involved in the manual procedure. Secondly, there would be opportunity to make the model move, change expression and even appear to talk.

1.5.6 Computer Generated Characters

Over the past decade there has been unparalleled growth in the computer modelling and animation of human figures. They have become a standard feature of advanced research into animation techniques [3, 4, 5, 6, 7, 51, 82, 95, 97, 181, 184, 187]. Certain aspects of human communication are especially important as models on which to base synthetic actors in computer generated films [100, 132]. The stories depicted by such films require actors with human-like characters and personalities. People tend to want to see human characters in action since the overuse of special effects can quickly become tiresome. Research into human figure animation has been concentrated in the three major areas of face, hand and body animation. Face and hand modelling need special attention because they entail specific problems.

The modelling of human characters breaks down into two fields; in one characters are created anew, for example *Nester Sextone* and *Dozo* [93], while in the other characters are resurrected from the dead, for example *Marilyn Monroe* and *Humphrey Bogart* [100] and more recently *Elvis* [114]. There are moral, ethical and copyright issues associated with the second [148], while both share in the quest to model accurately the physical and behavioural characteristics of humans.

The ultimate goal is an integrated system for the creation and animation of synthetic actors. According to Magnenat-Thalmann and Thalmann [99, 161] synthetic actors should ideally: appear and behave like real people; have their own personality; be conscious of their environments; be capable of obeying task-level commands; and have faces and bodies that deform naturally.

These ambitious goals are likely to be undermined without successful modelling of the expressiveness of the face. It is only necessary to consider the number of close-ups of a character's face that are currently used in films to convey to the viewer some deep emotion that the character is feeling.

1.6 Research Management

A variety of techniques were used throughout the duration of the research project. These ranged from information gathering activities to writing computer software for testing and evaluating particular ideas.

1.6.1 Multi-Disciplinary Approach

It was clear from the outset that the nature of the face would demand a multi-disciplinary approach to the conduct of the research. The areas covered include biology, anatomy, forensic science, criminology, anthropology, psychology and art, as well as computer graphics and computer animation. Literature searches were undertaken in pertinent areas of each of these fields in order to grasp fundamental results, limitations and problems that still remained.

1.6.2 Interviews

Several personal interviews were conducted with the following people: Tony Kitson of the *Scientific Research and Development Centre* of the Home Office, with regard to the Photofit and Electronic-fit systems; Richard Neave, Director of Medical Illustration at the *University of Manchester Medical School*, in connection with facial reconstruction; and Johnathan Musgrave of the *University of Bristol*, in his capacity as lecturer in anatomy and anthropology. Each of these gentlemen imparted valuable information which influenced the development of FACES.

1.6.3 Hardware and Software

FACES is implemented in ANSI C for portability. Underlying the software for the system is a graphics library, called *Gigalib* [189], which was developed at the *University of Bath.*

The hardware configuration that was available during the development of the system comprised: two HLH Orion Super-Minicomputers (running Berkeley UNIX 4.2); two

Mitsubishi C9918NE 8-bit colour displays (1280 × 1024 resolution); two Summagraphics MM1102 digitising tablets with puck; a Howtek Scanmaster; a Matrix Camera and Cine Recorder; and an Apple II laser printer.

1.7 Summary and Conclusions

The face fulfills two major purposes; these concern identification and communication. In addition, the appearance of the face may range from stylistic to realistic to specific.

There is no doubt that the face is a diverse and fascinating structure which is used in numerous and varied applications. All of these applications, which cover a broad spectrum, from advertising to criminology, could benefit from a computer based system such as FACES.

The applications considered suggest that an approach based on reality may hold promise in achieving the objectives of the FACES project. The first goal concerns the modelling of a variety of faces, an area which has to date received relatively little attention in comparison with the animation of faces. The second objective requires realistic animation of any face modelled, while the third goal is to develop appropriate control facilities for the user. An additional requirement is the integration of all the objectives mentioned into one system. In order to develop an approach that uses reality as a frame of reference, it was necessary to undertake a research strategy based on the multi-disciplinary nature of the face.

Although caricatures are suitable for some applications, there is a major requirement for more realistic and expressive faces which are plausible. We shall therefore take the view that an animation system should provide flexible tools and leave the development of a character's personality to skilled animators. As Thomas and Johnston explain [162], to produce good quality animation, whether by traditional methods or with a computer, necessitates familiarity and understanding of the principles of animation [87]. Such an understanding can come only from a thorough knowledge of what actually happens in reality; in other words it is necessary to have a benchmark from which abstractions can be made. It is only then that the user can develop a character by selectively exaggerating or de-emphasising particular attributes.

Chapter 2

Anatomy of the Human Head

2.1 Introduction 12
2.2 Skin 13
2.2.1 Structure 13
2.2.2 Movement 13
2.3 Muscles 14
2.3.1 Structure 14
2.3.2 Contraction 14
2.3.3 Movement 15
2.4 Bone 17
2.4.1 Race, Gender and Age 17
2.4.2 Arrangement 18
2.4.3 Movement 19
2.5 Eyes 19
2.5.1 Structure 19
2.5.2 Movement 20
2.6 Summary and Conclusions 20

2.1 Introduction

In this chapter we are concerned with the anatomy and physiology of the head and face. Bone, muscle, skin tissue and surface features, such as the eyes, form the major components of the head. In addition, the structure and motion of each of these fundamental elements have great influence over both the appearance and movement of the face. Consequently, the *histology, myology* and *osteology* which correspond to skin, muscle and bone respectively were studied, together with the eyes [62, 147], as an essential precursor to the development of FACES.

The structure and movement of skin is discussed is section 2.2. Major muscles of the face and their characteristics of movement are described in section 2.3. The skull, which forms the armature for the head and face, is presented in section 2.4. Eyes, which are important for both appearance and facial expression, are considered in section 2.5.

2.2 Skin

Facial skin is important because it covers a large part of the face. This makes skin a particularly memorable attribute of the face. Skin is significant in determining both the appearance and the movement of the face.

2.2.1 Structure

Human skin comprises several layers; it consists of the *dermis* and the *epidermis* which covers it. Beneath the *dermis* lies a layer of loose irregular connective tissue which forms the *superficial fascia* also known as the *hypodermis* or *subcutaneous* surface. This arrangement allows the skin considerable freedom of movement over the muscles and bones that lie underneath.

The visual appearance of facial skin is dependent on several factors including texture, depth and colouration. Texture depends mainly on the glands contained within the skin. Depth of the *dermis* varies over different areas of the face, for example the lips are very thick while the eyelids are thin and delicate. Colour is determined by blood circulation, the presence of pigments and health. Skin-tone also varies greatly from infancy to adulthood, as well as between males and females, and between different races.

2.2.2 Movement

As well as being tough, flexible and highly elastic, the varying thickness of the *dermis* also affects the motion characteristics of the skin. Tension in the skin can be caused by the action of muscles, which is also likely to affect colouration. For example, states of fear and embarressment can cause paleness or blushing respectively.

The outer surface of the skin, the *epidermis*, is marked by three main types of markings known as tension and flexure lines, and papillary ridges. Tension lines form a network of linear furrows of variable size. Flexure lines or skin joints correspond to folds in the *dermis* associated with habitual joint movement and to lines of attachment to the underlying deep *fascia*. Papillary or friction ridges appear only on the palms of the hand and the soles of the feet.

Ageing reduces the elasticity of both the *dermis* and *epidermis*. This results in wrinkles, such as those under the eyes, and deepening of flexure lines from constant joint movement, for example around the region of the mouth.

2.3 Muscles

The form of the face is to a large extent dependent on the size, thickness and shape of the major muscles. However, the most significant rôle that muscles play is in determining facial movement.

2.3.1 Structure

Muscles lie between the bone and skin. Attachment at the bone end is known as the *origin* while the connection into the *fascia* of the skin is called the *insertion*. All facial muscles, with the exception of the *orbicularis oris*, emerge or have origins on the underlying bone and insert into the skin. The origins and insertions of the major muscles are described in Appendix A.

2.3.2 Contraction

All facial actions occur as a result of muscular contraction. Consequently, it is the interaction of various muscles that causes expressions to appear on the face.

Muscles occur in a large variety of sizes, shapes and complexity, but they are essentially bundles of fibres that operate in unison. It is the arrangement of these fibres that determines both the relative strength and range of movement that a muscle can produce. For example, muscles with only a few long fibres have great range of movement and less strength, while muscles with many but short fibres have greater power and less range of motion. As a result, six general forms of muscle can be identified according to the bundling of their fibres, they are called: *quadrilateral, fusiform, triangular* or *pyramidal, rhomboidal* or *penniform, bipenniform* and *spiral.*

A majority of facial muscles can be classified into one of three groups [62]: *parallel,* *oblique* and *spiralized*. Names given to the muscle categories are indicative of the type of contraction that they cause; the orientation of the fibres that make up the muscles determine the direction of pull at their attachment. In each case the contraction is towards the attachment to bone. There are two types of muscular contraction *isotonic*, during which the muscle shortens and *isometric*, during which it develops tension but does not shorten; all facial muscles contract isotonically.

2.3.3 Movement

Facial muscles are in general thin, voluntary and subcutaneous. They also occur in pairs so that there is one for each side of the face. Muscles tend to be considered according to the region in which they occur [62, 147], below we adopt a similar approach in examining their movement with respect to the forehead, eyelids, nose, mouth, cheeks and chin.

Forehead

At the top of the head the *epicranius* consists of the *occipitofrontalis* and the *temporoparietalis* [62, 147]. Covering the dome of the skull, the muscular layer of the *occipitofrontalis* consists of four parts. The two *occipitals* lie at the back while the *frontals* cover the forehead. The *occipital slips* draw the scalp backwards while the *frontal slips* act from above to raise the eyebrows and the skin covering the root of the nose. The frontal muscles come into play during actions such as glancing upwards and in expressions which involve raising of the eyebrows, such as surprise, horror or fright. Years of expressing emotional reactions in this manner gradually leaves horizontal wrinkles across the forehead.

Eyelids

A broad, flat, elliptical muscle called the *orbicularis oculi* forms the eyelids; this surrounds the eye socket and extends into the *temporal* region and cheek area. The muscle comprises three sections which are known as the *orbital*, *palpebral* and *lacrimal* parts. The *palpebral* portion closes the eyelids gently when sleeping or blinking. The act of blinking keeps the eyes lubricated to stop them from drying out. The *lacrimal* part compresses the *lacrimal sac* to restrain the flow of tears. The *orbital* portion retracts the skin of the forehead, temple and cheek, which causes the eyelids to close firmly to protect the eyes in an emergency. In strong contractions the entire *orbicularis oculi* is involved in causing the skin to be thrown into folds or 'crows feet', radiating from the lateral angle of the eyelids.

Raising the upper eyelid and exposing the front of the eyeball, the *levator palpebrae superioris* is an antagonist of the *orbicularis oculi*. The *corrugator supercilli* is a small pyramidal muscle located at the medial end of the eyebrow and goes deep into the frontal part of the *occipitofrontalis* and the *orbicularis oculi*. This muscle draws the eyebrow downwards and in a medial direction to produce vertical wrinkles in the forehead. The *corrugator* muscle also draws the eyebrow downwards in bright sunlight, as well as during a frown.

Nose

The *procerus, nasalis* and *depressor septi* form the muscles of the nose. The *procerus* is located at the bridge of the nose. Working together with the *corrugator* muscles, it produces transverse wrinkles in the skin between the eyebrows. Two sections known as the *transverse* and *alar* parts form the *nasalis*. The *transverse* part compresses the *nasal aperture* at the junction of the *vestibule* with the *nasal cavity*. The *alar* part draws the *ala* downwards and laterally, assisting in the widening of the *anterior nasal aperture*; these actions become apparent during deep respiration. The *depressor septi* is attached to the *maxilla* above the central incisor tooth and runs deep into the mucous membrane of the upper lip. This muscle assists the *alar* part of the *nasalis* in widening the *nasal aperture* during deep inspiration.

An antagonist to the *nasalis* is the *dilator nares* which dilates the nostril. Another muscle, which runs down from the sides of the bridge of the nose and past the nostrils to the upper lip, is called the *levator labii superioris alaeque nasi*. This muscle raises the upper lip and the wing of the nose. Together with the *zygomaticus minor* it forms the *nasolabial furrow* which extends from the side of the nose to the upper lip.

Mouth, Cheeks and Chin

Numerous muscles around the mouth and lips make this the most flexible area of the face. It is a region which is extremely important for articulation during both speech and mastication. One of the most prominent muscles in this region is the *orbicularis oris*, a thick sphincter muscle that makes up the bulk of the lips. The main function of the *orbicularis oris* is to close the lips. With the use of deep and oblique fibres the muscle can compress the lips against the teeth, bring them together and protrude them.

During sucking and whistling, the *buccinator* or trumpeter muscle pulls the cheek inward. The muscles that cause smiles to be displayed are the *zygomaticus major* and *minor*, see Plate 6.5. Of the two, the major muscle pulls the corner of the mouth upward and backward, while the lesser muscle draws the upper lip upward and outward. The angle of the mouth during a grin is retracted by the *risorius*. The *levator anguli oris* raises the corner of the mouth, while the *depressor anguli oris* draws the corner of the mouth downward in an expression of sadness, as shown in Plate 6.5. Contraction of the *levator labii superioris* raises the upper lip. When working together with the *zygomaticus* the *depressor labii inferioris* draws down and everts the lower lip.

The front of the chin is covered by a small, thick muscle called the *mentalis*. It raises and protrudes the lower lip as well as wrinkling the skin and chin when expressing doubt for instance.

2.4 Bone

Major differences in the form of individual heads occur due to several factors including age, race, gender and hereditary reasons [54, 104, 110]. In addition, variability in the size, shape and relative placement of the major bones is responsible for the overall shape and proportions of the head and face. These factors together with deviations in fat and fascia tissue are extremely important in determining the appearance of the face.

2.4.1 Race, Gender and Age

With respect to race, the nature of the skull broadly varies between Negroid, Caucasoid and Mongoloid. The most salient racial differences occur with regard to nasal aperture breadth, orbital placement and shape, and alveolar prognathism [146].

A distinction that is often made concerns the structure of the mandible. Some groups, such as Polynesians for example, have mandibles whose lower border tends to be convex rather than straight. The result is that when they are placed on a flat surface, they rock and have therefore come to be known as 'rockers' as opposed to 'non-rockers' [70, 104].

Further variations occur due to gender and age. The male skull has larger proportions than that of the female. In particular, it has a larger dental arch; rounded orbits and heavy bossing over the forehead. The female skull has similar characteristics to those of a child; a small face with large eyes. It further has a light jaw and a pointed chin. The effects of ageing on the skull result mainly from 'bonery absorption', leading to a shrinking in the size of the *maxillae* and *mandible*, due to loss of teeth, as well as a thinning of the cranium. Such effects are however not as apparent as the dramatic loss of elasticity in the facial skin.

2.4.2 Arrangement

The front elevation of the skull has an oval profile which is wider at the top than at the bottom. A closer examination reveals that the skull consists of two parts, the *cranium* and lower jaw which is also known as the *mandible* [62, 147]. The upper part of the cranium protects the brain and is called the *calvaria.* The rest of the skull is known as the *facial skeleton,* of which the upper part is fixed to the calvaria and the lower part is the freely moving mandible.

Numerous bones make up the structural support provided by the skull, however only the most significant of these need to be considered for our purposes. The lower part of the skull, or facial skeleton, is very irregular and consists of several major bones and a number of discrete sections.

The forehead consists of the *frontal* bones which form the eyebrow ridge, and the upper parts of the eyeball sockets. Eye sockets form the recesses of the orbits which contain the eyeballs and their associated muscles. Each orbit has a conical shape which is made up of several major bones, notably the: *ethmoid, lacrimal, maxilla, zygomatic* and the *sphenoid.* The eye sockets determine the extent of the separation of the eyes to ensure binocular vision.

The prominent *zygomatic* bones form the cheek bones. They originate in the *squamous part* of the *temporal* bone which lies above the ear and extend around the facial skeleton to form the base of the orbit.

In the upper jaw the *maxillae* form the roof of the mouth and the floor of the nasal cavity and orbit. The upper teeth are embedded in these bones. The nasal bones separate the *frontal processes* of the two *maxillae* and form the upper boundary of the *nasal aperture.*

Forming the largest and strongest bone of the face, the *mandible* has a curved horizontal *body* which is convex forwards and two broad *rami*, which project upwards from the posterior end. The upper border of the bone, known as the *alveolar* part, contains sixteen sockets for the roots of the lower teeth.

Relative sizes, shapes and distances between these major bones are infinitely variable. This variation is responsible for the unique characteristics of each face. For example, growth of the *maxillae* is responsible for elongation of the face, while prominence of the *zygomatic* bones determines the roundness of the face. These characteristics aid visual recognition of the face.

2.4.3 Movement

The *lower mandible* is the only bone in the head which is capable of movement. This bone articulates with the *temporal* bone to form the *temporomandibular joint*. The upper jaw remains fixed while the lower jaw moves in a downward direction to open the mouth.

Speaking and chewing involve three main types of movement comprising elevation, depression and side-to-side grinding. Two strong muscles called the *masseter* and the *temporalis* act as elevators to close the jaw when biting or clenching the teeth. The *medial pterygoids* close the jaw and cause the *mandible* to protrude. A side-to-side grinding movement is produced when the *lateral pterygoid* muscles act alternately. Lowering of the jaw is caused by simultaneous contraction of these muscles.

2.5 Eyes

Eyes are a feature in themselves since they have a distinct form and specific motion characteristics. They are a particularly noticeable attribute and therefore significant to both the appearance and motion of the face.

2.5.1 Structure

The eyeball consists of three distinct elements [62, 147], namely: the white of the eye known as the *sclera*; the *iris*; and the *pupil*. The *sclera* is a dense, hard membrane which serves to maintain the shape of the eyeball. The *iris* is a delicate and adjustable diaphragm which surrounds the pupil orifice. The *pupil* itself controls the amount of light entering the eye. As a result of this function, the size of the pupil changes so that pupil dilation is apparent in emotions such as anger, surprise and fear which are depicted in Plate 6.5.

2.5.2 Movement

In addition to the muscles that raise and lower the eyelids, which have already been described in section 2.3.3, there are six pairs of muscle that are attached to the eyeballs and which move them within their sockets. These muscles are known as *extrinsic* muscles of the eyes. *Intrinsic* muscles modify the shape of the lens during focusing and regulate the amount of light entering the eye.

Four of the six pairs of extrinsic muscles are 'straight' muscles and are named according to their position relative to the eyeball and the actions that they perform. The muscles are known as the *medial, lateral, superior* and *inferior rectus* muscles. The remaining two muscles have a slanting or oblique arrangement and are called the *superior* and *inferior oblique* muscles.

When the eyes are functioning properly, both eyeballs move in coordination so that stereoscopic vision is achieved and maintained. Each eyeball can be considered to move around three axes. These are *anteroposterior* for rotational movements; *vertical* for abduction, adduction; and *transverse* for elevation and depression.

2.6 Summary and Conclusions

An anatomical study of the human head has revealed that each of the layers: bone, muscle and skin has a profound influence over the conformation and motion characteristics of the face. Although all human faces have the same physcial structures, as far as form and appearance is concerned, the bones determine the overall size and proportions of the head and face; they vary with race, gender and age; the size, shape and thickness of muscles are influential factors which depend on the health of an individual; and the texture and colour of skin, as well as surface features are important because they are particularly visible attributes of the face. The motion of the face is dependent on: muscular contractions, of which three types can be identified; the elasticity of skin, which varies with age and health; and movement of features and organs such as the eyes and tongue.

Given the major objectives of the FACES project, it is clear that a three-layer computer model of the head is required for the research project to be viable and credible in meeting the goals outlined in section 1.2. To attain any degree of realism it is necessary to take account of the structure and motion of skin, muscle and bone since complex interactions occur between the layers. However, since each of the layers has very different characteristics in terms of structure and motion, it is necessary to treat them as separate entities for the purpose of modelling.

Chapter 3

Representation of the Head

3.1 Introduction 22
3.2 Facial Modelling 23
3.2.1 Morphology 23
3.2.2 Movement 24
3.3 Representation of Skin 24
3.3.1 Techniques for Extraction of Polygonal Data 26
3.3.2 Definition of Facial Topology 26
3.3.3 Acquisition of Topological Data 27
3.4 Representation of Muscle 28
3.5 Representation of Bone 28
3.6 Extensions to the Mesh Model 30
3.6.1 Neck 30
3.6.2 Eyebrows 30
3.6.3 Eyelids 31
3.6.4 Eyeballs 31
3.7 Eye Focusing and Tracking 32
3.8 Summary and Conclusions 33

3.1 Introduction

In this chapter we identify the characteristics which are necessary of a head model to facilitate conformation modelling and animation. Although an exact simulation is not a requirement, the model does need to be relatively realistic or 'life-like' in order to be comprehensible. It is necessary that the visible surface reflects changes in the underlying muscle and bone structure. This feature is particularly important when the face is animated. Dynamic consistency is required so that the face remains recognisable while appearing to behave in a human manner.

Anatomical considerations such as those discussed in the previous chapter dictate the development of a three-layer model comprising skin, muscle and bone. However, as established in chapter 2, bone, muscle and skin have differing characteristics in terms of both form and motion [62]. The head consists of a rigid arrangement of bone called the skull. This is not only covered, but interleaved by an intricate layer of tissue or the muscles. The muscles are in turn overlaid by a semi-elastic deformable membrane, known as the skin. The head is therefore partially rigid and partially flexible. The implications of this on techniques for the modelling of the face are described in section 3.2.

A head model in FACES comprises a surface skin, referred to as the face, together with an underlying representation for the skull. The muscle layer connecting the bone and skin is modelled in terms of muscle vectors as proposed by Waters [175]. Development of each of these three layers is described in sections 3.3, 3.4 and 3.5. A model based on the anatomy of the head provides a method of unifying the three functions of creation, modification and animation of faces.

Careful representation and structuring of the three main anatomical layers is necessary since the model must also allow integration of modelling and animation functionality within one system. The aim is to be able to animate any face that is constructed within FACES. It is therefore also necessary to develop a representation that will accommodate user control in a natural manner. Extensions to the head model are discussed in section 3.6 while the model for eye movement is described in section 3.7.

3.2 Facial Modelling

In order to provide facilities for modelling differing faces together with their animation, two types of modification are required. Conceptually, these can be regarded as either *inelastic* changes or *elastic* changes.

3.2.1 Morphology

As explained in chapter 2, differences in the morphology of the head result from variations in bone, muscle, skin and surface features. A separate, but related issue is that of growth and ageing. Such changes affect the structural characteristics of the head and are considered permanent or long term changes. In modelling terms these changes to the structure of the head model can be viewed as **inelastic** modifications. Inelastic deformation can be effectively employed to develop a particular head and face, in effect defining its 'neutral' state without expression. Therefore, the CONSTRUCT and MODIFY sub-systems, which are described in chapter 5, utilise inelastic changes to allow the user to model distinct faces. Inelastic changes enable a face to have unique features which faciliate identification.

3.2.2 Movement

The other major function of FACES concerns the modelling of facial movement. In chapter 2 we saw that all such movement occurs as a result of muscular contraction which involves bone, muscle, skin and surface features of the face. This type of change is considered to be temporary or short term in nature. With respect to modelling facial movement such changes are regarded as **elastic** changes. Muscle models which simulate muscular contraction are therefore used to cause transitory deformations of the face. A detailed description of the muscle models is provided in chapter 4.

3.3 Representation of Skin

The research project required an empirically derived 'average' head and face model on which to base the computer model. However, measurement of a large sample population of faces was prohibitive in terms of both time and finance. The data was therefore obtained from a plastic mask which is illustrated in Plate 3.1.

The mask provided a suitable alternative to an average face in two respects. Its use avoided reliance on a human face which would have resulted in the data being specific to a particular individual's face. In addition, the mask neither has pronounced features nor does it exhibit any expression, thereby making it an ideal representation from which other faces can be moulded and developed.

In the head model developed, the facial skin is represented using a polygonal mesh, see Figure 3.1. This approach was adopted for several reasons. A polygonal mesh allows irregular shapes such as the facial features to be represented easily. The technique also enables both global and local geometric transformations to be applied. This allows the head to be moved, scaled or rotated as one structure, as well as permiting local deformations to be applied to specific areas of the mesh to create facial expressions. Manipulation

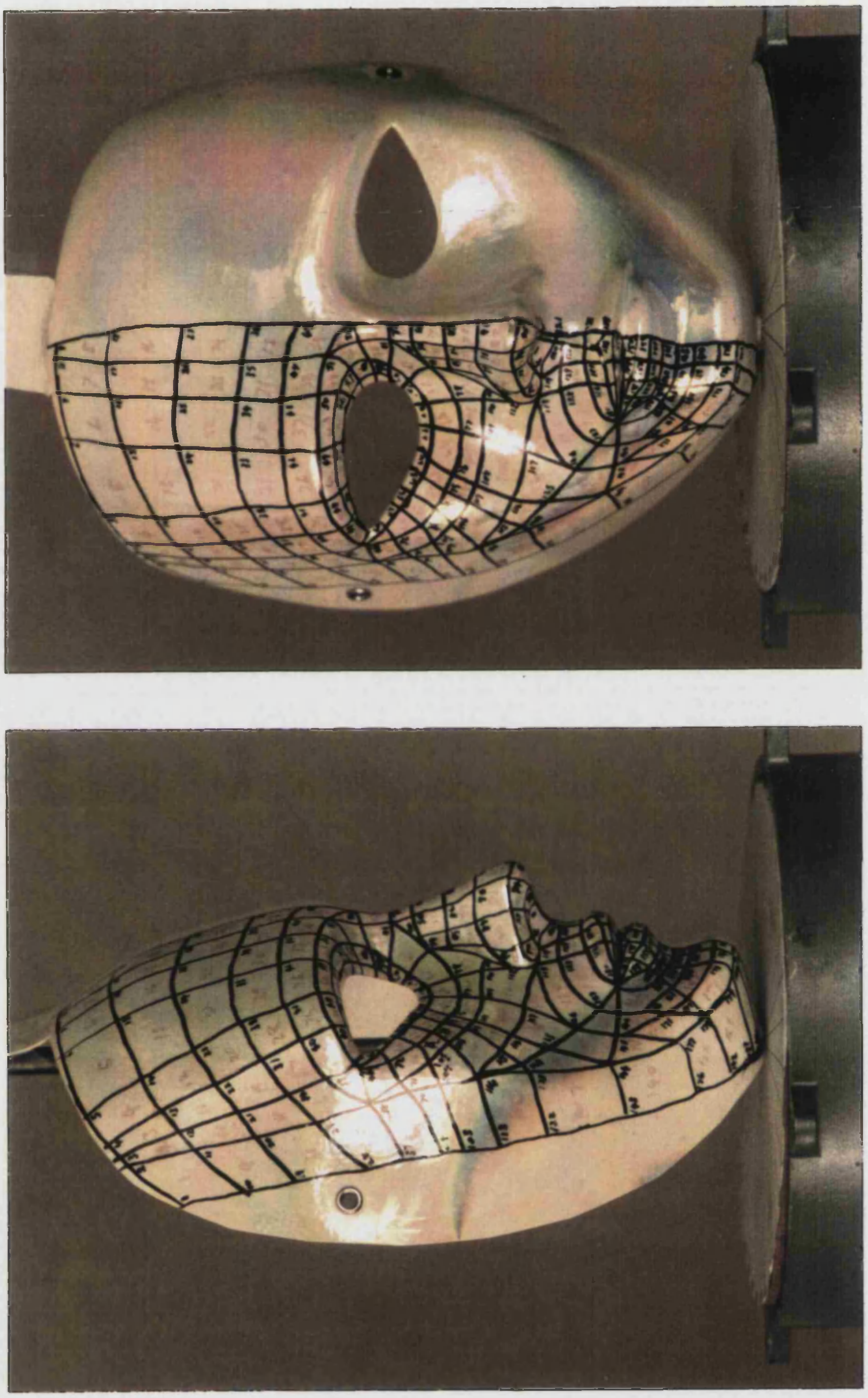

Plate 3.1: Plastic Mask

Data was acquired from a theatre mask rather than a real human face to avoid use of a mesh topology specific to one particular person. A mesh was drawn on the mask which was later digitised through stereo photogrammetry.

of a polygonal mesh is not as computationally expensive as that of surface patches. Furthermore, well-established, fast and efficient algorithms can be used for clipping, hidden surface removal and shading of polygonal meshes [21, 50, 60, 111]. In addition, hardware support with respect to rendering and matrix operations for rigid transformations is now widely available.

3.3.1 Techniques for Extraction of Polygonal Data

Lofting techniques, laser or light beam scans, photogrammetry and manual digitisation are all approaches which can be used to obtain a polygon or patch mesh. Lofting methods are often used in medical imaging. More specifically, Computed Axial Tomography (CAT) is used for hard tissue and Electro-Magnetic Resonance (EMR) for soft tissue [168]. Lofting techniques involve the extraction of 'slices' of 2D data. These cross-sections are later stacked and linked together to recreate a full 3D head model. Coarseness of the data can be varied by regulating the intervals at which slices are taken.

Laser and light beam scanning are capable of providing a high resolution of the surface detail. However, this is only possible if a large number of points are recorded, often amounting to as many as 70,000 [93]. Such large quantities of data pose problems regarding storage, manipulation and control of the data.

It was infeasible to use either lofting or laser scanning to digitise the mask since both approaches require specialised equipment. The two remaining alternatives were photogrammetry and manual digitisation; both of which first require a mesh to be drawn on the surface of the mask. Stereo photogrammetry was eventually used to record the topology of the mask. This technique was adopted because points recorded using a manual 3D digitiser proved inaccurate due to insufficient rigidity of the mask and due to restrictive movement of the tool-tip.

3.3.2 Definition of Facial Topology

Approximation of a face involves sampling the surface at various points. The points are then connected to adjacent vertices to form a polygonal surface skin. The face is assumed to be approximately symmetrical about the central vertical meridian. A line was therefore drawn to represent this plane of symmetry on the mask. The right side of the mask was used to provide data for the right side of the mesh. This was then reflected about the vertical plane onto the left side to create a complete symmetrical face model.

The human face is composed of many prominences and depressions. It is therefore necessary to select vertices at such topographical extremes to avoid losing the relief and consequently 'smoothing out' the features of the face. Major features such as the eye, nose and mouth were subsequently outlined to enhance their general shapes and to ensure that polygonal boundaries would coincide with both creases and colour boundaries of particular features. This technique avoided a 'flattening' of features during the shading stage when colours were interpolated between the boundaries of the polygons. Mesh lines were next constructed within the outlined features whilst maintaining continuity with the rest of the face. A larger number of polygons were used in areas of high curvature, in particular around the eyes, nose and mouth to capture the topography of the face.

Each vertex and facet which forms the mesh was uniquely numbered to facilitate easy referencing. The number of polygons was minimised to enable effective manipulation of the mesh and to avoid a 'data explosion' which can occur with laser scanning techniques [49]. Despite being labour intensive, such an approach does allow explicit control over those facets and vertices that are required. A further advantage of the approach is that various features such as the eyes and mouth can be isolated into discrete structures which can then be manipulated individually.

3.3.3 Acquisition of Topological Data

Once a polygonal mesh had been defined, the mask was photographed from two orthogonal views, one from directly in front and the other from the side, 90° to the first. The views selected enabled the coordinate axes to define the symmetry of the face. Having obtained photographs of the mesh from both the front and side elevations, a program was written to enable the photographs to be manually digitised using a puck and tablet.

The first two points which were recorded established the coordinate frame and bounding box of the half face viewed in the photograph. All subsequent points sampled were mapped from the coordinate system shown in the photograph into screen device coordinates. Plate 3.1 shows the front elevation view taken at 0°, to provide the x-y coordinates of the facial mesh, while a side elevation at 90°, produced the y-z coordinates. According to Magnenat-Thalmann et al. the difference in angles should be greater than 15° and less than 165° [100]. Measurements of vertices were recorded carefully to ensure that the coordinate systems for both views corresponded with each other. A routine was then written to merge the data-sets to produce a 3D polygonal mesh model of the face mask.

Despite the benefits of photogrammetry, the technique has two main shortcomings. Firstly, photographs are perspective projections and not orthographic. The resulting distortion in the coordinates was minimised through use of a long focal lens attached to the camera. The coordinates did nonetheless need some adjustment to counteract the effects of perspective distortion. Secondly, several points were not visible in both of the views; in such cases a best guess estimate was required for at least one of the coordinates.

The data-file for one half of the face model consists of a list of point indices and 3D coordinates. This data is followed by connectivity information for each facet, which consists of the number of points that form the facet, followed by the corresponding vertex indices. After extensions to the original mesh model, which are described in section 3.6, the complete face model comprises a total of 494 points and 459 facets, see Figure 3.1.

3.4 Representation of Muscle

A muscular layer, which comprises the most influential muscles of the face, lies between the skull and the surface skin mesh. The muscles are represented using muscle vectors [175]. The head of a vector is connected to the bone which represents the origin of a muscle, while the tail represents the insertion into the skin. The muscle models used to simulate muscle contractions are described in section 4.5.2.

During conformation modelling it is necessary for the user to be able to modify the physical form of the head without affecting the functionality of the muscle models. This requirement can be achieved by either of two methods. The first would enable the user to interactively implant muscles into a head model after it has been modified. The second approach would permit the system to modify automatically muscle vectors in accordance with changes to the structure of the head as specified by the user. This latter approach has been adopted initially since the former would require a sophisticated editing capability. Therefore the difference between inelastic and elastic modifications is modelled such that conformation changes affect both the origins and insertions of muscles, whereas expression deformations displace muscle insertions alone.

3.5 Representation of Bone

Major bones of the head are modelled in the form of a polygonal mesh. At present a scaled-down version of the surface skin mesh, discussed in section 3.3, is used for convenience.

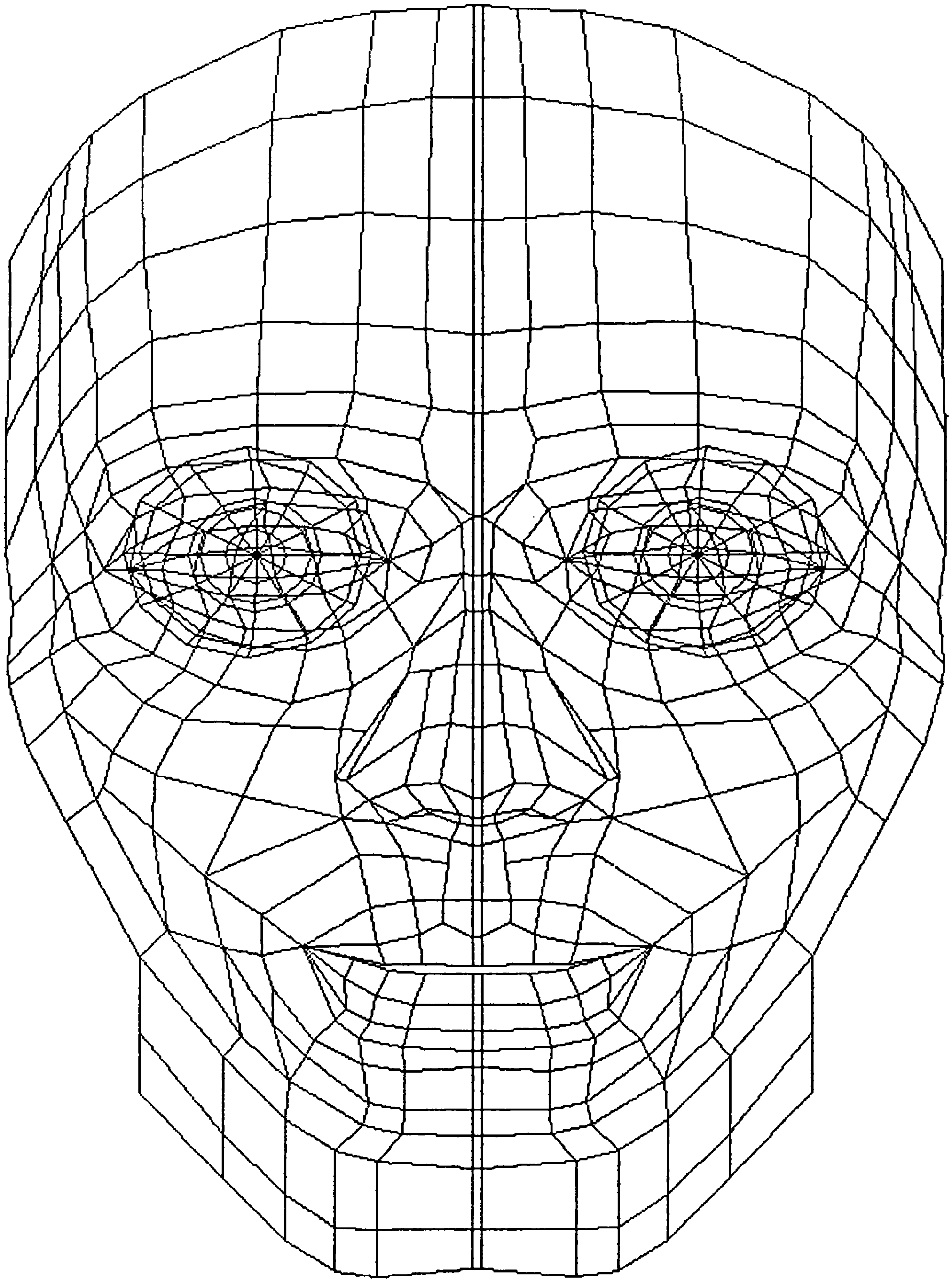

Figure 3.1: Mesh Model

The full mesh model, derived from the mask illustrated in Plate 3.1, comprises a total of 494 points and 459 facets. The left side of the face is a reflection of the right, so that the two sides are symmetrical.

The skull mesh is further distinguished from the skin mesh through a lack of eyeballs and through colour. In particular, the skull is displayed in white while a skin-tone is used for the facial skin surface.

In order to provide modelling control over specific bones, it is necessary to group particular vertices to represent each bone. Grouping of vertices for this function is specific to the skull mesh and totally independent of the facial skin mesh.

3.6 Extensions to the Mesh Model

Several essential features for facial animation include the neck, eyebrows, eyelids and eyeballs. These elements were incorporated into the initial polygonal skin mesh since the theatre mask lacked these features. Each feature was constructed for one side of the face while the corresponding feature on the other side is a mirrored reflection. Expansion of the model caused difficulties with respect to the maintainence of a 3D representation since two orthogonal views of the features were unavailable. As a result, coordinates were composed according to visual aesthetic judgement, see Figure 3.1.

3.6.1 Neck

The neck facilitates all movement of the entire head such as nodding and turning from one side to the other. Therefore the mesh model was extended to incorporate a neck consisting of two rows of vertices.

3.6.2 Eyebrows

Eyebrows convey a variety of information, according to vertical movement, as well as according to their overall shape and relationship to each other [41]. To include these features into the model several live faces were examined to determine their location on the forehead. Consequently, the eyebrows appear on the ridge of the forehead just above the eyes, as illustrated in Figure 3.1. Furthermore, they are slightly raised in the z-dimension, above the facial mesh to give the impression that they protrude above the skin. The shape of the eyebrows was determined by following the vertices around the upper part of the eye socket and the base of the forehead. Thickness of the brows was determined by visual aesthetic reasoning.

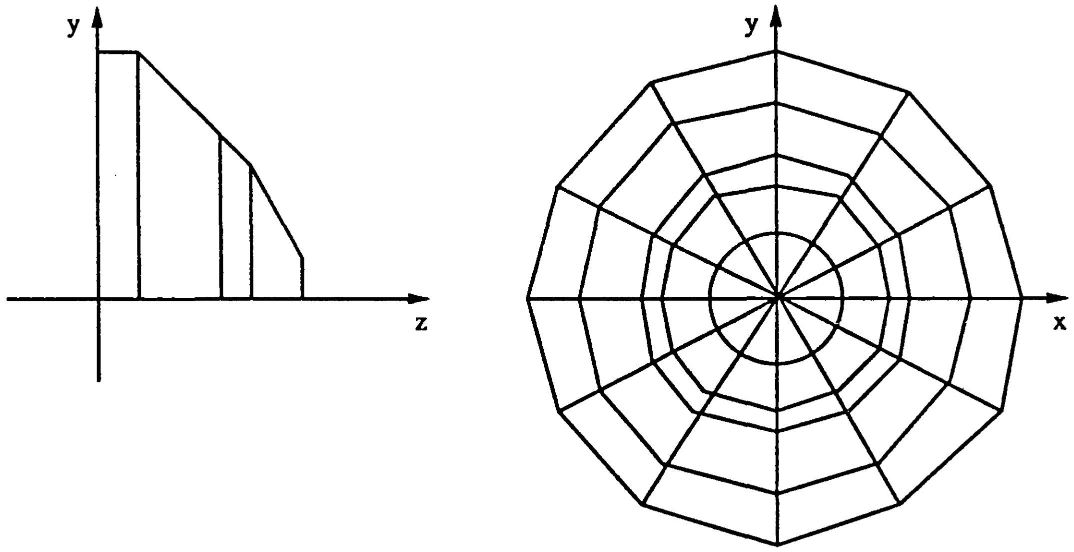

Figure 3.2: Construction of Eyeball

Eyeballs are produced procedurally through a surface of revolution technique. The profile is first defined in the y-z plane and then rotated about the z-axis to sweep out a hemi-sphere.

3.6.3 Eyelids

Realistic eyelids are particularly awkward to model since they consist of a very flexible and deformable skin that stretches over the eyeball. Tapering at each end to blend into the corners of the eye causes further difficulty. Consequently, eyelids were examined on several real faces to determine how they operate. Observation revealed that the top eyelid moves to a much greater extent than the lower eyelid. As a result, a convex mesh was constructed and scaled to conform with the eye sockets, see Figure 3.1. A row of vertices in the centre of the mesh was duplicated in order to facilitate opening of the eyelids.

3.6.4 Eyeballs

The eyeballs were formed procedurally by the use of a surface of revolution technique. A cross-sectional profile consisting of points forming the outer perimeter, mid-ring, iris-ring, iris and pupil was first defined in the y-z plane, as illustrated in Figure 3.2. This profile was then rotated about the z-axis to sweep out a hemispheric eyeball in 3D space. The number of points that make up each ring is dependent on a parameter which is passed into the routine. Controls for pupil dilation were incorporated in the form of a parameter to determine the size of the radius of the pupil. The resulting eyeball was then scaled and translated to fit into the eye sockets of the mesh model, as shown in Figure 3.1.

3.7 Eye Focusing and Tracking

The model which has been developed, enables the eyes to both focus and track an object in 3D space. The technique employed to facilitate these operations is similar to the approach used by Parke [116] and Waters [176]. The method functions in accordance to the 3D geometric model shown in Figure 3.3. Realistic results are achieved since the eyes actually

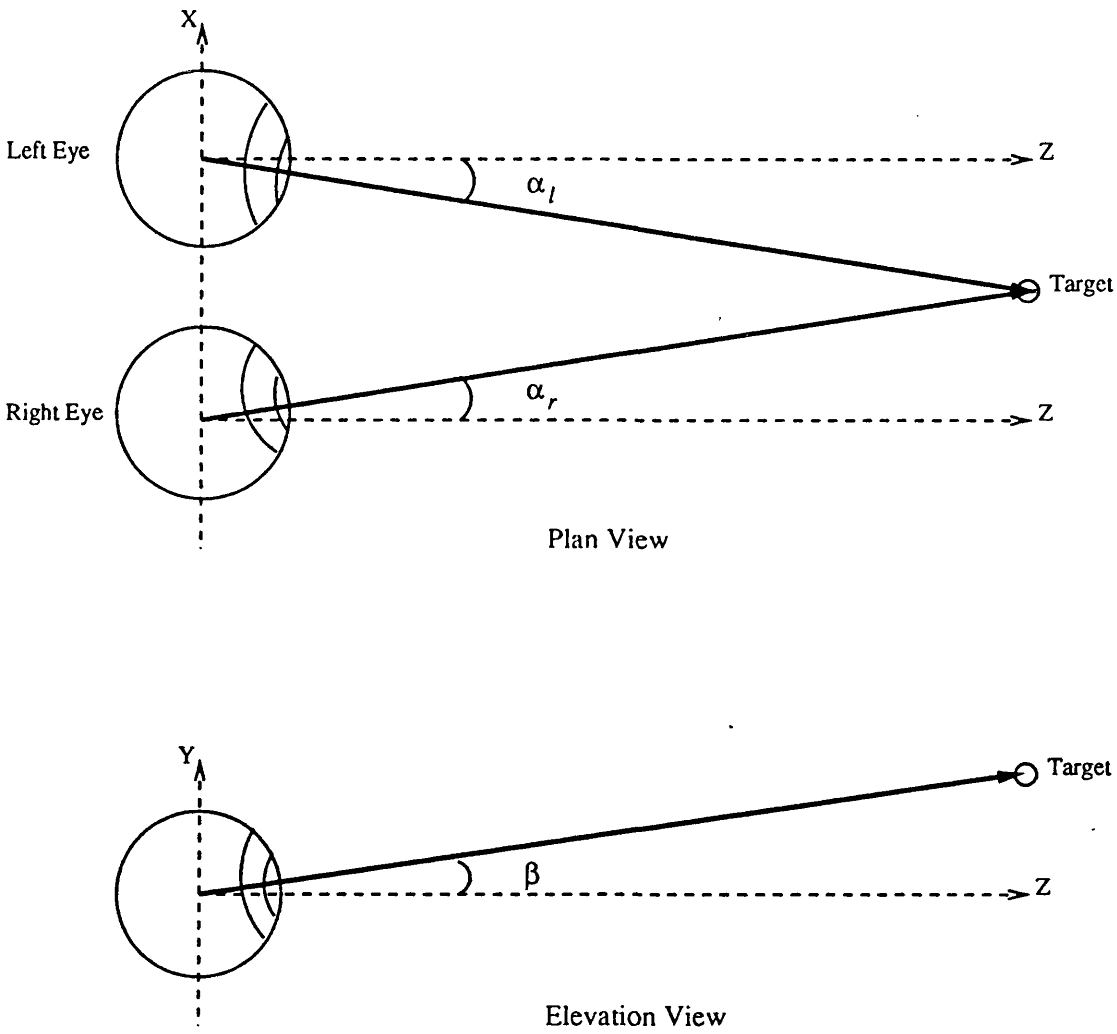

Figure 3.3: Eye Focusing and Tracking Mechanism

The geometry associated with eye focusing and tracking of objects in 3D space. The eyeballs can each rotate both horizontally and vertically, to provide two degrees of freedom.

focus on an object rather than staring aimlessly into oblivion. Such realism is particularly noticeable when a target object is close to the face, requiring each eyeball to rotate by differing amounts. To provide greater flexibility for the user, the system allows separate control over each eyeball so that special effects such as 'fish-eyes' and eyeball rolling can be achieved, see Plates 6.1 and 6.2.

Central coordinates of the left and right eyeballs form translation vectors which are used to position the eyeballs within the head. The target object has coordinates (x_t, y_t, z_t) and needs to be located in front of the face, otherwise it cannot be observed. The eyeballs can each rotate both horizontally and vertically, providing two degrees of freedom. Horizontal movement of each eyeball in the x-z plane is given by a rotation about the y-axis. The angle of rotation required in the x-z plane for the left eyeball to focus on the target is α_l, which is obtained from equation 3.1. Similarly, the angle for the right eyeball to focus on the same point is α_r, see equation 3.2. The two angles are determined as indicated in Figure 3.3.

$$tan(\alpha_l) = \frac{x_t - x_l}{z_t - z_l} \tag{3.1}$$

$$tan(\alpha_r) = \frac{x_t - x_r}{z_t - z_r} \tag{3.2}$$

The z-axis is taken as being at 0° so that a positive angle results in anti-clockwise rotation and a negative angle in clockwise rotation. Vertical movement in the y-z plane is given by rotation about the x-axis. The corresponding angles are β_l and β_r which are determined according to formulae 3.3 and 3.4.

$$tan(\beta_l) = \frac{y_t - y_l}{\sqrt{(x_t - x_l)^2 + (z_t - z_l)^2}} \tag{3.3}$$

$$tan(\beta_r) = \frac{y_t - y_r}{\sqrt{(x_t - x_r)^2 + (z_t - z_r)^2}} \tag{3.4}$$

3.8 Summary and Conclusions

It is evident that realistic and accurate conformation modelling and animation necessitates the development of a three-layer head model. Conformation modifications can be thought of as structural or inelastic changes, while expression generation results largely from muscular contraction which can be modelled as elastic deformation.

To incorporate both types of modelling into FACES, a three-layer anatomical model of skin, muscle and bone has been developed. The topology of the skin mesh was derived from a plastic mask in order to acquire a face model without any particular traits; a face which can be subsequently modified by the user as desired. The resulting polygonal mesh was extended to include constituent components of the head such as the neck, eyebrows, eyelids and eyeballs. To make the research project tractable, other elements such as hair, ears, teeth and tongue have been omitted. This is not because such elements are unimportant, but rather because they are not germane to the issues currently being addressed as part of the project.

The three-layer model caters for construction and animation in a natural manner since it is based on real anatomical structures. The importance of the head model constructed is that it not only serves as a foundation which integrates the functionality of modelling and animation in one system, but that the model also enables varying strategies to be used for investigation of the form and movement of the face. It is proposed that Parke's approach of using parameters for both conformation and animation [118] is too restrictive and that the two issues need to be examined separately.

Chapter 4

Dynamics of the Head

4.1	Introduction		36
4.2	Facial Movement in FACES		36
4.3	Kinematic Models		37
	4.3.1	Shape Interpolation	38
	4.3.2	Parameter Interpolation	39
	4.3.3	Rotoscopy	40
	4.3.4	Performance Driven Animation	41
4.4	Dynamic Models		41
	4.4.1	Tension Nets	42
	4.4.2	Simulation of Muscles	43
	4.4.3	Abstract Muscle Actions	43
4.5	Modelling Facial Dynamics		44
	4.5.1	Elasticity of Skin	45
	4.5.2	Muscle Contraction	45
	4.5.3	Muscle Interaction	49
	4.5.4	Parameters of Muscle Models	49
4.6	Facial Action Coding System		52
	4.6.1	Suitability	52
	4.6.2	Limitations	53
4.7	Time-Variance of Facial Movement		54
	4.7.1	Path of Motion	54
	4.7.2	Duration of Motion	56
	4.7.3	Amount of Movement	57
4.8	Summary and Conclusions		57

4.1 Introduction

Creation of a three-layer static head model was described in the previous chapter; next we turn our attention to the motion dynamics of that same model. Simulation of both naturalistic and realistic motion of complex objects, such as the face, remains a major issue even though this area has been the subject of much research in recent years [183, 184, 185].

Facial expressions are an inherent part of human characteristics. However, simulating facial movement is not a straightforward task. The surface skin is extremely elastic, flexible and deformable.

The way in which motion is generated is independent of the manner in which it is specified. Modelling of motion is examined in this chapter whilst issues relevant to motion specification and control are addressed in chapter 6. The method used to generate motion has implications for the type of control that can be made available to the user. In particular, trends towards greater complexity in animation increases the importance of developing techniques for convenient and automatic motion control.

We begin, in section 4.2, by examining the major problems that were encountered during the development of an animation system for faces. This is followed by a review of techniques for modelling facial movement which have been used in the past in sections 4.3 and 4.4. We shall then proceed to examine the approach used within FACES to model the motion characteristics of the face in sections 4.5, 4.6 and 4.7.

4.2 Facial Movement in FACES

Three major issues were addressed during the developmental stages of an animation system for faces. Firstly, **how** is facial movement to be generated? In particular, which mechanism is to be used to distort the mesh model? Secondly, **which** parts of the face need to move and in which combinations to create recognisable and meaningful expressions? Finally, **when** should particular parts move in relation to each other to generate plausible and convincing motion?

It is instructive to undertake a review of approaches that have already been devised, with the aim of identifying the advantages and disadvantages of each. Due to the irregular nature of the face, most models are based on a polygonal mesh representation, although there are the odd exceptions [123, 171]. Once a mesh model has been constructed, production of synthetic expressions and their animation involves the simulated motion of the mesh. Various methods of creating such movement have been developed over the years. The most realistic of these are based on the physical structure of the face. Modelling of motion falls into two broad categories [180, 182, 186, 187], they are known as *kinematic* and *dynamic* techniques. Characteristics of these two techniques for modelling motion are examined in sections 4.3 and 4.4.

Facilitating expression animation in a generic manner is a major requirement of FACES. More specifically, the system should allow any face constructed within it to be animated. This requirement necessitates that the solutions to the problems outlined be general and applicable to all faces.

From an investigation into the anatomy of the face we are aware that all facial movement occurs as a result of muscular contractions, see section 2.3. A natural corollory of this observation is that muscular movement should be simulated in order to distort the mesh model. The major advantage of this approach is that it has generality since all human faces comprise the same set of muscles. The techniques used to model elastic facial distortion are discussed in section 4.5.

An answer to the second issue, regarding which parts of the face are related, is provided by the *Facial Action Coding System* which is better known as FACS [44]. It is a comprehensive system for coding facial expressions. The benefits and drawbacks of the system are described in section 4.6.

Timing and duration of facial movement are critical to the message that is eventually conveyed to the observer. This particular issue poses difficult problems since time-variance data for non-verbal communication is currently unavailable. The question of timing is discussed in more detail in section 4.7.

4.3 Kinematic Models

In kinematic motion emphasis is placed on the production of believable visual effects; the illusion of realistic motion without having to model complex physical laws. Kinematic models produce motion from positions, speeds and accelerations. Many animation systems are based on this approach since less computation time is required than for dynamic models. Furthermore, kinematic techniques are more intuitive to use. Four major categories of kinematic model have been identified, they are *interpolating, procedural, recorded* and *natural* models [134].

Interpolating models represent motion by a set of scenes each at a different location in time. A method of interpolation is used to simulate the process of inbetweening normally carried out by assistant animators in conventional 2D animation. In procedural models motion is described by an algorithm together with a set of input data. This technique is used to implement movements such as oscillations and circular motion, or motion specific to a particular application. Recorded models consist of stored data which has been obtained from precomputed or real motion, such as *rotoscopy* [29] for example. With natural models, motion is described using commands from a near natural language notation. A set of rules is applied to the parameters of the commands to produce the desired motion.

In conventional 2D animation the techniques of 'squash and stretch', arcs in the motion and 'slow-in and slow-out' are used to imitate the effects of forces on mass. The equivalent of such effects in the context of computer animation are shape deformation and non-linear interpolation such as use of the cosine function to model inertia. Kinematic methods that have been used for the generation of facial movement include *shape interpolation, parameter interpolation, rotoscopy* and *performance driven animation.*

4.3.1 Shape Interpolation

Parke's original research, involving facial animation, was based on shape interpolation [115]. A polygonal mesh was produced using stereo photogrammetry in a similar manner to that described in section 3.3.3, but through the use of photographs of a real person taken from two different views [117]. Extreme poses, which formed keyframe data were also acquired using the same technique; the two photographs had to be taken simultaneously in order to capture an expression on a live face at a particular instance. Data-files were created for each of a number of different expressions to represent key expressions. Animation of the face was achieved through the use of a cosine interpolant to generate intermediate facial expressions.

For shape interpolation to be successful, the topology must be the same for both extremes of the mesh representation. In particular, the number of vertices defining the surfaces and their interconnections must be identical. This is not always the case and additional points may need to be calculated for one of the extremes until the conditions are satisfied [94].

Parke observed that shape interpolation produced good results when working in 2D as for cartoons, however the approach proved inefficient for 3D facial animation. In particular, the method required the complete specification of a large amount of data; it was necessary to collect data for each expression of each face and each keyframe had to be explicitly defined and stored. Another major drawback of shape interpolation is that the range of expressions which can be displayed is limited by the number of keyframe poses available.

Variations of the shape keyframe approach are nevertheless popular. It is possible to generate realistic and effective animation provided that a sufficient number of keyframes are used. MacNicol cites an example in which 18 frames out of every 30 were keyframes for each second of animation [93]. Another example is the computer animated short film *'Tony de Peltrie'*, in which a Character Bank of Standard Expressions was developed through digitisation of real human expressions. Facial expressions for the computer character were achieved by mapping expressions or blends of several expressions onto the mesh model using the TRANNA and DADS systems [48]. An alternative approach was taken by Kleiser-Walczak who successfully animated two characters, called *Nestor Sextone* and *Dozo*, by using the upper face largely for expressions and the lower parts for speech [119].

4.3.2 Parameter Interpolation

Prompted by the data-intensiveness and the associated restrictions of the shape interpolation approach, by 1974 Parke had managed to develop a parameterised method [116]. This enabled the animator to both create a face and configure expressions on it by specifying an appropriate set of parameter values. This model was also based on a collection of polygons, but this time the polygons were manipulated by a set of parameters which were interpolated. The objective was to create a model from which a wide range of both faces and expressions could be generated through a limited set of input parameters.

The rationale behind the use of parameters is that a set of differentiation or specification criteria is required to distinguish between the members of a particular set. A 'complete' set of criteria is one which allows any member of an object class to be specified by selection of appropriate parameter values. The issue becomes one of the *quality* and *scope* of such parameters. More specifically, it is necessary to investigate the appropriateness and efficiency of the chosen parameters, together with the range of faces and expressions from the 'universe' of faces and expressions that can be generated using those parameters.

The parameters selected by Parke were largely based on observation. They were subdivided into two main classes known as *conformation* and *expression* parameters. Those parameters concerning conformation dealt with attributes that vary from one individual to another, while expression parameters described variations between expressions.

Parameterised models are convenient for the user since only a limited amount of information needs to be manipulated. Such models are more flexible than those that use shape interpolation since vertices are grouped together into sub-structures which may be manipulated as a whole and independently of other groups of vertices. In contrast, shape interpolation requires a whole keyframe for even minor changes to the model.

The problem with Parke's method of parameter interpolation is that specific vertices need to be identified for manipulation of particular parts such as the eyebrows and jaw for example. Since such actions are dependent on a particular mesh topology the vertices have to be built into the software, thereby reducing the generality of the system.

Pearce et al. [123] describe extensions to Parke's model by development of a key-word interface to assist the user in creating expressions from which a library of partial expressions can be constructed. Each partial expression is specified through a set of key-words to describe the: part of the face to be manipulated; type of movement; initial and final frame numbers; a value for the parameter at the final frame; and the type of interpolation required, the default being linear.

4.3.3 Rotoscopy

Rotoscopy and variations on the approach have produced the most realistic results to date [93, 188]. The technique is a way of extracting complex motion by taking tracings from live-action footage on a frame-by-frame basis. Traditional animators, such as those at the Disney Studios, used rotoscopy to achieve realistic animal motion in 2D [162]. Recording live-action motion with three cameras to capture the x, y and z views has also enabled the method to be used in 3D animation [29].

An alternative method of obtaining data for realistic motion is by the use of electro-gonyometers [24]. This method allows both kinematic and dynamic data to be obtained without the tedious process of tracing individual frames. Yet another approach derives control points from a video sequence of a live performance and spatially maps this information to conform to a synthetic face [122]. Mapping takes into account differences in proportions between the two faces. The model is animated by deforming the texture and geometry of the synthetic face in the region of the control points.

Rotoscopy has also been used for lip synchronisation. The lip shapes are traced off and either exaggerated or used directly [116, 175]. The biggest problem with rotoscopy is that it can only be used where actions have first been performed live. This severely reduces the scope for creating special effects. In addition, it is impossible to get motion descriptions for imaginary beings or conditions for which recordings cannot be obtained. Furthermore, the camera records everything with an impartial lack of emphasis, which for caricature animation is often inadequate.

4.3.4 Performance Driven Animation

Interactive real time performance models for face animation are a recent approach developed by deGraf-Wahrman Productions [151]. Akin to the parameter interpolation approach, this method permits various facial actions to be predefined and later controlled interactively [140, 188]. Facial motions can be either autonomous or controlled by a puppeteer. This technique is analagous to traditional puppet models which can be 'performed' rather than animated. The method, which is also known as *animation by enactment*, provides animators with fluidity and flexibility of expression as well as spontaneous improvisation which have been neglected in other approaches.

4.4 Dynamic Models

Dynamics facilitates naturalistic movement patterns to imitate realism, which is the major benefit of this approach. The technique involves prediction of motion through analysis of the effects of forces and torques on mass. To apply dynamics models, objects to be animated need to be defined in terms of mechanical elements which include: material, mass, joints, rods and springs.

However, there are three major problems with using dynamics techniques [181]. Firstly, the equations tend to be complex and compute-intensive; they are solved numerically to find the position for each moving point. This makes interaction using present day systems difficult. Secondly, due to the interaction between moving points, dynamics equations are usually coupled and must be solved as a system of equations. This necessitates numerical techniques which are iterative and computationally expensive; in addition, the equations are often numerically unstable, giving widely ranging results. Finally, motion control is extremely difficult because it involves specification of the duration, magnitude and direction of forces; these are entities which tend not to be intuitive to control.

Major dynamics based methods that have previously been used in facial animation include: *tension nets, simulation of muscles* and *abstract muscle actions*. These techniques are described below.

4.4.1 Tension Nets

Platt and Badler's model is based on the anatomy of the head and comprises a three level representation to model the skin, muscles and bones [131]. Points at the three levels are connected by elastic arcs forming 3D networks or *tension nets*. Arcs connect a point to all adjacent nodes. Muscles are represented by arcs which stretch from each fibre-point to a point on the underlying bone and from the muscle to one or more skin points. Since different sections vary in terms of elasticity, information such as length parameters are stored on the arcs.

The basic action performable on the network is the application of a force, or tension, to selected parts of the network. When a force df is applied to a point p, the change in location is computed using equation 4.1.

$$dl = \frac{df}{k} \tag{4.1}$$

Here k is the sum of the spring constants, or elasticities at the point p. When a simple fibre contracts, a force is applied to a muscle point in the direction of the tail, or bone, point. This causes lesser forces to be propagated out from the initiating point across the face. Animation is achieved by dividing the contractions into n, each of force $\frac{1}{n}$. This assists in the distribution of the effects of simultaneous pulls over intermediate areas of the face.

The model has several disadvantages. Since an iterative technique is used, initial values are required for the length and elasticity of the arcs connected to both the skin and bone. Dynamics with respect to the lower part of the face, notably the mouth and jaw are not addressed. It is not clear to what extent complexities such as these can be dealt with by the model. Platt found that since the model attempts to emulate the low level intricacies of the face, the approach was unacceptably expensive in computational terms [130]. Although the model does provide a close relation between the causes of facial actions and their simulation, a more feasible approach is to simulate the most significant characteristics of facial muscles.

4.4.2 Simulation of Muscles

Also based on the anatomical structure of the face, Water's model concentrates on the actual motivators of the dynamics in the face, that is the muscles [175]. The aim was to develop a general and flexible parameterised muscle model which would allow facial control without having to predefine the performable actions based on the topological representation of the model. Muscles, which exist in all human faces, are modelled so that the method is independent of any particular facial topology.

Facial skin is represented as a mesh in which each node has a finite degree of mobility. *Linear* and *parallel* muscles that pull, together with *sphincter* muscles, which squeeze, are modelled. Positioning of muscles is achieved through identification of key nodes on the face and enabling correspondence of these points to the 3D computer model. Animation occurs as a result of parameter interpolation. The model has recently been extended to a physically-based one, in which all three layers of skin, the *epidermis*, *dermis* and *subcutaneous facia* are modelled [159].

The approach has limitations in that it requires specific parameters such as muscle tension and elasticity to drive it. Such measurements are difficult to both acquire and use as a basis for expression synthesis. Research is currently being undertaken to derive automatically muscle parameters from video sequences of live performers [160].

4.4.3 Abstract Muscle Actions

Abstract Muscle Action or *AMA procedures* are part of the *Human Factory System* which has been used for modelling human characters [96]. An AMA procedure is a specialised routine which simulates the specific action of a face muscle rather than modelling general types of muscle as in Waters' case. This means that the user is provided with a more detailed and accurate level of control. The procedures operate on specific areas of the face, each of which must be defined when the face is constructed. Each procedure is responsible for actions which correspond to a particular muscle.

The AMA procedure level is considered to be the lowest level of access to the face. Two further higher levels, known as the *expression* and *script* levels, are also defined. Facial movements comprise phonemes and emotions both of these consist of combinations of parameters to the AMA procedures. The expression of a synthetic actor is characterised by: a unique face for the actor; a set of regions; and a set of parameter values. An actor's personality can therefore be defined as the set of expressions for that actor.

A script is a collection of multiple tracks [52], where each track is a chronological sequence of keyframes. There is one track for each facial parameter or AMA procedure. A track may be modified and combined with the rest of the facial expression, making changes straightforward. Animation itself is performed by applying spline interpolation [80] to the mixed track. Consequently, the user works with expression interpolation rather than parameter interpolation, which provides a higher level of access to facial animation. The *Human Factory System* has been used to generate the computer animated sequence entitled *'Rendez-vous á Montréal'* in which two synthetic actors emulate *Marilyn Monroe* and *Humphrey Bogart*.

4.5 Modelling Facial Dynamics

At the lowest level generation of facial expressions in FACES is governed by muscle models. The models correspond to three types of muscle found in the face: *parallel, oblique* and *spiralized* [62] and are based on models which were first formulated by Waters [175].

There are several reasons why the muscle models are considered to be the most appropriate for use in FACES. To date, the models have produced the most convincing results observed; this is because they are based on reality. The technique is also topology independent which is necessary since FACES caters for the creation and modification of a multitude of faces. Topological independence is essential when modelling a variety of faces since each is likely to require a distinct topology for the purpose of recognition. More specifically, it is necessary to use an animation model that is applicable to any face; the muscle models operate in terms of anatomical muscles which form the basis of all human faces. The muscle models are also ideal as a foundation upon which FACS can be implemented; FACS describes facial expressions in terms of combinations of muscular contractions. The approach can also accommodate a three-layer anatomical model which has been deemed necessary for a viable and complete system for modelling and animating faces, see chapter 3.

The head model in FACES differs from the model developed by Waters. A complete anatomical model was not considered by Waters since the focus of his research was facial animation rather than the creation and modification of faces. Waters therefore concentrates on the skin and muscular layers alone, which resulted in the omission of the explicit definition of the underlying skull [176].

4.5.1 Elasticity of Skin

Skin needs to be modelled accurately to produce realistic results because it is a particularly visible feature of the face. The difficulty with modelling skin is that it has a non-linear, anistropic stress-strain relationship. Furthermore, the stress-strain relationship is time-dependent because skin is also viscoelastic. Extensive literature on the mechanical properties of skin exists, but even the most sophisticated skin models are greatly simplified because of complexity and computational constraints [86]. Waite and Komatsu argue that a B-spline representation adequately models the required elasticity [81, 171]. However, more sophisticated models, based on Finite Element Methods [61, 63, 86], have also been developed.

In FACES skin is modelled as an elastic mesh which deforms with the application of muscle contractions. This technique avoids the complexity and the computational requirements of a non-linear viscoelastic model which would require compute-intensive and iterative numerical techniques and would therefore be unsuitable for an interactive system.

4.5.2 Muscle Contraction

Mathematical modelling of the muscle models used in FACES is described in this section. Details regarding the development of the original models can be found in Waters' doctoral thesis [176].

Facial muscles can be classified according to the orientation of their individual fibres. Consequently, three types of muscle can be identified as the primary motion muscles, these are linear, sphincter and sheet. Linear or parallel muscles pull in a single straight direction. Elliptical or circular muscles, also known as sphincter muscles, squeeze the skin to pull it in towards a central point. Sheet muscles behave in a similar manner to a series of parallel linear muscles spread over an area.

Muscles are modelled in terms of vectors, using a direction and a magnitude. Direction is towards the point of attachment on the bone while magnitude is the amount of contraction. The disturbed area of skin known as the *zone of influence*, increases with muscle tension. FACS provides an approximation of this area of disturbance.

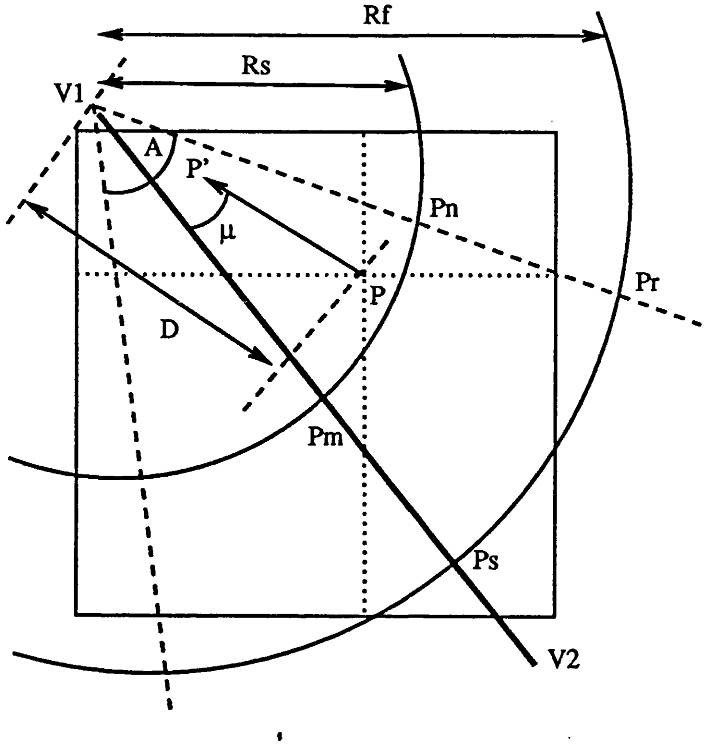

Figure 4.1: Geometry of Linear Muscle Model

The linear muscle contracts surrounding skin towards a static *origin* on the bone until the force dissipates to zero [176]. Only a proportion of the force is effective along the line of contraction since fibres become oblique in relation to the node of attachment.

Linear Muscles

A linear muscle contracts the surrounding skin towards a static node on the bone until, at a finite distance away, the force dissipates to zero, see Plate 4.1. The displacement can be approximated by multiplying the length of the muscle fibre by the cosine of the angle of the muscle fibre to the tendon [62].

Major parameters of the linear muscle model include the: area of flesh influenced by the muscle contraction; length of the muscle; and position of the muscle in 3D space relative to the underlying bone. The model assumes that there is no displacement at the point of attachment to the bone, but a maximum deflection at the point of insertion into the skin. The geometry of the model is illustrated in Figure 4.1. To calculate the displacement of a node P on the mesh to P' within the segment $V_1 P_r P_s$, a displacement towards V_1 along the vector $\vec{PV_1}$ is created. This gives

$$P' \propto \mathcal{F}(K, A, R, P)$$

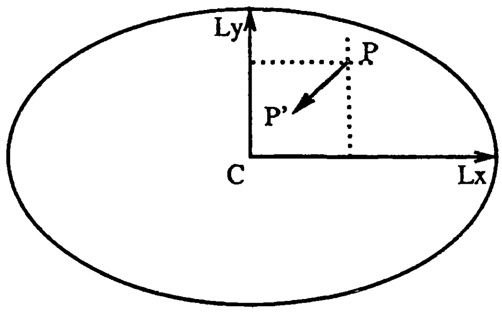

Figure 4.2: Geometry of Sphincter Muscle Model

The sphincter muscle causes the tissue of the skin to be drawn together around a single point [176]. The points are squeezed together according to the major and minor axes.

Here P' is a function of K, the muscle spring constant; an angular displacement

$$A = cos(\mu) = \frac{\vec{V_1V_2} \cdot \vec{V_1P}}{\mid \vec{V_1V_2} \mid\mid \vec{V_1P} \mid}$$

and a radial displacement factor

$$R = \begin{cases} cos\left(\frac{1-D}{R_s}\right) & \text{if } P \text{ in } V_1P_nP_mV_1 \\ cos\left(\frac{D-R_s}{R_f-R_s}\right) & \text{if } P \text{ in } P_nP_rP_sP_m \end{cases}$$

Sphincter Muscles

The sphincter muscle causes the tissue of the skin to be drawn together around a single point, see Plate 4.1. Since the points are squeezed together in a uniform manner, the angular displacement becomes redundant, and a major (L_x) and minor (L_y) axis are used, as shown in Figure 4.2. Now,

$$P' \propto \mathcal{F}(K, L_x, L_y, P)$$

Here K is the muscle spring constant. The displacement of P to P' is given by

$$F = 1 - \frac{\sqrt{{L_y}^2{P_x}^2 + {L_x}^2{P_y}^2}}{L_xL_y}$$

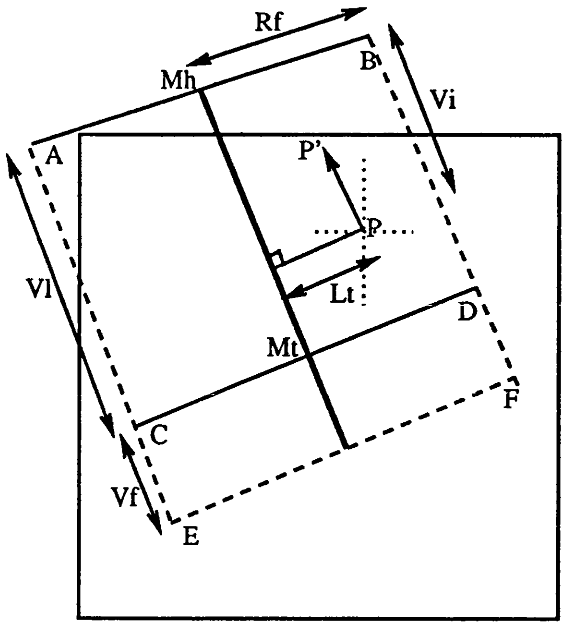

Figure 4.3: Geometry of Sheet Muscle Model

Sheet muscles may be described as a series of almost parallel, linear fibres spread over an area [176]. The muscle model requires definition of a displacement parallel to the direction of the central muscle vector.

Sheet Muscles

Sheet muscles may be described as a series of almost parallel fibres spread over a region, see Plate 4.1. A particularly good example is the *frontalis major* which lies on the forehead and is responsible for raising the eyebrows.

The geometry of the sheet muscle model is shown in Figure 4.3. The muscle requires definition of a displacement parallel to the direction of the central muscle vector, so that:

$$P' \propto \mathcal{F}(K, d, P)$$

Where K is the muscle spring constant and d is the dissipation of the force depending on which zone the node lies in

$$d = \begin{cases} cos\left[1 - \frac{L_t}{R_f}\right] & \text{for } P \text{ in } ABDC \\ cos\left[1 - \frac{L_t}{R_f}\left(\frac{V_i}{V_l} + V_f\right)\right] & \text{for } P \text{ in } CDFE \end{cases}$$

4.5.3 Muscle Interaction

Muscle contractions are by no means mutually exclusive, in fact it is seldom that a single muscle will act in isolation. Physically, some muscles tend to merge together and over-lay others so that their zones of influence overlap. In a dynamic environment, muscles will interact often pulling in conflicting directions and with differing amounts of force. Such actions need to be resolved in a consistent manner. The complication of simultaneous muscle pulls is automatically resolved by averaging out the tension during an action.

To simulate this process in FACES muscle vectors displace points from their current position rather than from their rest position. The effect of opposing muscles is therefore to move the node points to a location which is the average of several muscle displacements on a frame-by-frame basis. Plate 4.1 shows interactions that occur between two linear, two sphincter and two sheet muscle contractions.

4.5.4 Parameters of Muscle Models

Having decided to treat the conformation and animation aspects of the face separately, it was realised that a similar approach could be taken with regard to the specification of muscles. The parameters are therefore partitioned into those which influence structural characteristics and those which affect expression animation. All muscle data, with the exception of tension, is read in from a file for each head model. This data includes elasticity information which varies both between individuals and according to age. Such characteristics are concerned with conformation aspects rather than animation issues. Separation of conformation and animation parameters means that the user need only be concerned about control over animation parameters, such as tension, whilst generating moving sequences. Table 4.1 shows the parameters of the three muscle models and their ranges.

In FACES it is necessary to use nodes that lie on the skull and facial skin meshes, as heads and tails of vectors, since modifications in the CONSTRUCT and MODIFY sub-systems lead to changes in the coordinates of both meshes. This approach is in contrast to that of Waters who uses absolute coordinates to represent muscle origins and insertions.

Parameter	Abbreviation	Range
Linear Muscle Model		
Head of muscle vector	h	—
Tail of muscle vector	t	—
Start of fall-off radius	R_s	$0.0 \leq R_s \leq 1.0$
End of fall-off radius	R_f	$0.0 \leq R_f \leq 1.0$
Zone of influence	θ	$0^\circ \leq \theta \leq 180^\circ$
Tension	T	$0.0 \leq T \leq 1.0$
Muscle Spring Constant	K	$0.0 \leq K \leq 1.0$
Muscle Elasticity	E	$0.0 \leq E \leq 1.0$
Sphincter Muscle Model		
Centre	C	—
Length of major axis	L_x	$0.0 \leq L_x \leq 1.0$
Length of minor axis	L_y	$0.0 \leq L_y \leq 1.0$
Tension	T	$0.0 \leq T \leq 1.0$
Muscle Spring Constant	K	$0.0 \leq K \leq 1.0$
Muscle Elasticity	E	$0.0 \leq E \leq 1.0$
Sheet Muscle Model		
Head of central muscle vector	h	—
Tail of central muscle vector	t	—
Right-angle range	R_f	$0.0 \leq R_f \leq 1.0$
Tension	T	$0.0 \leq T \leq 1.0$
Extension	X_t	$0.0 \leq X_t \leq 1.0$
Muscle Spring Constant	K	$0.0 \leq K \leq 1.0$
Muscle Elasticity	E	$0.0 \leq E \leq 1.0$

Table 4.1: Parameters of the three Muscle models

The muscle models provide an approximation to the biomechanics of muscle activity. Parameters are required only for the three major types of muscle found in the face, namely: linear, sphincter and sheet.

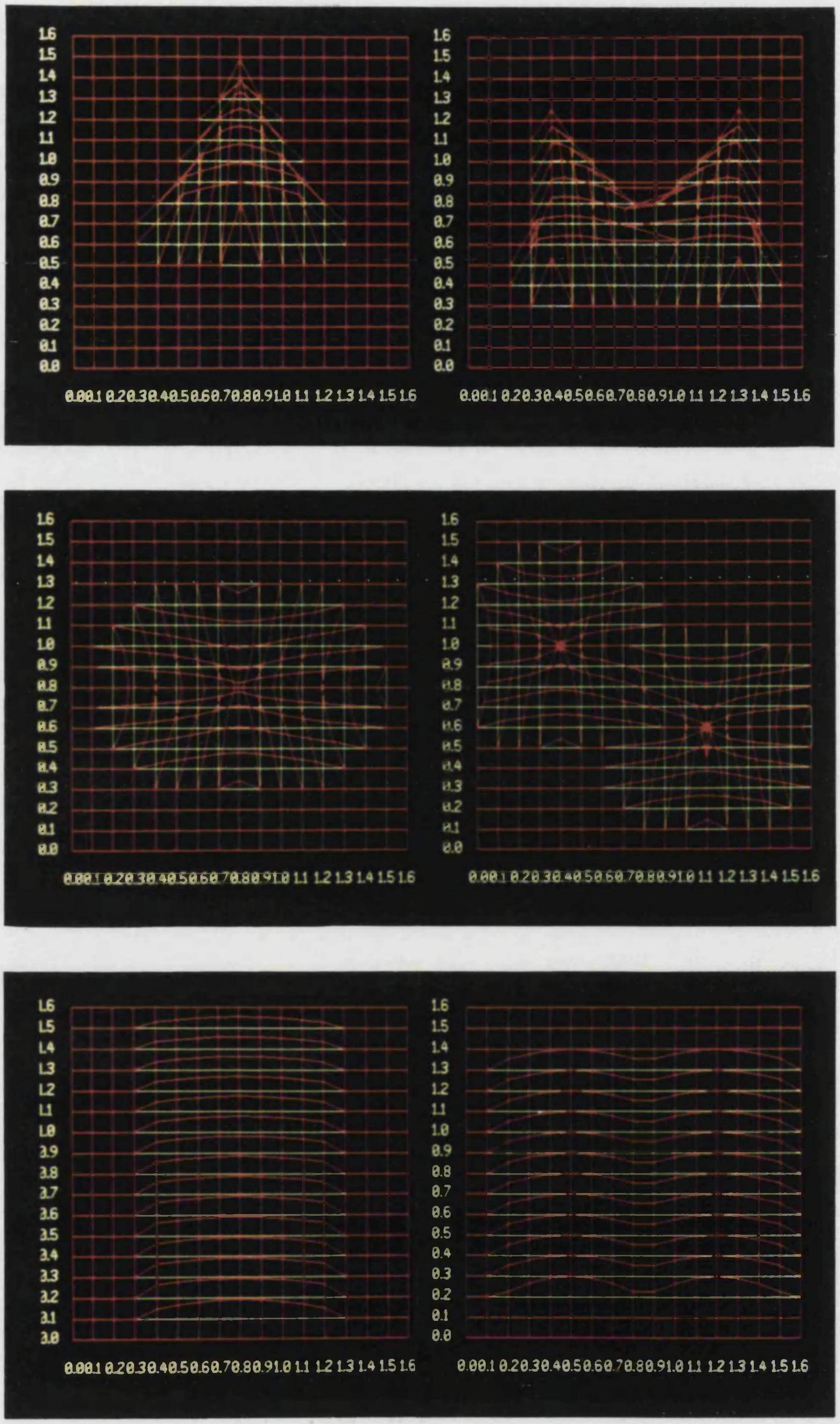

Plate 4.1: Linear, Sphincter and Sheet Muscles

From top to bottom, along the left of the plate, examples of skin deformation caused by single linear, sphincter and sheet muscle model contractions. From top to bottom, along the right of the plate, interactions that occur between two linear, sphincter and sheet muscle contractions.

4.6 Facial Action Coding System

Use of muscle movement as a mechanism for simulating facial dynamics is realistic, but inadequate and inconvenient for creating meaningful expressions; it is also a very low-level approach to facial animation. A detailed analysis of real faces and their behaviour is necessary to determine: which particular muscles need to contract; by what amounts; and in which combinations to cause recognisable expressions to be generated.

A study undertaken by psychologists Ekman and Friesen, has resulted in the *Facial Action Coding System* or FACS [44, 45], which has already been mentioned. The system differentiates between all possible visually distinguishable facial movement and is free of any cultural interpretations placed on facial behaviour. FACS describes the set of all possible basic actions performable on the human face; actions such as raising the inner brow, or lowering the corners of the lips have been categorised.

Each basic movement is called an *Action Unit*, or AU. Such actions are based on the anatomy of the face, each being caused by either a single muscle or a small set of closely related muscles. Each AU consists of a minimal action in the sense that it cannot be broken down into smaller actions. AUs interact in order to build up a complete expression. FACS identifies 58 AUs which, separately or in various combinations, are capable of characterizing any human expression [44, 58]. The AUs classified within FACS are presented in Appendix B.

4.6.1 Suitability

FACS is the most suitable system for use in FACES for several reasons. FACS operates in terms of generic facial actions which are applicable to all human faces. Photographs which are presented in the FACS manual exemplify, but do not typify facial expressions. This form of generality is ideal for use in FACES since a major requirement of the system is that it enable the manipulation of a host of different faces. Actual expressions generated will therefore depend to a large extent on the conformation of the particular face.

AUs are realistic since the system has been derived from a detailed analysis of real faces. A model which emulates the effect of real muscles in the face ensures that the resultant facial expressions are within the range of natural human expressions.

Modelling realistic facial movement is difficult because of simultaneous activity and interdependence of numerous muscles. Combinations of specific muscles can be grouped together to form AUs which provide a higher level of control than individual muscles. Consequently, the user is insulated from the requirement to learn the muscular basis of the face.

FACS is a modular system which decomposes facial movements into basic components which serve as construction blocks to enable the user to build up composite expressions. This approach also resolves the problem of formulating complex parameter sets to represent particular facial expressions. An expression can be viewed as a group of parameter values and AUs which together transform a neutral face into an expressive one. Having a series of AUs that need to be activated for an expression is particularly suitable for computer manipulation. FACS is not graphically oriented as are *Labanotation* [72], *Sutton Notation* [158] and Birdwhistell's communication notation system [10].

In terms of the practicalities of developing a facial animation system, FACS has a further point in its favour. Use of a modular interface to facial movement allows AU routines to be either simple or sophisticated without affecting the top-level interface. Such an arrangement caters for a great deal of experimentation in the low-level routines to accommodate the requirements of particular applications.

4.6.2 Limitations

Despite the many benefits, FACS does have a few limitations. It should be borne in mind that the system was conceived as a notation for recognising and grading facial expressions rather than for their generation. A trained observer is supposed to view a static face and complete a score sheet to rate the expression in terms of AUs and their levels of activation. In order to make use of the system in FACES it was necessary to convert AUs into causes rather than measurements of movement. Fundamental problems are encountered when using FACS in a generative context which requires precise definitions since regions used to describe AUs and AUs themselves are both informally defined.

A further problem is that FACS deals solely with facial actions that reflect muscular movements. The system does not cater for aspects such as blushing, paleness or tear-filled eyes. These are important components of expressions which communicate a great deal about the emotional state of a person.

4.7 Time-Variance of Facial Movement

Timing of facial actions is critical to non-verbal communication. Bruce and Valentine [19] have found that the motion attributes of expressions provide far more information than that from a static pose or photograph. Static images do not clearly reveal unique and subtle messages, it is therefore necessary to take their variance over time into account.

However, the creation of believable facial motion is difficult since the task involves coordination of many simultaneous activities. In addition, there appears to be a dearth of important time-variance information in the literature. The problem has therefore been decomposed into three aspects in order to provide adequate control for the user. More specifically, the issues concern the **path** of motion, **duration** of motion and **amount** of movement. Here we examine the principles behind the approach adopted; the mechanism for specification and control of facial movement is elaborated in chapter 6.

4.7.1 Path of Motion

Our highly developed skills for interpretation of facial movement raises the problem that viewers will easily detect unnatural motion. AUs and their associated muscles have both differing rates of contraction and duration times. Yet, accurate time-variance data for muscular contractions is presently unavailable. In addition, FACS was developed as a system for coding static facial expressions and therefore does not describe the dynamics of muscle movement; it merely identifies the static position of muscles in a facial expression. In effect, the system provides 'keyframe' data but gives no indication of the motion dynamics that should be used to generate naturalistic movement.

In FACES, this gap is filled through generation of motion using a parameterised keyframe approach. Specification of differing paths of motion is facilitated by the provision of four laws of motion, see Figure 4.4. At each frame, interpolated parameters are input to particular AU routines which pass the parameters on to the relevant muscles. The muscles in turn create contortions on the face mesh through the muscle models described in section 4.5.2.

Consider the value of a parameter p at frame t which lies between two keyframes s and e, then

$$p_t = p_s + \frac{t}{n} * l * (p_e - p_s)$$

where n is the total number of frames to be generated and l is the law of motion. The value

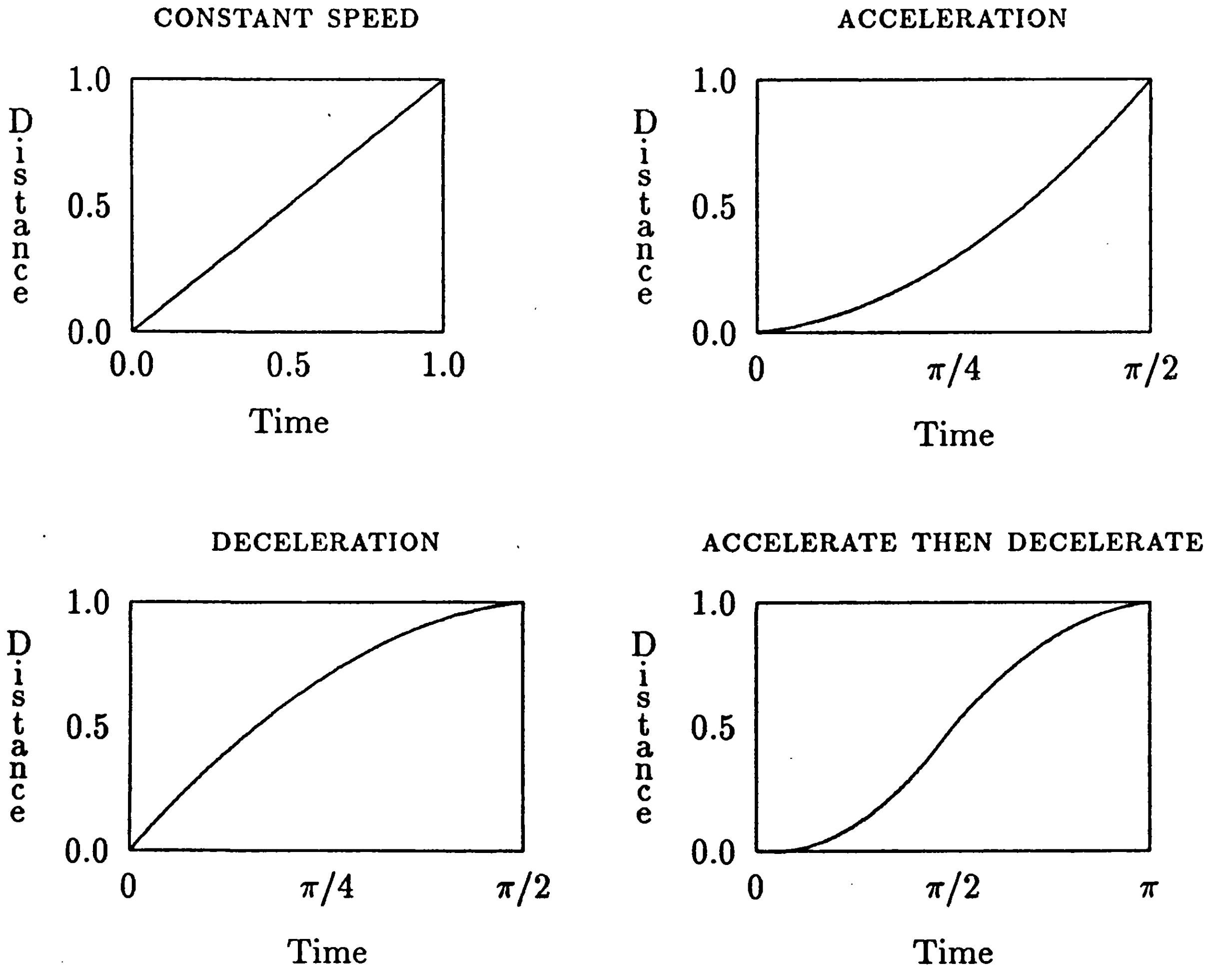

Figure 4.4: The Four Laws of Motion

These four laws of motion represent techniques for performing interpolation. They simulate linear or constant speed, acceleration, deceleration and acceleration followed by deceleration.

of p_t depends on the law l. There are four common methods for interpolating between keyframes [94], they are shown in Figure 4.4. Constant speed or *linear* interpolation is the simplest case; l has a constant value of one for all frames, so that smooth motion with uniform velocity is achieved. This however is not very realistic. Conventional animators spend a considerable amount of time ensuring that movements are correctly "faired", that is accelerated to and from rest. Fairing involves varying the **rate** at which the parameter value changes depending on how far the frame is between the two key frames.

With *acceleration* the interpolated value begins by changing slowly and speeds up as the end value is approached. The law used for modelling acceleration is given by equation 4.2

$$1 - cos\left(\frac{\pi}{2}\frac{t}{n}\right) \tag{4.2}$$

Deceleration can be modelled using the law expressed in equation 4.3, which slows down the interpolation as it reaches the end value.

$$sin\left(\frac{\pi}{2}\frac{t}{n}\right) \tag{4.3}$$

To model *acceleration followed by deceleration* equation 4.4 can be used.

$$\frac{1 - cos(\pi\frac{t}{n})}{2} \tag{4.4}$$

Attempts are currently being made to derive automatically numerical data regarding intensity and time-variance of facial actions. This is being pursued through the analysis of video sequences of a live person and use of the information to drive a computer model [159]. This principle was first advocated by Platt and Badler [131] who envisaged a camera-processor that could identify the activation of AUs.

4.7.2 Duration of Motion

The onset and duration of facial actions are important factors in the creation of the subtle nuances that inject so much information into facial movement. A major problem is that there appears to be no usable time-variance information available at present with respect to non-verbal communication. In FACES, control over duration is therefore provided by allowing the user to specify the start and end frames for each action as explained in section 6.6.3. Such regulation over individual facial actions provides the user with considerable flexibility and control.

4.7.3 Amount of Movement

Another problem is that of determining what amount of tension to apply to the AUs in order to acquire various degrees of facial movement. FACS again does not provide any numerical data; instead, it relies on a human coder to make measurement judgements on a qualitative scale of comparison relative to the neutral face. The system defines six levels of activation for AUs, they are *slight, marked, pronounced, severe, extreme* and *maximum*. To provide flexibility, in FACES a continuous scale is used rather than the discrete levels listed in the FACS manual. The user therefore specifies a value ranging between zero and one, see section 6.6.3.

4.8 Summary and Conclusions

Previous research has shown that it is not easy to create realistic, believable computer generated facial animation. Three particular problems have been identified, they are concerned with: how facial movement should be generated; which parts of the face need to move; and when specific parts should move in relation to each other.

Methods that have been used by previous researchers to simulate the movement of the face can be classified into either kinematics or dynamics approaches. Kinematics techniques are concerned with the generation of convincing visual effects without the burden of complex physical mathematical laws. A major benefit of dynamics models is that it is possible to achieve naturalistic and realistic movement. In the quest to attain realistic computer generated facial motion, a dynamics based approach has been adopted within FACES.

The dynamics of the face are modelled on the underlying muscles which are the instigators of facial movement. The muscle models provide an approximation to the biomechanics of muscle activity. They are also an abstraction in that all faces have the same set of anatomical muscles. As a result, parameters are required only for the three major types of muscle found in the face: linear, sphincter and sheet.

The issue of providing the user with a higher and more meaningful level of control than through manipulation of individual muscles has been resolved by the use of FACS. AUs encapsulate related muscles to provide a meaningful interface over facial movement. In addition, FACS has been derived from real expressions and also provides generality.

Timing of facial movement is critical in order to convey specific messages. Due to a lack of readily available quantitative information in this area, detailed and flexible control is provided over individual facial actions to assist the animator in achieving the required time-variance over facial movement.

Chapter 5

Creating Faces

5.1 Introduction 60
5.2 Determinants of Facial Form and Appearance 61
5.2.1 Psychology of Recognition 61
5.2.2 Criminology and Identification 62
5.2.3 Facial Reconstruction and Anthropology 65
5.2.4 Medicine and Anatomy 66
5.2.5 Portraiture, Sculpture and Proportion 67
5.3 Computer Representation of the Face 68
5.3.1 Digitisation 68
5.3.2 Shape Interpolation 69
5.3.3 Local Deformation 70
5.3.4 Kit of Parts 70
5.4 Conformation Modelling in FACES 71
5.5 The CONSTRUCT Sub-System 73
5.5.1 Global Control 74
5.5.2 Regional Control 75
5.5.3 Local Control 78
5.6 The MODIFY Sub-System 81
5.6.1 Global Control 81
5.6.2 Regional Control 84
5.6.3 Local Control 84
5.7 Summary and Conclusions 87

5.1 Introduction

In this chapter we concentrate on the factors that distinguish human faces from one another, making them unique to each and every person and serving the principal function of identification. It is necessary to establish such factors in order to determine the type of modelling facilities that are required to provide effective control over the form of the head and face.

Note that there are two strands to the research in this area of facial processing. The first is concerned with using an automated method, normally based on image processing techniques, to perform the identification task itself. The second involves generation of facial images, using computers, which are subsequently to be recognised by humans. In the following sections we shall delve into the second of the two aspects.

Interactive modelling of different faces is an area which has received relatively little attention in comparison with the animation of faces. Within the area of modelling of arbitrary faces there are two issues which need to be addressed. The first concerns the representation of the face. This is not straightforward because faces are irregular structures; they are not composed of regular geometric primitives which can be easily represented through parameters, for example. Secondly, faces tend to vary from person to person. Unfortunately, there appears to be a lack of definitive data to indicate what it is that makes a face unique to a person. The deficiency of adequate and usable information regarding conformation of the face makes development of modelling facilities difficult.

However, the face is a popular object of study in many disciplines other than computer animation. Information that has been established within these fields is likely to have important implications for how faces should be modelled and represented using the computer as a medium. In section 5.2, we therefore begin by reviewing valuable lessons that have been learned by psychologists, criminologists, medical practitioners and artists. It is also necessary to examine computer based modelling techniques that have been developed to represent irregular structures, this is accomplished in section 5.3. Finally, we shall study the functionality provided for conformation modelling in the CONSTRUCT and MODIFY sub-systems and the motivation behind its provision in sections 5.5 and 5.6.

Although colour is an influential factor in determining the appearance of the face, we restrict our attention to the form of the face in this chapter. Issues regarding colouration are addressed in chapter 7.

5.2 Determinants of Facial Form and Appearance

Variation in facial structure between individuals is much less understood than the ways in which a face varies from expression to expression. Nonetheless, it is possible to derive some rules and heuristics regarding the form of the face from several fields of study including: psychology, criminology, facial anatomy and art. Each of these areas is discussed in the following sections with a view to the identification of information pertinent to conformation modelling of the head and face.

5.2.1 Psychology of Recognition

Humans can distinguish very effectively between faces, even though all faces have the same basic features which appear in more or less the same relative positions. Results from research undertaken by psychologists indicate that memory for pictures is superior to that for words. Furthermore, memory for faces is better than that for other visually observed objects [194]. This is because memory for pictures involves *recognition memory* whereas memory for words involves *recall memory*. Humans are better at recognition than at recall. Such results have led researchers to investigate the processes that underly face recognition.

Representation in Memory

One hypothesis tested by Rhodes et al. [139] is whether the distinctive features of the face are exaggerated and encoded in terms of a caricature. They discovered that while caricatures are not identified more accurately than realistic line drawings, they are identified twice as quickly. Furthermore, drawings in which distinctive aspects had been exaggerated by approximately 16 per cent were judged to be the best likenesses.

Another question asked by the same researchers was whether faces are represented and recognised by holistic comparisons with representations in memory, or by feature analytic comparisons. They used a caricature generator, developed by Brennan, which is based on a holistic theory of caricature [38]. Since caricatures were found to be more identifiable than accurate line drawings, there is some evidence to suggest that faces may be encoded and remembered through a holistic comparison with an 'average' face.

In addition, Laughery determined that a subject of a given race is a better recogniser of people of that same race than of other races [89]. These results are consistent with those of Carey et al. [25] which suggest that this 'average' face varies for each individual and is established through the process of viewing a large number of faces.

Classification of Significant Features

Attempts have also been made to discover which features are particularly important in facial recognition. Zavala found that the features most frequently used in identification are the nose, eyes, face shape, hair colour and the chin [195]. An interesting question is whether the techniques used by good identifiers differ from those employed by poor identifiers. Zavala's study indicated that good identifiers concentrate on a small number of features; they also tend to notice facial marks such as moles, beauty marks and freckles. These observations help to drastically reduce the number of alternative faces to be considered. In addition, poor identifiers tend to use features which are 'complex', while good identifiers appear to use features that are not particularly complicated.

Harmon raised a number of questions whilst investigating facial recognition, amongst them [67]: why are faces so readily recognisable? how accurately can a face be formally described? given a verbal description, how easily can a particular face be identified? what kinds of image degradation most seriously affect recognition? and how can a computer be made to recognise a human face? The last question led Harmon to classify 21 facial features which provide the most information in distinguishing one face from another [67]. These characteristic features of the face are presented in Table 5.1.

It should be noted that although evidence regarding the shape, position and size of features is useful for computer synthesis, it is by no means the only information that is used in recognising faces. Bruce and Young have identified seven distinct types of information that can be used in the process [20], they are: pictorial, structural, visually derived semantic, identity-specific semantic, name, expression, and facial speech codes.

5.2.2 Criminology and Identification

The motivation behind a substantial amount of the psychological research into the processes of recognition and identification has been to assist victims of crime in the identification of perpetrators. Despite the findings of Rhodes et al. [139], as ascertained in section 5.2.1, caricatures have not been used to represent faces. For the identification to be effective, it is necessary that the face is familiar to the witness. This however, tends to be rare in cases of crime. One technique that is in use, can be thought of as caricature

FEATURE	CHARACTERISTIC	DESCRIPTION				
Hair	Coverage	*Full*		*Receding*		*Bald*
	Length	*Short*		*Average*		*Long*
	Texture	*Straight*		*Wavy*		*Curly*
	Shade	*Dark*	*Medium*	*Light*	*Gray*	*White*
Forehead		*Receding*		*Vertical*		*Bulging*
Eyebrows	Weight	*Thin*		*Medium*		*Bushy*
	Separation	*Separated*		*Meeting*		
Eyes	Opening	*Narrow*		*Medium*		*Wide*
	Separation	*Close*		*Medium*		*Wide*
	Shade	*Light*		*Medium*		*Dark*
Ears	Length	*Short*		*Medium*		*Long*
	Protrusion	*Slight*		*Medium*		*Large*
Cheeks		*Sunken*		*Average*		*Full*
Nose	Length	*Short*		*Medium*		*Long*
	Tip	*Upward*		*Horizontal*		*Downward*
	Profile	*Concave*		*Straight*		*Hooked*
Mouth	Upper Lip	*Thin*		*Medium*		*Thick*
	Lower Lip	*Thin*		*Medium*		*Thick*
	Lip Overlap	*Upper*	*Neither*	*Lower*		
	Width	*Small*		*Medium*		*Large*
Chin		*Receding*		*Straight*		*Jutting*

Table 5.1: Characteristics of Facial Features

Attempts have been made to discover which features are particularly important in facial recognition. Harmon classified the 21 features, listed above, as those that are most useful in discriminating one face from another [67].

generation of a type; a police sketch artist can produce a face from a victim's description of the salient features of the criminal's face. An alternative approach is to use a formal classification system based on feature identification such as the *Identikit* and *Photofit* systems.

Identikit and Photofit Systems

The Identikit and Photofit systems share the same principle in that a face can be composed from separate sets of features. However, the Identikit system involves composition of a face from line-drawn features, whereas photographs of features are used in the Photofit system.

The Photofit system consists of sets of features for the forehead and hairline, eyes, nose, mouth and chin. Three such 'kits' are available for male caucasians, female caucasians and male afro-asians. Each kit comprises a catalogue of features together with a library of individual features mounted on strips of cardboard. Once chosen from the catalogue, features are selected from the library to create a composite face in a glass frame.

Despite being based on much psychological research, results from the Photofit system have been far from satisfactory. Investigation undertaken by Ellis et al. indicated that people have difficulty in making up a reconstruction of a face using Photofit [46, 47]. Performance of subjects seemed largely independent of memory factors. For instance, subjects were unable to construct better likenesses even when the target face remained in view! The researchers found that people differ markedly in their ability to reconstruct faces, so that those subjects who were adept at reconstructions made Photofit faces which were more easily identifiable. In addition, subjects found it difficult to compare isolated, individual features with those embedded in an appropriate context, that is, a face.

The results imply that the major impediment resides in the Photofit system itself. A possible reason for this failure may be due to an inadequate sampling of the population for each feature. Another reason may be because the system requires people to undertake a task which they are not familiar with; that is, fragmentation of a holistic precept or gestalt. More specifically, subjects frequently expressed difficulty in searching through numerous isolated features for an approximation to one seen within a total physiognomic context.

An examination of the Photofit system itself revealed other short-comings [78]. For example, features appear on strips of cardboard and fit into 'slots'. Consequently, the capability to experiment through variation of distances between features is limited.

Electronic-Fit System

Recently a project was undertaken by *Cadcentre Ltd.* in collaboration with the Home Office's *Scientific Research and Development Centre* to implement an *Electronic-Fit* or *E-Fit* system. The aim of this computer based photomontage system was to resolve some of the problems that had become apparent from the manual Photofit system.

Numerous experiments have shown that the identification performance of a witness deteriorates both with delay in the search process and with the number of other faces observed in the meantime [1, 194]. In particular, identification performance deteriorates after approximately fifty photographs have been viewed, so that the longer the series of pictures, the poorer the performance of the witness. Such results called for a rapid prescreening device to reduce the number of 'mug-shots' that the witness is required to view.

The E-fit system achieves this objective by taking a verbal description of the target face from the witness and using this to discard all irrelevant faces prior to viewing. Additionally, Harmon had identified that recognition can be enhanced by filtering out high frequencies such as those which result from sharp edges [67]. To improve the realism of composite photographs, the E-fit system merges the edges of features into the face. To a limited extent, the system also allows features to be scaled and moved within the context of a facial image.

5.2.3 Facial Reconstruction and Anthropology

Facial reconstruction is the scientific art of sculpting a face onto a skull. While Identikit, Photofit, E-fit and police sketches are useful for identification of live missing persons or fugitives, this technique aids in the identification of unknown skeletal remains. The method is commonly used in archeology as well as forensic sculpture. The successfulness of the method has been demonstrated in several cases of criminal investigation which have involved unidentified skeletal remains [54, 107, 141].

Facial reconstruction involves taking a cast of the skull, usually in plaster. Plastic spheres are then inserted into the eye sockets. The next stage involves drilling small holes at specific anatomical sites, into which wooden pegs are inserted. The lengths of the pegs correspond to average tissue thickness measurements taken from predefined tables [138]. The muscular layers of the face are then built up, using clay, on the plaster cast support with the pegs as guidance.

The skull is therefore regarded as the armature of the face. From the bone structure can be determined the disposition of the facial features including the mouth, eyes and nose [141]. This method facilitates construction of the face from the surface of the skull outwards. Although the technique provides an indication of the overall size and shape of the head, together with the position of major features, it still remains extremely difficult to determine the exact shape of features since there is no live image for reference. The lower nose, mouth and ears tend to cause the greatest problems [54, 141], although there are various heuristics which can help with the construction of these features [55, 56, 74, 153].

Several computerised approaches based on facial reconstruction have been developed. Most of these involve photographic or video superimposition [37, 169, 170]. Once a skull has been digitised, prescanned photographic images of people are mapped onto the model to determine whether the face 'fits' the skull.

Vanezis et al. describe a system which 'grows' facial tissue onto the digitised skull model to provide a facial mask [169]. The process probably consists of interpolation between tissue thickness at specific sites and subsequent 'smoothing' of the resultant surface. Predigitised features similar to those used in the Photofit system are then chosen and mapped onto the model of the skull. Features can be adjusted to provide a better correspondence with the skull through scaling and translation techniques. Voci gives an account of another very similar system which was used to create a specific face [170].

These computer based systems emphasise the point at which science stops and art takes over. The systems can handle well-defined aspects up to the point of creation of a facial mask from known relationships between the face and the underlying bone structure. However, the task of choosing specific features for the mask relies on personal judgement. It is readily acknowledged that a large number of faces could conform to a particular skull [27, 169, 170], consequently there are just too many variations involved.

5.2.4 Medicine and Anatomy

The study of the face is widespread in the world of medicine; any amount of information that can be gleaned is likely to be pertinent to areas such as facial surgery and orthodontics. Many x-ray type techniques, such as use of lateral craniographs and cephalometric radiography [55, 56, 153], have been devised in an attempt to understand the relationship between cranial features and soft tissue. As a result, some useful principles have been established [56, 153]. However, most such techniques operate on 2D or profile data, which tends to be of limited use when working in 3D.

Although computer technology has been in use for some time in planning facial surgery [191], simulations have been constrained because of the lack of quantitative analysis of 3D head and face data. To address this problem, Moss et al. have recently developed a 3D method which can provide objective, qualitative and quantitative descriptions of the face [103]. The system has been used to analyse the effects of reconstructive surgery on the face [34, 35, 103]. Various measures can be established by comparison of classifications of pre and post operative laser scans of the head. The method works by rendering the head data into a z-buffer. Depth values in the z-buffer are then processed to generate Gaussian curvature and mean curvature at each pixel. This data is used to classify the surface at each pixel into one of eight 3D primitive shapes.

The same method has also been used to establish average male, female and non-gender heads based on a sample population of ten male and ten female laser scans [34]. The technique has great potential for use in the measurement of growth and its effects on the morphology of the head.

5.2.5 Portraiture, Sculpture and Proportion

Many an artist and sculptor has studied the variations and proportions of the human figure. Well-established rules indicate a relationship between age and the size of an individual [150]; human figures are measured in terms of 'heads'. As the body grows in height, the proportions change accordingly. For example, a one year old child is four heads high; at nine years of age the child is six heads high; and when at the age of fifteen the teenager is seven heads in height.

The Greeks considered proportion to be equivalent to beauty and based much of their work on the Golden Section [13]. In accordance with this a 'perfect face' would be one which has its brow a third of the way down from the hairline, while the mouth would be one third of the way up from the chin. Dürer and Da Vinci, Renaissance artists of the 16th Century, were also preoccupied with proportion [113, 154] and established guidelines to relate different parts of the body based on divisions of sevenths. Such heuristics were also used to determine the relative proportions of the face and facial features.

5.3 Computer Representation of the Face

In recent years a number of methodologies have been employed to represent and model irregular surfaces. The major computer based techniques that have been used to represent the human form include: *digitisation, interpolation, local deformation* and *composition from parts*.

5.3.1 Digitisation

There is no doubt that the computer generated animation short film, *'Tony de Peltrie'*, which premiered at SIGGRAPH in 1985 [48] was a landmark in facial animation. However, the film, which lasts seven minutes and fifty seconds, was a culmination of four man-years of effort by a group composed of four artists and programmers. Tony's head, as well as every other prop that appeared in the film had to be tediously and painstakingly digitised from a physical representation.

Many other researchers have also used digitised data [48, 97, 116, 175]. As discussed in section 3.3.1, several approaches can be used, but all of them require a physical representation of the model and most of the methods require a grid to be either drawn or superimposed on the face. Lofting and laser scans both require hardware which is specialised and expensive. Lofting tends to be used primarily in medical applications which need precise data.

Laser and light beam scanning of human heads has recently become extremely popular [103, 122, 160]. This technique is probably the most appropriate when a recognisable representation of a specific individual is required to overcome the viewers expectations of observing a realistic face. Although the method can be fast, with speeds of 15 seconds for obtaining 20,000 points [35], such large amounts of data cause problems with respect to storage, manipulation and control.

Photogrammetry and manual digitisation using a 3D device both entail the labour-intensive process of first drawing relevant facets on the surface. Several additional problems arise with manual digitisation of human faces. Firstly, an expression has to be held fixed and unchanged for a considerable length of time. Secondly, many parts of the face are soft and therefore liable to distort during contact with the tool-tip, to cause inaccuracies in the recorded coordinates. Thirdly, some mechanical digitisers are limited in movement to certain degrees of freedom which causes difficulties during digitisation.

Despite the potential accuracy, the process of digitisation can be extremely tedious and time-consuming. This is particularly apparent when the model is to be animated at a later stage, requiring a manual delineation of specific regions. Although this may not be a major problem for one face, it would become tiresome if the process had to be repeated for a large number of faces.

Terzopoulos and Waters have recently devised an automatic technique to perform just such a task on laser-scanned data [159]. In a process known as *adaptive meshing*, salient points are extracted from the scanned data. The number of points extracted is increased around areas of high curvature. Kurihara and Arai [83] describe an alternative approach which matches a 3D canonical model with photographs of individual faces. In both of these cases the model can be subsequently animated.

A major drawback of acquiring data from human heads is that the data is specific to a particular face. This is fine when it is necessary to develop models of existing people and indeed it is useful for testing animation models to check whether they produce similar expressions to those displayed by real people. However, a problem arises when it is necessary to create totally new characters with distinctive faces.

5.3.2 Shape Interpolation

3D shape interpolation between two predigitised faces can be used to generate new intermediate facial structures [100]. This method can provide very effective results, but there are difficulties due to the irregularity of the face. Individual faces differ to such a large extent that to represent them accurately and recognisably requires a differing number of vertices, facets and topology for each face. As a result, the interpolation process becomes complicated and requires a 1 to n correspondence between vertices and facets in the two structures to be interpolated [9, 69, 77].

There are two possible solutions to the problem. The first involves making facets and vertices in one face appear or disappear in order to make them correspond with the other face [9, 77, 100]. With the second method the two faces are re-structured according to a set of facets and vertices common to both [69, 100].

The greatest drawback of the shape interpolation approach is that two predigitised faces are necessary before a new model can be created. In addition, it is not possible to control local areas independently of the rest of the face.

5.3.3 Local Deformation

From the outset, Parke was aware of the need to model interactively a large number of facial shapes. His use of parameterisation was intended to enable concise criteria to encapsulate every member of the 'universe' of faces. Parke used the method to address both face creation and animation [118]. Based on observation, the parameters were divided into the two categories of *conformation* and *expression*. Conformation parameters dealt with those aspects that vary from one face to another, as opposed to those that vary between expressions. Several techniques were used to implement operations involving conformation parameters. For example, interpolation was used to vary the forehead from sloping to bulging; scaling determined the aspect ratio of the face; and translation facilitated movement of the chin in various directions.

Experiments were also undertaken to derive a general topology for the face, one which would be suitable for representing a large number of faces. The major reason for studying this issue was that the model required predefinition of regions, in terms of groups of vertices on which parameters operated. Only by maintaining the same topological mesh would the parameterisation remain applicable to many faces.

Magnenat-Thalmann et. al. describe several techniques for interactive selection of facial regions to be transformed [100]. This resolves the issue of having to predefine the regions. Four methods for implementing the transformation of regions are also described. These vary from manipulation of individual vertices to operations on groups of vertices.

The major restriction of using local deformation is that the number of vertices and facets remains constant. It is doubtful that one topological mesh can adequately represent a large number of human faces; faces which are totally unique and individual as people themselves.

5.3.4 Kit of Parts

The WHATISFACE system was an early face composition system which operated on static 2D line-drawn faces [57]. The face of a male caucasian alone was dealt with. Sets of features were extracted from the work of psychologists and supplemented with statistical tests on 255 photographs. These same photographs were used to establish an average 2D face, which could be modified by the user to create a new face. Editing facilities consisted largely of automated changes based on yes/no type answers to predetermined questions. The system made use of 17 line-drawn features of which 7 were left-right pairs. The major features used included: cheek lines, chin lines, ears, eyebrow outlines, lower eyelids including crow's feet, upper eyelids, eyes comprising iris and pupil, forehead lines, hair outline, upper lip, lower lip, mouth lines, naso-labial lines, neck, lower and upper parts of the nose.

Despite the fact that E-fit is a more recent computerised system, the images produced are still static and 2D. In addition, the facial composites are totally devoid of expression. Since E-Fit uses image processing techniques, it is difficult to manipulate the faces in 3D. A volumetric approach would probably serve better than the surface techniques that are currently used.

However, a 3D approach based on the principles used in the Identikit and Photofit systems would first require a large set of features to be digitised, or otherwise represented, possibly through the use of parameters. Facilities would then be required to extract particular sections of a face and to insert other parts in order to form a composite face.

Magnenat-Thalmann et al. have demonstrated a technique for the composition of various parts of the human body to create a synthetic actor [98]. The method uses 'brothers' which are points at which two body parts are to be joined together. Three non-colinear points on the boundaries of parts are also required to ensure that the two parts correspond in size. A similar technique could be developed for the composition of 3D faces.

5.4 Conformation Modelling in FACES

One of the greatest problems with computer generated facial animation is that the task of obtaining a 3D computer model must precede any other work. The problem lies not in the task itself, but in the subtle diversities in the form of the face. Unlike many other objects, the irregular nature of facial features has defied precise geometric description at a macro level and led researchers to acquire data for new models completely from scratch.

The usual way of obtaining face data is through the process of digitisation, which entails several problems, as explained in section 5.3.1. The major drawback is that the data acquired is specific to the particular physcial representation used in the digitisation process. There is therefore a major requirement for the interactive construction of models of arbitrary faces.

A major objective of the CONSTRUCT and MODIFY sub-systems is to overcome the problem of having to resort to digitisation whenever a new face model is required. This goal can be achieved by providing the user with a model of the head which can be interactively modified. Such a capability enables models which represent arbitrary faces to be created; the goal is to animate such models later on.

In FACES the user modifies the structure of a predefined head model in terms of changes to bone, muscle, skin and surface features. The head model notionally defines an 'average' head and face as explained earlier in section 3.3. There are several reasons for the provision of a predefined head model. The validity of such an approach is substantiated by Carey et al. who have found evidence to suggest that we encode and remember facial information, relevant to a particular face, by comparison with an average face [25], see section 5.2.1.

Provision of an initial head model also gives the user a 'head start'. It is clear that we are adept at recognising faces, but most people experience great difficulties in drawing or otherwise creating a face. In addition, a predefined head model provides a context within which modifications can be viewed. A major problem with systems such as Identikit and Photofit is that people are required to choose isolated features and then to compose these into a recognisable face. This task has proven to be extremely difficult to perform [46, 47], see section 5.2.2.

From a detailed investigation of the anatomy of the head and face, as described in chapter 2, it is apparent that although there is individual variability, all faces have a similar general form. In addition, the features appear in the same relative positions. These observations suggest that generic modelling is an appropriate technique to use in modelling the conformation of the head and face. A survey of the major disciplines that are concerned with the face has further indicated that modelling should be considered in terms of three levels; global, proportional and localised changes need to be accommodated.

Techniques used in facial reconstruction, as discussed in section 5.2.3, demonstrate that the rigid substructure of the skull determines the overall shape and proportions of the face [110, 146]. Although muscles and soft tissue change radically throughout life, it is the structure of the skull that determines the general shape that we recognise as the head. The skull should therefore form the starting point for creating faces. Indeed, the problems encountered by Pixar Ltd., during the production of *'Tin-Toy'*, a short animated film featuring the modelling of a baby's facial expressions, were largely attributed to the lack of an underlying bone structure [88].

A three-layer model based on the anatomy of the head is used in FACES because bone,

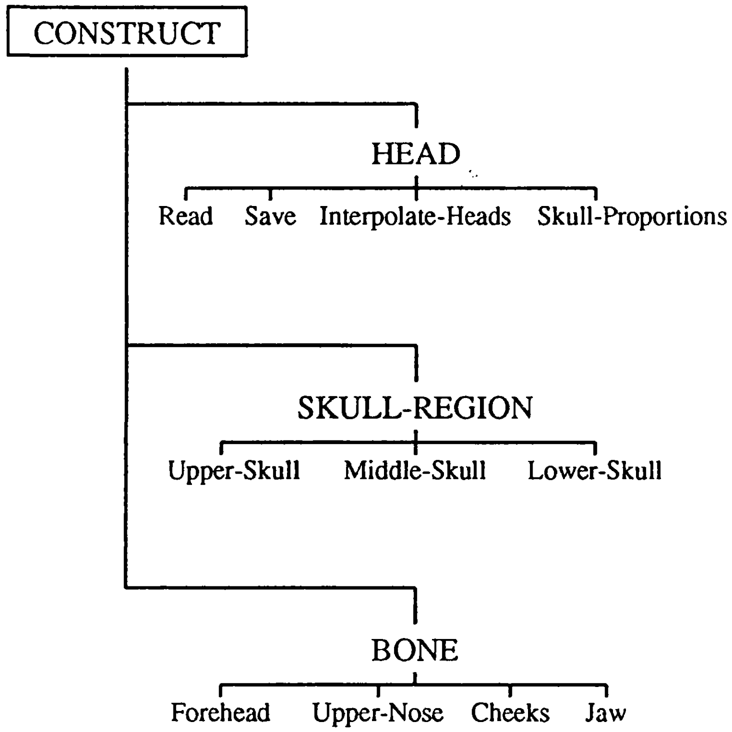

Figure 5.1: The CONSTRUCT Sub-System

Structure and organisation of the CONSTRUCT sub-system in FACES. Modifications to the bone structure may be performed at three levels: global, regional and local.

muscle and skin all form an integral part of the modelling process. Modelling facilities are subdivided between the CONSTRUCT and MODIFY sub-systems such that the user works with the bony structure of the head in the CONSTRUCT part and the muscle, skin and surface features in the MODIFY sub-system.

5.5 The CONSTRUCT Sub-System

Since the skull forms the structural support for the head and face, it is insufficient to modify the surface facial skin alone as in the *Human Factory System* [96, 100]. Within the CONSTRUCT sub-system, the user can modify the structure of the skull to define the general shape and proportions of the head and face. Changes to the skull are automatically reflected in the muscle layer and the facial skin mesh, which wraps around the underlying skeletal base. Alterations to the bone structure lead to changes in the location of muscle origins, while the corresponding changes in the face mesh result in displacements to muscle insertions. Modifications to the skull therefore have implications for the dynamics of the muscle models; they are responsible for variations in the motion of differing head models.

The complete structure of the CONSTRUCT sub-system is shown in Figure 5.1, while the sub-system's context within FACES is illustrated in Appendix E. To guide the reader through the hierarchical organisation of this part of FACES, menu structures are used to indicate clearly the options that are available at a particular stage.

At the highest level the CONSTRUCT sub-system consists of three major sections which enable modifications to be made to the skull at three levels of control. The sections are accessed through the following menu:

HEAD
SKULL-REGION
BONE

The HEAD option caters for operations on the overall and global nature of the head, see section 5.5.1. SKULL-REGION facilitates proportional changes as explained in section 5.5.2. The BONE option enables modifications at a local level to individual bones, see section 5.5.3.

5.5.1 Global Control

The HEAD part of this sub-system provides two types of functionality. Firstly, it enables new head models to be brought into the system and secondly, it allows global amendments to be made to the structure of the head. The HEAD menu-item itself allows access to the following options:

Exit
Read
Save
Interpolate Heads
Skull-Proportions

Read and *Save* provide a communication link between FACES and disk storage. The *Read* option retrieves head model data from disk files, while *Save* writes head structures to disk files. The user is prompted for the names of data-files in both cases.

The *Interpolate Heads* option is a method for generating new head models through interpolation between two existing models. Evidence to support such a strategy can be found in Carey et al. [25] who claim the existence of a *face-space*, such that interpolation between two faces results in another face which belongs to the domain of all faces. In FACES, the process of generating new head models has been made flexible through provision of several interpolation techniques based on the laws of motion described in section 4.7.1. Intermediate models such as those depicted in Figure 5.2 are displayed on the screen. From these heads it is possible to select a head model and read in the corresponding data-file so that the selection becomes the 'current' head that the user is working with.

At present such conformation interpolation is only possible between two models having the same topology. However, algorithms for interpolating between objects with differing topologies are available [9, 69, 77] and could be incorporated into the system at a later date.

A special aspect of working with the face is that it is essential to establish a correspondence between similar physical regions. For instance, a nose must be mapped to another nose, otherwise the resultant structure will become unrecognisable as a face.

Selection of the *Skull-Proportions* option facilitates global changes to the proportions of the skull and thereby to the head model. As discussed in section 5.2.1 the overall shape and proportions of the head and face are important factors in facial recognition [57, 194]. A 3D scaling facility enables alterations to the whole of the skull structure in each of the x, y and z dimensions, see Figure 5.3. In the case of an effect not being that required, an *undo* facility allows the change to be discarded.

5.5.2 Regional Control

SKULL-REGION, a major option in CONSTRUCT, facilitates modifications to relative proportions of the facial skeleton. The structure of the skull has been subdivided into three regions known as the *upper-skull, lower-skull* and *middle-skull*. Each of these sections can be either stretched or compressed through the use of valuator bars as shown in Figure 5.4.

In the x-y plane, an upward movement of a bar stretches the region below the bar by a corresponding amount, while at the same time compressing the area above the bar. A downward movement of the bar reverses the effect. In the y-z plane movement of a valuator bar to the left causes compression, while that to the right results in stretching of the region. A stack-based *undo* operation allows the user to reverse changes that are not required.

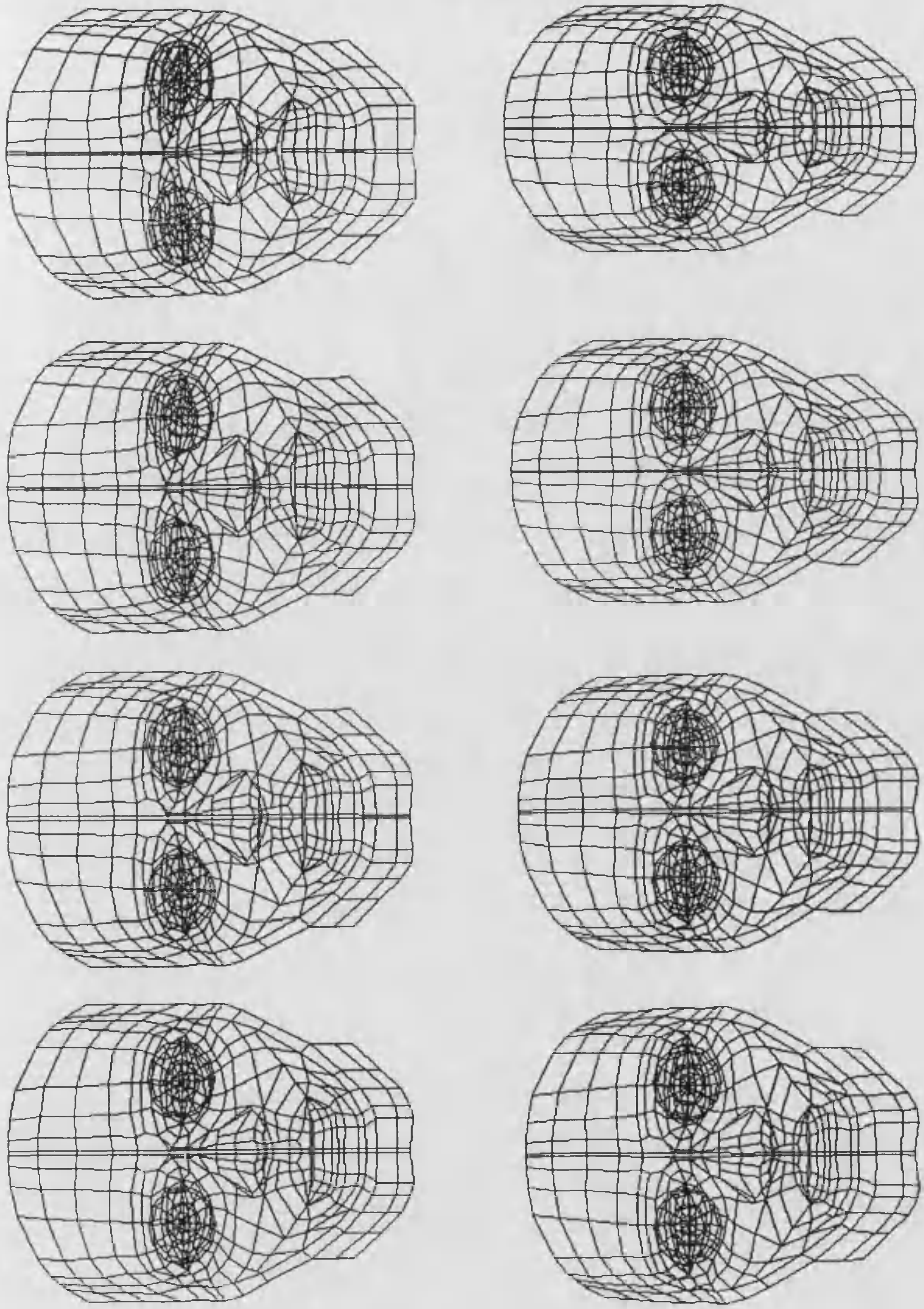

Figure 5.2: Generation of New Heads

Generation of new head models through interpolation between two existing head structures. It is possible to select any of the new head structures as the head model to be worked upon.

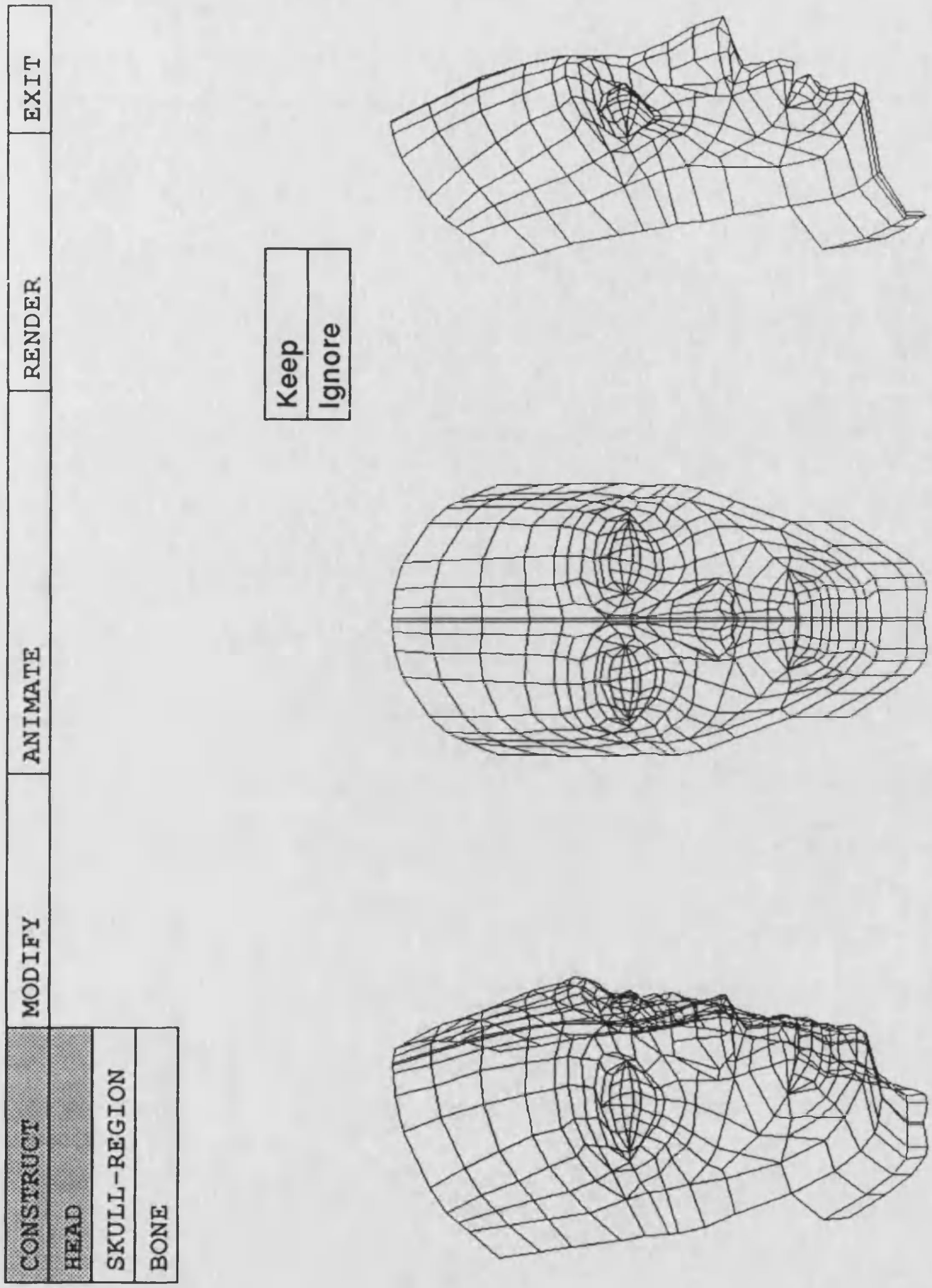

Figure 5.3: Global Scaling of Skull

Global changes to the entire structure of the head through scaling. The overall shape and proportions of the head and face are important factors in facial recognition. In this case the skull has been reduced in width.

5.5.3 Local Control

At the lowest level of the CONSTRUCT sub-system, invokation of the major option BONE enables amendments to be made to more specific parts of the skull. The human skull comprises fourteen major bones [62, 147]. Of these, control is currently provided over the: *frontal bone* or forehead; *nasal bone* or upper-nose; *zygomatic bone* or cheeks and *mandible* or lower jaw. Each bone may be operated on through the menu shown below:

Exit
Forehead
Upper-nose
Cheeks
Chin

Facial reconstruction artists argue that such bones determine the size and position of the major facial features [54, 110], see section 5.2.3. In addition, bones are rigid structures which are not capable of flexible deformation. An appropriate method for provision of control over such structures is through the use of rigid geometric transformations. Adjustments to individual bones can therefore be made through transformations such as *Shift*, *Scale* and *Rotate*, see Figure 5.5.

A particular aspect of working with the face is that some bones need to be treated as 'pairs' and others as 'single' bones. Paired bones are those that are separate and distinct on the two sides of the face, they include the *Cheeks* and the *Forehead*. Single bones are those that are joined at the central vertical meridian such as the *Upper-nose* and the *Chin*. For paired bones the user may specify one of *Left*, *Right* or *Both* to constrain modifications to either one of the sides or to apply changes to both sides simultaneously.

Modifications to paired bones result in opposing actions for certain types of geometric deformations. For example, a *Shift* of the *zygomatic* bone outwards in the x-dimension involves movement of the left and right bones to cause the face to widen on both sides at the cheeks. Further, *Rotate* operations about the depth-axis cause one bone of a pair to be rotated clockwise and the other anti-clockwise. In the case of single bones, a *Shift* operation in the horizontal plane, on the nasal bone for example, causes both sides of the bone to move in the same direction. Similarly rotation about the depth axis results in a turning of both sides of the bone in the same direction. Whilst modifying particular bones it is possible to ignore alterations through the use of an *undo* operator.

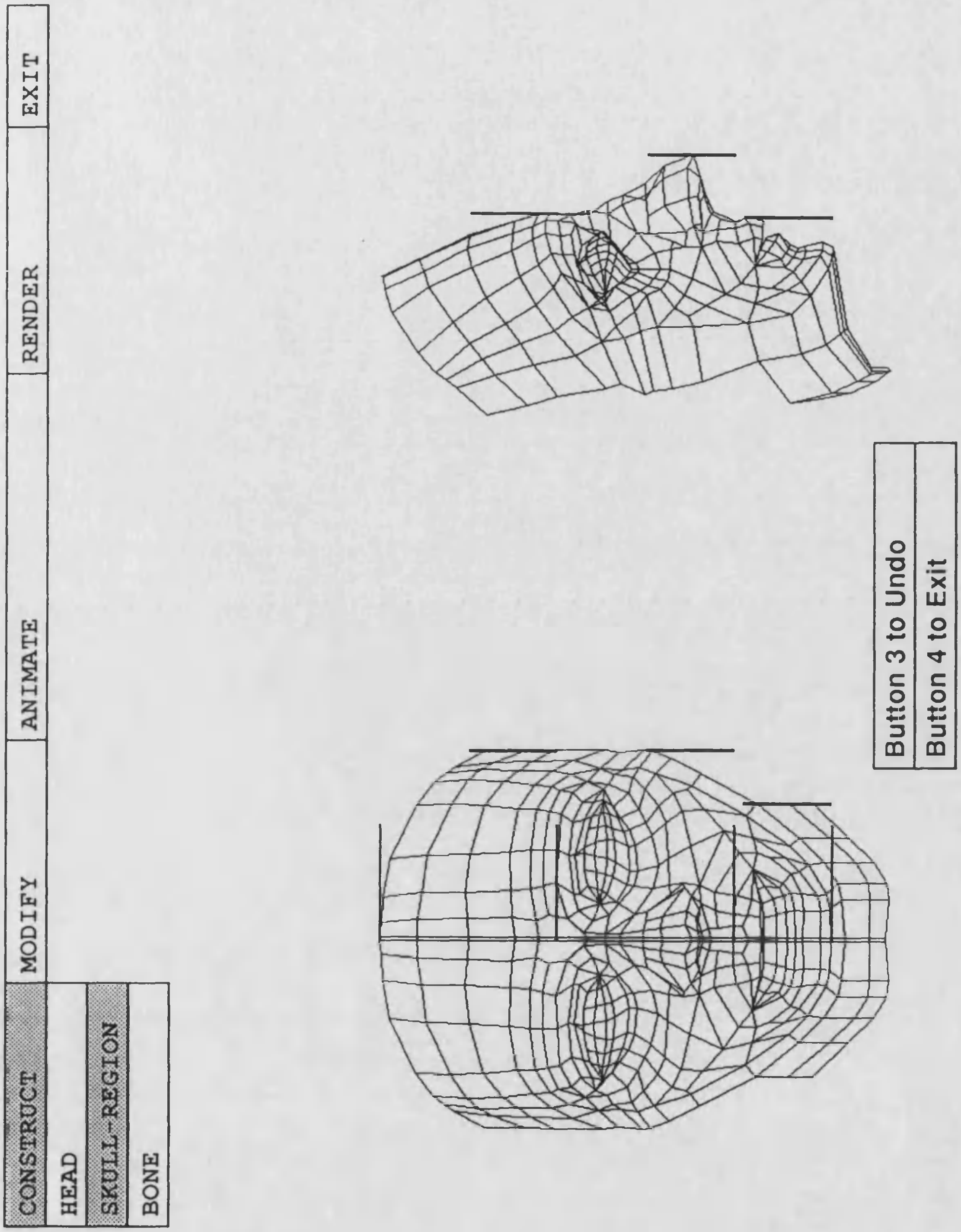

Figure 5.4: Changes to Relative Proportions of the Head

The representation for the skull has been subdivided into three sections. The ***upper-skull***, ***lower-skull*** and ***middle-skull*** can each be stretched or compressed in the x-y and y-z planes.

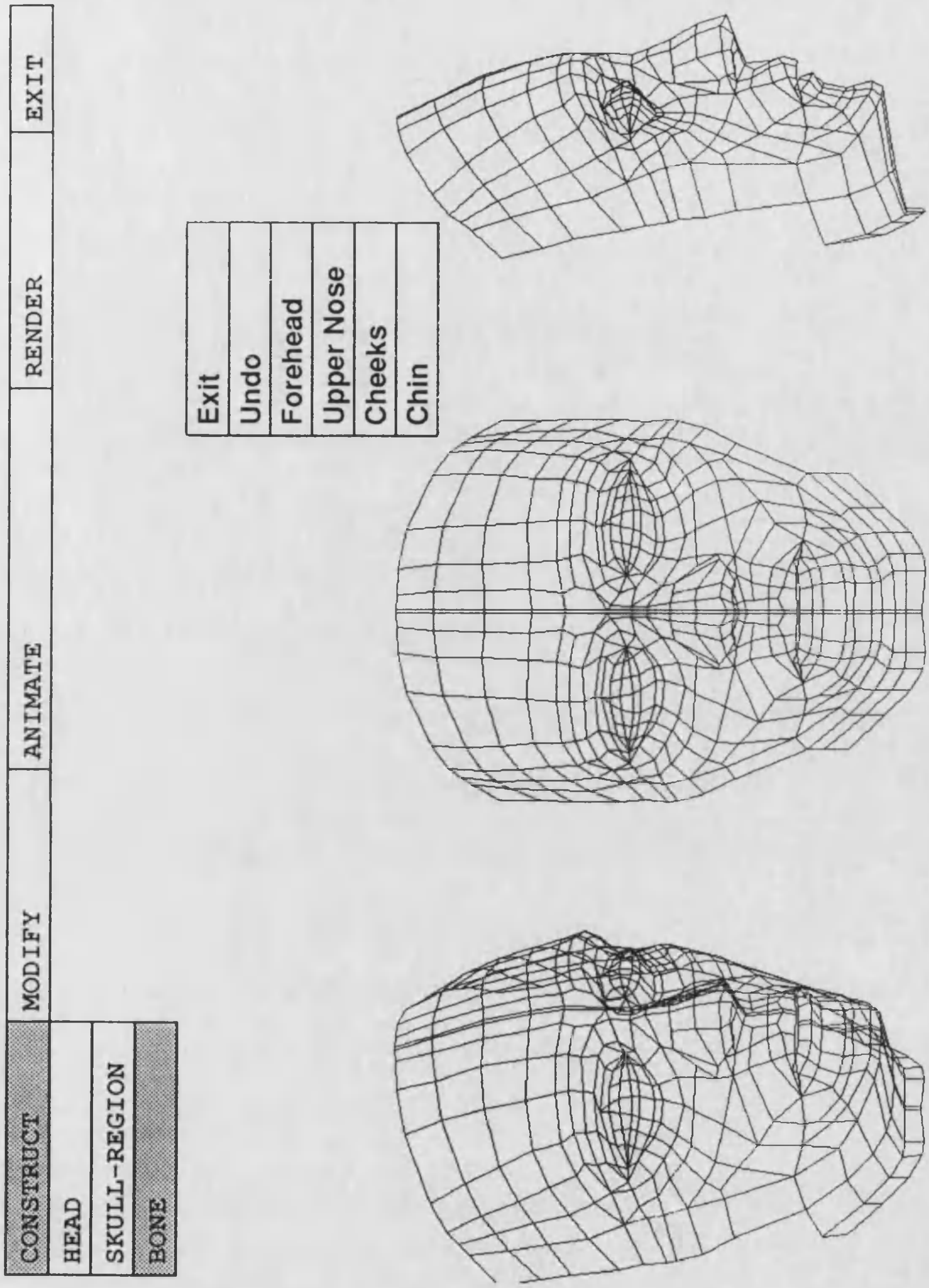

Figure 5.5: Modifications to Bone

At a local level changes can be made to specific bones. These determine the size and position of the major facial features. Bones may be treated as either paired or single.

5.6 The MODIFY Sub-System

While it is the skull that determines the general proportions of the head, there are variations in the surface attributes of the face which cannot be accurately determined from the underlying bone structure [54, 141]. More specifically, these differences relate to muscle thickness and *soft* features such as the lower nose, lips, eyes and ears, as explained in section 5.2.3. Such characteristics vary significantly between individuals and form the major reason why faces are perceived to be so different from one another. In addition, the soft features tend to be more important than bone structure since they are more visibly significant.

The MODIFY sub-system addresses conformation modelling with respect to variations in both muscle and surface characteristics. The structure of this sub-system is presented in Figure 5.6, while its position within FACES is shown in Appendix E.

When displayed on the screen, the skin mesh is distinguished from the skull mesh by its larger size, colour and the inclusion of the eyeballs. Whereas modifications to the skull in the CONSTRUCT part of the system indirectly affect the muscles and the facial features, the MODIFY sub-system caters for more subtle amendments by influencing only the characteristics of the surface skin and muscle insertions. To facilitate the creation of distinct faces, the sub-system has been organised into three major sections which provide global, regional and local control over the face. The three sections may be invoked through the following menu:

FACE
FACE-REGION
FEATURES

Facilities in the FACE option enable changes to muscle thickness, as explained in section 5.6.1. FACE-REGION provides a method of altering relative proportions of the facial skin, as demonstrated in section 5.6.2. The FEATURES option which caters for localised changes to specific soft features is discussed in section 5.6.3.

5.6.1 Global Control

Thickness of facial tissue is an important factor in fleshing out a face. As explained in chapter 2, muscle thickness varies between males and females [54], as well as with time and health. Global control over muscle thickness in provided within the FACE option. Using a

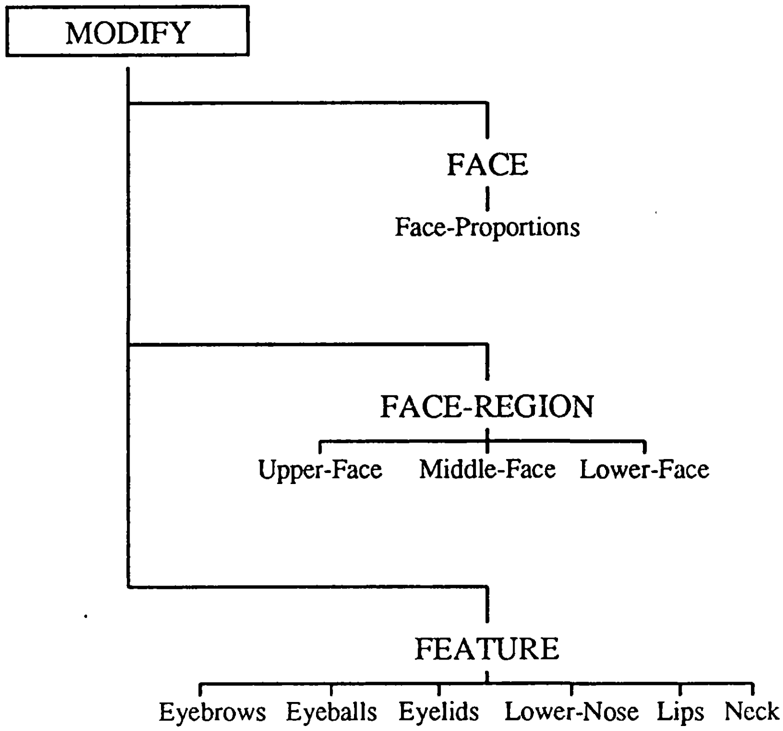

Figure 5.6: The MODIFY Sub-System

The structure and organisation of the MODIFY sub-system in FACES. Variations in the surface attributes of the face cannot be accurately determined from the underlying bone structure. Changes to skin, muscle and surface features may be made at global, regional and local levels.

scaling function which operates in 3D, the user can change the underlying distance between the skull and the facial skin meshes. Such changes have an indirect effect on overall muscle thickness which is represented through the length of muscle vectors. Adjustments to the thickness of muscles has implications for the dynamics of the head during animation, see section 4.5.2. Subsequent to observing the effects of global changes to the face, the user may either *Keep* or *Ignore* the changes made.

Craniofacial measurements for the average thickness of soft tissue at particular sites on the skull are available from tables [138]. This data would be useful in facilitating control over individual muscles, but has yet to be incorporated into FACES.

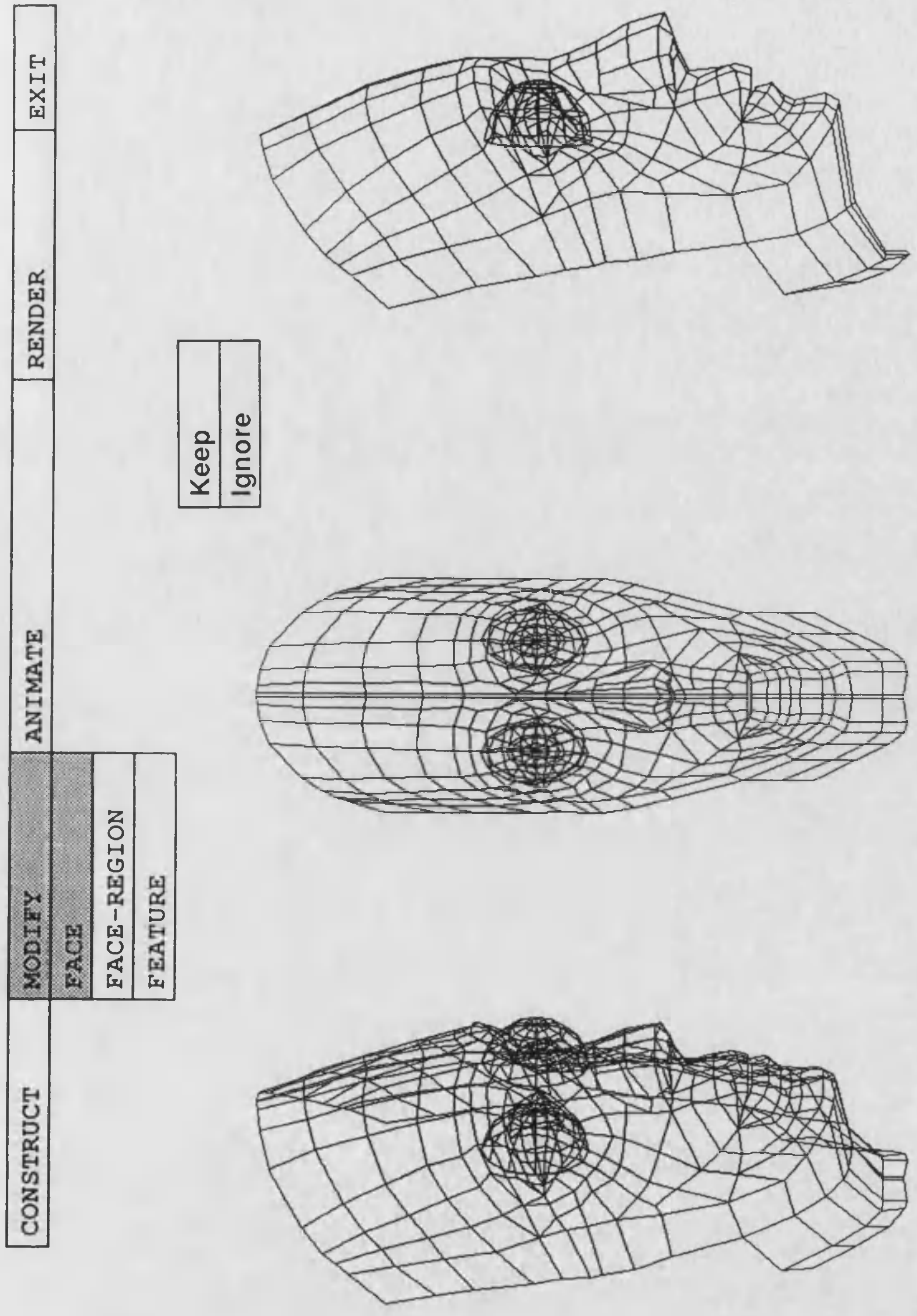

Figure 5.7: Changes to Overall Proportions of the Face

Global changes to the proportions of the face do not affect the underlying skull structure, but have an indirect effect on overall muscle thickness. Modifications can be made in each of the x, y and z dimensions.

5.6.2 Regional Control

The major option FACE-REGION, caters for alterations in the relative proportions of the face. Sculptors and portrait artists have for years used a variety of rules that relate parts of the face to each other, see section 5.2.5. The facial skin mesh has been divided into three sections known as the *upper-face, middle-face* and *lower-face*; these are based on the regions used in FACS [44] since FACS is used as the basis for generating expressions in FACES. The upper face consists of the eyebrows and forehead; the lower face is the area from the chin to the upper lip; and the middle region comprises the area from the upper lip to the eyes.

As illustrated in Figure 5.8 valuator bars similar to those employed in SKULL-REGION are used to make changes to relative proportions of the face. In the context of regional modifications to the face, a stack-based *undo* operation is available to enable successive changes to be discarded when necessary.

5.6.3 Local Control

The FEATURES option comprises the final part of the modelling process. This involves refinement in the form of scaling and positioning of the soft features of the face, see Figure 5.9.

According to Ekman, individuals differ greatly with respect to the size, shape and location of their facial features [45]. Harmon, however, considered that it is not the shape and size of features that is critical, but the relative distances between them on the face [67]. From psychologists' concern with the process of face recognition, it has been established that people notice differences in facial shape, eyes, eyebrows, nose, mouth and lips, and the chin [195]. But how these features differ from person to person has not yet been fully established. Therefore to cater for experimentation, geometric deformations are employed to modify soft features through the following menu:

Eyebrows
Eyelids
Eyeballs
Lower-nose
Upper-lip
Lower-lip
Neck

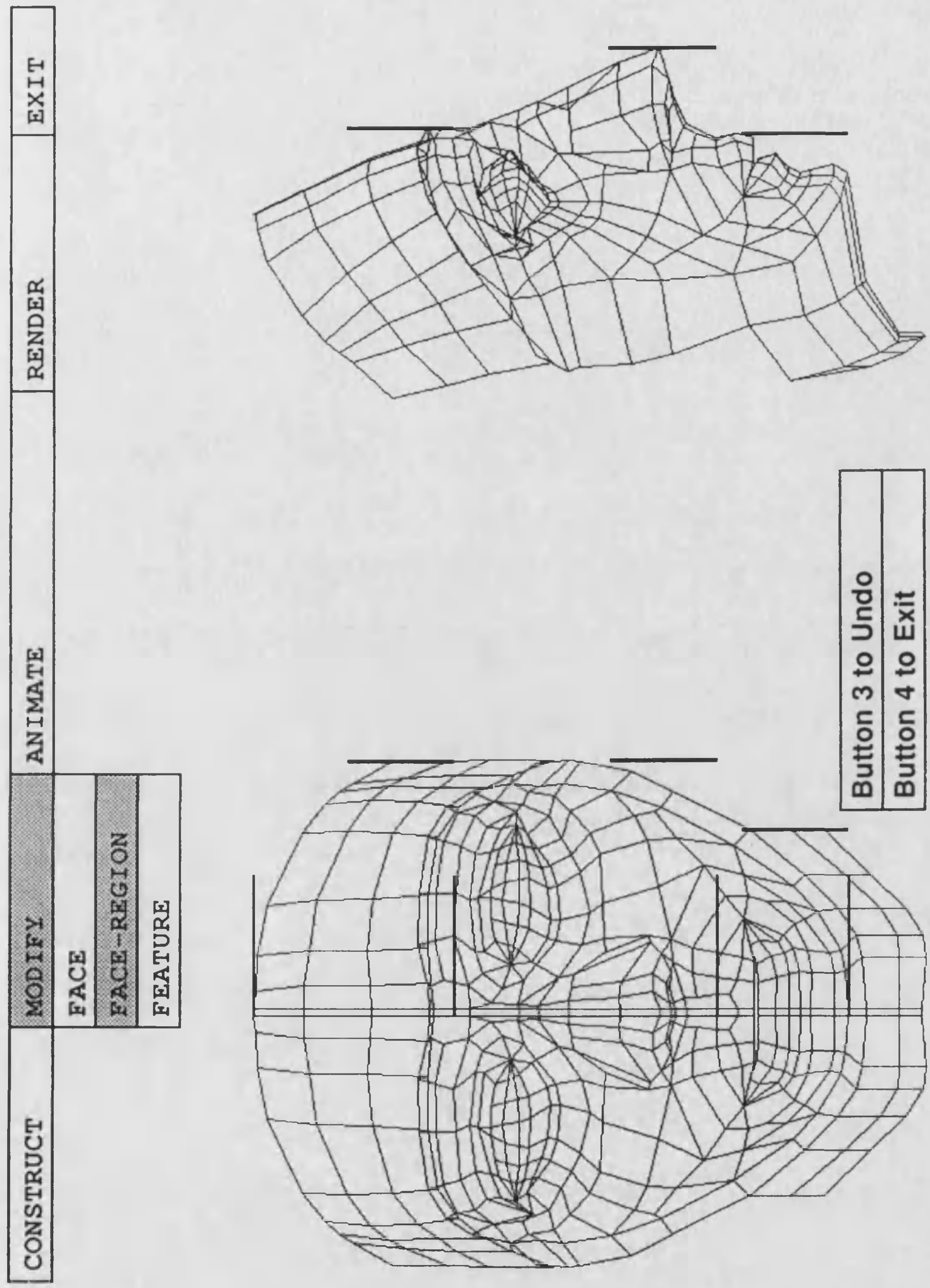

Figure 5.8: Changes to Relative Proportions of the Face

Relative proportions of the face can be altered through regional changes to the *upper-face*, *middle-face* and *lower-face*. Each section can be stretched or compressed in the x-y and y-z planes.

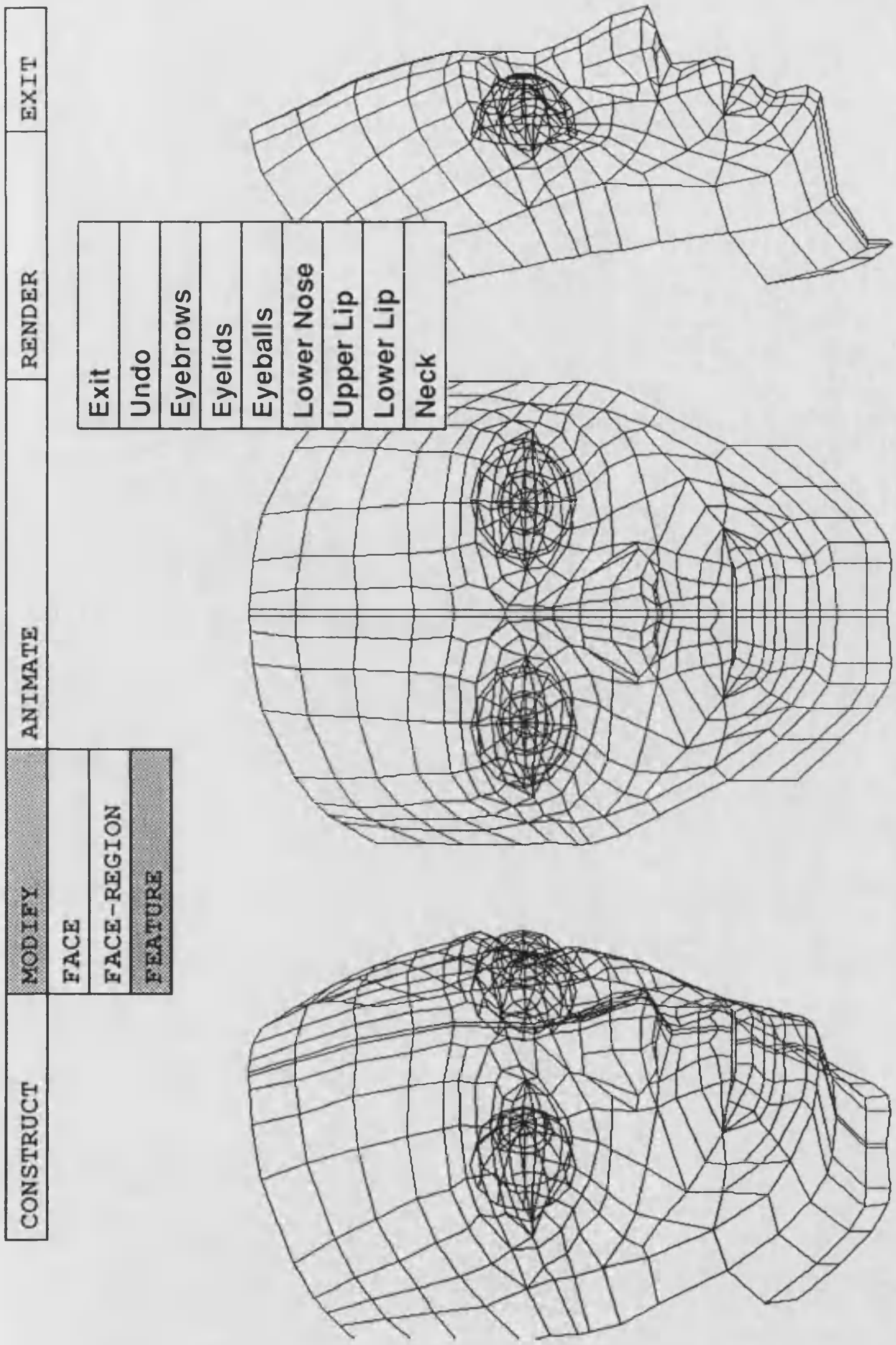

Figure 5.9: Modifications to Features

Local control enables changes to individual surface features which can be either paired or single. According to Ekman, individuals differ greatly with respect to the size, shape and location of their facial features [45]. Harmon, considered that it is not the shape and size of features that is critical, but the relative distances between them on the face [67].

As in the BONE section of the CONSTRUCT sub-system, some of the features are treated as pairs while others are single elements. For example, *Eyelids, Eyebrows* and *Eyeballs* are paired features over which independent control is available for the *Left* and *Right* sides, while the *Lower-nose, Upper-lip, Lower-lip* and the *Neck* are regarded as single features. With regard to adjustments to paired and single features, similar comments to those made under the BONE option, in section 5.5.3, apply. Undo functionality allows changes to be reversed when necessary.

5.7 Summary and Conclusions

The functionality provided for modelling faces gives a considerable amount of control for personification of the subtle variations that occur in facial form. Plate 5.1 illustrates examples of faces that have been generated using FACES.

Although the recognition and generation of faces appear to be different sides of the same coin, we are adept at one, but not particularly skilled at the other. To assist the user in the interactive modelling of arbitrary faces, it is necessary to provide facilities that will allow a wide variety of faces to be represented. However, despite years of research there appears to exist little definitive data regarding the characteristic determinants of facial form.

Knowledge gained from several areas of study has been of importance in the development of the CONSTRUCT and MODIFY sub-systems. Psychologists concerned with the process of face recognition have identified that we remember faces by comparing them with an 'average' face. Research has also revealed the major facial features that are used to distinguish faces from one another. Work undertaken in criminology has demonstrated that a major drawback of systems such as Identikit and Photofit is that features are chosen and put together out of context. From techniques used in facial reconstruction we learned the importance of the skull and bones in determining the appearance of the face. Artists and sculptors have unveiled the significance of variation in relative proportions of the face.

The information gleaned from this range of disciplines has not only influenced the facilities provided in the CONSTRUCT and MODIFY sub-systems, but also the structure of the sub-systems themselves. It has been found that extensive and often subtle variability of the form of the human face makes it necessary to have a range of modelling controls at global, regional and local levels.

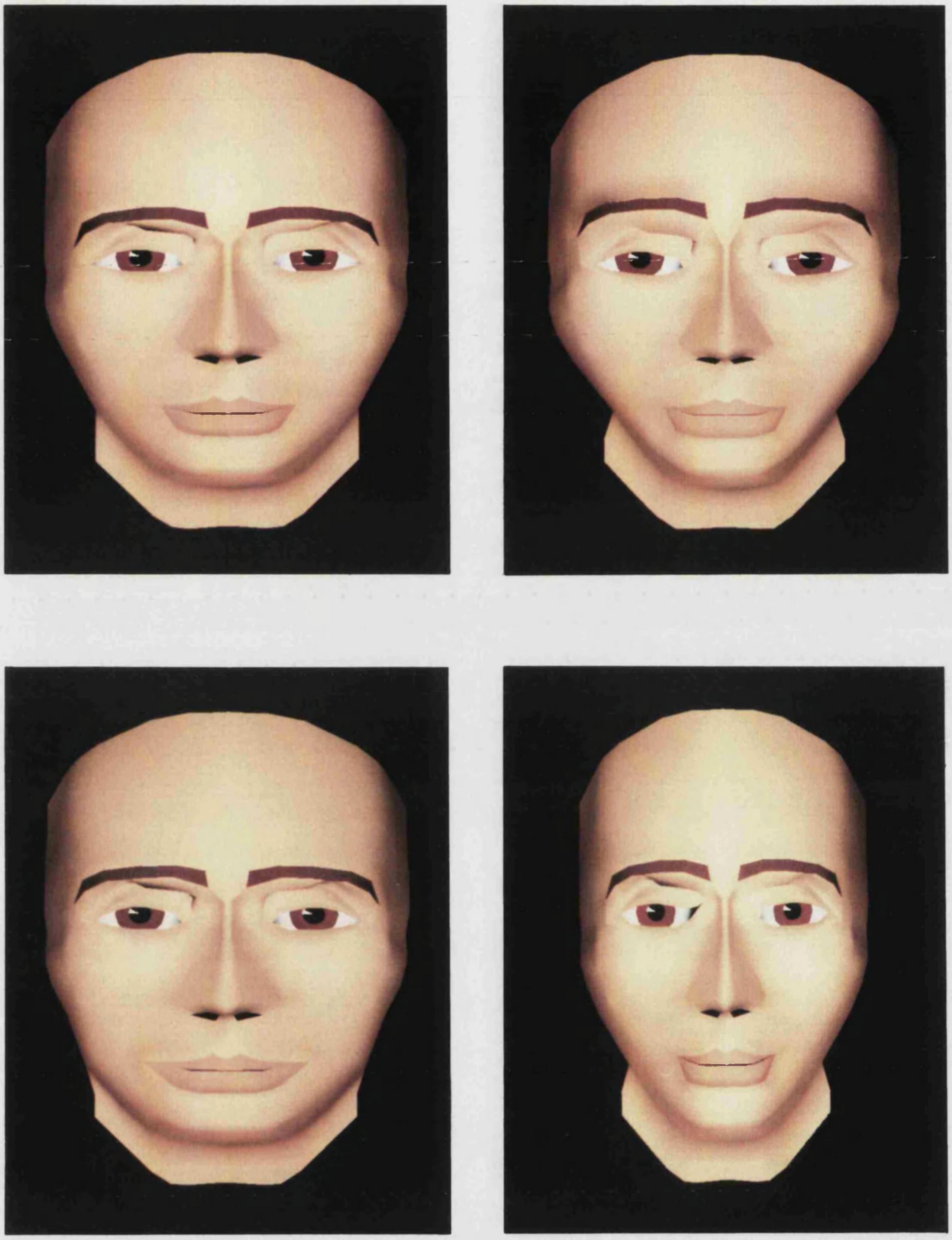

Plate 5.1: Modelling a Variety of Faces

Global, regional and local modelling control within the CONSTRUCT and MODIFY subsystems enables the creation of many different faces. The top-left image is that of the 'neutral' face, without any modifications. The bottom two images demonstrate characteristics typical of male and female faces.

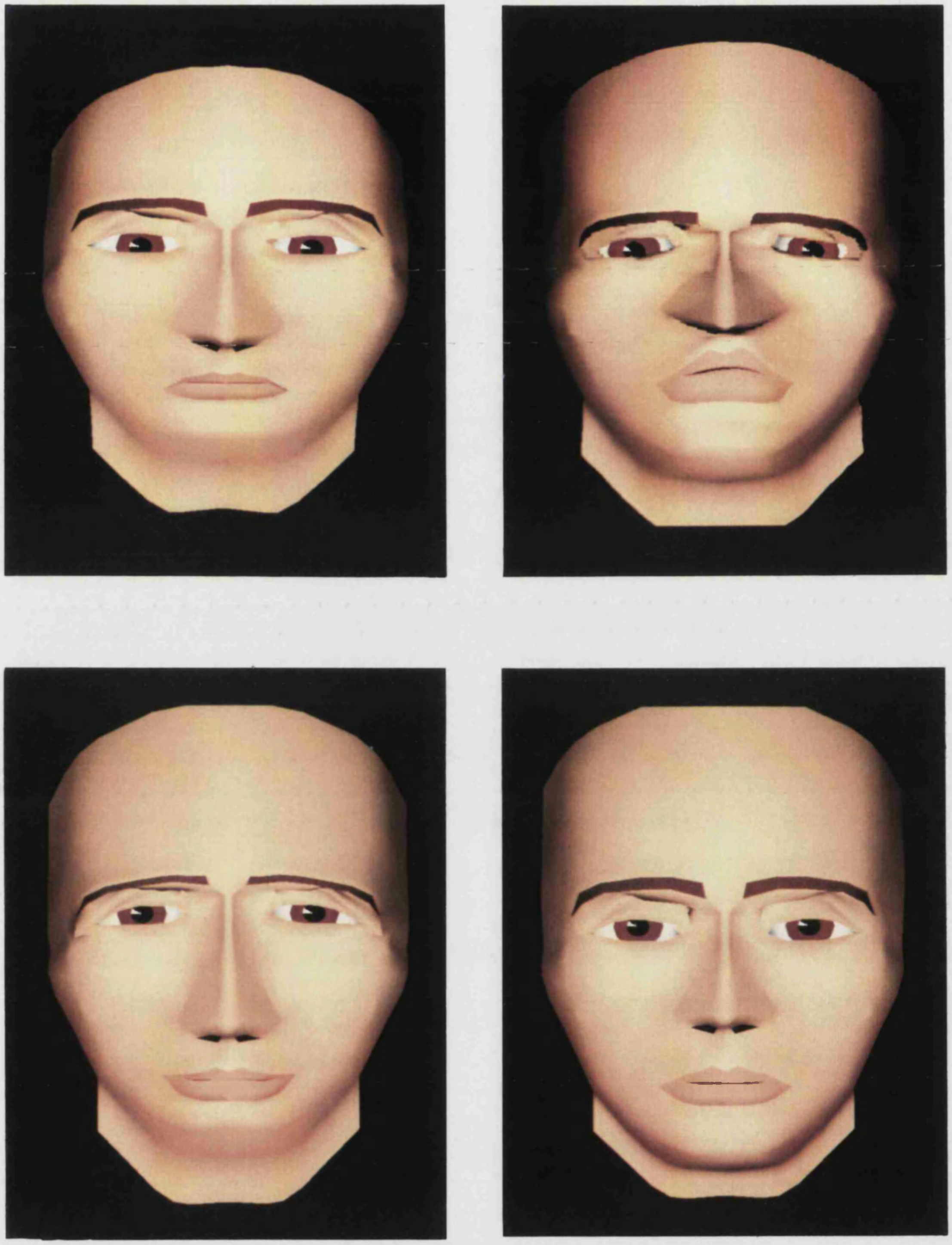

Plate 5.1: Modelling a Variety of Faces (Continued)

Within the CONSTRUCT sub-system the user works with the bone structure. Alterations to muscle thickness and surface features can be made through the MODIFY sub-system. Adequate control exists for the creation of many subtle effects.

An investigation was also undertaken to identify techniques that are commonly used for the representation of irregular surfaces and in particular the face, in the general area of computer modelling. The benefits and the problems associated with digitisation, interpolation, local deformation and composition methods have had a bearing on the techniques used to implement facilities for conformation modelling of the head and face in FACES.

Chapter 6

Animating Faces

6.1 Introduction 91
6.2 Animation in FACES 92
6.3 The ANIMATE Sub-System 94
6.4 Creation of Animation Sequences 94
6.5 Conformation Animation 96
6.6 Facial Animation 97
6.6.1 Expression and Gesture in Emotion 97
6.6.2 Components of Facial Movement 98
6.6.3 Expression Animation 100
6.6.4 Auxiliary Movement 103
6.6.5 Generic Expressions 107
6.6.6 New Expressions 107
6.6.7 Bodily Requirements 111
6.7 Storage and Retrieval of Sequences 111
6.8 Generation of Frames 113
6.9 Summary and Conclusions 118

6.1 Introduction

In the preceding chapter we considered the modelling facilities provided for the user to enable creation of static head models. Next we turn our attention to facilities for making such heads move, an issue which is addressed within the ANIMATE sub-system.

Motion control remains a central issue in the general area of computer generated

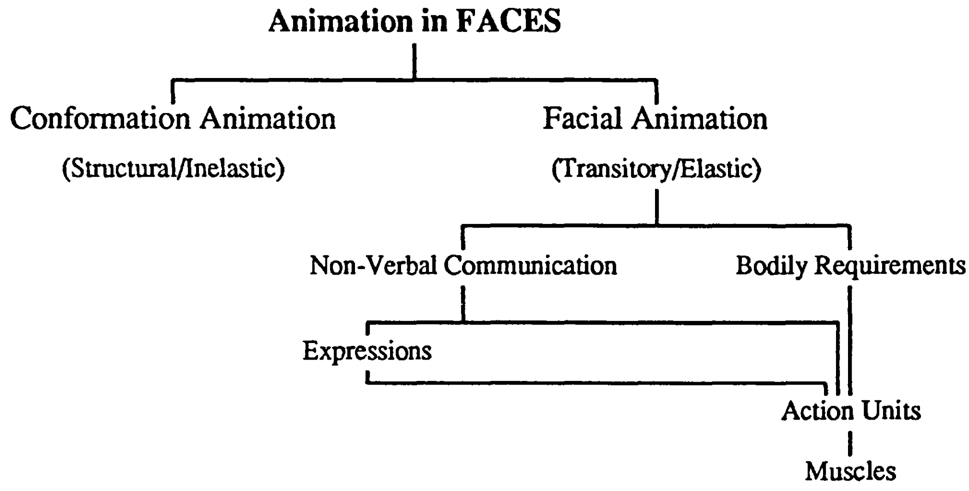

Figure 6.1: Animation in FACES

Two types of animation sequence may be generated. Conformation animation gives the effect of a metamorphosis between faces. Facial animation is concerned with transitory movements of the face, that is temporary changes that reflect expressions.

animation [180, 182, 185, 197]. However, control over movement is even more important in facial expression animation since the timing and motion of facial changes are critical to the information that is eventually conveyed to the viewer.

The rationale behind the approach adopted towards the animation of faces and the motivation for the functionality in the ANIMATE sub-system are first presented in section 6.2. These are followed by a detailed study of the facilities provided in the sub-system for the creation, storage and retrieval of sequences consisting of moving faces.

6.2 Animation in FACES

Within FACES it is possible to generate two types of animation sequence, see Figure 6.1. The first type is known as *conformation* animation. This is useful for creating sequences which represent a metamorphosis of one character into another. Demand for such special effects is demonstrated by the newspaper clip shown in Figure 6.2 [75]. Conformation animation is further discussed in section 6.5.

The second type of animation sequence is referred to as *facial* animation, although this includes head and eye movement as well as transitory changes to the face, see section 6.6.

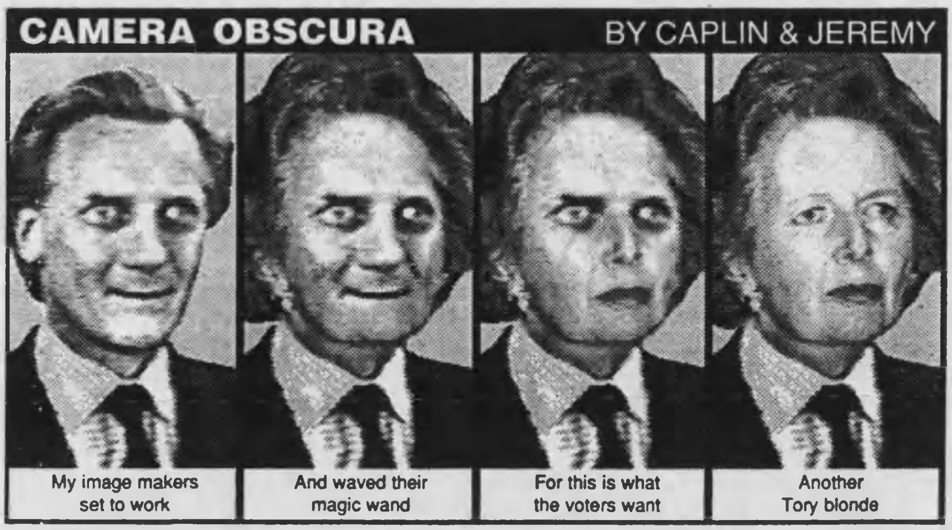

Figure 6.2: Interpolation between two well-known Faces

There is a demand for special effects such as this transformation of Michael Heseltine into Margaret Thatcher. This effect, courtesy of Steve Caplin [75], was achieved through an image processing technique.

Facial animation subdivides into movement pertaining to non-verbal communication and that concerning 'bodily requirements'. Bodily functions such as blinking are necessary simply to make the face appear 'alive' when expressions are not being displayed.

The structural and movement aspects of the face are treated separately. This is to provide generality so that the user may create animation sequences which are equally applicable to different head models.

An animation system further needs to incorporate amenities for the longer term storage of sequences than the duration of one session. It is therefore possible to store sequences in disk files. A textual format with keywords is used in order to make the files easy to both understand and modify. Sequences that have been saved may later be read back into the system as explained in section 6.7.

Creation of an animation sequence comprises a two stage process. A sequence is first defined in terms of the activities that are to take place together with their time-variance. Individual frames that represent the sequence are generated at a later stage as described in section 6.8. Animation sequences therefore have two manifestations, a definition and a series of frames generated from that definition. Generation of individual frames is performed through the process of interpolation which was discussed in section 4.3. A purely kinematic technique, that of shape interpolation, is used for conformation sequences, while a hybrid approach involving both kinematics and dynamics, in the form of parameter interpolation for muscular contraction, is employed for facial animation.

Within the ANIMATE sub-system emphasis is placed on the creation of facial animation sequences. Use of individual muscles as a control interface to the user is regarded as being unacceptably low-level since such an interface would require the user to be familiar with the anatomy of the face. As an alternative strategy, two layers of control are provided in the form of basic facial actions and the higher level of expressions. Consequently, there is a three-layer representation which underlies the control of facial movement; this consists of muscles, AUs and expressions.

6.3 The ANIMATE Sub-System

The overall structure of the ANIMATE sub-system is shown in Figure 6.3, while its context within FACES is illustrated in Appendix E. In order to navigate the reader through the hierarchical organisation of the ANIMATE sub-system, menus are used where appropriate to clarify the options available at particular stages.

At the top-most level, the sub-system has been divided into three major sections which facilitate the creation of animation sequences, their storage and retrieval, and the generation of frames which correspond to the definition of an animation sequence. The three major options are shown in the following menu:

MOTION SPECIFICATION
SCRIPTS
PENCIL-TEST

Sequences are first defined using the facilities in MOTION SPECIFICATION and later generated through the PENCIL-TEST option. Animation sequences can be stored to and retrieved from disk files using the SCRIPTS option. The data corresponding to a sequence is saved in the form of a text file known as a *script*.

6.4 Creation of Animation Sequences

MOTION SPECIFICATION facilitates the creation of both conformation and facial animation sequences. Several further options are available within this major option, these are listed

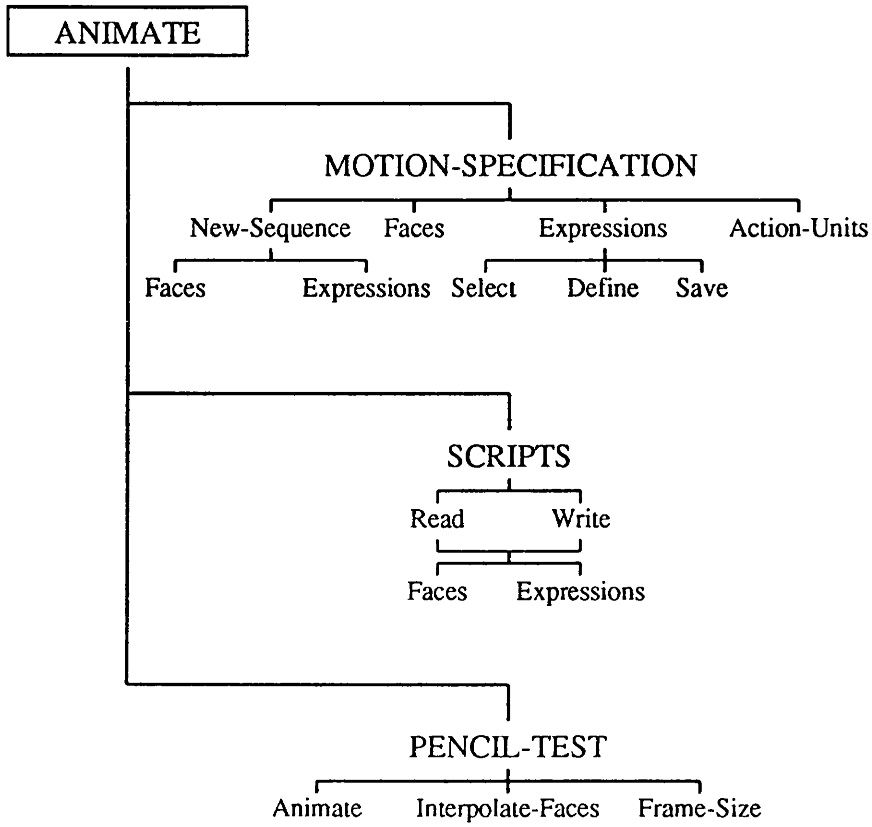

Figure 6.3: The ANIMATE Sub-System

The structure and organisation of the ANIMATE sub-system in FACES. Three major sections cater for the specification, storage and retrieval, as well as generation of animation sequences.

in the menu below:

Exit
New-Sequence
Faces
Expressions
Action Units

The *New-Sequence* option initialises the system for the specification of a completely new animation sequence. This facility is of practical use whenever the user needs to start afresh. A further menu enables the user to decide whether it is the 'current' conformation or facial animation sequence which is to be replaced. Ordinarily, sequences are persistent and continue to exist until FACES is exited, or *New-Sequence* is chosen, or a script is read in.

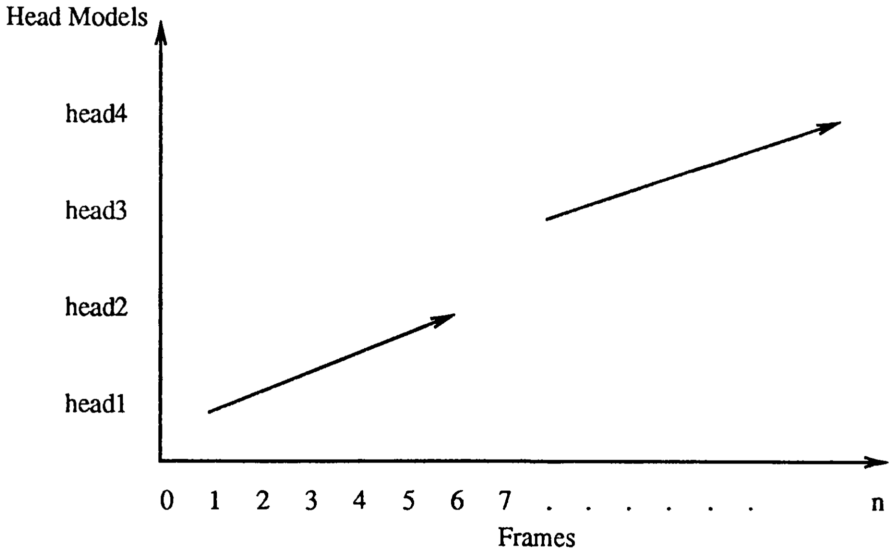

Figure 6.4: Specification of a Conformation Animation Sequence

Specification of data for a metamorphosis sequence involving a transformation of head1 into head2, followed by a transformation of head3 into head4. The user determines the duration and the law of interpolation for each metamorphosis.

Invokation of the *Faces* option enables conformation animation sequences to be defined, while the *Expressions* and *Action Units* options cater for the specification of facial animation sequences.

6.5 Conformation Animation

Selection of the *Faces* option in MOTION SPECIFICATION allows the user to develop a conformation animation sequence. This defines a series of physical transformations of one head model into another and corresponds to shape interpolation between the structure of head models.

For each metamorphosis in the sequence, the user is prompted for three pieces of information, see Figure 6.4. Firstly, the names of the data-files containing the start and end head models are required. For either the start or end head it is possible to specify 'current' which is a generic term to represent the head being used in FACES at that particular time. Secondly, it is necessary to specify the frames over which a metamorphosis is to take place; this is indicated in the form of initial and final frame numbers. The last piece of information required concerns the type of technique to be used during the shape interpolation process which is required to generate intermediate head structures. Here, the user may choose from a selection of four methods based on the laws of motion which are described in section 4.7.1.

6.6 Facial Animation

Facial movement can be associated with both non-verbal communication and bodily requirements. Within FACES particular emphasis has been placed on facial movement for expression animation, however as explained in section 6.6.1, this is only one component, albeit an important one, of non-verbal communication.

An encouraging approach to resolving the problem of the complex nature of facial movement is to develop a modular or 'kit-of-parts' technique for creating facial movement. This is discussed in section 6.6.2.

The task of creating convincing facial movement involves several aspects including the generation of meaningful expressions and plausible movement. Facilities for achieving these goals are described in section 6.6.3.

Additional movements involving the head and eyes are discussed in section 6.6.4. A higher level of control over facial movement is elaborated in sections 6.6.5 and 6.6.6 which are concerned with predefined expressions. Finally, section 6.6.7 views AUs in terms of bodily functions aside from expressions.

6.6.1 Expression and Gesture in Emotion

Facial expressions together with body gestures provide non-verbal communication to convey important information regarding a person's emotional state. Research is currently in progress to understand the connections between facial expressions, speech, hand and bodily gestures [68, 123, 125]. Ekman has identified two broad categories of expression which are known as *emblematic* and *conversational punctuators* [41]. Emblematic expressions have a verbal equivalent in common words or phrases, for example a cheeky wink. Conversational punctuators are facial signals which emphasise whatever is being said in words, such as nodding the head whilst saying 'yes'.

There is however more to emotional experience than that reflected by outward appearance. Emotional feelings also influence additional aspects including a person's: physiology, autonomic nervous system, brain reactions, verbal responses and memory [142, 143].

Emotions are highly personal and complex experiences. People often react differently to particular situations. The study of human behaviour is the domain of psychologists, we therefore limit our attention to providing the user with adequate tools for the creation of facial animation and view the task of injecting life and emotion into the character as one that can only be performed by the skilled animator through judicious use of these tools.

Although FACES provides a mechanism for generating facial distortion, it should be noted that there is still considerable debate regarding what constitutes a 'real' expression as opposed to a 'feigned' one. It has been found that spontaneous expressions are symmetrical while deliberate expressions are stronger on the left side of the face [40]. In the 19th Century a physiologist, named Duchenne, tried to derive information about individual facial muscles through the application of electrical currents to both guillotined and live heads. Although the approach was arbitrary in that it was difficult to isolate particular muscles, Duchenne nevertheless catalogued his results through observation of muscle movements and their associated facial distortions. These findings identified the difference between a 'true' smile and an artificial smile generated by systematically contracting individual muscles. Further work in this area has enabled Ekman to distinguish this 'Duchenne smile' from other types of smile [42].

6.6.2 Components of Facial Movement

As explained in chapter 4, to generate realistic facial movement, the motion characteristics of the face are based on the dynamics of muscle movement. To create animation sequences using individual muscles of the face would be an extremely cumbersome task. FACS is therefore used as the control interface between the user and the models that simulate muscular contractions.

FACS is a comprehensive system which defines fundamental basic actions known as *Action Units* (AUs). Each AU describes the contraction of one facial muscle or a group of related muscles. FACS defines 58 AUs which are presented in Appendix B. This repertoire of basic actions can be regarded as a 'kit' both for the creation of composite facial expressions and for the selection of actions to make the face appear to be alive.

To assist the user in the creation of facial animation sequences, 29 fundamental facial actions have been defined in FACES. They are based on the AUs of FACS. These primitive actions are presented in Table 6.1. Note that pupil-dilation is a supplementary action to those defined in the FACS manual [44]. Also, the facial actions implemented in FACES on

Facial Action	Action Unit	Parameter Range
Neutral-Face	—	—
Inner-Brow-Raiser	AU-1	$0.0 \leq Intensity \leq 1.0$
Outer-Brow-Raiser	AU-2	$0.0 \leq Intensity \leq 1.0$
Brow-Lowerer	AU-4	$0.0 \leq Intensity \leq 1.0$
Upper-Lid-Raiser	AU-5	$0.0 \leq Intensity \leq 1.0$
Cheek-Raiser	AU-6	$0.0 \leq Intensity \leq 1.0$
Lid-Tightener	AU-7	$0.0 \leq Intensity \leq 1.0$
Lips-Towards-Each-Other	AU-8, AU-25	$0.0 \leq Intensity \leq 1.0$
Nose-Wrinkler	AU-9	$0.0 \leq Intensity \leq 1.0$
Upper-Lip-Raiser	AU-10	$0.0 \leq Intensity \leq 1.0$
Lip-Corner-Puller	AU-12	$0.0 \leq Intensity \leq 1.0$
Lip-Corner-Depressor	AU-15	$0.0 \leq Intensity \leq 1.0$
Lower-Lip-Depressor	AU-16	$0.0 \leq Intensity \leq 1.0$
Chin-Raiser	AU-17	$0.0 \leq Intensity \leq 1.0$
Lip-Stretcherer	AU-20	$0.0 \leq Intensity \leq 1.0$
Lip-Pressor	AU-24	$0.0 \leq Intensity \leq 1.0$
Lips-Part	AU-25	$0.0 \leq Intensity \leq 1.0$
Jaw-Drop	AU-26	$0.0 \leq Intensity \leq 1.0$
Cheek-Puff	AU-34	$0.0 \leq Intensity \leq 1.0$
Cheek-Suck	AU-35	$0.0 \leq Intensity \leq 1.0$
Eyes-Closed	AU-43	$0.0 \leq Intensity \leq 1.0$
Turn-Left	AU-61	$0.0 \leq Intensity \leq 1.0$
Turn-Right	AU-62	$0.0 \leq Intensity \leq 1.0$
Head-Up	AU-53	$0.0 \leq Amount \leq 1.0$
Head-Down	AU-54	$0.0 \leq Amount \leq 1.0$
Tilt-Left	AU-55	$0.0 \leq Amount \leq 1.0$
Tilt-Right	AU-56	$0.0 \leq Amount \leq 1.0$
Eyes-Track	AU-61 to AU-66	$-1.0 \leq x, y \leq 1.0$, $z \geq 0.0$
Pupil-Dilation	—	$0.0 \leq Amount \leq 1.0$

Table 6.1: Facial Actions available in FACES

Listed above are 29 fundamental facial actions which have been defined within FACES. With the exception of pupil-dilation, all the actions are based on those defined in the FACS manual [44].

occasion correspond to several FACS AUs; these are indicated where appropriate. Despite these differences, in the interests of clarity the low-level facial actions are referred to as AUs within FACES.

The AUs are implemented as calls to the lower level muscle models which are described in section 4.5.2. Such a scheme makes the muscular basis of the face transparent to the user. AUs are particularly suitable as a means of choosing which parts of the face to incorporate into a sequence. They provide a higher level of abstraction than muscles, so that the user does not need to understand the anatomy of the face before using FACES.

As illustrated in Table 6.1, AUs also have descriptive names, which the user can relate to. An important implication of the fact that AUs are generic and therefore applicable to all human faces, is that the user can create facial animation sequences which can be applied to any FACES head model. Distortions caused by AUs are naturalistic and believable since FACS was derived from a detailed analysis of real faces. This also means that AUs provide viable and realistic constraints on facial movement.

6.6.3 Expression Animation

There are two aspects to the generation of convincing, expressive sequences involving facial animation. First, the expressions need to be meaningful and second, motion should be plausible.

Meaningful Expressions

It is generally accepted that the eyes, eyebrows, eyelids and mouth are the most expressive areas of the face [118, 162, 174]. However, no particular single element is paramount in the perception of an expression, rather the whole is greater than the sum of the parts. In FACES, the generation of expressions involves selection of particular AUs which cumulatively distort the face to represent a meaningful message. Various combinations of the complete set of AUs are capable of representing more than 7000 expressions [58].

FACS divides the face into three regions, notably the upper, lower and middle face. AUs which operate in the upper area are concerned with the forehead and eyebrows. There are several visibly distinct actions involving the eyebrows [41]. AUs involved in the middle part affect the visible appearance of the face from the eyes to the top of the upper lip. The major actions include: raising the cheeks; lowering the brows; tightening of the eyelids; and contraction of the skin around the nose and the corner of the mouth.

Dominating activities in the lower face arise from actions involving the lips and the jaw. This is exemplified by speech and mastication which involve the majority of muscles in this region.

Plausible Movement

The face is often the mirror of the soul; it frequently flashes messages for which words cannot be found. The speed with which facial movement occurs and the time over which the motion endures provide vital cues to vast amounts of information which is present in facial expression. Facial movement normally involves minor displacements, however the movement tends to occur in a series of rapid successions. Moreover, many actions last only a fraction of a second.

Unfortunately, as explained in section 4.7, data relating to the timing and duration of non-verbal communication is currently unavailable, making it impossible to incorporate such information into the system at present. Consequently, the facilities in the MOTION SPECIFICATION part of the sub-system are aimed at providing the user with detailed control over facial actions in conjunction with the flexibility to experiment in order to achieve the required motion and timing.

Creation of Facial Animation Sequences

To build up a facial animation sequence over several frames, the user selects AUs, from those presented in Table 6.1, through the *Action Units* option, see section 6.4. For each action chosen, it is necessary to specify information relating to duration, the law of motion and any parameters that the AU requires. A list of AUs and their associated parameters is given in Table 6.1. Note that for most AUs, only one parameter is required, that of intensity and amount.

Figure 6.5 summarises the specification of a facial animation sequence in terms of combinations of AUs and their corresponding timing. Duration defines the frames over which an action will take place; the user is prompted for the start and end frames. Intensity provides control over the degree of movement in the form of a parameter which ranges between zero and one. A value of zero indicates that the AU is to remain inactive, while a value of one is equivalent to maximum activation. The user specifies the initial and final intensity values for the start and end frames respectively. Finally, four laws of motion enable specification of differing paths of motion for each AU selected; a choice may be

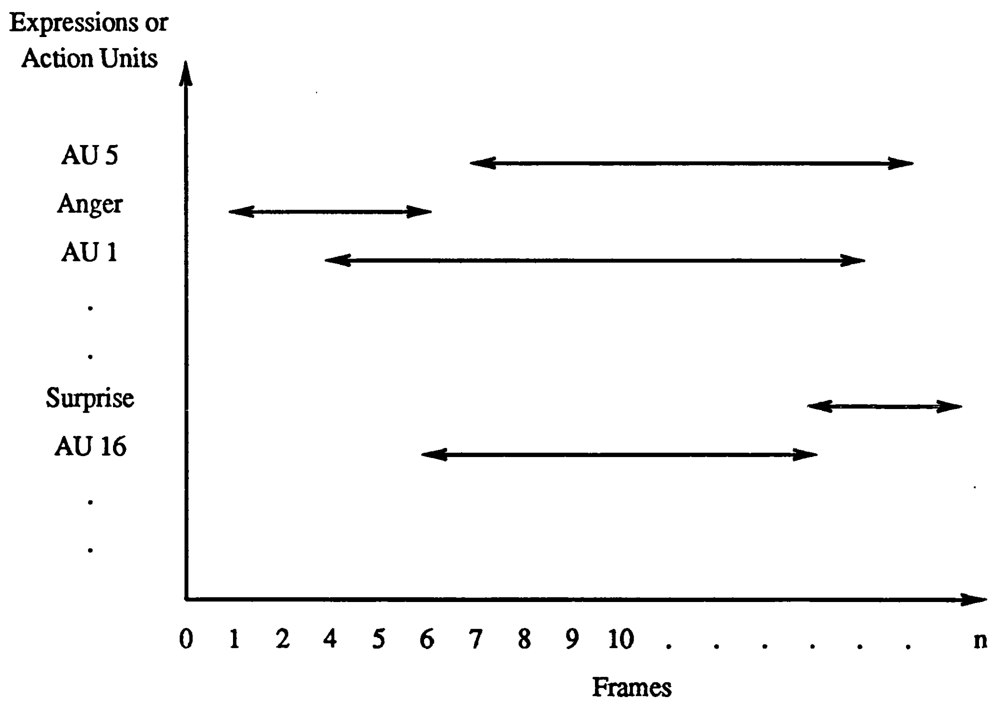

Figure 6.5: Specification of a Facial Animation Sequence

Specification of a series of AUs and predefined expressions to define a facial animation sequence. It is possible for actions to overlap in time, so that they may be active over the same frames, allowing parallel and synchronised motion to be achieved.

made from one of the following:

linear
accelerate
decelerate
accelerate-decelerate

A detailed description of the laws of motion and their effects are presented in section 4.7.1. The laws enable motion effects such as 'ease-in and ease-out', or 'fairing', to be simulated mathematically and correspond to four different techniques for interpolation between the start and end values of AU dependent parameters.

A choice of motion technique means that there is considerable scope for experimentation in order to achieve the correct timing. In addition, some AUs are capable of operating on one side of the head independently of the other side. For such AUs the user needs to specify one of *Left, Right* or *Both*.

It is also possible for AUs to overlap in time, that is they may be active over the same frames, as demonstrated in Figure 6.5. This allows parallel and synchronised motion to be achieved. If conflicting AUs are active over the same time interval the resulting expression will be a 'blended' one. The scheme adopted gives the user fine control, both over the areas of the face to be involved and over the relative timing and duration of AUs, to facilitate creation of the subtle nuances that make facial movement so expressive.

Our familiarity with the movement of the face raises the possibility that the observer may easily detect unnatural motion. Consequently, to generate convincing motion it is often necessary to meticulously refine a sequence. To aid in this process, AUs do not need to be specified in time or frame order, making it straightforward to insert new actions into an existing sequence. This facility offers an explanation for why sequences are persistent in FACES. It is also the reason for the adoption of a two stage process for the creation of animation sequences. The user may experiment with the definition of a sequence independently from the generation of the frames.

6.6.4 Auxiliary Movement

Eyes have consistently been cited as an important source of expressive information [118, 137, 162, 175]. More specifically, the eyeballs play a significant rôle in maintaining attention during conversation and provide clues as to where a person is 'attending' [127]. The capability to track a moving focal point in 3D space in the form of an AU for eye-tracking is therefore included. It is also possible to model movement such as horizontal and vertical eyeball rolling as illustrated in Plate 6.1. In fact, the user has separate control over the eyeballs to enable each to focus on a different point for special effects such as modelling 'fish-eyes', see Plate 6.2. Control over pupil dilation is also available since this is an important component of expressions such as surprise and fear [43], which are illustrated in Plate 6.5.

Since many AUs offer independent control over each side of the face, it is possible to create asymmetric effects such as winking shown in Plate 6.3. The ability to control the left and right sides of the face offers additional flexibilty to model behavioural quirks. Individuality of a face depends on behavioural characteristics as well as differences in physical attributes.

Both verbal and non-verbal communication involve substantial amounts of head movement or prosodic nodding. Movements such as turning, tilting and nodding cannot be excluded if naturalistic expression animation is required, see Plates 6.4 and 6.5

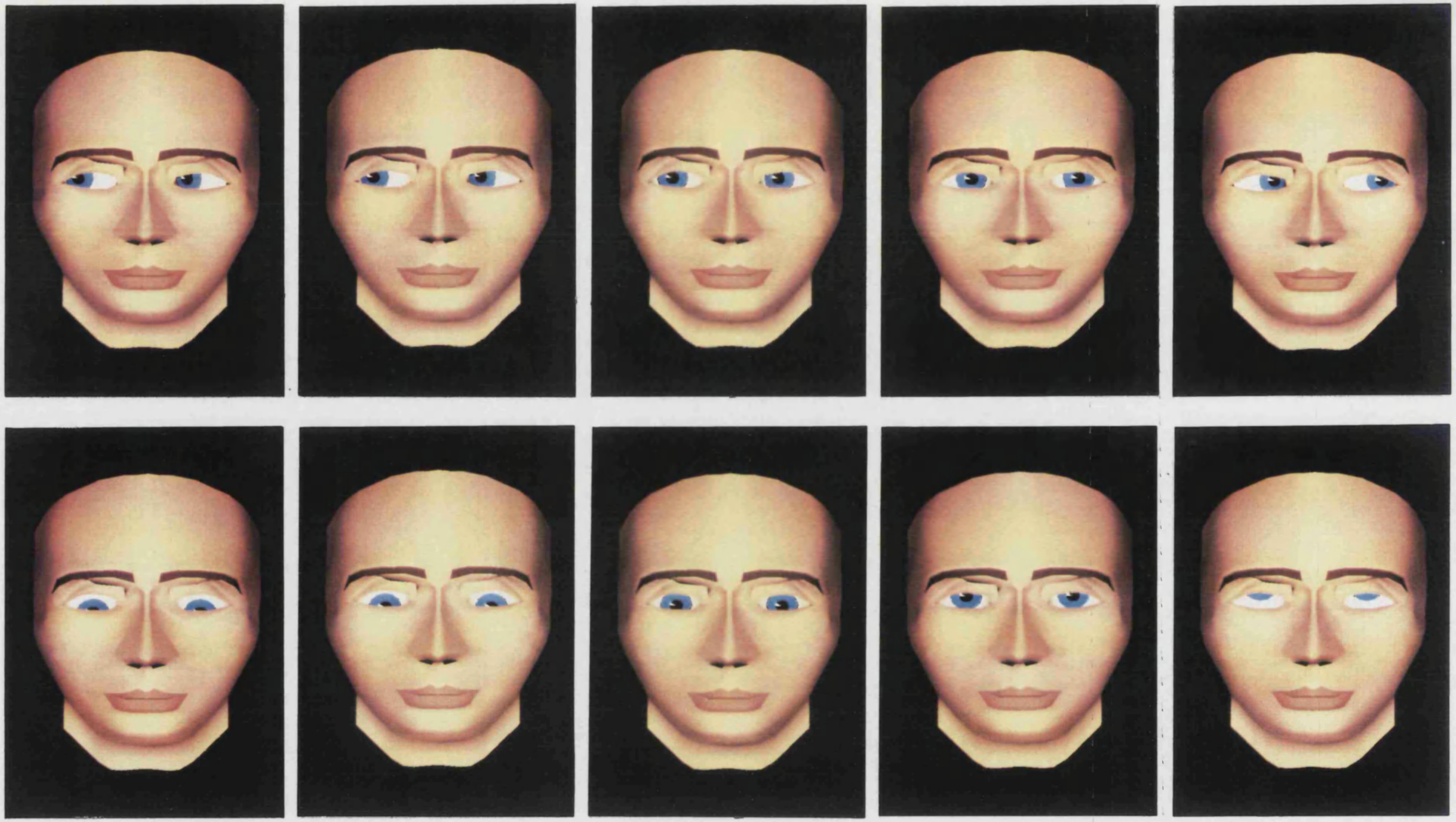

Plate 6.1: Eyeball Rolling

The eyes play a significant rôle in maintaining attention during conversation. Eyes are also frequently used for creating special effects such as those illustrated above. The top sequence of frames demonstrates horizontal eyeball rolling, while the bottom sequence depicts vertical eyeball rolling.

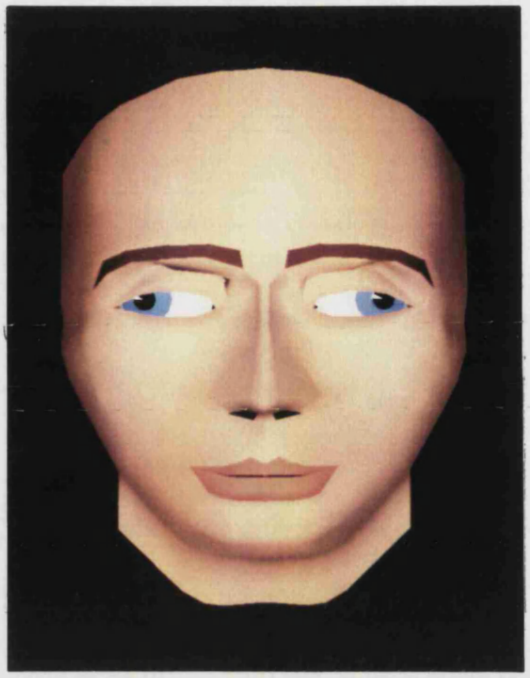

Plate 6.2: Fish-Eyes

Separate control over each individual eyeball enables special effects such as these 'fish-eyes' to be generated. It is possible for each eyeball to track a different target point.

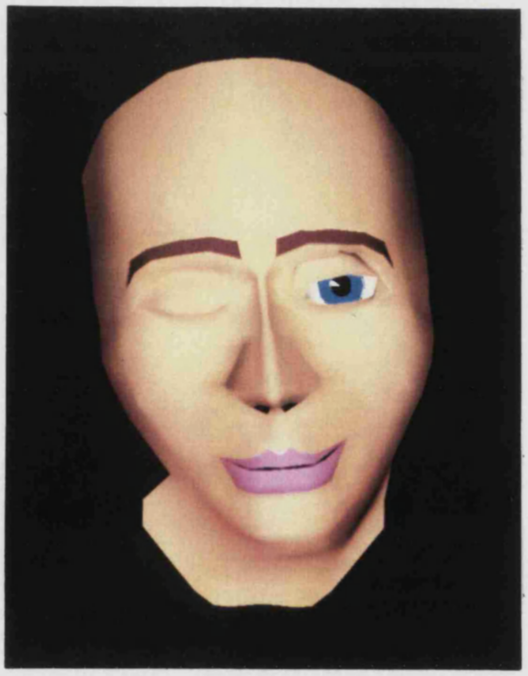

Plate 6.3: Winking

Independent control over each side of the face allows asymmetric effects such as winking to be created. Many AUs are capable of operating on either one side or both sides of the face simultaneously.

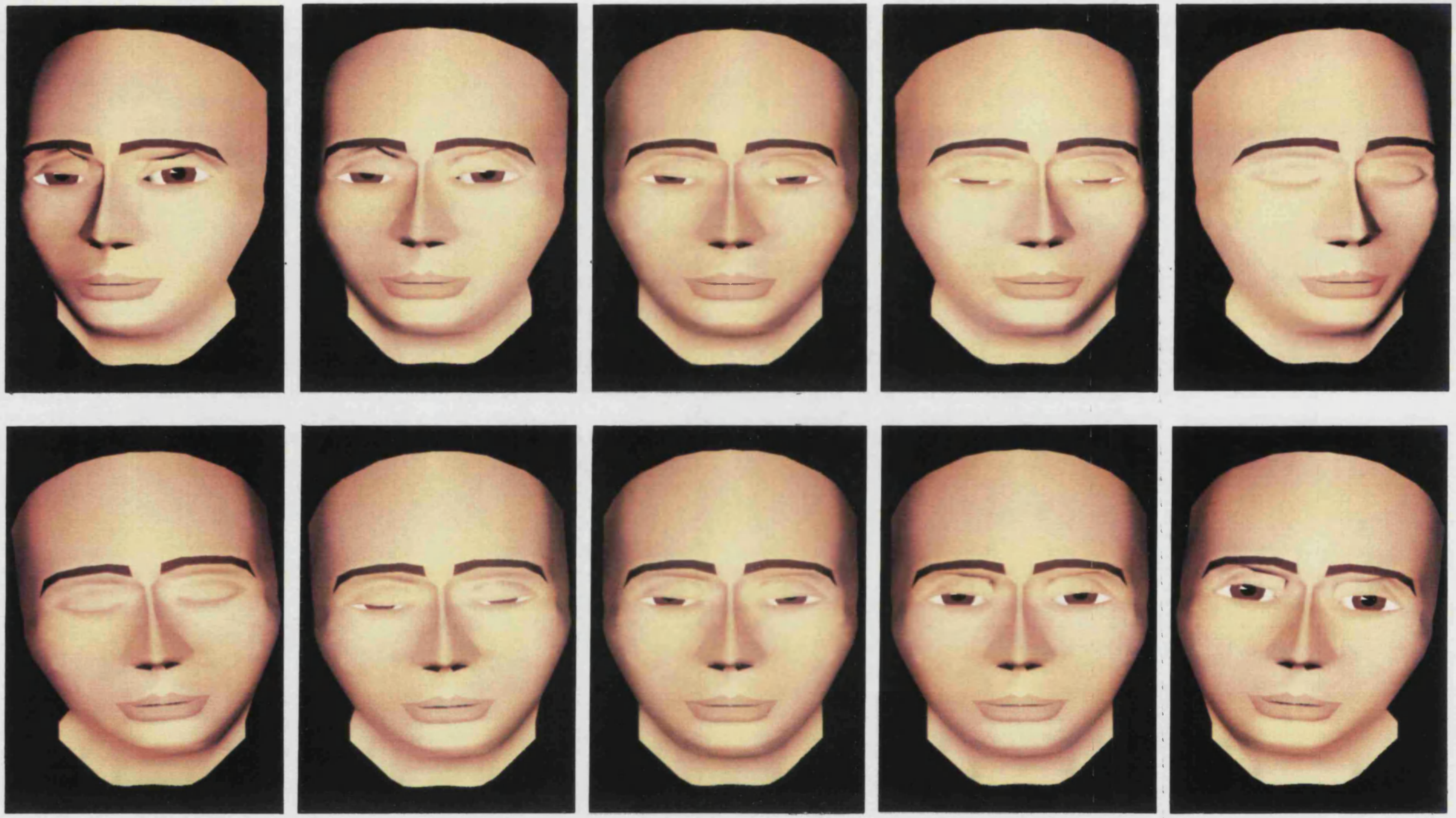

Plate 6.4: Head Movement

Both verbal and non-verbal communication involve substantial amounts of head movement also known as prosodic nodding. The head, in conjunction with the neck, plays a significant rôle in movements such as turning from one side to the other, tilting and nodding. Such movements are necessary for naturalistic facial motion to be achieved.

6.6.5 Generic Expressions

Although AUs facilitate great subtlety in creating facial expressions, they are considered to work at too low a level for some purposes. Certain facial patterns are very common and can be easily identified as evidence of basic emotions. The user is therefore given the option to work at a higher level with predefined expressions. Such expressions can be considered to correspond directly to recognisable and meaningful emotions.

Through extensive research, psychologists Ekman, Friesen and Ellsworth categorised six fundamental facial expressions which are commonly understood by all humanity [58]. The expressions are called *happiness, sadness, surprise, anger, disgust* and *fear*. They are provided in FACES as part of a predefined expressions library. Ekman and Friesen have outlined the essential characteristics of the face in showing the six generic expressions [43]. Figure 6.6 and Plate 6.5 illustrate the typical distortions employed to represent these expressions in FACES.

Predefined expressions are encoded in terms of AUs and their respective parameter values. For instance, anger comprises AU-4, AU-5 and AU-24, while sadness is composed of AU-1, AU-4 and AU-15. The consistency and generality of FACS makes it possible to apply such predefined expressions to any head model.

Access to high level expressions is provided through the *Select Expression* option, which in turn is accessed through the *Expressions* option under MOTION SPECIFICATION, see section 6.4. The user is provided with control over each expression in terms of duration, intensity and law of motion, in a similar manner to control over AUs, see section 6.6.3. Note that selections from the two sets of AUs and expressions may be freely intermixed, as demonstrated in Figure 6.5. Thus it is possible to choose a high level expression and amend it using lower level AUs.

6.6.6 New Expressions

In addition to having access to the predefined generic expressions described in the previous section, the user can construct and save static expressions and thus establish a personal library of predefined high level expressions. More specifically, the user has the capability to combine primitive AU operations and represent the collection as a higher level expression. The *Expression* option allows access to several other options shown in the menu below:

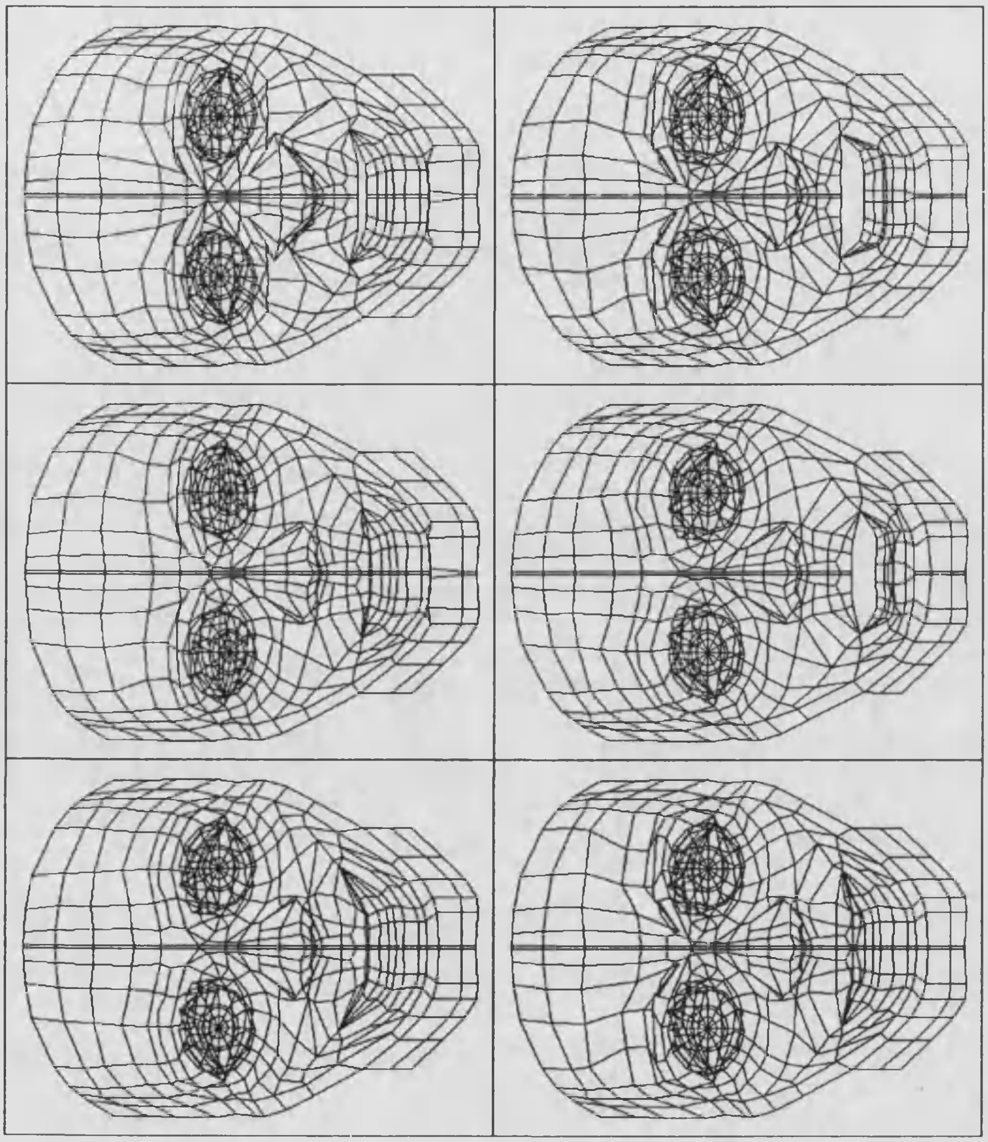

Figure 6.6: Wireframe Representation of Generic Expressions

From top-left to bottom-right, *happiness, sadness, disgust, anger, surprise* and *fear,* are universally recognised. Ekman and Friesen have outlined the essential characteristics of the face in showing the six generic expressions [43]. The expressions are predefined within FACES in terms of Action Units to provide a higher, emotional level of control. This diagram illustrates the typical distortions employed to represent the expressions.

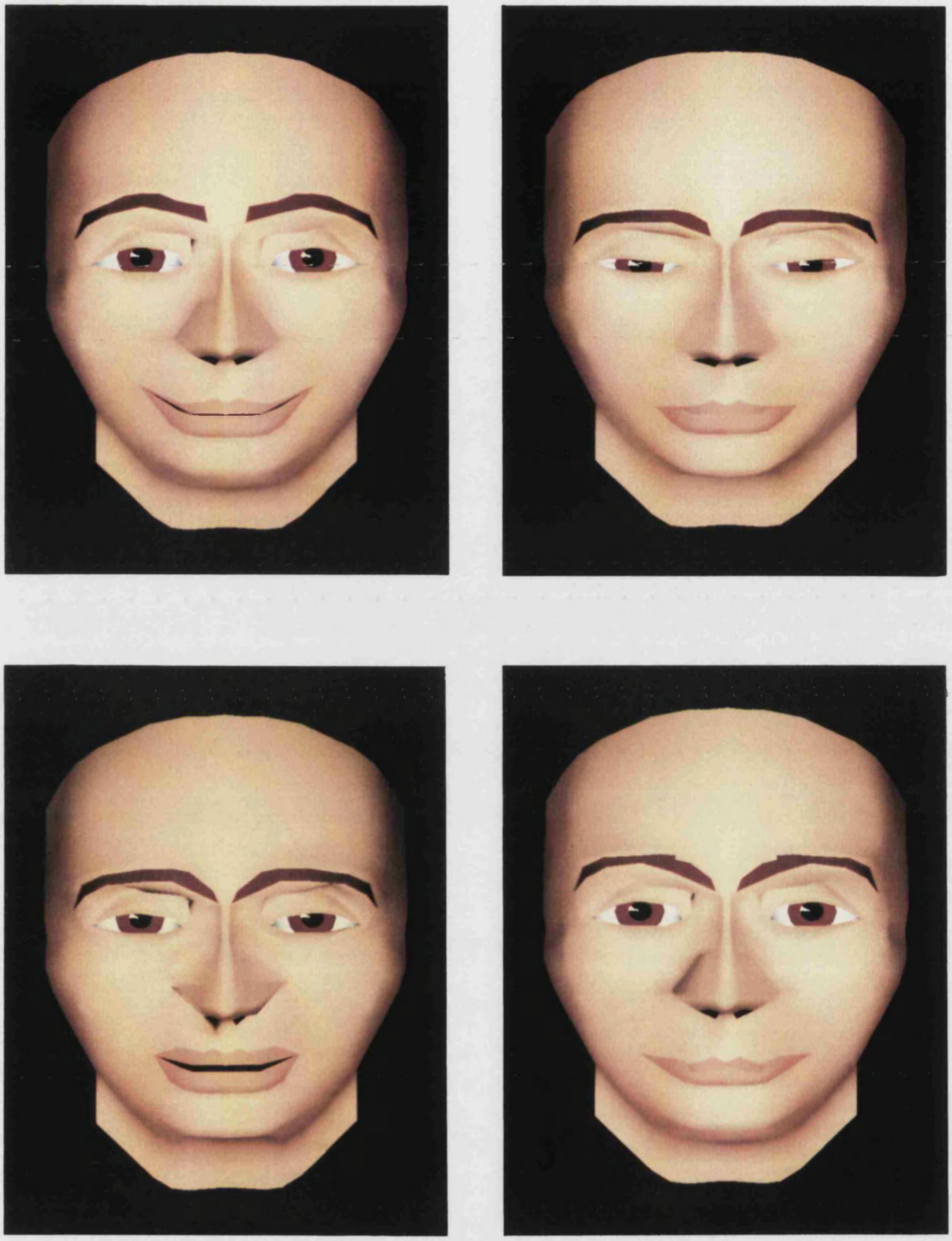

Plate 6.5: Generic Expressions of Emotion

Six expressions are universally recognised. **Happiness**: eyebrows rise (AU1, AU2); eyelids compress (AU6); corners of lips rise and widen (AU12). **Sadness**: inner eyebrows draw together (AU1, AU2, AU4); eyes cast downwards (AU61); corners of mouth pull down (AU15). **Disgust**: upper lip and flanges of nose rise (AU10); lower lip rises (AU17); eyebrows lowered (AU4). **Anger**: eyebrows lowered and drawn together (AU1, AU4); eyelids wide open (AU5); lips pressed firmly against teeth (AU24).

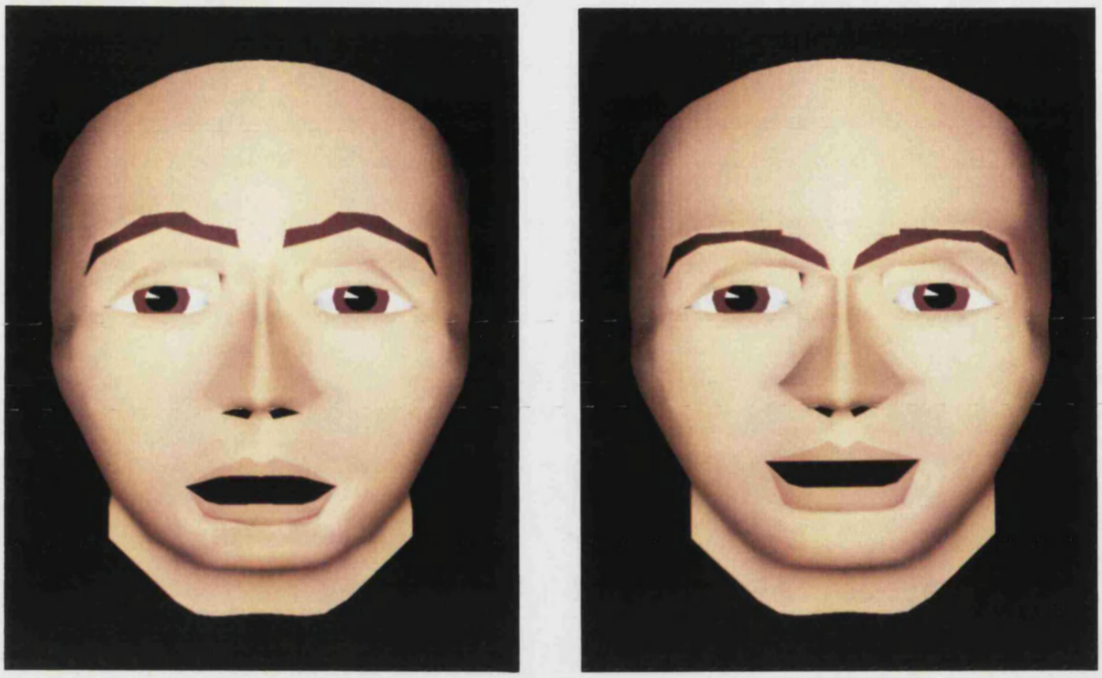

Plate 6.5: Generic Expressions of Emotion (Continued)

Surprise: eyebrows raised high (AU1, AU2); eyelids wide open (AU5); jaw drops (AU26).
Fear: eyebrows rise and draw together (AU1, AU2, AU4); eyelids wide open (AU5); corners of lips drawn backwards (AU20).

Exit
Backup
Select Expression
Define Expression
Save Expression

Define Expression enables the user to interactively 'sculpt' an expression on the face by selecting AUs from the same repertoire as that used in the creation of animation sequences, see Table 6.1. The process of creating a new static expression is illustrated in Figure 6.7.

For each AU selected the system prompts for relevant parameter values. In the majority of cases this involves specification of only one value, the intensity. Since a static expression is being developed, duration and time-variance information is unnecessary.

A stack based undo operator is available to facilitate experimentation during the creation of new expressions. Once complete, the expression can be saved using the *Save*

Expression option, when it must be given a name. As soon as the expression has been saved, it immediately becomes part of a predefined expressions database. From this database the expression can be selected for use in facial animation sequences through the *Select Expression* option as described in the previous section.

6.6.7 Bodily Requirements

Although most of the AUs implemented form significant components of facial expression, there are some actions which also serve the functional requirements of the face. These actions are important elements in maintaining the ebb and flow of facial movement, that is in making the face appear 'alive' even when expressions are not being displayed.

The eyes consist of eyelids and eyeballs, each of which provide different functions. Eye blinking is essential as a means of keeping the eyeball moist to stop the eyes from drying out. The eyelids blink once every 2–10 seconds during which time the eyes remain closed for an average of 0.15 second.

Eyeballs serve the major function of enabling vision. Each eyeball must be capable of rotational movement to allow the pupil to focus on a point in 3D space and thereby achieve and maintain stereoscopic vision. The pupil controls the amount of light which enters the eye, making dilation of the pupils an important aspect of vision.

Movement of the lower jaw is important in mastication and speech as well as in non-verbal communication. The head, in conjunction with the neck plays a significant rôle in movements such as turning from one side to the other, tilting and nodding, see Plate 6.4. Such movement does not necessarily form a part of non-verbal communication.

6.7 Storage and Retrieval of Sequences

Having created an animation sequence it is useful to save it and read it back into FACES at some later stage. The major option SCRIPTS, in the ANIMATE sub-system, enables the user to save and retrieve both conformation and facial animation sequences. SCRIPTS allows access to the following further operations:

Exit
Backup
Read
Write

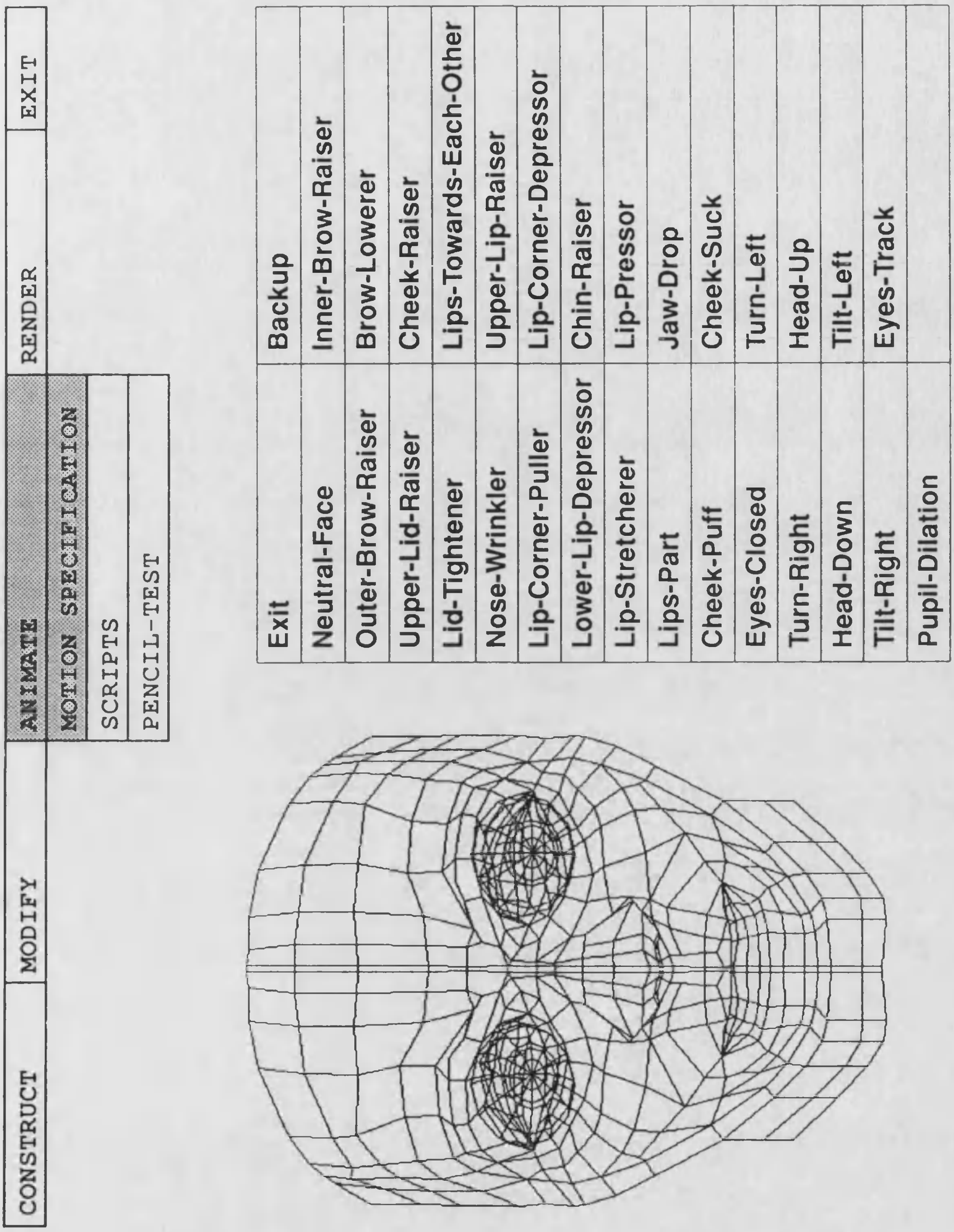

Figure 6.7: Creation of New Expressions

New static expressions may be created through interactive sculpting using the same repertoire of AUs as is available for generating facial animation sequences. Such expressions can be saved to form a library of predefined expressions.

The *Write* option caters for the storage of animation sequences to disk files. Through an additional menu, the user may select which type of sequence, conformation or facial animation, is to be saved. The user is then prompted for a name by which to call the corresponding disk file.

Facial animation scripts are stored in terms of AUs and expressions, which are exactly the entities that the user works with. These scripts contain information regarding: an expression or AU; the frames over which it is active; the start and end parameter values; and the law of motion to be used for interpolation. Facial animation scripts therefore provide a definition of the timing, duration and combination of AUs and expressions that comprise a sequence.

Conformation animation scripts consist of the: names of data-files corresponding to the start and end head models; initial and final frames; and interpolation technique to be used for inbetweening. Conformation animation scripts therefore define a sequence of physical transformations of one head into another. Both types of sequence are stored as textual data making scripts easy to read, understand and modify with the use of a standard text editor.

The *Read* option facilitates retrieval of animation sequences from script files. When a script is read back into FACES, two options become available. If a sequence of the same type already exists in the system, the old sequence may be either overwritten by the new sequence, or the new sequence may be incorporated into the old. This feature allows several scripts to be merged into one.

6.8 Generation of Frames

Once the definition of an animation sequence has been created, it is necessary to generate the individual frames that comprise the sequence. Although real time playback is desirable, the state of accessible technology does not permit 15–25 frames to be rendered every second. A facility is therefore provided for previewing wireframe representations. Even the production of wireframe images cannot be achieved in real time, however it does give the user an indication of what the sequence will look like in its final form.

Generation of frames is facilitated by a menu associated with the PENCIL-TEST option which presents several additional operations:

Exit
Animate
Interpolate Faces
Frame-Size

Invokation of the *Animate* option causes generation of a sequence of frames through the application of a facial animation sequence to the current head model, see Plate 6.6. A hybrid approach which draws on both kinematics and dynamics methods is used. Parameter values at key frames define extreme facial movement while intermediate frames are generated through interpolation of parameter values. At each frame, for each active AU or expression, parameters are interpolated according to the path of motion for that action. Dynamic simulation of muscle contractions causes the actual distortion that represents facial movement.

In the *Human Factory System*, a face, expressions and scripts are considered to be an integral part of a particular synthetic actor [100]. In contrast, in FACES, head models and facial animation sequences are totally independent of each other. As a result, a facial animation sequence can be applied to any head model constructed within FACES.

The *Interpolate Faces* option enables generation of frames corresponding to a conformation animation sequence. This describes a sequence consisting of interpolations between the structure of head models. In this case, intermediate head models are generated through shape interpolation.

From one conformation sequence definition it is possible to generate two different sets of frames, as illustrated in Plates 6.7 and 6.8. When a facial animation sequence has been defined in the system, facial movements are generated on intermediate faces. However, when a facial animation sequence does not exist within the system, interpolation is performed without the application of facial movement so that inbetween faces have a 'neutral' state.

A conformation animation sequence usually consists of references to particular head models that are to be used in the process of metamorphosis. However, whenever 'current' is encountered instead of the name of a data-file, the head model that is in use in FACES at that particular time is substituted.

The *Frame-Size* option allows interactive control over the size of frames to be generated. This facility is of practical use during the development of sequences since the rendition time of a frame increases with its size.

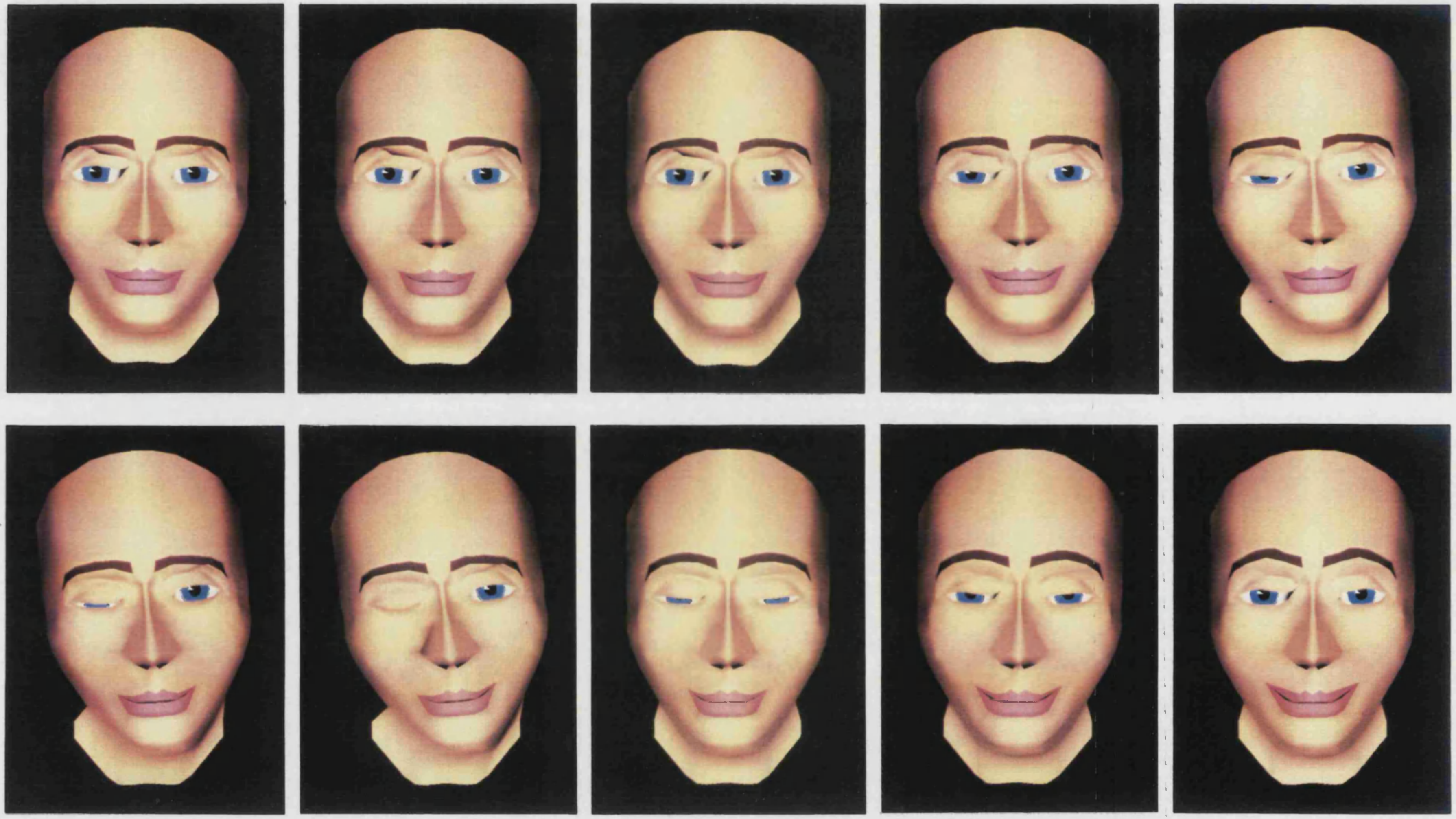

Plate 6.6: Facial Animation

A facial animation sequence showing winking followed by happiness on a face that would typically be perceived as being that of a woman. Generality within FACES allows one facial animation sequence to be applied to many different head models. This is possible because of the generic nature of the Actions Units defined in FACS.

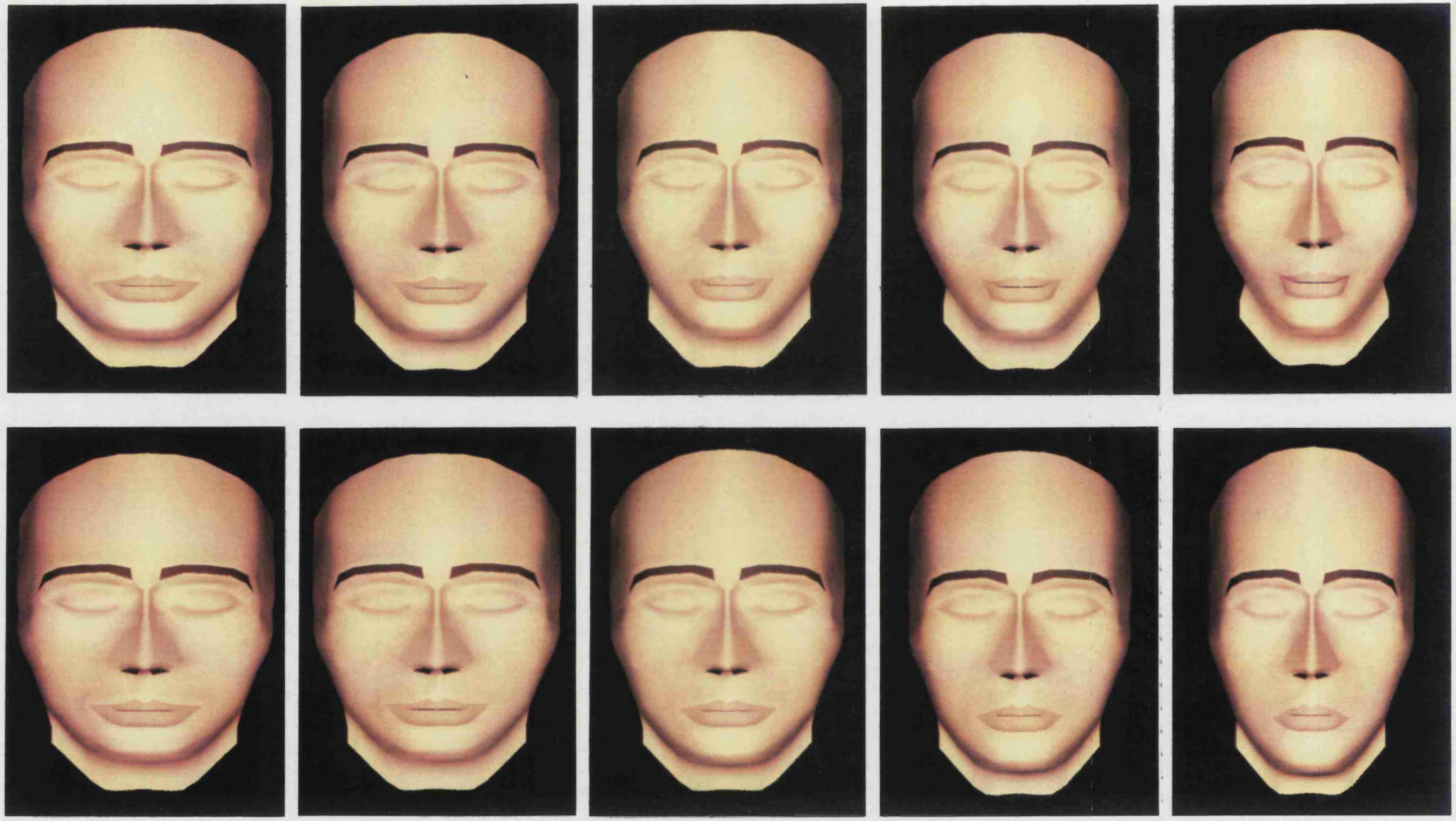

Plate 6.7: Conformation Animation without Facial Movement

Conformation animation enables generation of sequences of metamorphoses between head structures. This type of animation is achieved through *shape interpolation*. The top sequence shows the head of a man turing into a woman, while the bottom sequence demonstrates the same man transforming into another man with a longer face.

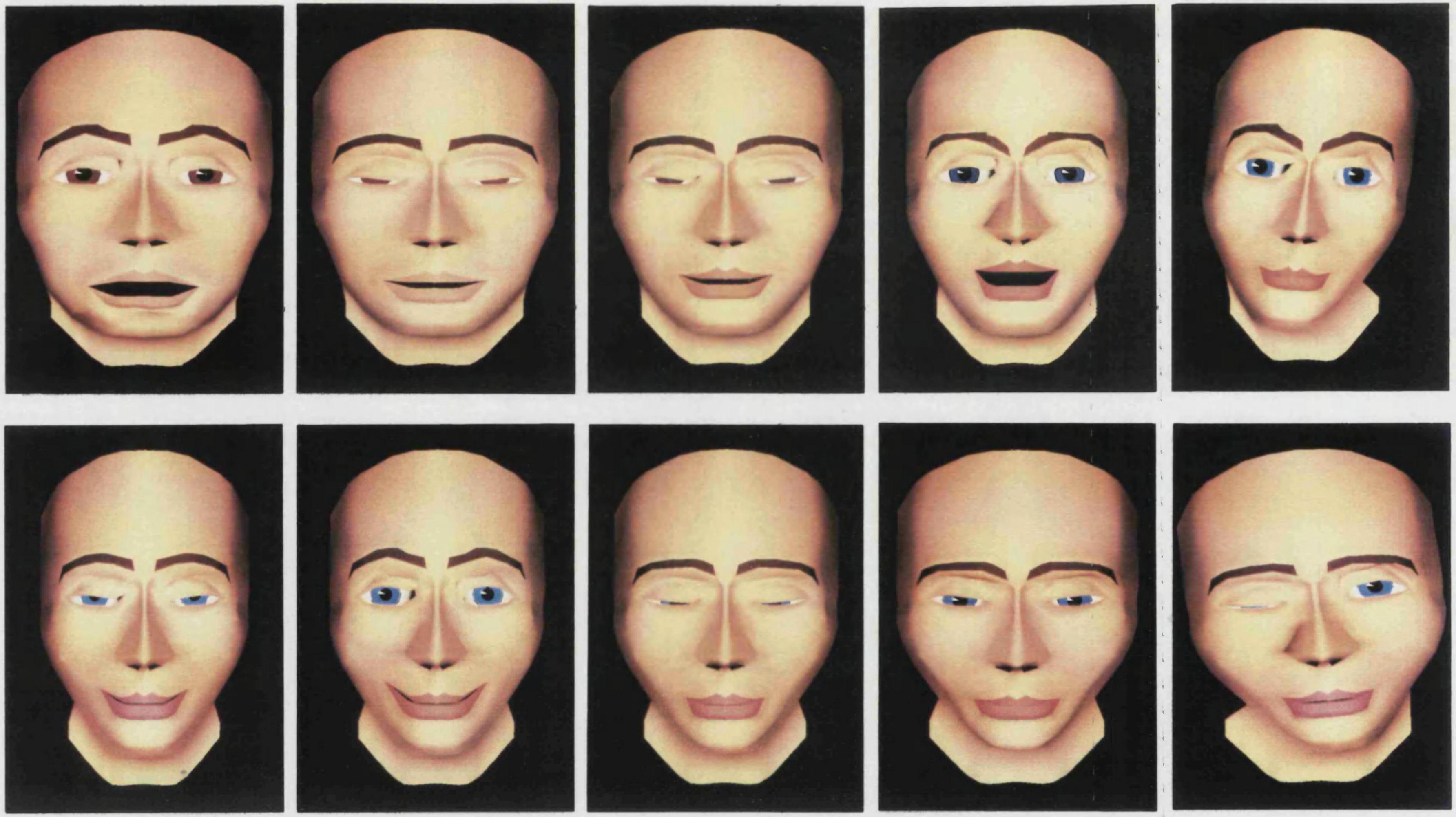

Plate 6.8: Metamorphosis with Facial Animation

Selected frames from two sequences in which conformation and facial animation have been combined. The top sequence is the face of a man transforming into a woman, with facial expressions overlaid. The bottom sequence illustrates metamorphosis of a female face into a different female face, again with facial movement applied to the physical transformation of the head model.

6.9 Summary and Conclusions

The ANIMATE sub-system addresses issues related to the animation of static head models. This sub-system enables two types of animation sequence to be generated. The first represents a sequence of physical transformations or metamorphoses of head models. The second is concerned with facial and head movement and can be further subdivded into movements which are used for non-verbal communication and movements such as blinking for the functional requirements of the face. Both types of sequence are first defined or created using the MOTION SPECIFICATION part of the ANIMATE sub-system. The definition is later used to generate individual frames through the PENCIL-TEST option. Examples of both types of sequence are presented in video sequences, see Appendix D.

In the absence of usable time-variance data for non-verbal communication the user is provided with detailed control over both low level facial actions and higher level expressions, together with the ability to experiment. Control is available over the duration, intensity and path of motion of each facial action. The use of FACS as a control mechanism provides a natural interface which promotes the creation of believable expressions.

Animation sequences can be saved to and retrieved from disk files in the form of human readable scripts. Use of a textual recording technique makes script files readily modifiable by the user. An independent treatment of conformation and facial animation sequences provides convenience and flexibility. More specifically, facial animation sequences can be applied to either a single head model or to a sequence of interpolated heads as a result of the generality in FACES.

Chapter 7

Rendering Faces

7.1 Introduction 119
7.2 Rendering in FACES 120
7.3 The Rendering Pipeline 121
7.3.1 Viewing 121
7.3.2 Visible Surface Determination 121
7.3.3 Scan Conversion 123
7.3.4 Illumination Model 123
7.3.5 Shading Polygonal Meshes 125
7.4 The RENDER Sub-System 126
7.5 Colouration of Features 127
7.5.1 Appearance 128
7.5.2 Emotion 129
7.6 Examination of Frames 129
7.7 Evaluation of Motion 129
7.8 Summary and Conclusions 136

7.1 Introduction

Facilities described so far in the CONSTRUCT, MODIFY and ANIMATE sub-systems all display a wireframe representation of the head model. Once a satisfactory model of the head together with its motion has been developed, it is necessary to generate sequences which appear more realistic in visual terms. Within the RENDER sub-system there are operations which cater for rendering, monitoring and motion evaluation of animation sequences.

Motivation for the facilities provided in the RENDER sub-system are described in section 7.2. The process of generating shaded images comprises several stages known as the *rendering pipeline*, these are discussed in section 7.3. This is followed by an explanation of the organisation of the RENDER sub-system in section 7.4 and a detailed description of the functionality that it has to offer in sections 7.5, 7.6 and 7.7.

7.2 Rendering in FACES

The RENDER sub-system in FACES provides for more than the term 'render' normally implies in computer graphics. As well as facilities for the generation of shaded images, an animation system should also provide a means both for viewing groups of frames and for evaluating the motion of animated sequences. Furthermore, an effective system should allow some control over which particular parts of a sequence are operated upon.

Although we have a 3D representation of the head and face, in order to view the model it is necessary to project it onto a 2D display screen. This process is known as *viewing* and involves mapping from *modelling coordinates*, in which the model is defined, into the coordinate system of the output device.

Viewing is followed by a journey through the *rendering pipeline* to produce a realistic 3D image of the face. This process involves determination of visible parts of the model together with their shading. Both of these processes serve to enhance the visual realism of the computer generated face. Visible surface determination is important for the elimination of ambiguity that wireframe models can present, while shading makes apparent the shape of the face, as well as conveying relative depth. Calculations involved in the rendering process are time-consuming. The impetus for adopting certain algorithms and techniques is that a major requirement of FACES is that of speed; the system needs to be interactive, making rapid visual feedback vital.

Speed is also of the essence in evaluating the motion characteristics of a sequence of frames. Timing is especially important in non-verbal communication where rapidly changing expressions can indicate extremely subtle meanings. Consequently, once a sequence has been generated, it is instructive to test the motion using a variety of speeds.

Colour is an important factor in the determination of the appearance of the face. Interactive control over the colouration of facial features is a major requirement for modelling different faces.

7.3 The Rendering Pipeline

A realistic rendering of the 3D head model requires several stages of processing which include: viewing, visible surface determination, shading and application of an illumination model. The state of accessible technology does not currently permit realistic rendering in real time which requires 30 frames every second for film and 25 frames per second for video. Frames comprising an animated sequence are therefore rendered through an interactive batch process following a wireframe preview using facilities in the ANIMATE sub-system. As each frame is rendered, it is output to a graphics display to provide feedback to the user.

7.3.1 Viewing

Viewing models based on the 'synthetic camera' can be elaborate [50]. Such sophistication is considered unnecessary in FACES. It is adequate to view the head model from the front using a parallel, or *orthographic*, projection [26]. This corresponds to a viewer position at infinity on the z-axis. Note that a right-handed coordinate system is used in FACES.

Coordinates obtained from digistisation of the mask can be regarded as a representation of the skin mesh in either modelling coordinates or *world coordinates.* When read into FACES these coordinates are converted into a logical or normalised coordinate system such that each x, y and z coordinate value lies in the range -0.5 to 0.5. Whenever it is necessary to display parts of the head model, such as the skull or surface skin, the logical coordinates are transformed into the coordinate system of a *canvas* [189], which is a rectangular array of pixel values. Once the model has been rendered into the canvas, it can be displayed at any location on the output device.

7.3.2 Visible Surface Determination

The problem of determining which surfaces are visible from a particular viewpoint is also known as *hidden surface removal.* In general, the process has two components. Firstly, surfaces must be sorted to find those which are closest to the observer. Secondly, an evaluation needs to be performed to discover which parts of surfaces are obscured by others closer to the observer. The calculations involved are time-consuming, therefore many algorithms concentrate on improving efficiency to increase speed.

A classification of hidden line and hidden surface removal algorithms due to Sutherland et al.[157] has been widely adopted [50, 111]. This categorisation is based on whether an algorithm deals with object definitions directly or with their projected images. Three classes of algorithm are identified: *Image Space, Object Space* and *List-Priority.*

Image space algorithms work with projected data, in the coordinate space of the output device and capitalise on the limited resolution of the device. Such algorithms are particularly popular for raster devices and include scan line methods, the z-buffer algorithm and Warnock's algorithm. Object space algorithms work in the coordinate system used to define the objects and involve methods such as the Weiler-Atherton algorithm and octree techniques. List-Priority algorithms tend to perform depth sorting in object space and scan-conversion in image space. Examples include the painter's algorithm and the depth-sort or Newell, Newell and Sancha algorithms.

Execution Efficiency and Memory Requirements

The execution efficiency of visible surface determination algorithms is largely dependent on efficient sorting algorithms. Methods can be distinguished by the order of the dimensions in which they perform the sorting phases. For example, the depth-sort algorithm first sorts on z, then x followed by y. Scan line algorithms first sort on y, then x and finally on z. Area-subdivision techniques such as the Warnock and Weiler-Atherton algorithms do a parallel sort on x and y, and then a search in z. The z-buffer algorithm does no explicit sorting and searches only in z.

Another major consideration in the selection of a visible surface algorithm is the amount of memory it requires. The z-buffer algorithm, for example, has a large appetite for memory. The method needs a pixel-buffer and a z-buffer each of two dimensions to represent the resolution of the display screen.

Hidden Surface Removal in FACES

Advance knowledge of the properties of the images to be generated and the nature of the head model were taken into consideration during the development of the visible surface determination technique used in FACES. The structure of the head model led to the adoption of a hidden surface removal technique based on the list-priority category of algorithms.

Since it is only necessary to display the visible parts of the face, the sorting phase can be predetermined and therefore optimised. More specifically, the eyeballs are rendered first, followed by the facial skin mesh on top of which the individual features such as the eyebrows and lips are displayed. This technique amounts to a variation on the painter's algorithm which involves over-painting of primitives.

A major drawback of the painter's algorithm is that intersecting primitives and cases of *cyclic overlap* are dealt with erroneously. These problems do not arise with the head model in FACES since it comprises polygonal mesh representations, the components of which can be ordered correctly. A prior knowledge of the structure of the head model has therefore circumvented the need to precede the painter's algorithm by a time-consuming sorting phase such as that required in the depth-sort algorithm. Such a sorting process is necessary for more general scenes containing an unknown amount of complexity.

7.3.3 Scan Conversion

The entire head model consists of polygons. The basic process of scan-converting polygons has been well-established for some time [11, 50, 111, 121, 179]. The overall strategy is to create and fill the polygon one scan line at a time, from the top to the bottom. It is necessary to determine which pixels on the scan line are within the polygon in order to set them to the appropriate intensity. However, calculation of the intersections between a scan line and a polygon can be time-consuming since each edge must be tested for intersection with each scan line. *Scan line coherence*, also known as *edge coherence*, is used to speed up the process since many of the edges intersected by scan line i are also intersected by scan line i-1. Coherence is thus used to exploit regularities and converts costly absolute calculations into less expensive incremental ones.

7.3.4 Illumination Model

Once a surface has been established as being visible it is then necessary to apply an *illumination* or *reflection* model. It is the task of this shading model to determine the quantity and quality of light which is reflected to the viewer from a visible point on a surface. This is a function of: the light source; its direction and strength; the viewer's position; and the surface properties of the object.

Light sources which illuminate an object consist of two basic types; *light-emitting* and *light-reflecting.* Examples of light-emitting sources include candles, light bulbs and the Sun. Light-reflecting sources are the illuminated surfaces of other objects, such as the walls of a room, that are near the object being viewed. Many illumination models have been developed based on different types of light sources together with their particular characteristics, such as the distribution of light rays [112, 173].

In most illumination models, the light reflected from a surface is considered to comprise two components: *diffuse* reflection and *specular* reflection. The diffuse component itself consists of two parts: *ambient* light and that light due to specific sources. The ambient part is 'background' light or the overall base level of illumination. Such light is uniformly incident and uniformly reflected.

Diffusely reflected light from specific sources is similar to ambient light, but is not uniformly incident. Light landing at a point on a surface varies in strength according to its direction. This is normally modelled using *Lambert's Cosine Law* [50, 111]. Whereas diffuse reflections are uniformly distributed, specular ones are highly directional. Only those rays reflected towards the observer will be seen. Specular reflection is apparent on shiny surfaces in the form of highlights and can be calculated using *Phong's Model of Specular Reflection* [21].

To simplify and minimise calculations in FACES, only one point light source at infinity is incorporated, so that all light rays are assumed to be parallel and directional. In addition, it is only necessary to consider ambient and diffuse reflection for the surface of the face, specular reflection is required only for highlights on the pupils.

Quest for Visual Realism

Illumination models traditionally used in computer graphics have little formal theoretical basis, but are popular because they produce attractive results with a minimum amount of computation. Other more sophisticated models have been developed in the pursuit of accurately rendered three dimensional images, visual realism being the ultimate goal. The *Torrance and Sparrow Model* was developed theoretically, based on the physics of light reflection [166]. This technique takes into account the concept of light reflected from tiny micro-facets. The *Cook and Torrance Model* accounts for the relative brightness of different materials and light sources in the same scene [33].

Global illumination techniques such as ray-tracing and radiosity account for the interaction of light between all the surfaces in a scene. Ray-tracing is renowned for the simple manner in which it allows effects such as cast shadows, reflection, refraction, transparency and translucency to be incorporated into the basic algorithm [179]. However, the method is also well-known for the exorbitant amount of processing time that it requires. Radiosity eliminates the need for the ambient component by providing a more accurate treatment of inter-object reflections [32, 59, 73]. Luxurious effects such as interobject reflection, shadows and transparency are considered unnecessary for rendering faces and therefore have been omitted.

7.3.5 Shading Polygonal Meshes

Polygonal meshes are primarily used to model irregular and curved surfaces, such as the head model in FACES. The polygon is a basic primitive in computer graphics, it is frequently used for constructing entire scenes. Consequently, techniques for shading objects defined by polygonal meshes are well-established [21, 50, 60, 111]. There are three categories of method, the simplest of which is *constant* shading. This technique is also known as *flat* or *faceted* shading and involves the evaluation of an illumination model to determine a single intensity value which is used to shade an entire polygon. As a result, facets are visually distinguishable from each other, as demonstrated in Plate 7.1.

The need for a more curved appearance has prompted the development of *smooth-shading* techniques such as those of *Gouraud* and *Phong*. Gouraud shading tries to eliminate intensity discontinuities for curved surfaces approximated by planar polygons [60]. Vertex normals are evaluated by averaging the surface normals of all facets that share the vertex. Intensities at these vertices are calculated using an illumination model. A polygon is then shaded by a linear interpolation of the vertex intensities along each edge, and then between edges along each scan line. Since intensities are interpolated, specular reflections which are dependent on the view direction are forfeited. Phong's method of shading polygons interpolates the averaged surface normal vector rather than the intensity at every pixel [21]. This method allows specular reflections to be exhibited, but increases the computational cost of applying the illumination model.

Within FACES, Gouraud shading is used to display the smooth contours of the face because specular reflections would result in an artificial and shiny appearance. An added benefit is that Gouraud shading requires both fewer and less compute-intensive calculations than Phong shading. At present highlights on pupils are rendered by extraction of particular facets since the position of the light source cannot be varied. An advantage of implementing shading through an interpolation technique, is that shading calculations can be incorporated into the process of scan conversion.

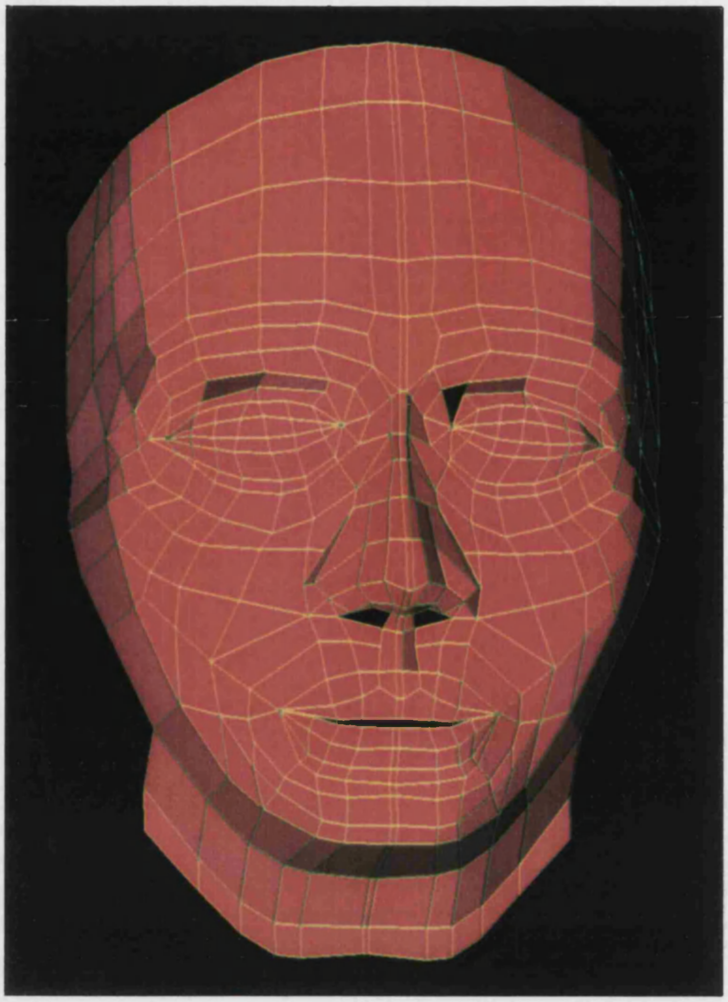

Plate 7.1: Constant Shaded Face

Constant, or faceted, shading involves evaluation of only one intensity for each facet [50, 110]. As a result of this, facets are visually distinguishable from each other.

7.4 The Render Sub-System

The complete structure of the Render sub-system is illustrated in Figure 7.1 while its context within FACES is presented in Appendix E. As with the descriptions of the Construct, Modify and Animate sub-systems, the Render sub-system is also explained through the use of menus where appropriate. The sub-system comprises three major options, which are listed below:

SHADE
STORYBOARD
PLAYBACK

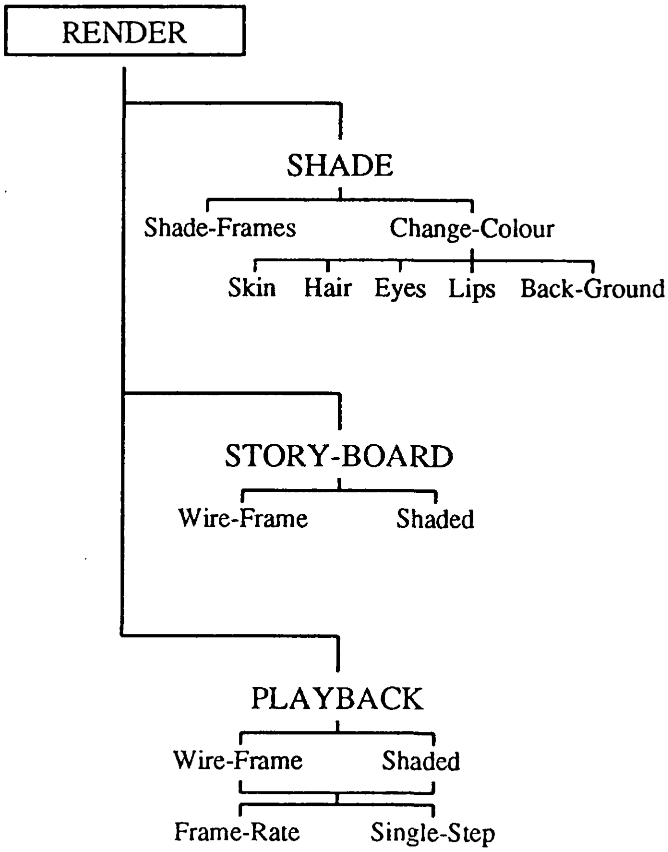

Figure 7.1: The RENDER Sub-System

The structure and organisation of the RENDER sub-system in FACES. Three major sections cater for shading, examination and near real time playback of animation sequences.

The SHADE option provides facilities for creating a realistic rendition of individual frames in conjunction with interactive control over the colour of various parts of the face; this is discussed in section 7.5. The STORYBOARD option, which is described in section 7.6, enables visual changes between successive frames to be monitored, while the PLAYBACK facility caters for the evaluation of motion as explained in section 7.7.

7.5 Colouration of Features

It is important to realise the significance of colour in the visualisation process [167]. As well as enhancing the image, colour clarifies the information presented and helps in making apparent features that would be obscure in black and white pictures.

Colour images of the face are displayed on an output device which has a resolution of

1024 by 1280 pixels. This monitor has a colour palette of 4096, however with 8 bit-planes, only 256 of these may be displayed at any one time. Advance knowledge of the contents of images enabled an optimisation of the colours that are loaded into the video look-up table or VLT. Colour entries are necessary for the functional requirements of the system, in terms of background and menu colours, as well as for shading the face. Altogether there are 8 different colours which must remain constant, leaving a total of 248 VLT entries available for the face model.

Shading images of the face requires only five basic colours to cater for: skin tone, hair, iris, sclera and lips. However, each of these features needs gradations of one hue. The colours of the smaller features, which include the lips, eyebrows, iris and sclera, are set up in bands of 32 gradations. Skin covers the greatest surface area of the face and therefore requires a larger spectrum of shades; 120 VLT entries are therefore allocated to skin tone.

The SHADE option in the RENDER sub-system caters for both frame rendering and interactive modification of the colour of parts of the face. Default colour values are read in from a textual data-file which is provided as an argument when FACES is first invoked, see Appendix C.

The user may change the colour of skin, hair, eyes, lips and the background by specification of the hue in terms of values for red, green and blue components. An RGB colour cube model [50, 111] is used to specify the colour of the face. Each of the red, green and blue components range between zero and one. These values are mapped into a range of either 32 or 120 gradations depending on the feature of the face selected.

At present hair colour is restricted to affect the eyebrows alone. With regard to the eyes, control is available over the colour of the iris but not the pupil or sclera. All humans have a black pupil and white sclera, which are kept constant.

In addition, the user has control over both the first and last frames to be shaded. Such control enables experimentation with colours prior to a whole sequence of frames being rendered. Since the rendering process is optimised, FACES generates and outputs each Gouraud shaded image within a few seconds.

7.5.1 Appearance

Colouration of features is an important component in distinguishing between races and between gender. For instance, we normally associate: yellow with Far Eastern cultures; white with Caucasions; a light tan with Europeans and people from the Middle East;

brown with Asians; and black with Africans. Plate 7.2 demonstrates that a large range of skin tones can be represented in FACES.

Both hair and eye colour are also important in modelling the appearance of the face. For example, we tend to affiliate blue eyes and blonde hair with Europeans and dark eyes and dark hair with Asians and Africans. Colour is also a significant factor in our portrayal of gender; we associate red lips with females, for instance. Plate 7.3 shows colourations that are typically used to project images of male and female faces.

7.5.2 Emotion

The colour of the skin can reflect various states of emotion, for example blushing is associated with embarrassment, see Plate 7.4 and video sequence 5 in Appendix D. Another example is paleness due to shock or fear, as illustrated in Plate 7.5 and video sequence 6. Skin tone can also give an indication of the physiological states of the body, such as red for hot, blue for cold and green for envy.

7.6 Examination of Frames

Story boards have traditionally been used in 2D cell animation to give the animator an impression of the key moments in the sequence [162]. The STORYBOARD option in the RENDER sub-system serves a slightly different purpose. It allows examination of consecutive frames of a sequence by displaying individual frames next to each other, as illustrated in Plate 7.6. This facility is of invaluable help to the animator throughout the development of a sequence. It enables monitoring of changes that occur from one frame to the next. The user controls the first and last frames to be displayed, so that long sequences can be divided and inspected in groups of consecutive frames. STORYBOARD is capable of operating on both wireframe and shaded sequences of frames.

7.7 Evaluation of Motion

It is essential that the correct speed and duration of facial movement be achieved if a particular message is to be conveyed to the observer in an unambiguous manner. Testing the smoothness of generated motion is therefore of paramount importance.

The PLAYBACK option caters for near real time display of short sequences under user

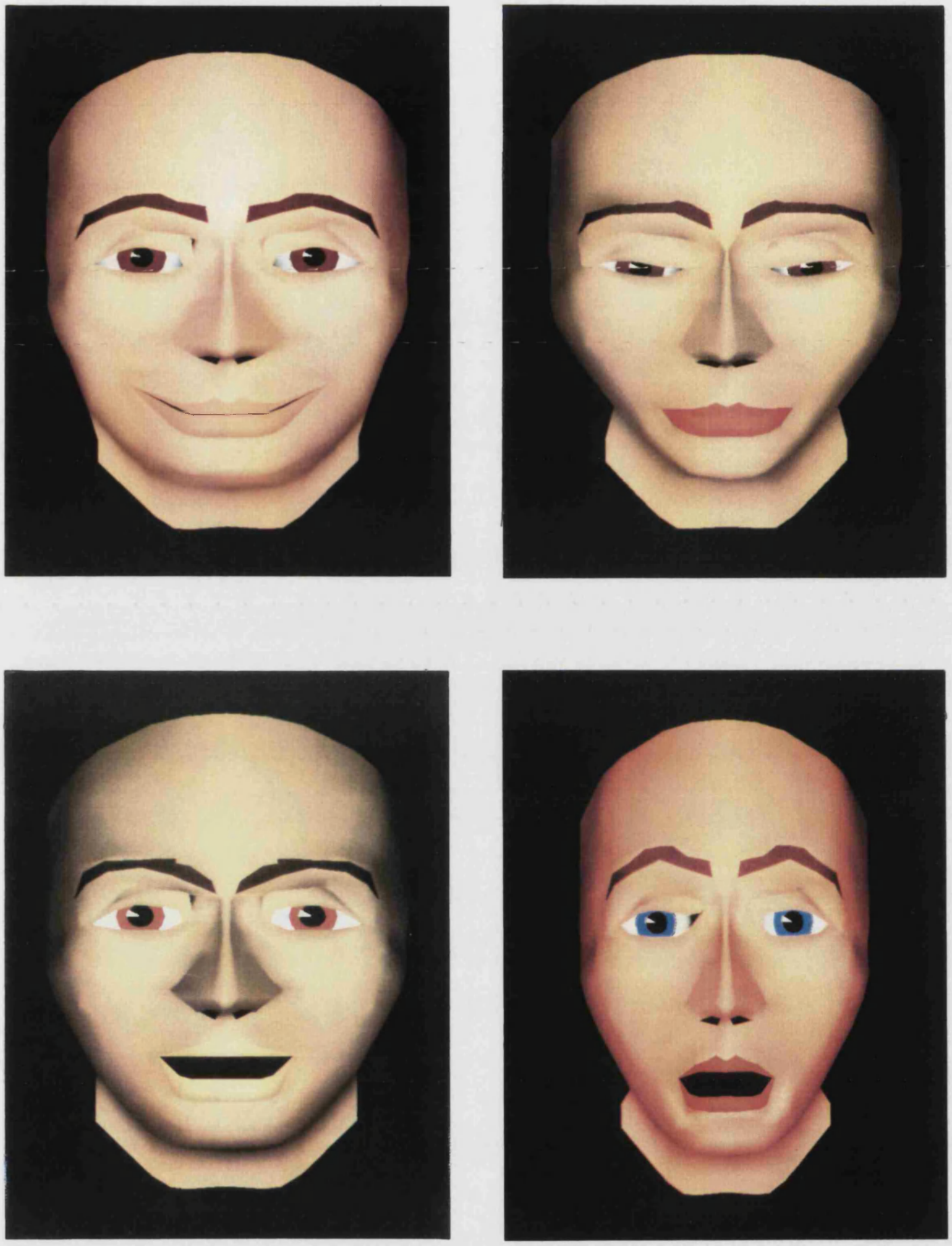

Plate 7.2: Examples of Skin Tones

Colouration of facial features is important for distinguishing between races. This plate illustrates a variety of skin tones in conjunction with differing head models and facial expressions.

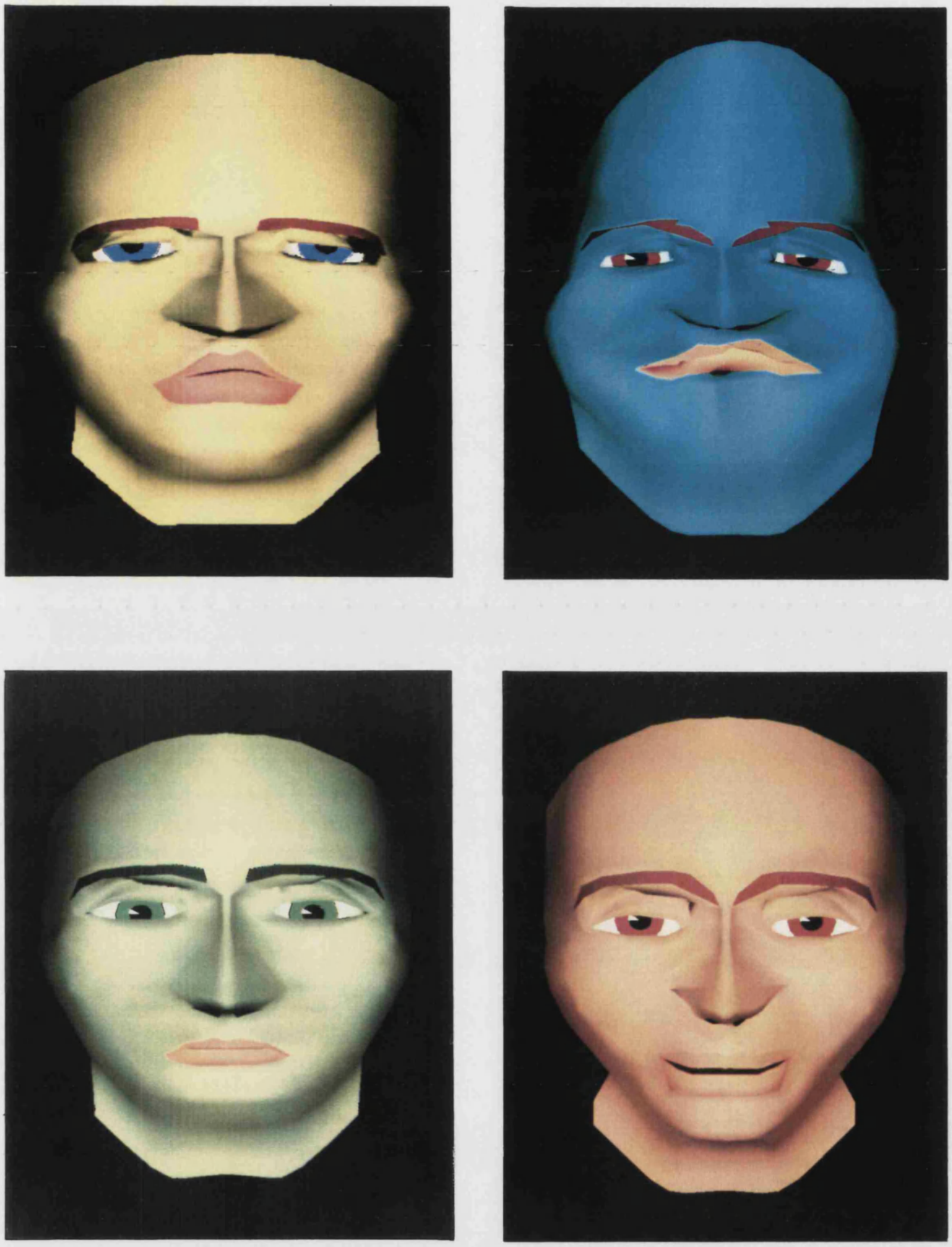

Plate 7.2: Examples of Skin Tones (Continued)

Modelling facilities available within FACES, together with control over facial colouration, provides the capability to create both realistic and stylistic faces.

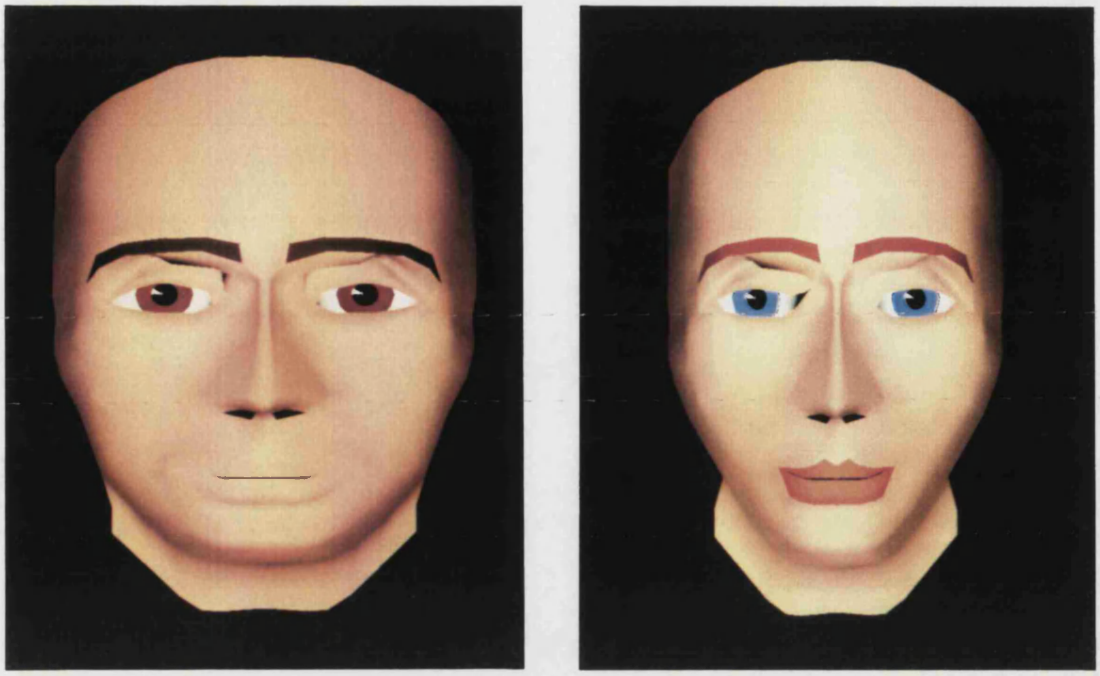

Plate 7.3: Colouration of Features

Our portrayal of gender often includes variation of the colouration of particular features. For example, we tend to associate smooth skin and red lips with female faces.

control. Individual static images of the face are over-laid in rapid succession to provide the illusion of motion. In a similar manner to the STORYBOARD facility, it is possible to operate on both wireframe and shaded sequences, as well as to select sub-sequences of an animation sequence.

The user may select from two modes of display, either *Frame-Rate* or *Single-Step*, in order to test the motion of a sequence. Within the *Frame-Rate* option there are additional controls, such as over the pause rate between frame updates and whether the sequence is to be cycled through in reverse order when the last frame is reached. Such control provides the animator with a considerable degree of flexibility to experiment with the timing of particular actions and to test the motion at varying speeds before refinements are made to the animation script. In *Single-Step* mode, the user has total control and frames are updated with the press of a puck button. Under both modes frames continue to be displayed cyclically until the user decides to exit from the option.

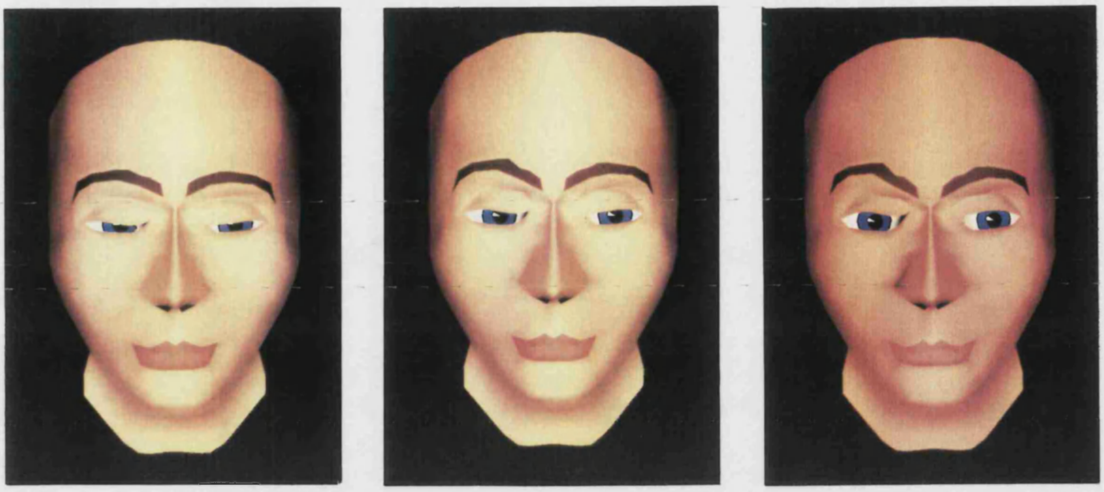

Plate 7.4: Blushing

Colour is important for conveying the emotional state of a person. This sequence demonstrates variation of skin tone over an animation sequence to produce the effect of blushing.

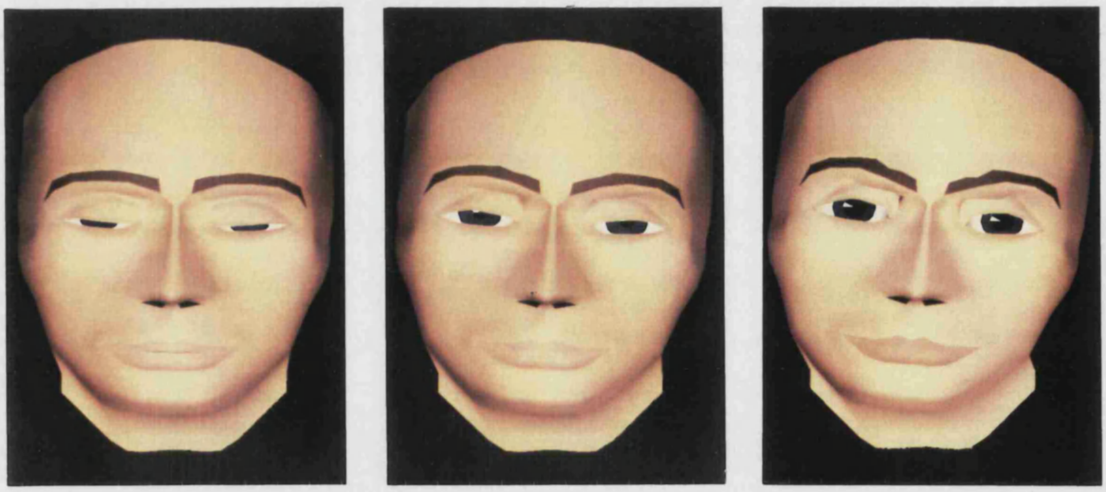

Plate 7.5: Paleness

This sequence shows the variation of skin tone through an animation sequence to produce the effect of paleness or pallor, which could be associated with shock or fear.

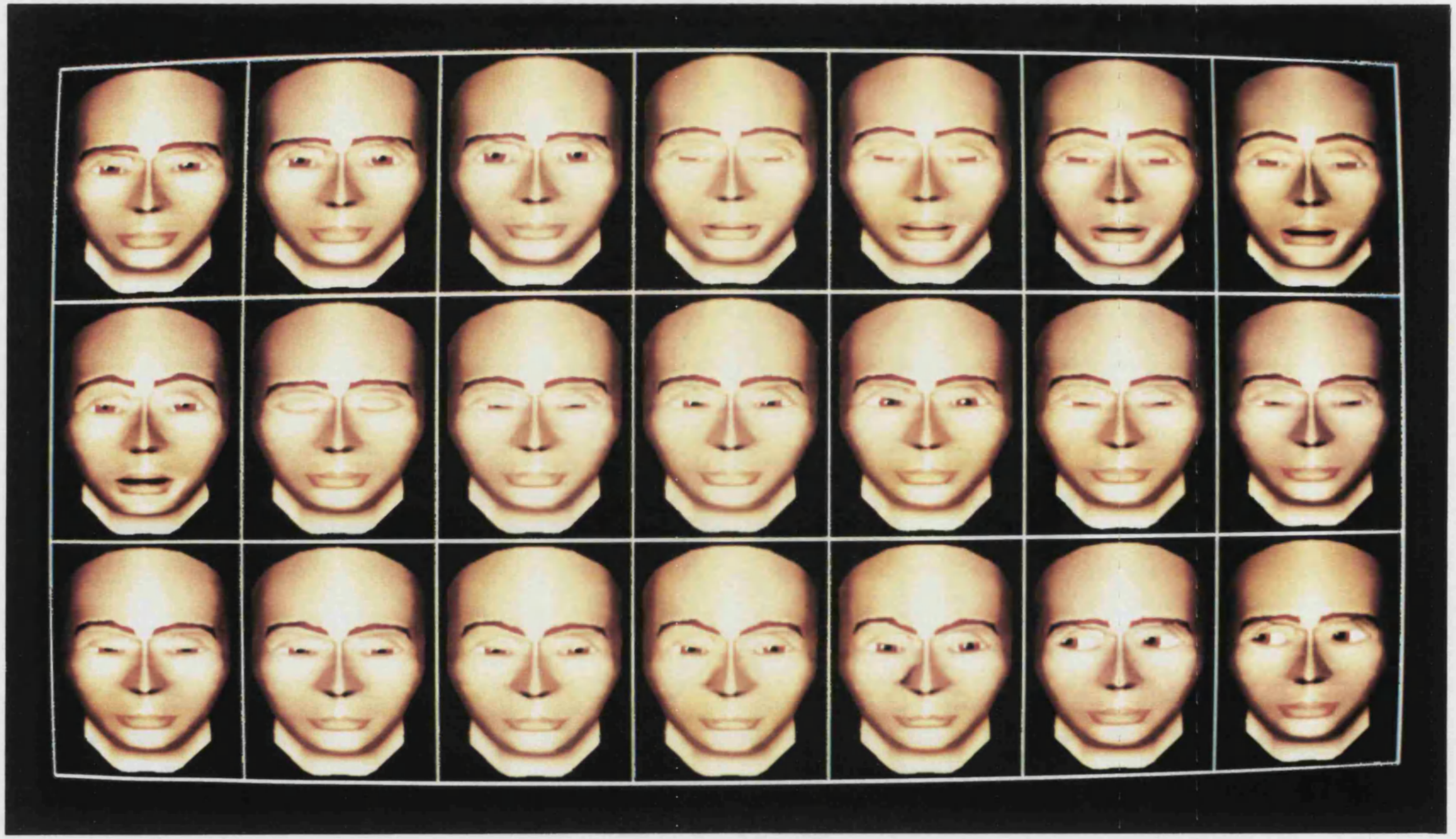

Plate 7.6: The STORYBOARD Facility

The STORYBOARD facility can be of invaluable help during the development of a sequence; consecutive frames are displayed next to each other. This facility allows animators to closely examine the changes that occur between successive frames of a sequence. Control over the first and last frames to be displayed enables sub-sequences to be selected.

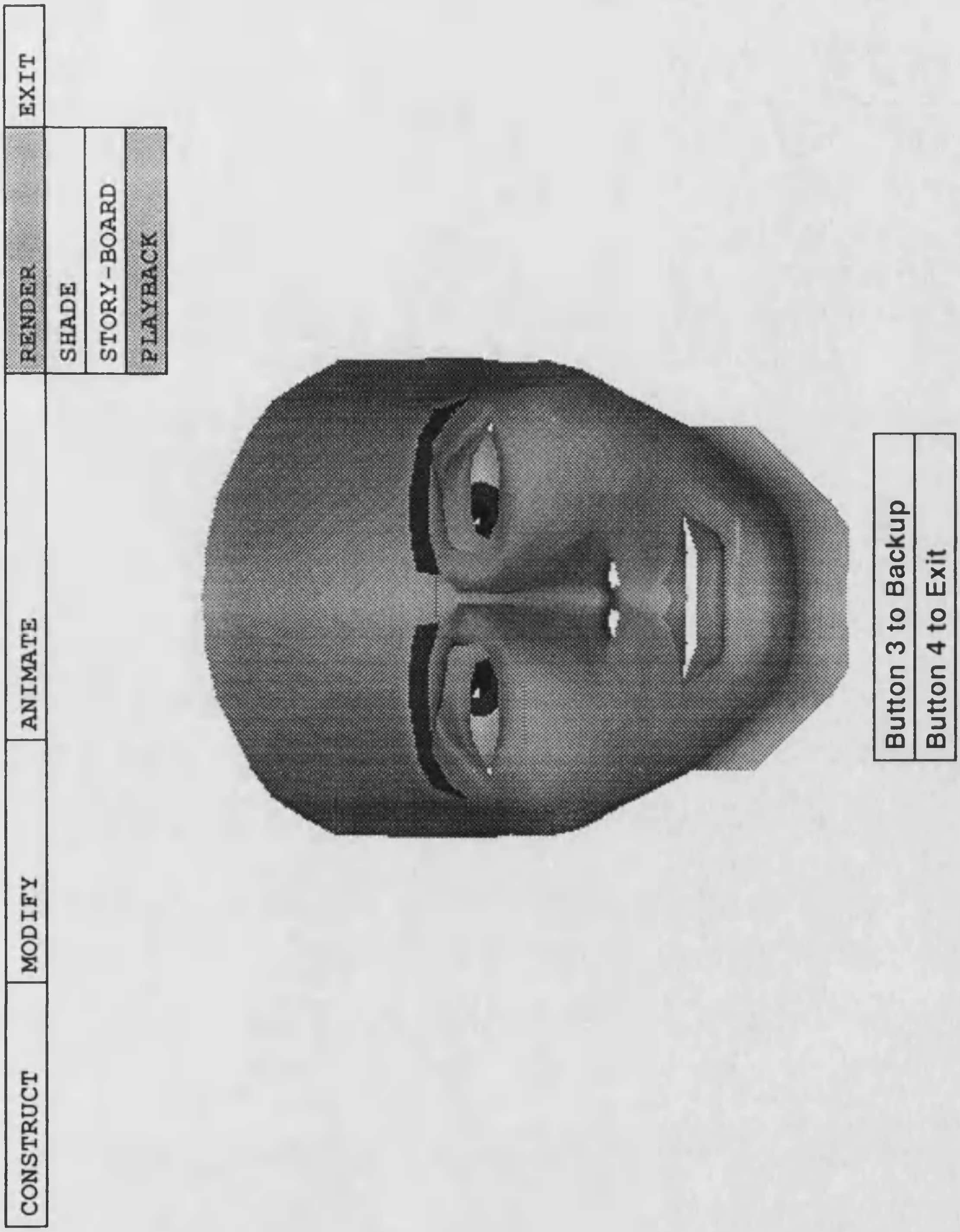

Figure 7.2: The PLAYBACK Facility

This facility is used to appraise motion characteristics of an animation sequence. Two modes of operation, *Frame-Rate* and *Single-Step* provide controlled evaluation.

7.8 Summary and Conclusions

Although it is highly desirable to allow the user to work with realistic shaded images throughout FACES, the state of accessible technology does not permit this at present. In fact, rendering of each 3D image involves several time-consuming stages before the image eventually appears on the screen. With regard to visible surface determination, careful choice of sorting and coherence methods are vital since these two processes have important implications for the execution efficiency of an algorithm. Advance knowledge about the face and the type of primitives to be rendered has enabled considerable computational savings to be made in each of the processes of hidden surface removal, illumination and shading.

The user is provided with colour control over various parts of the face. This can be important in modelling appearance, in terms of race and gender, as well as emotional states such as embarrassment, shock and fear.

The STORYBOARD option is extremely useful as a means of tracing changes that occur from one frame to the next, while PLAYBACK puts control over motion evaluation in the hands of the user. Both STORYBOARD and PLAYBACK are capable of operating on either wireframe or shaded sequences. Viewing of wireframe images is useful for debugging purposes and for checking the motion before frames are rendered. These types of facilities are crucial for the user since timing and movement are the essence of animated work.

Chapter 8

FACES for the User

8.1 Introduction 137
8.2 The User's Conceptual Model 138
  8.2.1 Control for Modelling 139
  8.2.2 Control for Animation 140
  8.2.3 Control for Rendering 142
8.3 Design of the User Interface 142
  8.3.1 Interaction 142
  8.3.2 Feedback 144
  8.3.3 Experimentation 145
  8.3.4 Practicality 146
8.4 Summary and Conclusions 146

8.1 Introduction

In this chapter we examine the system from the user's point of view through consideration of the facilities which are provided, their appropriateness and practicality. It is necessary not only to consider the graphical operations to be performed but also how those operations are to be made available to the user.

The goal is to provide flexible tools which give the animator as much control as possible, without the burden of having to be concerned with the low-level mechanics of computer animation. The animator needs the ability to create expressive, suggestive and idiosyncratic sequences. Creativity must be enhanced, not stifled. Furthermore, simplicity and ease of use are essential to make the system user friendly.

For the system to be readily usable, it is necessary in the first instance to provide the required functionality for a particular application; this issue is discussed in section 8.2 in terms of the amount and levels of control available for modelling, animation and rendering. However, provision of the required functionality in itself is rarely adequate, it is also necessary to present that functionality to the user in an effective manner. This issue involves careful design of the user interface and is further elaborated upon in section 8.3.

8.2 The User's Conceptual Model

A user model serves to define concepts which relate to how a system operates in terms of: application concepts; the type of objects that can be displayed; and how those objects can be manipulated. The model therefore provides a context within which a *dialogue* can be developed between the user and the computer system.

Communication between people is rich, but often sparse, comprising of gestures and subtle expressions as well as spoken words. Richness of the interaction comes from a sharing of common knowledge, assumptions and environment or world. In order to duplicate such efficiency in human-computer interaction, it is necessary to model an application based on the user's conceptual world. The interaction language should be as natural as possible, with operations and objects specified in terms normally used within the application area, so that the user does not need to learn new concepts and in particular those that relate to computer science.

The user is concerned with the functionality that the system can provide. With respect to FACES this amounts to the range of faces that can be modelled and the range of expressions that can be generated. A central issue here is that of control. As in other areas of computer animation [22, 185, 197], control in facial animation comprises two components. Firstly, what **amount** of control should be provided for the user? Secondly, at what **levels** should that control be made available? As the complexity of animation models has increased, there has been a trend towards the incorporation of as much a priori knowledge into the software as possible, in order to aid the animator. However, the consequence is often that the user loses control over the low-level details.

Control can be viewed as Parke suggests in terms of the specification and regulation of parameters as a function of time [120]. In our case, we ascribe a general meaning to the word 'parameter', which is here used to represent characteristic criteria which facilitates control over specific aspects of both face modelling and animation. According to Parke, control can be divided into two main issues. The first concerns the development of 'universal' control parameterisations which incorporate conformation and animation together with speech, to enable an animator to 'easily' specify any individual face with speech and expression. Secondly, the development of optimal techniques to implement facial animation based on these universal control parameters.

The user is interested in: which parameters are available; whether they are adequate and appropriate; and how the parameters are manipulated. Quality of control, in conjunction with range, complexity and whether the parameters are intuitive to use are also of concern to the user. Implementors on the other hand need to establish: which parameters should be provided; what type of user interface to the parameters should be made available; and which low-level implementation techniques and algorithms are suitable. In FACES, conformation modelling and animation are considered separate processes which require different types of control.

8.2.1 Control for Modelling

The face is a popular object of study in many disciplines such as medicine, anthropology, portraiture and criminology. Research undertaken in such fields indicates that several types of control are required to model arbitrary faces and heads. However, as explained in chapter 5, the derivation of parameters to control the form of the head and face are tenuous and less well understood than those for facial expression. This leads to problems in the development of a 'universal' set of parameters for conformation modelling.

In the CONSTRUCT and MODIFY sub-systems modelling capability is based on the anatomy of the head, to provide for naturalistic and intuitive control as well as realistic results. The head model itself is a three-layer structure which consists of bone, muscle and skin; these are the three major components of the human head which influence appearance, see chapter 2.

The user is provided with several types of control at a variety of levels over the three layers. Global control facilitates changes to the overall shape of the head and face. Regional modifications cater for changes to relative proportions and local control allows alterations to specific bones and features, as explained in chapter 5. Modelling functionality in FACES caters for a range of controls and enables subtle variations to be made to the predefined head model as shown in Plate 5.1.

8.2.2 Control for Animation

Development of motion is amongst the most time consuming activities of the animation process. It is possible for designers and animators to produce quality work using low-level techniques, but only as a result of spending hours to define meticulously a path or shape. Impressive animation sequences such as those involving the synthetic characters *Nestor Sextone* and *Dozo* have been produced, but they have required 18 out of every 30 frames in a second of animation to be keyframes [92].

More automated methods should allow the production of sophisticated animation with less user effort [185]. Techniques for controlling motion can be classified into the three general types known as *Guiding, Animator* and *Task* levels [134, 197]. At the Guiding level, motion is described explicitly at a low-level, by moving individual vertices of a model to define a keyframe, for example. Parameter interpolation also falls into this category, although it is at a higher level than shape interpolation. The problem at the guiding level is that complex movements require the animator to provide large amounts of data which can become tedious.

At the Animator level, the user may describe behaviour algorithmically in a programming notation. Use of programming languages is a powerful tool; it enables the animator to control the motion of any element, or set of elements, at different levels of abstraction. However, the user does need some knowledge of computer science, which is a drawback. Furthermore, some movements are difficult to specify algorithmically.

The Task level is a higher level of control in which motion is specified in terms of events and relationships such as 'walking', 'grasping' and 'talking'. High level commands are employed to perform predefined or computable movements through the use of low-level motor programs [196].

Zeltzer argues for the need to integrate all three levels of control since no single mode can provide complete yet economical control [197]. The ideal system in Calvert et. al's opinion is also one which provides varying levels of control [23] through the use of: natural language input; detailed scripts which can be edited by the animator; and low-level movement instructions for 'fine-tuning'. Control is necessary at all levels, with the ability to transcend levels of control because good animation often requires extensive fine-tuning and 'tweaking'.

Animation in FACES

The functionality provided in the ANIMATE sub-system enables the user to create the two types of animation sequence, known as conformation and facial animation, see chapter 6. In both cases the user has flexibility in defining a sequence since the activities that comprise a sequence do not need to be specified in a temporal order.

With respect to facial animation, the complexity of animating models which are as intricate as the face has meant that work has focused on making such control manageable. Use of parameter interpolation minimises the amount of user input required. The user has two levels of access to facial movement, through the AUs of FACS and through predefined expressions.

A major advantage of using FACS is that naturalistic and accurate facial expressions are generated. High level expressions can be considered to correspond to the task or goal-direction level of control, which enables the user to work at an 'emotional' level. The animator can therefore decide which actions are to occur without having to worry about the low-level details of how they happen.

Use of both levels of control has been simplified by minimising input parameters. Most AUs and all expressions require only one parameter, that of intensity, which is also intuitively easy to control since it ranges between zero and one. In a similar manner to predefined expressions, AU routines can also be viewed as 'black-boxes' with internal strategies which need not concern the user. Simulation of the dynamics of muscular contraction frees the animator from the task of ensuring that the motion generated conforms with the laws of nature.

Control is also available over the duration and law of motion for each AU and expression selected. This provides the user with a substantial amount of control over the particular parts of the face that need to be activated as well as the exact type of motion that the parts need to undergo in order to create expressive sequences, see section 6.6.3.

Furthermore, animation scripts are capable of being edited since they are stored in a human readable textual format, with keywords to aid comprehension. An additional benefit is that FACS provides generality, so that once animation sequences have been created, they can be applied to any head model. Thus the control provided for the animation of faces is flexible and the facilities in existence ensure that the low-level mechanics of computer animation are hidden from the user.

8.2.3 Control for Rendering

Creation of realistic faces in terms of both appearance and motion is necessary to make animation sequences convincing. Colour is an important component in determining appearance, while generation of believable motion requires an amenity that allows the testing and appraisal of motion. Control over both of these aspects is provided within the RENDER sub-system, which enables interactive control over the colour of various parts of the face, as well as near real time playback of sequences, see chapter 7.

8.3 Design of the User Interface

Ultimately, whether a system is used or not will depend on the quality of the user interface provided. Development of an effective interface requires a multi-disciplinary approach which takes account of all relevant human factors and ergonomics.

Major goals in the design of an interface for an interactive system are to: increase the speed of learning and speed of use; reduce the error rate; encourage rapid recall of how to use the interface; and increase the attractiveness of the system to potential users [50]. A good user interface should make all possible actions clear and guide the user through the application. It should be informative without being distracting; it should anticipate errors and provide mechanisms for graceful recovery from them. The interface should also provide a good visual design which promotes clarity, consistency and an attractive appearance.

8.3.1 Interaction

The user interface to FACES integrates the four components of the system to provide a uniform manner in which the animator is allowed access to the sub-systems. A limited amount of screen space has necessitated careful design of menus and other output in order to provide visual effectiveness. Three basic components need to be taken into account, they include: a work area; a menu area; and a region in which to output error messages and text to prompt the user for input. In FACES, error messages and other textual messages are output to a terminal while graphical output and interaction with menus takes place on a graphics display, through the use of a puck and tablet.

Use of menus relieves the user of the burden of remembering input sequences and options since presentation of a menu invokes the user's recognition memory rather than recall memory [194]. Menus also provide a way of making only valid options available at any particular stage and therefore reduce the probability that the user will make a mistake. Additionally, it is possible to control both the precise placement and timing of appearance of menus.

Two types of menu are in use, they are known as fixed or permanent and pop-up or movable menus. In order to maximise the screen region available as a work area, a permanent horizontal menu-bar, located at the top of the screen, acts as the gateway which allows interactive access to the four sub-systems that comprise FACES:

CONSTRUCT	MODIFY	ANIMATE	RENDER	EXIT

Whenever a top-level menu item, which represents a sub-system, is selected the relevant menu-box is highlighted and a list of the major menu options in that sub-system appears underneath the menu-item, see Figure 8.1. This pull-down menu remains on the screen until its parent menu-item is de-selected. The horizontal menu-bar and the pull-down menus associated with each of the four sub-systems are fixed to encourage the user to make a habit of choosing particular selections at a specific location.

The lower levels of the system make liberal use of pop-up menus. These movable menus have several advantages. Firstly, they can be placed near the current position of the screen cursor so that menu selection involves a minimum amount of hand and eye movement. Secondly, such menus can be made to disappear off the screen when they are not required, to release valuable screen space. Thirdly, since pop-up menus are transitory, it is unnecessary to erase the work area, which allows visual continuity to be maintained.

The top-down approach adopted is reflected in the hierarchical organisation of the system, see Appendix E. This approach prevents the user from being saturated with irrelevant detail; it is only necessary to be concerned with the particular task at hand. However, a hierarchical organisation necessitates the nesting of menus to a depth of several levels. The tedium of backtracking to higher levels is resolved by the provision of an *Exit* option in the pop-up menus. This helps a user to escape from a menu hierarchy. Furthermore, whenever it is available, the *Exit* option consistently appears as the first menu-item.

Permanent and pop-up menus enable parts of the system to be *modeless* and other parts to be *guided* [50, 111]. The user is able to choose freely from the menus associated with the sub-systems and their major options. Such a choice requires explicit activation, while pop-up menus appear automatically at relevant stages, to guide the user. This

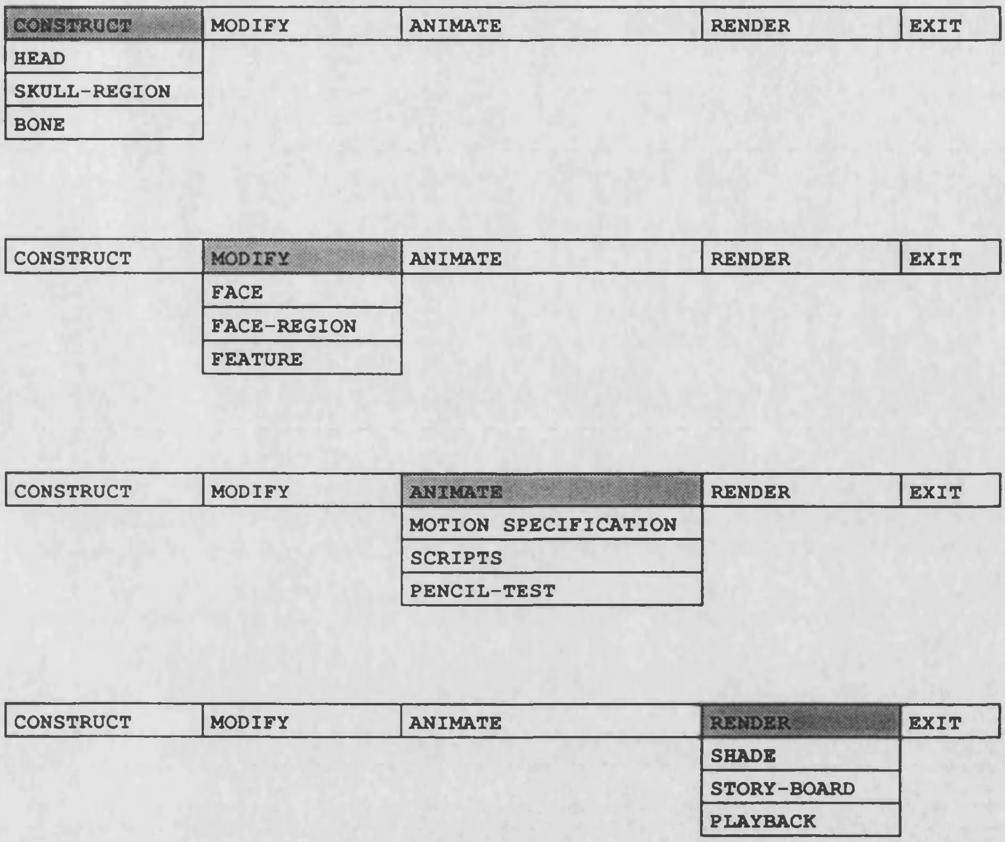

Figure 8.1: Top-level Menus

In order to maximise the screen region available, a permanent horizontal menu-bar, at the top of the screen, acts as the gateway which allows access to four sub-systems. Facilities within each of the sub-systems are revealed through pull-down menus.

strategy provides considerable flexibililty in the way that a user works, while at the same time ensuring that FACES has all the relevant information in order to perform a particular function.

As far as possible the system checks to ensure that preconditions are met for correct operation. For example, selection of the *Animate* option in PENCIL-TEST is meaningless if an expression animation sequence does not yet exist within the system and the user will be so informed.

8.3.2 Feedback

Feedback is an important component in any interactive system. Each input needs to be followed by a clear message to indicate to the user that input has been received and the information is being processed. To develop a continuous interactive dialogue the system must inform the user of what is happening at each stage.

With regard to menu selection, a feature in FACES is that a selected option remains highlighted until that option is exited. For example, selection of MOTION SPECIFICATION within the ANIMATE sub-system will cause the menu item to remain highlighted throughout the creation of an animation sequence. This is useful feedback as a constant indication of the state of the system. When the menu item reverts to the background colour, the system is ready for the next option to be selected.

Modelling operations on the head model and the creation of static high-level expressions result in 'visual echoes' in the form of an up-dated appearance of the model being displayed on the screen. During the creation of animation sequences, selections of AUs, expressions and associated parameter values are printed out on the user's terminal. Error messages and mistakes, such as invoking the undo operation when an undo-stack is exhausted, result in audio feedback as well as textual messages in order to draw the user's attention to the problem.

8.3.3 Experimentation

A system which provides the ability to 'back up' and recover from errors permits a user to explore confidently the capabilities of the system in the safe knowledge that the effects of any mistakes can be erased without damage. Error recovery can be implemented in several forms such as undo, abort, cancel and correct [50].

In FACES, the main method of recovery from errors is through an undo facility. This allows the effects of previous actions to be cancelled and removes the need for continual confirmation of possibly dangerous actions. The undo operation is equivalent to an 'implicit-accept, explicit-reject' strategy. An interactive system together with such an operation is a powerful method of catering for experimentation during both modelling and animation.

FACES provides two types of undo operation. Whilst making regional modifications in the CONSTRUCT and MODIFY sub-systems and defining new expressions in the ANIMATE sub-system, the operation is stack-based so that the user may back-track as required. During global and local changes to the head and face, only one level of undo is provided. In addition, graceful recovery from erroneous menu selections is available through the *Exit* option in pop-up menus.

Experimentation is also of importance during motion generation, particularly since this is an area in which FACES provides limited built-in assistance, see chapters 4 and 6. It is essential that an animation system allows flexible, easy and rapid testing of motion; preferably in real time. Such facilities are available within the RENDER sub-system through the PLAYBACK option.

8.3.4 Practicality

Practicality is a very important factor in determining how often a system is eventually used. Several additional facilities are provided within the system to make the user's life a little easier. FACES saves modified head structures and animation scripts to default files behind the scenes. This is to guard against a user's memory problems as well as system faults!

A stand-alone version of the rendering process enables the user to generate shaded frames which comprise a sequence as a batch process separate to the system. The STORYBOARD and PLAYBACK facilities also exist as stand-alone utilities which are totally separate from FACES. This allows both wireframe and shaded images to be examined and animated sequences to be displayed independently of FACES, thereby avoiding the computational overheads associated with the system's start-up.

8.4 Summary and Conclusions

There are two aspects to encouraging the user to make effective use of an interactive system. Firstly, it is necessary to provide the correct functionality that a particular application requires in a manner which conforms with the user's conceptual model of that application. Secondly, it is necessary to facilitate effective use of the system by designing an interface which is simple, intuitive, flexible and which does not obstruct the user. Both aspects entail careful consideration of many human factors.

Within FACES functionality is provided for modelling faces, animating and rendering them. Control over these operations is provided at several levels. The anatomical basis of the head model in conjunction with dynamic simulation of facial motion provides for naturalistic and intuitive control in addition to realistic results. Final assessment of the system will pertain to: the range of faces that can be modelled; the range of movements that can be generated; and the range of colouration that can be used to portray the face. Substantial control over each of these factors is available within FACES. Emphasis has been placed on making the complex and intricate nature of the face easily controllable.

Implementation of the functionality in FACES has been influenced by a number of issues relating to ergonomics. Interaction with the system is made straightforward through the use of menus. Feedback is provided rapidly to maintain a continuous dialogue and error recovery encourages experimentation.

Chapter 9

Further Development

9.1 Introduction 149
9.2 Head Model 149
9.2.1 Surface Patches 149
9.2.2 Additional Features 149
9.2.3 Modelling Deformation 150
9.3 Muscle Dynamics 150
9.4 Modelling Faces 151
9.4.1 Skull Prototypes 151
9.4.2 Growth and Ageing 151
9.4.3 Face Composition 152
9.4.4 3D Sculpting 153
9.5 Animating Faces 154
9.5.1 Skin Wrinkles 154
9.5.2 Smoother Motion 154
9.5.3 Timing for Facial Movement 155
9.5.4 A Parser for Action Units 155
9.6 Rendering Faces 155
9.6.1 Illumination and Shading 155
9.6.2 Skin Tone and Texture 156
9.6.3 Colour for Emotion 156
9.6.4 Hair 157
9.7 Speech Synthesis 157
9.8 User Interface 158
9.9 Summary and Conclusions 158

9.1 Introduction

FACES consists of many facilities and desirable features that cater for the modelling, animation and rendering of the face. In addition, development of the system has revealed many new avenues which could be explored. Rudimentary facilities are provided which could be developed further in order to accommodate particular applications.

In this chapter we identify the major areas that could benefit from further research and development in the context of: the head model; dynamics of the face; modelling, animation and rendering facilities; speech synthesis and the user interface.

9.2 Head Model

The polygon based three-layer representation of the head and face, which was described in chapter 3, has been adequate for experimental purposes. However, there are several enhancements which could be beneficial.

9.2.1 Surface Patches

Although use of a polygonal mesh is a popular method for representation of the face, the technique remains a crude approximation to the smooth curves of a real face. Furthermore, shading algorithms for polygonal meshes smooth the centre of the mesh but leave a faceted silhouette edge. This problem becomes more apparent in an animated model than in a static image. Consequently, a surface patch representation, such as B-spline surfaces [71, 171], may be more appropriate even though a surface patch model will make the system more compute-intensive and therefore slower.

9.2.2 Additional Features

At present, the head model consists of only those components and facial features that are directly relevant to the research being undertaken. For a more complete system the model needs additional elements such as a tongue, teeth, ears and hair. There is also a further demand for superficial items such as spectacles, beards, moustaches and scars. Such details are important for increasing the range of faces that can be modelled as well as improving realism.

9.2.3 Modelling Deformation

Modelling controls in the CONSTRUCT and MODIFY sub-systems, which are described in chapter 5, provide considerable flexibility. However, for even more effective conformation modelling, the head model could facilitate not only 'reasonable' alterations to major parts of the face, but also accommodate more outrageous changes. For example, if a user decides to enlarge the eyes to take up half the size of the face, the surrounding regions would need to 'give' to maintain an acceptable and recognisable structure. Such considerations would enhance the facilities for creating stylistic or caricature faces.

9.3 Muscle Dynamics

The muscle models, described in section 4.5.2, simulate the major characteristics of linear, sphincter and sheet muscles. There are however, two areas in which the models implemented do not provide a precise simulation of actual muscle and skin deformation. The first issue is concerned with the modelling of furrows, bulges and wrinkles in the skin. These features are caused by forces pressing the skin towards each other, resulting in a 'buckling' action. Ekman and Friesen have identified that forehead lines and certain wrinkles are a significant part of some expressions [43].

The second problem is that only surface deformation is modelled and therefore no account is taken of volumetric changes which occur within muscles. Such a consideration is necessary to accurately model facial deformation.

Both of these problems have recently been addressed by Waters and Terzopoulos through extension of Waters' original anatomical model into a physically-based model. The recently developed model provides a more accurate simulation of both the structure and behaviour of skin [159]. Related work has been undertaken by Pieper, with particular emphasis on the manner in which skin 'bulges' and 'buckles' [129].

A less elegant, but also less compute-intensive method for the generation of creases in the skin is through the creation of discontinuities at salient facet edges, through duplication of edges so that normal vectors will not be averaged during smooth-shading [116].

9.4 Modelling Faces

There are several ways in which the modelling facilities in the CONSTRUCT and MODIFY sub-systems could be extended to provide the user with further flexibility and control for the creation of a wide variety of faces.

Faces are a variation on a theme. In the case of realistic faces, the distinctions need to be of degree rather than kind, while for stylistic faces there is more scope for diversity in exaggerating particular features. In the absence of a 'universal' set of parameters for conformation modelling, a large variety of tools of varying nature are necessary to have any hope of capturing the diversities of the facial form.

Laser and light beam scans are probably the only reliable methods of acquiring a realistic physical likeness of a person. Such techniques need to be integrated into the system. Other types of facilities such as interpolation and face composition are more appropriate for creating new characters and caricatured faces.

9.4.1 Skull Prototypes

For a comprehensive system it is necessary to incorporate information regarding race, gender and age. At present, a scaled down version of the facial skin mesh is used to represent the skull, see section 2.4. While this has served a useful purpose for experimentation, it would be practical to have prototype skull models which represent race dependent bone structures based on the Caucasoid, Negroid and Mongoloid skull types. Furthermore, for each of these races there would need to be two models to represent each gender. Finally these models could be modified as a function of age.

9.4.2 Growth and Ageing

Time alters the countenance of every face. Both the processes of growth and ageing have a profound influence over the changes that take place in the structure of the head and the appearance of the face. Moreover, it is apparent that faces age at different rates and intensities. An automated ageing facility would be extremely useful. However, all the studies undertaken to date have been concerned with 2D metrics which are difficult to apply in 3D systems. A number of people have investigated the processes of growth and ageing.

Thompson was one of the first to realise that geometric transformations could be employed to model morphological changes in the head and face. In his classic work, *On Growth and Form*, he used geometric and mathematical transformations [163], but offered no explanation for their use, preferring to emphasise the visual perception of the results obtained.

In 1959 Subtelny and Rochester [155] concluded that growth of the human head follows a somewhat orderly pattern, but is complex, with different parts of the head growing at different rates and at different times. Furthermore, all parts of the soft tissue do not directly follow the underlying skeletal profile.

Walker and Kowalski developed a method for the analysis, measurement and prediction of craniofacial growth by marking 177 anthropometric landmark points on cephalograms [172]. However, the method required substantial amounts of data before predictions could be made regarding the growth pattern of an individual.

Todd et al. [165] discovered that the *cardioid strain* transformation seems to correspond closely to the perceived growth sequence of the head from infancy to adulthood. Experiments were performed using a 2D profile of the head and several types of transformation including, the Cardioid Strain: $\theta' = \theta$; $R' = R(1 - kcos\theta)$ and Revised Cardioid Strain: $\theta' = \theta$; $R' = R(1 + k(1 - cos\theta))$ which provides improved results.

More recently, Bookstein has experimented with mathematical and statistical techniques in 2D based on orthogonal grid distortions after Thompson [12]. Results from such investigations could form the basis for the development of a computerised method for ageing faces.

9.4.3 Face Composition

The three-layer anatomical model, described in chapter 3, has proven to be extremely flexible. However, it does have implications for the implementation of a 'kit-of-parts' method for creating new faces. The process of fitting different parts into a face model becomes more complex since muscle and bone need to be taken into account in conjunction with the surface skin. Nonetheless, an Identikit type of approach is necessary to overcome the problems associated with representing grossly differing features with a fixed number of facets and vertices.

Despite the problems involved, a 'kit-of-parts' approach holds tremendous potential. For example, users could be allowed, not only to incorporate predefined features, but also to create their own parts with techniques such as surfaces of revolution [36, 128, 164] and *Non-Uniform Rational B-Splines* (NURBS). NURBS allow the concise and precise representation of quadric primitives such as spheres, cylinders and cones, as well as the representation of free-form surfaces [128, 164]. Such a facility would enable the construction of both realistic and caricature faces.

A major issue here includes the management of intersecting surfaces [84, 126] and how these are blended together to form a 3D composite face. Problems similar to these are to be found in Computer Aided Geometric Design (CAGD) [66, 102]. Control over the modelling process could be subsequently enhanced through the provision of additional parameters for more subtle refinement to the features.

9.4.4 3D Sculpting

Although conformation modelling is provided at global, regional and local levels, facilities for low-level modelling in terms of modifications to individual vertices and facets are currently missing from the system. It would be desirable to have a facility which mimics the process of sculpting in 3D, but which operates through changes to facets and vertices [2]. One can envisage such a process that makes use of virtual reality and its associated 3D interaction devices, such as the data-glove, to allow users to 'sculpt' faces from a 3D ovoid, for example.

Some types of curves such as B-splines allow changes to be made locally without affecting the whole model. This is an extremely useful attribute for interactive modelling. Huitric and Nahas show that through the use of the Oslo algorithm [30], B-splines can be used to model shapes in the same way that a sculptor models clay or plasticine [105].

An alternative would be to use techniques that are already used in CAD/CAM/CAE applications which utilise *solid modelling*. Through the use of *winged-edge* data-structures, in conjunction with *Euler* operators as a facility for amendments, the topological integrity of the model could be constantly maintained [101, 190].

There are other methods which could be investigated, such as free-form deformation [144], fuzzy and soft object modelling, for instance. However, the aim should always be to keep the computational representation of the model transparent to the user, who should only be aware of working with a face in terms of bone, muscle, skin and surface features. Provision of certain rule-based modifications in the form of heuristics [55, 56, 153] would further improve the modelling capability available to the user.

9.5 Animating Faces

The facilities described in chapter 6 provide substantial flexibility and control over the generation of facial animation sequences. Nonetheless, the functionality provided in the ANIMATE sub-system would benefit from the development of several areas, in particular those relating to the creation of expressive facial animation.

9.5.1 Skin Wrinkles

Temporary skin wrinkling occurs naturally when someone laughs, grimaces or frowns. It is an important cue for discrimination between expressions [41]. For example, the forehead wrinkles when showing fear and 'crows-feet' radiate from the eyes when a person smiles. The muscle models implemented do not cater for such effects.

This shortcoming has recently been addressed by Terzopoulos and Waters [177]. They have turned to physically-based modelling and consider all three layers of the skin, which include the *epidermis*, *dermis* and *subcutaneous layer*. A further consideration would be the provision of wrinkles as part of an animated 'ageing' process [165] as opposed to transitory wrinkles which form a part of expressions.

9.5.2 Smoother Motion

Good motion control is vital if any subtlety of expression variation is to be achieved. Linear interpolation lacks smoothness due to velocity discontinuities which may occur in both the path of motion and speed of motion. In general, it is insufficient to use a non-linear law since this provides non-linearity in time only and not in space. The resulting movement could be jerky, uneven and unnatural in appearance.

Refinement of motion control in the form of spline interpolation between keyframe parameters is desirable to give smoother motion [135, 152]. Approximating splines such as B-splines and β-splines have second order continuity and therefore produce very smooth results [79, 80, 128], however such techniques do require more time and computation than linear interpolation. Nonetheless, use of spline interpolation would make it possible to change the nature of the motion merely through modification of the values of bias, tension and continuity, to give the animator finer control.

9.5.3 Timing for Facial Movement

As explained in chapters 4 and 6, there is a distinct lack of information with regard to the timing of facial movement. A further aspect of timing is that facial expressions are not generally a series of fixed poses; the face is continually mobile. Even when not involved in speech and expression the ebb and flow of facial movements must remain plausible at all times. Mere interpolation between static expressions is probably inadequate. The problem can be viewed as one analagous to that of co-articulation in speech synthesis [17, 68, 156] and requires further attention.

9.5.4 A Parser for Action Units

In the interests of providing a general system with flexibility and scope for experimentation, checks are at present not made on the combinations of AUs that are active simultaneously. One enhancement that would be of great use to applications which require realistic constraints on the behaviour of the face, is an AU parser. This would manage conflicting AU activations in an 'intelligent' manner by checking the validity of combinations of particular AUs.

Another aspect of grouping various muscles to represent particular AUs is that the relative contractions of muscles are predetermined within the software. An alternative is to develop a scheme whereby muscles and their contractions can be altered. For example, through use of a textual file containing such information which can be edited. At run time, FACES would then establish the relationships between muscles and AUs by reading in the data-file.

9.6 Rendering Faces

Shading operations provided in the RENDER sub-system enable a large range of faces to be represented, see Plate 7.2. The facilities provided could, nevertheless, be enhanced to provide more visually realistic images of the face. Rendering of realistic faces entails paying special attention to skin, hair and eyes.

9.6.1 Illumination and Shading

To minimise computations, the shading strategy and illumination models used in traditional computer graphics have been employed. For more accurate rendering of the face it would be necessary to take account of subtle effects such as motion blur [133] and possibly shadows under the eyes.

Illumination methods which model light reflection more accurately already exist, for example the *Torrance-Sparrow* [166], *Cook-Torrance* [33] and *Hall-Greenberg* [64] models. Such models could be incorporated into FACES for a more realistic image of the face. Note however, that the calculations tend to be much more compute intensive.

9.6.2 Skin Tone and Texture

At present Gouraud shading [60] is used to generate realistic faces. However, both Gouraud and Phong shading can result in an artifical smoothness which gives the impression that the model is made of plastic rather than natural skin. This is particularly disconcerting for faces that are meant to be human in nature and even more so for male faces.

All physical surfaces have a detailed structure visible to the human eye; this texture provides a tremendous amount of information about the nature of a substance. A means of conveying the texture of skin is required to communicate whether it is smooth or rough, its colour, reflectance and pigmentation.

The most popular method of resolving this problem is through the use of texture-mapping [193] which is an effective way to increase perceived detail. This is becoming a widely used method of overcoming the viewers expectations of seeing a 'real' face [83, 177] and would enhance the present facilities which are available for shading.

9.6.3 Colour for Emotion

One aspect of generating realistic faces is that skin colour tends to change depending on the emotional state of an individual. Opportunities for shading selected parts of the face are currenty limited, only eyebrows, lips, eyes and skin may be varied, placing a restriction on the effects that can be achieved.

The existing facility needs to be extended so that the colour of selected portions of the face can be varied with time to cater for emotional visual cues such as blushing due to embarrassment, or pallor due to fear or anger. A possibility is to use a variation of radiosity [31, 32, 73, 145] which would cater for colour-bleeding between parts of the face.

9.6.4 Hair

Rendering of hair is a major research area in its own right [28, 105, 149] and has therefore not been addressed within the FACES project. Aspects include texture, styles and growth as well as the shape of eyebrows, eyelashes, beards and moustaches.

Hair is often simulated using texture-mapping, but the result does not tend to be realistic. One complication is that highlights on the hair need to change as the head moves from one position to another. An alternative is to use particle systems [136] or *anisotropic reflection models* [76] which have proven effective for rendering hair and fur. Hair is an important factor in distinguishing between individuals, a solution to the problem of rendering realistic hair in all its forms would greatly enhance the system.

9.7 Speech Synthesis

Although speech synthesis has not been a goal of the FACES work to date, it is possible that the system could serve as a vehicle for investigation in this area of research. It would be necessary to extend the head model in FACES by incorporating a tongue and teeth since these components are important for speech analysis and synthesis; in particular for determining mouth shapes.

When modelling synthetic characters, a major enhancement would be to permit them to talk. At present, traditional cartoon animators take a simplified view and use a phonetic break-down of the sound-track as a basis for standard mouth shapes [178]: a, e and i are open vowels, requiring the lips and teeth to be apart; m, b and p are consonants produced with a closed mouth; u, o and w need the mouth to make an oval shape; and f and v sounds require the lower lip to be tucked up under the top set of teeth. This approach is adequate and effective for two main reasons. Firstly, the staging of bodily gestures emphasises the dialogue, so that attention is drawn away from the face and lips. Secondly, cartoon characters are not meant to be realistic and as a result, the observer's expectations are reduced.

Realistic speech animation however, requires more than stylised lip and jaw movements. An inability to deal adequately with variation, intonation, co-articulation and rythm has often produced unnatural results. However, information on how to control the lips, jaw and tongue movements is difficult to acquire. In addition, it is not obvious how it should be applied [15, 17, 53, 68, 123, 156].

Several automated approaches for speech animation have been based on some knowledge of the relationship between speech sounds and the configuration of the lips [68, 123]. In the *Human Factory System* [96] speech synthesis is modelled in terms of 28 basic phonemes each consisting of lip motion and tongue positions.

9.8 User Interface

There is ample room for development of the user interface, described in chapter 8, in terms of improved interaction techniques, feedback, practicality and methods for experimentation. More appropriate input tools could also be provided. For example, many parameters are still input using the keyboard. A more friendly system would result if user aids such as iconised slide-bars and dials were to be provided on the screen. Icons or pictorial representations of actions would improve the interface as well as making it more universally understood.

In the longer term, FACES could itself contribute to the improvement of user interfaces for interactive systems. One can readily envisage a 'talking face' interface that would some day make the keyboard and other input devices obsolete.

9.9 Summary and Conclusions

There are many ways in which FACES could be extended. At a fundamental level improvements which affect the functionality that the system provides could be incorporated. Extensions at a more superficial level could influence how particular operations are presented to the user.

The three-layer head model could be improved by the adoption of a surface patch representation, additional features and modelling capabilities. Simulation of muscle dynamics could be enhanced by taking account of volumetric changes in muscle mass. Facilities for conformation modelling would benefit from the incorporation of skull prototypes, a facility for ageing, a kit-of-parts and a tool for 3D interactive sculpting. The motion aspects of facial movement could be improved through the incorporation of wrinkles in the skin, smoother motion through spline interpolation, additional timing information and an AU parser. Enhancement of rendering facilities could be effected through more sophisticated shading techniques, texture-mapping, an ability to control the colour of selected parts of the face with time and extensions to provide for the rendering of hair. Although speech synthesis entails specific problems, incorporation of a capability to mimic speech would in a sense 'complete' the system.

However, as the system currently stands, the areas that need earliest attention are those of conformation modelling and timing of facial movement for non-verbal communication. Further development of these aspects would increase the range of faces that can be modelled as well as improving the facilities for generating expressive facial animation.

Chapter 10

Concluding Remarks

10.1 Introduction 160
10.2 Synthetic Faces 161
10.3 Anatomy of the Head and Face 161
10.4 Representation of the Head 162
10.5 Dynamics of the Head 162
10.6 Creating Faces 163
10.7 Animating Faces 164
10.8 Rendering Faces 166
10.9 User Interface 167
10.10 Further Development 168
10.11 Summary and Conclusions 169

10.1 Introduction

For years people have been obsessed with reproducing the complex and diverse nature of the face. In this thesis has been described a comprehensive computer based system with facilities for the creation, animation and rendering of faces.

In this concluding chapter major issues and achievements are summarised. Section 10.2 comprises a reminder of the major objectives of the research project and the philosophy behind the approach adopted.

Major points of interest from an anatomical study of the head and face are discussed in section 10.3, while significant characteristics of the representation of the head are summarised in section 10.4. Issues pertinent to the dynamics of facial movement are highlighted in section 10.5.

Sections 10.6, 10.7 and 10.8 are concerned with the functionality that is provided within FACES for modelling, animation and rendering faces. Salient features relevant to the user interface of FACES are addressed in section 10.9, while section 10.10 is concerned with the further development of the system. Finally, section 10.11 concludes the thesis by drawing together the major achievements of the research project.

10.2 Synthetic Faces

The major benefit of a computer based system for modelling and animating faces, is that once the rudimentary facilities have been established, it is possible to adapt the system to many different applications. Moreover, other types of media, such as paint on canvas and plaster, do not lend themselves to modification and animation of the 'model'.

Through an examination of the state of contemporary computer graphics and animation regarding the face, three overall objectives were established for the project. The first concerned conformation modelling of the head and face which is important for identification. The second involved generation of naturalistic facial movement, largely for non-verbal communication. The final goal was to investigate appropriate facilities for the user to enable flexible, efficient and intuitive control over the creation and animation of the face.

The need for generality has played a major rôle in the development of the system. It has been necessary to integrate functionality for the three major objectives of the project into one system, so that a large variety of faces can be created and subsequently animated. Use of reality as a frame of reference from which abstractions can be made has proven to be efficacious, as has the multi-disciplinary approach adopted in carrying out the research.

10.3 Anatomy of the Head and Face

A detailed anatomical study of the human head and face indicated that each of the fundamental substances of bone, muscle and skin affect both appearance and motion. Despite the fact that all human faces have the same general physical structures, there is variability in the size, shape and relative distance of the major bones. This together with variations in the size and shape of muscular tissue determines the gross shape and proportions of the head and face. The skin and surface features hold sway in the recognition process because they are visibly significant parts of the face. The thickness and elasticity of skin varies over the face. Furthermore, elasticity of skin tends to diminish with age resulting in wrinkles.

The study also revealed that the major structures behave in a predefined manner, subject to local variations dependent on individual faces. All facial movement results from muscular contractions. Elasticity of skin and movement of features such as the eyes are also important factors in facial dynamics.

10.4 Representation of the Head

Provision of the envisaged modelling and animation facilities necessitated development of a three-layer anatomical model of the head and face. Incorporation of bone, muscle, skin and surface features provides a basis for the creation of realistic faces and believable facial movement.

The three-layer model described in chapter 3, imparts several major advantages to the FACES project. The model enables the user to operate at an intuitive and abstract level involving use of anatomical descriptions rather than terms used in computer science. In addition, such a model is sufficiently flexible to represent and animate stylistic and realistic faces, see Plate 5.1.

The head model also acts as a foundation which is capable of accommodating both inelastic modifications for changes to the structure of the model and elastic deformations for modelling transitory facial movement. Furthermore, it enables unification of both types of functionality in one software system. The greatest advantage of the three-layer anatomical model is that it has allowed differing strategies to be used for implementing facilities for conformation modelling and animation of the face.

10.5 Dynamics of the Head

Previous research has demonstrated that the generation of convincing facial movement is a difficult task. Early methods were based on kinematics techniques while more recent approaches have employed dynamic analysis to improve the realism of facial motion. The type of model used to generate motion influences the type of control that can be provided for the user.

Development of an animation system for faces necessitated that three issues be addressed. The first required a mechanism for distorting the head model and was resolved through dynamic simulation of muscle contractions. The second issue concerned the provision of a higher and more meaningful level of control than manipulation of individual muscles; this has been solved by the use of FACS. The remaining problem concerned time-variance of facial actions and has been overcome by providing detailed and flexible control over individual actions.

The benefits of using muscular contractions to distort the face model are several. Firstly, muscles are the causes of facial movement which should therefore lead to believable facial movement. Secondly, muscles provide generality since all faces have the same anatomical muscles. Thirdly, it is only necessary to develop models and parameters for the three major types of muscle found in the face.

Many advantages arise from using FACS as the control interface between the user and the low-level muscle models. To start with, it is viable to apply AUs to any face model; such generality is vital to FACES. Secondly, AUs are described in terms of muscular contractions which can be grouped together to make the muscular basis of the face transparent to the user. Thirdly, FACS is a modular system, as a result combinations of AUs can be used to create composite expressions; a convenient and descriptive form of control over the areas of the face to be manipulated.

Relative timing of facial actions is crucial in non-verbal communication. Since hardly any time-variance information is available, it was necessary to decompose this problem so that adequate facilities could be provided for the user. As a result, for each facial action, the user has control over the path of motion, the duration of the motion and the degree or intensity of the movement. These facilities provide considerable and subtle control over facial movement.

10.6 Creating Faces

Although facial animation has received considerable attention from the research community, the requirement for a capability to produce a wide range of faces has been neglected in comparison. Most researchers use a method of digitisation to acquire data for a face model. This entails use of either specialised equipment or tedious and time-consuming manual digitisation or photogrammetry. The greatest drawback however, it that the data acquired is specific to the particular physical model used in the digitisation process.

The problem of modelling a variety of faces is analogous to the general problem of representation and modification of irregular 3D structures. A lack of definitive data to indicate the determinants of facial form appeared to be a major obstruction to the development of suitable modelling facilities.

Nevertheless, exploitation of valuable information from a wide variety of disciplines has enabled functionality for effective conformation modelling to be provided. The information derived has not only influenced the facilities provided, but it has also determined the structure of the CONSTRUCT and MODIFY sub-systems. In particular, it was found that a means of modifying the structure of the head model at global, regional and local levels was required; this is reflected in the configuration of the CONSTRUCT and MODIFY parts of FACES. Within the CONSTRUCT sub-system, modelling facilities are available to cater for changes to the structure of the skull, while within the MODIFY sub-system there is functionality to effect changes to muscles, skin and surface features.

The user is provided with a predefined model of the head and face for several reasons. The first is to provide an initial model which can be modified as desired. Secondly, a predefined model overcomes the problem of having to select isolated features; changes can be viewed within the context of a face. Thirdly, the head model notionally represents an 'average' head.

Controls for the composition and modification of faces are abstractions of the processes used in the reconstruction of realistic faces [141]. Computer reconstruction systems such as that of Vanezis et al. [169] are aimed at reconstruction and identification of real people rather than the creation of new faces. They therefore require a physical representation of the skull which needs to be digitised before prescanned images of people are mapped onto the model. Whereas reconstruction of real human faces involves using the skull as the starting point, FACES goes back one stage to allow the skeletal basis of the face to be altered as a foundation for the creation of completely new faces.

Furthermore, all known automated attempts at realistic reconstructions cater only for static images. In FACES any face constructed is capable of being animated. This could serve as a useful stimulus in the identification task itself.

10.7 Animating Faces

Functionality contained within the ANIMATE sub-system caters for the creation, storage and retrieval of animation sequences. Two types of sequence may be created; they are known as conformation and facial animation sequences. Conformation animation sequences define the metamorphosis of one head model into another, while facial animation sequences are concerned with head and face movement. An animation sequence is first defined and later used to generate a set of frames that correspond to that sequence.

Two levels of control are available over facial actions; these are low-level AUs and high-level expressions. Both of these are eventually converted into muscle contractions which cause distortions on the head model. Creation of a facial animation sequence involves: selection from a set of AUs or expressions; specification of the duration of the chosen action; input of initial and final values for associated parameters; and a choice of law of motion. Such a scheme provides detailed and subtle control over: each area of the face required; the speed of its motion; the degree of movement; and the motion path of the action. This amount of control is necessary due to a lack of readily available time-variance data for facial movement. The approach adopted provides the user with considerable control over the relative timing and duration of facial actions to promote creation of subtle nuances which contribute so much to expressive animation sequences.

In addition to having access to predefined generic expressions, it is also possible for the user to create new expressions by grouping specific AUs together. From the definition of a facial animation sequence it is possible to generate the frames that represent the sequence. This is achieved through a hybrid method involving kinematics for parameter interpolation and dynamics for the simulation of muscular contractions. Individual frames for a conformation animation sequence are generated through the kinematic technique of shape interpolation.

The two types of animation sequence may be saved to disk files for later use. The advantage of using a keyword based textual format is that the technique enables the scripts to be readily understood and modified by the user. It is also possible to either overwrite an old sequence by reading in a new one, or to merge the old and new sequences to form the new current sequence.

Conformation animation sequences and facial animation sequences exist as separate entities, so that a facial animation sequence may be applied to either a single head model or to a sequence of interpolated head models. Such generality is available because of the generic nature of FACS.

10.8 Rendering Faces

Functionality in the RENDER sub-system is concerned with more than mere shading of facial images. Operations are provided which cater for colouration of facial features, examination of groups of frames corresponding to animation sequences and appraisal of motion.

A head model must pass through several time-consuming stages of the *rendering pipeline* before a shaded version of the face can be displayed. Many of the calculations involved in the processes are computationally intensive. An overriding concern in the implementation of FACES has been execution efficiency since the system was intended to be interactive. An advanced knowledge of the contents of images, the structure of the head model and the primitives to be rendered has enabled a substantial reduction to be made in the amount of computation required. Consequently, each image is rendered and displayed within a few seconds to provide rapid visual feedback to the user.

Colour is an important factor for conveying both appearance and emotion. Interactive control is therefore available over the colour of skin, hair, eyes and lips; hair only affects eyebrows at present. In addition, it is possible to regulate which particular frames are to be rendered by specifying the first and last frame numbers. This enables the user to experiment with the colouration of the face before a large number of frames are shaded.

An ability to view groups of consecutive frames in a sequence is of tremendous worth during the production of an animation sequence. It enables monitoring of changes that occur between successive frames of animation. Control is also available over sub-sequences to allow long sequences to be inspected.

Timing of facial movement is critical to non-verbal communication. It is therefore essential to provide a means of testing and evaluating the motion characteristics of a sequence. Once again, for flexibility, it is possible to invoke this facility on sub-sequences of a sequence. Two modes of display enable either near real time or single-step display of frames, with further controls in the *Frame-Rate* option for varying the speed of update. Both the STORYBOARD and PLAYBACK facilities are capable of operating on wireframe images as well as shaded frames. The advantage of this is that it is possible to check the sequence and its motion before individual frames are rendered.

10.9 User Interface

User interface design can be contemplated at two levels. Firstly, at a fundamental level it is necessary to provide facilities which are appropriate to a particular application in a manner which complies with the user's conceptual model of that application. At the second level it is necessary to consider how the operations can be presented to the user so as to provide effective control. Development of an efficacious user interface encompasses consideration of many human factors.

Functionality has been provided for creating different faces and for generating a large range of expressions on them. The major issue regarding these operations is the amount of control available to the user.

Use of an anatomical head model together with the use of dynamics to simulate movement provides for naturalistic and intuitive control as well as realistic results. Modelling control is provided over the skull, muscles, skin and surface features. Several types of control at varying levels are available. Changes to the broad nature of the head and face are possible through global control. Relative proportions can be altered through regional modifications, while changes constrained to individual bones and features provide local control.

With regard to the animation of faces, it is possible to create two types of sequence, namely conformation and facial animation. Controlling the motion of models which are as complex as the face is liable to be a tricky and time-consuming process. Emphasis has therefore been placed on making control for the user as compliant as possible.

Creation of sequences is simplified since activities do not need to be specified in frame or time order. Use of parameter interpolation and minimisation of parameters reduces the amount of user input required. Facial movement may be specified using descriptive basic actions or through an 'emotional' level using predefined expressions. Consequently, the muscular basis of the face is hidden from the user. Additonal direction is available over the interval of duration, law of motion and degree of movement of each facial action. Such detailed capabilities give the user generous control over regions of the face and their motion characteristics.

Other factors which are considered to be important to the user interface of FACES are interaction, feedback, experimentation and practicality of the system. Interaction with FACES is accomplished through fixed and pop-up menus. This strategy has several advantages. Menus remove the need for the user to remember input sequences and options, and thereby reduce the chance that the user will make a mistake. Use of two types of menu allows the system to be configured as partly modeless and partly guided to provide flexibility in the manner in which the user works.

Feedback to the user is vital in any interactive system. To help the user to keep track of processing, a menu selection remains highlighted until the option is exited. Modelling facilities for creating faces and high-level expressions all display a wireframe version of the model.

Error recovery is an important component in an interactive system. In FACES this is largely achieved through an undo operator and an exit option in pop-up menus. For practicality, FACES automatically saves modified head structures and animation scripts to default files behind the scenes. In addition, there are utilities separate to the system which cater for rendering, story board and playback operations.

10.10 Further Development

Functionality for modelling, animating and rendering the face is extensive, but by no means exhaustive. There are many ways in which the system could be enhanced and developed further. Potential courses of action were outlined in chapter 9 for each of the major areas of research that FACES addresses.

Expression and behaviour animation can help in the development of a character's personality. The use of a topology independent model for simulating facial movement opens up the possibility to endow human emotions and personalities on inanimate objects for character animation.

The natural path of development should lead towards the integration of facial movement, speech, hand and body gestures. So that, it may one day be possible to give some measure of autonomy to the characters we simulate by defining a repertoire of skills that they can carry out. For example, Breen and Wozny describe an approach to choreography in computer animation, in the form of a message-passing technique [14], in which *actors* possess their own behaviour characteristics. The rôle of the user will then become analagous to that of a director rather than an animator, with a task manager accepting high-level task descriptions [196].

10.11 Summary and Conclusions

FACES is an interactive system aimed at rapid feedback to increase productivity. The underlying philosophy is that of a system for rapid prototyping. This approach has the potential to save the user a considerable amount of both time and effort.

The system is based on the principle that software should not attempt to replace the skills of the animator, but provide a variety of flexible tools to assist in the creation of numerous effects. The computer should be a medium which animators can exploit and use according to their own individual skills. Any system that helps someone to perform a task which would otherwise be beyond their capabilities must surely be beneficial. FACES makes it possible for an untrained user to create expressive faces, as well as enhancing the talents of already skilled animators.

It has been confirmed that the use of a three-layer model based on the anatomy of the face is a flexible and appropriate method of integrating the dual functions of modelling and animation. Conformation modelling and facial animation are treated as being separate and distinct. This separation of functionality allows differing strategies to be used for modelling and animation, while using the three-layer representation as a common foundation.

FACES is a hybrid system in the sense that a parameterised approach is used for conformation modelling while dynamics is used to simulate facial motion. The trend is now towards physically based modelling to provide naturalistic, realistic and accurate effects which can be incorporated into computer software. This trend is necessary if the eventual goal is the creation of convincing synthetic actors.

FACES encapsulates a considerable amount of information regarding facial shape, movement and expression. It has tremendous scope; the applications are numerous and vary greatly. Man's attempts to create a likeness of himself has been a starting point for many different disciplines such as art, portraiture, sculpture, medicine, robotics and recently computer animation. Our virtual worlds would be desolate indeed without characters that we can relate to and understand.

Bibliography

[1] J.F. Alexander. Search factors influencing personal appearance identification. In *Personal Appearance Identification*, pages 14–27. Springfield, Illinois, USA, 1972.

[2] J.R. Allen, B. Wyvill, and I.H. Waite. A methodology for direct manipulation of polygon meshes. *Proceedings Computer Graphics International '89*, pages 451–470, 1989.

[3] N. Badler. Digital representation of human movements. *ACM Computing Surveys*, 11(1):19–38, March 1979.

[4] N. Badler. Human body models and animation. *IEEE Computer Graphics and Applications*, 2(9):6–7, November 1982.

[5] N. Badler. Positioning and animating figures in a task-oriented environment. *The Visual Computer*, 1(4):212–220, 1985.

[6] N. Badler. Articulated figure animation. *IEEE Computer Graphics and Applications*, 7(6):10–11, June 1987.

[7] N. Badler, J. O'Rourke, and B. Kaufman. Special problems in human movement simulation. *Proceedings ACM SIGGRAPH, Computer Graphics*, 14(3):189–203, July 1980.

[8] P. Bahn. Skull fit for Mozart. The Sunday Correspondent, 22nd July 1990.

[9] E.W. Bethel and S.P. Uselton. Shape distortion in computer-assisted keyframe animation. *Proceedings Computer Animation '89*, pages 215–224, June 1989. Geneva, Switzerland.

[10] L. Birdwhistle. *Kinesics and Control.* University of Pennsylvania Press, Philadelphia, 1970.

[11] J.F. Blinn, L. Carpenter, J. Lane, and T. Whitted. Scan line methods for displaying parametrically defined surfaces. *Communications of the ACM*, 23(1):23–34, 1980.

[12] F.L. Bookstein. Principal Warps: Thin-plate splines and the decomposition of deformations. *IEEE-Transactions on Pattern Recognition and Machine Intelligence*, 11(6):567–585, 1989.

[13] C. Bragdon. *The Beautiful Necessity*. Reprinted Quest, London, 1939.

[14] D.E. Breen and M.E. Wozny. Message-based choreography for computer animation. *Proceedings Computer Animation '89*, pages 69–82, June 1989. Geneva, Switzerland.

[15] N.M. Brooke. Visible speech signals: Investigating their analysis, synthesis and perception. In *The Structure of Multimodal Dialogue*. Elsevier Science Publishers (North-Holland), 1989.

[16] N.M. Brooke. Computer graphics synthesis of talking faces. In *Tutorial Day on Speech Synthesis*, Autrans, France, September 1990.

[17] N.M. Brooke and Q. Summerfield. Analysis, synthesis and perception of visible articulatory movements. *Journal of Phonetics*, 11:63–76, 1982.

[18] V. Bruce and A.M. Burton. Modelling face recognition. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[19] V. Bruce and T. Valentine. When a nod's as good as a wink: the role of dynamic information in facial recognition. *Practical Aspects of Memory: Current Research and Issues*, 1:169–174, 1988.

[20] V. Bruce and A. Young. Understanding facial recognition. *British Journal of Psychology*, 77:305–327, 1986.

[21] Phong Bui-Tuong. Illumination for computer generated pictures. *Communications of the ACM*, 18:311–317, 1975.

[22] T.W. Calvert. The challenge of human figure animation. *Proceedings Graphics Interface '88*, pages 203–210, 1988.

[23] T.W. Calvert, J. Chapman, and A. Patla. The integration of subjective and objective data in the animation of human movement. *Proceedings ACM SIGGRAPH, Computer Graphics*, 14(3):198–203, August 1980.

[24] T.W. Calvert, J. Chapman, and A. Patla. Aspects of the kinematic simulation of human movement. *IEEE Computer Graphics and Applications*, 2(9):41–50, November 1982.

[25] S. Carey. Developmental changes in face recognition. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[26] I. Carlbom and J. Paciorek. Planar geometric projection and viewing transformations. *ACM Computing Surveys*, 10(4):465–502, December 1978.

[27] A.J.E. Cave. Head to head. The Times, 12th March 1988.

[28] H. Chen and E.-H. Wu. Radiosity for furry surfaces. *Proceedings EUROGRAPHICS '91*, pages 447–458, September 1991. Vienna, Austria.

[29] A. Christopher. The making of brilliance. *Computer Graphics World*, 8(7):32–36, July 1985.

[30] E. Cohen and T. Lyche. Discrete B-splines and subdivision techniques in CAGD and computer graphics. *Computer Graphics and Image Processing*, 14:87–11, 1980.

[31] M.F. Cohen and et al. A progressive refinement approach to fast radiosity image generation. *Proceedings ACM SIGGRAPH, Computer Graphics*, 22:75–84, 1988.

[32] M.F. Cohen and D.P. Greenberg. The hemi-cube—a radiosity solution for complex environments. *Proceedings ACM SIGGRAPH, Computer Graphics*, 19(3):31–40, 1985.

[33] R.L. Cook and K.E. Torrance. A reflectance model for computer graphics. *Proceedings ACM SIGGRAPH, Computer Graphics*, 15:307–316, 1981.

[34] A.M. Coombes, A.D. Linney, J.P. Moss, and R. Richards. A method for the analysis of the 3D shape of the face and changes in the shape brought about by facial surgery. *Proceedings SPIE Conference on Biostereometrics & Applications*, November 1990. Boston, MA.

[35] A.M. Coombes, J.P. Moss, A.D. Linney, and R. Richards. A mathematical method for the comparison of three dimensional changes in the facial surface. *European Journal of Orthodontics*, 13:95–110, January 1991.

[36] S. Couquillart. A control point based sweeping technique. *IEEE Computer Graphics and Applications*, 7(11):36–45, November 1987.

[37] S.M. Coveney. ME input helps investigators speed human identification. ASME News, May 1989.

[38] A.K. Dewdney. Computer recreations —the complete computer caricaturist and a whimsical tour of face space. *Scientific American*, 255(4):20–28, October 1986.

[39] Dorland. *Illustrated Medical Dictionary.* Saunders, 1981. 26th Edition.

[40] P. Ekman. Asymmetry in facial expression. *Science*, 209:833–834, 1980.

[41] P. Ekman. About brows: Emotional and conversational signals in Human ethology. *Journal of Phonetics*, 11:162–202, 1983.

[42] P. Ekman. The deceptive face. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[43] P. Ekman and W. Friesen. *Unmasking the Human Face.* Prentice-hall Inc., 1975.

[44] P. Ekman and W. Friesen. *Manual for the Facial Action Coding System.* Consulting Psychologist Press, Palo Alto, California, 1978.

[45] P. Ekman and W.V. Friesen. Measuring facial movement. *Environmental Psychology and Nonverbal Behaviour*, 1(1):56–75, 1976.

[46] H.D. Ellis. An investigation of the use of the Photofit technique for recalling faces. *British Journal of Psychology*, 66:29–37, 1975.

[47] H.D. Ellis. A critical examination of the Photofit system for recalling faces. *Ergonomics*, 21:297–307, 1978.

[48] A. Emmet. Digital portfolio: Tony de Peltrie. *Computer Graphics World*, 8(10):72–77, October 1985.

[49] E.J. Farrell, W.C. Yanga, and R.A. Zappulla. Animated 3D CT imaging. *IEEE Computer Graphics and Applications*, 6(7):26–31, July 1986.

[50] J.D. Foley, A. Van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice.* Addison-Wesley, Reading, Massachusetts, USA, 1990.

[51] L. Forest, N. Magnenat-Thalmann, and D. Thalmann. Integrating keyframe animation and algorithmic animation of articulated bodies. *Proceedings Computer Graphics '86*, pages 263–273, 1986.

[52] M. Fortin, J.F. Lamy, and D. Thalmann. A multiple track animator system for motion synchronisation. ACM Motion Representation and Perception, 1986.

[53] V. Fromkin. Lip positions in English American vowels. *Journal of Language and Speech*, 7:215–225, 1964.

[54] B.P. Gatliff and C.C. Snow. From skull to visage. *Journal of Biocommunication*, 6:27–30, 1979.

[55] R.M. George. The lateral craniographic method of facial reconstruction. *Journal of Forensic Science*, 32:1305–1330, 1987.

[56] M.M. Gerasimov. *The Face Finder.* Hutchinson of London, 1971.

[57] M.L. Gillenson. *The Interactive Generation of Facial Images on a CRT using a Heuristic Strategy.* PhD Thesis, Department Computer Science, Ohio State University, 1974.

[58] D. Goleman. The 7000 faces of Dr. Ekman. *Psychology Today*, 15:42–49, 1981.

[59] C.M. Goral, K.E. Torrance, D.P. Greenberg, and B. Battaile. Modeling the interaction of light between diffuse surfaces. *Proceedings ACM SIGGRAPH, Computer Graphics*, 18(3):213–222, July 1984.

[60] H. Gouraud. Continuous shading of curved surfaces. *IEEE Transactions on Computers*, 20:263–628, 1971.

[61] J.P. Gourret, N. Magnenat-Thalmann, and D. Thalmann. The use of finite element theory for simulating object and human body deformations and contacts. *Proceedings EUROGRAPHICS '89*, pages 477–488, September 1989. Hamburg, Germany.

[62] H. Gray, R. Warwick, and P.L. Williams. *Gray's Anatomy.* 35th Edition, 1973.

[63] B. Guenter. A system for simulating human facial expression. *Proceedings Computer Animation '89*, pages 191–202, June 1989. Geneva, Switzerland.

[64] R.A. Hall and D.A. Greenberg. A testbed for realistic image synthesis. *IEEE Computer Graphics and Applications*, 3(8):10–19, November 1983.

[65] N. Hammond. Researcher 'recreates' the face of Mozart. The Times, March 15th 1991.

[66] T. Harada, K. Konno, and H. Chiyokura. Variable-radius blending by using gregory patches in geometric modeling. *Proceedings EUROGRAPHICS '91*, pages 507–518, September 1991. Vienna, Austria.

[67] L.D. Harmon. The recognition of faces. *Scientific American*, 29(5):70–82, November 1973.

[68] D.R. Hill, A. Pearce, and B. Wyvill. Animating speech: an automated approach using speech synthesised by rules. *The Visual Computer*, 3(5):277–289, 1988.

[69] T.M. Hong, N. Magnenat-Thalmann, and D. Thalmann. A general algorithm for 3D shape interpolation in a facet-based representation. *Proceedings Graphics Interface '88*, pages 229–235, June 1988.

[70] P. Houghton. Rocker jaws. *American Journal of Physical Anthropology*, 47:365–369, 1964.

[71] H. Huitric and M. Nahas. B-spline surfaces: A tool for computer painting. *IEEE Computer Graphics and Applications*, 5(3):39–47, March 1985.

[72] A. Hutchinson. *Labanotation.* Theatre Art Books, New York, 1970.

[73] D.S. Immel, M.F. Cohen, and D.P. Greenberg. A radiosity method for non-diffuse environments. *Proceedings ACM SIGGRAPH, Computer Graphics*, 20(4):133–142, August 1986.

[74] M.Y. Iscan and K.A.R. Kennedy. *Reconstruction of Life from the Skeleton.* Alan R. Liss Inc., New York, 1989.

[75] Caplin & Jeremy. Camera obscura. Weekend Guardian, 17th November 1990.

[76] J.T. Kajiya. Anisotropic reflection models. *Proceedings ACM SIGGRAPH, Computer Graphics*, 19(3):15–21, 1985.

[77] A. Kaul and J. Rossignac. Solid interpolating deformation: Construction and animation of PIPS. *Proceedings EUROGRAPHICS '91*, pages 493–505, September 1991. Vienna, Austria.

[78] T. Kitson. Personal communication, February 1989. Scientific Research and Development Centre, Home Office, London, UK.

[79] D. Kochanch and R. Bartels. Interpolating splines for keyframe animation. *Proceedings Graphics Interface '84*, pages 41–42, 1984.

[80] D. Kochanch and R. Bartels. Interpolating splines with local tension, continuity and bias control. *Proceedings ACM SIGGRAPH, Computer Graphics*, 18(3):33–41, July 1984.

[81] K. Komatsu. Human skin model capable of natural shape variation. *The Visual Computer*, 3(5):265–271, 1988.

[82] J.U. Korein and N. Badler. Techniques for generating the goal-directed motion of articulated structures. *IEEE Computer Graphics and Applications*, 2(9):33–41, November 1982.

[83] T. Kurihara and K. Arai. A transformation method for modeling and animation of the human face from photographs. *Proceedings Computer Animation '91*, pages 45–58, April 1991. Geneva, Switzerland.

[84] J.M. Lane and R.F. Risenfeld. A theoretical development for the computer generation and display of piecewise polynomial surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(1):35–46, 1980.

[85] C. Langmaid. Bare bones to new faces. The Times, 8th March 1988.

[86] W.F. Larrabee. A Finite Element Method of skin deformation: I Biomechanics of skin and soft tissue: A review; II An experimental model of skin deformation; III The finite element model. *Laryngoscope*, 96:399–405, 406–412, 413–419, April 1986.

[87] J. Lasseter. Principles of traditional animation applied to 3D computer animation. *Proceedings ACM SIGGRAPH, Computer Graphics*, 21(4):35–44, July 1987.

[88] J. Lasseter. Personal communication, 3rd November 1989. International Animation Festival, Bristol, UK.

[89] K.R. Laughery. Photograph type and cross-racial factors in facial identification. In *Personal Appearance Identification*, pages 36–43. Springfield, Illinois, USA, 1972.

[90] J.P. Lewis and F.I. Parke. Automated lip-synch and speech synthesis for character animation. *Proceedings Human Factors in Computing Systems and Graphics*, pages 143–147, 1987.

[91] D.S. Lucinao, A.J. Vander, and J.H. Sherman. *Human Anatomy and Physiology: Structure and Function.* McGraw-Hill, 1978. 2nd. Edition.

[92] G. MacNicol. Animating motion. *Computer Graphics World*, 11(9):44–50, September 1988.

[93] G. MacNicol. Look Alive. *Computer Graphics World*, 12(4):63–68, April 1989.

[94] N. Magnenat-Thalmann and D. Thalmann. *Computer animation theory and practice.* Springer-Verlag, 1985.

[95] N. Magnenat-Thalmann and D. Thalmann. The direction of synthetic actors in the film 'Rendezvous á Montréal'. *IEEE Computer Graphics and Applications*, 7(12):9–19, December 1987.

[96] N. Magnenat-Thalmann and D. Thalmann. Abstract muscle action procedures for human face animation. *The Visual Computer*, 3(5):290–297, 1988.

[97] N. Magnenat-Thalmann and D. Thalmann. Construction and animation of a synthetic actress. *Proceedings EUROGRAPHICS UK '88*, pages 55–66, March 1988. Manchester, UK.

[98] N. Magnenat-Thalmann and D. Thalmann. Course notes: The problematics of human prototyping and animation. EUROGRAPHICS UK '89, 1989.

[99] N. Magnenat-Thalmann and D. Thalmann. Synthetic actors: The simulation of human movement. *The Computer Bulletin*, 1(1):12–14, January/February 1989.

[100] N. Magnenat-Thalmann and D. Thalmann. *Synthetic Actors in Computer Generated 3D Films.* Springer-Verlag, 1990.

[101] M. Mantyla and R. Sulonen. GWB: A solid modeler with Euler operations. *IEEE Computer Graphics and Applications*, 2(7):17–31, September 1982.

[102] R.R. Martin. Automatic generation of smooth surfaces. Department Computing Mathematics, University of Wales, College of Cardiff, Wales, 1990.

[103] J.P. Moss, A.M. Coombes, A.D. Linney, and J. Campos. Methods of three dimensional analysis of patients with asymmetry of the face. *Finnish Dental Journal*, 87(1), 1991.

[104] J. Musgrave. Personal communication, June 1990. Lecturer in Anatomy and Anthropology, University of Bristol, UK.

[105] M. Nahas, H. Huitric, and M. Saintourens. Animation of a B-spline figure. *The Visual Computer*, 3(5):272–276, 1988.

[106] R.A.H. Neave. Reconstruction of the heads of three ancient Egyptian mummies. *Journal of Audiovisual Media in Medicine*, 2:156–164, 1979.

[107] R.A.H. Neave. Facial reconstruction (abstract). *Journal of the Forensic Society*, 24(4):355, 1984.

[108] R.A.H. Neave. Facial reconstruction: A study of its accuracy and reproducibility. 69(2):244, 1986. Abstract.

[109] R.A.H. Neave. Reconstruction of the skull and the soft tissue of the head and face of 'Lindow man'. *Canadian Society Forensic Science Journal*, 22(1):43–53, 1989.

[110] R.A.H. Neave. Personal communication, June 1990. Director of Medical Illustration, University of Manchester Medical School, UK.

[111] W.M. Newman and R.F. Sproull. *Principles of Interactive Computer Graphics.* McGraw-Hill, 1979.

[112] T. Nishita, I. Okamura, and E. Nakamae. Shading models for point and linear sources. *ACM Transactions on Graphics*, 4(2):124–146, April 1985.

[113] Royal Academy of Arts. *Leonardo Da Vinci: Anatomical Drawings from the Royal Collection.* Royal Academy of Arts, 1977.

[114] A. Paouri, N. Magnenat-Thalmann, and D. Thalmann. Creating realistic three-dimensional human shape characters for computer-generated films. *Proceedings Computer Animation '91*, pages 89–99, April 1991. Geneva, Switzerland.

[115] F.I. Parke. Computer generated animation of faces. *Proceedings ACM National Conference*, 1:451–457, August 1972.

[116] F.I. Parke. A parametric model for human faces. PhD Dissertation, University of Utah, 1974.

[117] F.I. Parke. A model for human faces that allows speech synchronized animation. *Computers & Graphics*, 1:3–4, 1975.

[118] F.I. Parke. Parameterized models for facial animation. *IEEE Computer Graphics and Applications*, 2(9):61–68, November 1982.

[119] F.I. Parke. Parameterized facial animation revisited. SIGGRAPH 1990, Course Notes 26: State of the Art in Facial Animation, 1990.

[120] F.I. Parke. Control parameterization for facial animation. *Proceedings Computer Animation '91*, pages 3–13, April 1991. Geneva, Switzerland.

[121] M. Patel and R. J. Hubbold. A scanline method for solid model display. *Computer Graphics Forum*, 6(2):141–150, May 1987.

[122] E.C. Patterson, P.C. Litwinowicz, and N. Greene. Facial animation by spatial mapping. *Proceedings Computer Animation '91*, pages 31–44, April 1991. Geneva, Switzerland.

[123] A. Pearce, B. Wyvill, G. Wyvill, and D. Hill. Speech and expression: a computer solution to face animation. *Proceedings Graphics Interface '86*, pages 136–140, 1986.

[124] D.E. Pearson. The extraction and use of facial features in low bit-rate visual communication. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[125] C. Pelachaud. Linguistic issues in facial animation. *Proceedings Computer Animation '91*, pages 15–30, April 1991. Geneva, Switzerland.

[126] Q.S. Peng. An algorithm for finding the intersection lines between two B-spline surfaces. *Computer Aided Design*, 16(4):191–196, July 1984.

[127] D.I. Perrett. Functions and organisation of visual cell processing faces in the temporal cortex. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[128] L. Piegl and W. Tiller. Curve and surface constructions using rational B-splines. *Computer Aided Design*, 19(9):485–498, November 1987.

[129] S.M. Pieper. More than skin deep: Physical modeling of facial tissue. MSc Thesis, MIT Media Lab, Cambridge MA, 1989.

[130] S.M. Platt. A structural model of the human face. PhD Thesis, University of Pennsylvannia, Philadelphia, 1985.

[131] S.M. Platt and N.I. Badler. Animating facial expressions. *Proceedings ACM SIGGRAPH, Computer Graphics*, 15(3):245–252, August 1981.

[132] S. Porter. Made for the stage. *Computer Graphics World*, 13(8):60–69, August 1990.

[133] M. Potmesil and I. Chakravarty. Modelling motion blur in computer generated images. *Proceedings ACM SIGGRAPH, Computer Graphics*, 17(3):389–399, July 1983.

[134] X. Pueyo and D. Tost. Survey of computer animation. *Computer Graphics Forum*, 7(4):281–300, December 1988.

[135] W.T. Reeves. Inbetweening for computer animation utilizing moving point constraints. *Proceedings ACM SIGGRAPH, Computer Graphics*, 15(3):263–269, August 1981.

[136] W.T. Reeves. Particle systems —a technique for modeling a class of fuzzy objects. *Proceedings ACM SIGGRAPH, Computer Graphics '83*, 17(3):359–376, July 1983.

[137] W.T Reeves. Simple and complex facial animation: Case studies. ACM SIGGRAPH 1990, Course Notes 26: State of the Art in Facial Animation, August 1990.

[138] J.S. Rhine and C.E. Moore. Facial reproduction tables of facial tissue thickness of American Caucasoids in forensic anthropology. Technical series 1, Maxwell Museum, Alberque, New Mexico, 1982.

[139] G. Rhodes and S. Brennan. Identification and ratings of caricatures: implication for mental representations of faces. *Cognitive Psychology*, 19:473–497, 1987.

[140] B. Robertson. Mike, the talking head. *Computer Graphics World*, 11(7):57–59, July 1988.

[141] A. Roy. Making the face fit. Sunday Times Magazine, May 1990.

[142] K. Scherer and P. Ekman. *Expression and the Nature of Emotion, in Approaches to Emotion.* 1984.

[143] K. Scherer and P. Ekman. *Facets of Emotion: Recent Research.* Lawrence Erlbaum Associates, 1988.

[144] T.W. Sederberg and S.R. Parry. Free-form deformation of solid geometric models. *Proceedings ACM SIGGRAPH, Computer Graphics*, 20(4):151–160, August 1980.

[145] M-Z. Shao, Q-S. Peng, and Y-D. Liang. A new radiosity approach by procedural refinement for realistic image synthesis. *Proceedings ACM SIGGRAPH, Computer Graphics*, 22(4):93–102, August 1988.

[146] P. Shipman, A. Walker, and D. Bichell. *The Human Skeleton.* Harvard University Press, 1985.

[147] A. Silverstein. *Human Anatomy and Physiology.* Wiley & Sons Inc., 1980.

[148] G. Sims. Some like it hot, with graphics. The Times, 1st March 1990.

[149] R.L. Skinner. Modelling hair with structured particle systems. Technical Report UCSC-CRL-89-26, Computer Research Laboratory, University of California, June 1989.

[150] S. Smith, editor. *How to Draw and Paint.* Quatro Publishing Limited, 1981.

[151] P. Sorenson. Felix the cat: Real time computer animation. *Animation Magazine*, pages 16–17, Winter 1989.

[152] T. Spencer-Smith and G. Wyvill. Four dimensional splines for motion control in computer animation. *Proceedings Computer Animation '89*, pages 153–168, June 1989. Geneva, Switzerland.

[153] T.D. Stewart. The points of attachment of the palpebral ligaments: Their use in facial reconstructions on the skull. *Journal of Forensic Sciences*, 28(4):858–863, 1983.

[154] W.L. Strauss and A. Durer. *The Human Figure: The Complete 'Dresden Sketchbook'.* 1972.

[155] J.D. Subtelny and M.S. Rochester. A longitudinal study of soft tissue facial structures and their profile characteristics, defined in relation to underlying skeletal structures. *American Journal of Orthodontics*, 45(7):481–507, July 1959.

[156] Q. Summerfield. Audio-visual speech perception, lip-reading and artificial simulation. *Hearing Science and Hearing Disorders*, pages 63–76, 1986.

[157] I.E. Sutherland, F. Sproull, and R.A. Schumacker. A characterisation of ten hidden surface algorithms. *Computing Surveys*, 6:1–55, 1974.

[158] V. Sutton. *Sutton Movement Shorthand: Writing Tools for Research.* The Movement Shorthand Press, Newport Beach CA, 1978.

[159] D. Terzopoulos and K. Waters. Physically-based facial modelling, analysis and animation. *Journal of Visualization and Computer Animation*, 1:73–80, 1990.

[160] D. Terzopoulos and K. Waters. Techniques for realistic facial modeling and animation. *Proceedings Computer Animation '91*, pages 59–74, April 1991. Geneva, Switzerland.

[161] D. Thalmann. Motion control: From keyframe to task-level animation. *Proceedings Computer Animation '89*, pages 3–18, June 1989. Geneva, Switzerland.

[162] F. Thomas and O. Johnston. *Disney Animation—The Illusion of Life.* Abbeville Press, New York, 1981.

[163] D'Arcy W. Thompson. *On Growth and Form.* Cambridge University Press, London, UK, 1961. Abridged Version.

[164] W. Tiller. Rational B-splines for curve and surface representation. *IEEE Computer Graphics and Applications*, 3(6):61–69, September 1983.

[165] J.T. Todd, S.M. Leonard, R.E. Shaw, and J.B. Pittenger. The perception of human growth. *Scientific American*, 242:106–115, February 1980.

[166] K.E. Torrance and E.M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57:1105–1114, 1979.

[167] J.R. Truckenbrod. Effective use of colour in computer graphics. *Proceedings ACM SIGGRAPH, Computer Graphics*, 15(3):83–90, August 1981.

[168] J. Trux. The healing image. The Sunday Times Magazine, 19th March 1989.

[169] P. Vanezis, R.W. Blowes, A.D. Linney, A.C. Tan, R. Richards, and R.A.H. Neave. Application of 3D computer graphics for facial reconstruction and comparison with sculpting techniques. *Forensic Science International*, 42(1–2):69–84, 1988.

[170] P. Voci. Not-art digital images: An artist's perspective. Digital Images—Digital Cinema: ACM SIGGRAPH 1990 Art Show Catalogue, August 1990.

[171] C. Waite. The facial action control editor, FACE: A parametric facial expression editor for computer generated animation. MSc Thesis, MIT Media Lab, Massachusettes, January 1989.

[172] G. Walker and C. Kowalski. A 2D coordinate model for qualification description, analysis, prediction and simulation of craniofacial growth. *Growth*, 35:191–211, 1971.

[173] D.R. Warn. Lighting controls for synthetic images. *Proceedings ACM SIGGRAPH, Computer Graphics*, 17(3):13–21, 1983.

[174] K. Waters. Laugh, I almost cried. *Online Publications, Pinner*, 1986.

[175] K. Waters. A muscle model for animating three-dimensional facial expressions. *Proceedings ACM SIGGRAPH, Computer Graphics*, 21(4):17–24, 1987.

[176] K. Waters. A muscle model for animating three-dimensional facial expressions. PhD Thesis, Middlesex Polytechnic, UK, June 1988.

[177] K. Waters and D. Terzopoulos. The computer synthesis of expressive faces. In *Proceedings Discussion Meeting on Processing the Facial Image*, The Royal Society, London, U.K., July 1991.

[178] H. Whitaker and J. Halas. *Timing for animation.* Focal Press Ltd., London, 1981.

[179] T. Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23:343–349, 1980.

[180] J. Wilhelms. Towards automatic motion control. Technical Report UCSC-CRL-86-13, University of California, Santa Cruz, June 1986.

[181] J. Wilhelms. Using dynamic analysis for realistic animation of articulated bodies. Technical Report UCSC-CRL-86-14, University of California, Santa Cruz, June 1986.

[182] J. Wilhelms. VIRYA–a motion control editor for kinematic and dynamic animation. *Proceedings Graphics Interface '86*, pages 141–146, 1986.

[183] J. Wilhelms. Dynamics for everyone. Technical Report UCSC-CRL-87-6, University of California, Santa Cruz, May 1987.

[184] J. Wilhelms. MANIKIN: Dynamic analysis for articulated body manipulation. Technical Report UCSC-CRL-87-2, University of California, Santa Cruz, April 1987.

[185] J. Wilhelms. Toward automatic motion control. *IEEE Computer Graphics and Applications*, 7(4):11–22, April 1987.

[186] J. Wilhelms. Using dynamic analysis for realistic animation of articulated bodies. *IEEE Computer Graphics and Applications*, 7(6):12–27, June 1987.

[187] J. Wilhelms and B.A. Barsky. Using dynamic analysis to animate articulated bodies such as humans and robots. *Proceedings Graphics Interface '85*, pages 97–104, May 1985.

[188] L. Williams. Performance driven facial animation. *Proceedings ACM SIGGRAPH, Computer Graphics*, 20(3):235–242, August 1990.

[189] P.J. Willis. Gigalib—a family of graphics libraries. Computer Graphics Group, 1990. School of Mathematical Sciences, University of Bath, AVON, UK.

[190] P.R. Wilson. Euler formulae and geometric modelling. *IEEE Computer Graphics and Applications*, 5(8):24–36, August 1985.

[191] P.D. Wood. An interactive graphics system for planning reconstructive surgery. *National Proceedings Computer Graphics Association*, pages 130–135, 1986.

[192] J.K. Woodwark. *Computing Shape.* Butterworths, London, 1986.

[193] J.F.S. Yau and N.D. Duffy. A texture mapping approach to 3D facial image synthesis. *Proceedings EUROGRAPHICS UK '88*, pages 17–30, April 1988. Brighton.

[194] A. Zavala and J.J. Paley. *Personal Appearance Identification.* Springfield, Illinois, USA, Illinois, USA, 1972.

[195] R.T. Zavala. Determination of facial features used in identification. In *Personal Appearance Identification*, pages 44–54. Springfield, Illinois, USA, 1972.

[196] D. Zeltzer. Motor control techniques for figure animation. *IEEE Computer Graphics and Applications*, 2(9):53–60, November 1982.

[197] D. Zeltzer. Towards an integrated view of 3D computer animation. *The Visual Computer*, 1(4):249–259, 1985.

Appendix A

Origins and Insertions of Facial Muscles

Muscle	Origin	Insertion	Action
Muscles of the Scalp			
Occipitofrontalis: Occipital parts	Superior nuchal line of occipital bone	Galea aponeurotica	Draws scalp backwards
Frontal parts	Galea aponeurotica	Skin of eyebrows and root of nose	Elevates eyebrows; wrinkles forehead; draws scalp forwards
Muscles of the Eyelids			
Orbicularis oculi	Frontal, Maxillary and Zygomatic bones; Medial palpebral ligament	Skin around eye; Lateral palpebral ligament	Elliptical muscle that occupies eyelids, surrounds the orbit, and spreads onto the temporal region and cheek; closes eyelids
Corrugator supercilli	Brow ridge of frontal bone	Skin of eyebrow	Draws eyebrows together; forms vertical wrinkles in forehead above nose
Muscles of the Nose			
Procerus	Lower part of nasal bone	Skin between eyebrows	Forms horizontal wrinkles across bridge of nose
Nasalis	Maxilla next to incisor and canine teeth	Bridge and side of nose	Widens anterior nasal aperture, especially in deep inspiration
Depressor septi	Maxilla	Septum of nose	Draws septum down
Muscles of the Mouth			
Levator labii superioris	Maxilla	Upper lip	Raises upper lip and turns it outward
Zygomaticus Minor	Zygomatic bone	Upper lip	Elevates upper lip
Zygomaticus Major	Zygomatic arch	Muscle and skin at angle of mouth	Draws angle of mouth upward and laterally
Levator anguli oris	Canine fossa of maxilla	Muscles at angle of mouth	Raises angle of mouth
Mentalis	Incisor fossa of mandible	Skin of chin	Raises and protrudes lower lip; wrinkles skin of chin
Depressor labii inferioris	Mandible	Skin and muscles of lower lip	Draws lower lip downward and laterally

Table A: Origins and Insertions of Facial Muscles

All facial muscles, with the exception of the *orbicularis oris*, emerge or have *origins* on the underlying bone and *insert* into the skin. This table provides an indication of the location of the *origins* and *insertions* of the major facial muscles [147].

Table A: Origins and Insertions of Facial Muscles (Continued)

Muscle	Origin	Insertion	Action
Depressor anguli oris	Mandible	Muscles at angle of mouth	Draws angle of mouth downward and laterally
Buccinator	Mandible and Maxilla in region near molars	Muscles at angle of mouth	Compresses cheeks against teeth; provides a stable lateral wall to oral cavity for pressure in speech, sucking and mastication
Orbicularis oris	Maxilla and muscle fibres surrounding mouth	Fibres encircle mouth; some attach to skin and muscles at angle of mouth	Closes lips, presses lips against teeth, protrudes lips, and shapes lips in speech
Risorius	Fascia of Masseter muscle	Skin at angle of mouth	Retracts angle of mouth
Levator labii superioris alaequae nasi	Frontal process of maxilla	Skin of lip and ala of nose	Draws upper lip upward and widens nostril
Masseter	Zygomatic arch	Angle of mandible	Elevates mandible, closing jaw; small effect in lateral movements to same side or protrusion
Temporal	Temporal fossa	Coronoid process and ramus of mandible	Elevates mandible, closing jaw; draws mandible backwards after protrusion assists in lateral movements to same side
Lateral pterygoid	Greater wing of sphenoid bone and lateral pterygoid plate	Tissues of temporomandibular joint and neck of mandible	Assists in opening of mouth, protrusion of jaw, and lateral movements to the opposite side
Medial pterygoid	Pterygoid plate, palatine bone and maxilla	Ramus and medial surface of angle of mandible	Assists elevation and protrusion of mandible and lateral movements to the opposite side

Appendix B

Facial Action Coding System

Action Unit	FACS Name	Muscular Basis
AU1	*Inner Brow Raiser	Frontalis, Pars Medialis
AU2	*Outer Brow Raiser	Frontalis, Pars Lateralis
AU4	*Brow Lowerer	Depressor Glabellae; Depressor Supercilli; Corrugator
AU5	*Upper Lid Raiser	Levator Palpebrae Superioris
AU6	*Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
AU7	*Lid Tightener	Orbicularis Oculi, Pars Palpebralis
AU8	*Lips Toward Each Other	Orbicularis Oris
AU9	*Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
AU10	*Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
AU11	Nasolabial Furrow Deepener	Zygomatic Minor
AU12	*Lip Corner Puller	Zygomatic Major
AU13	Sharp Lip Puller	Caninus
AU14	Dimpler	Buccinnator
AU15	*Lip Corner Depressor	Triangularis
AU16	*Lower Lip Depressor	Depressor Labii
AU17	*Chin Raiser	Mentalis
AU18	Lip Puckerer	Incisivii Labii Superioris; Incisivii Labii Inferioris
AD19	Tongue Out	—
AU20	*Lip Stretcher	Risorius
AU21	Neck Tightener	—
AU22	Lip Funneler	Orbicularis Oris
AU23	Lip Tightener	Orbicularis Oris
AU24	*Lip Pressor	Orbicularis Oris
AU25	*Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
AU26	*Jaw Drop	Masseter; Temporal and Internal Pterygoid Relaxed
AU27	Mouth Stretch	Pterygoids; Digastric
AU28	Lip Suck	Orbicularis Oris
AD29	Jaw Thrust	—
AD30	Jaw Sideways	—
AD31	Jaw Clencher	—
AD32	Bite	—
AD33	Cheek Blow	—
AD34	*Cheek Puff	—
AD35	*Cheek Suck	—
AD36	Tongue Bulge	—
AD37	Lip Wipe	—
AU38	Nostril Dilator	Nasalis, Pars Alaris
AU39	Nostril Compressor	Nasalis, Pars Transversa and Depressor Septi Nasi
AU41	Lid Droop	Relaxation of Levator Palpebrae Superioris
AU42	Slit	Orbicularis Oculi
AU43	*Eyes Closed	Relaxation of Levator Palpebrae Superioris
AU44	Squint	Orbicularis Oculi, Pars Palpebralis
AU45	Blink	Relaxation of Levator Palpebrae and Contraction of Orbicularis Oculi, Pars Palpebralis
AU46	Wink	Orbicularis Oculi

Table B: Single Action Units in FACS

FACS defines the 58 Action Units [44] listed above. Action Units preceded by an asterisk (*) have been implemented in FACES. Those preceded by a plus-sign (+) are not listed in the FACS manual, but have been added into the system to provide greater control.

Action Unit	FACS Name	Muscular Basis
AD51	*Head Turn Left	—
AD52	*Head Turn Right	—
AD53	*Head Up	—
AD54	*Head Down	—
AD55	*Head Tilt Left	—
AD56	*Head Tilt Right	—
AD57	Head Forward	—
AD58	Head Back	—
AU61	Eyes Turn Left	—
AU62	Eyes Turn Right	—
AU63	Eyes Up	—
AU64	Eyes Down	—
AU65	Walleye	—
AU66	Cross-Eye	—
AD67	+Focus Eyes	—
AD68	+Pupil Dilation	—

Table B: Single Action Units in FACS (Continued)

Appendix C

Data Files for FACES

Default Settings

The following is an example of a textual data-file from which FACES sets up default values.

```
XMAX            250                 /* frame width               */
BCOLOR          0.0 0.0 0.0         /* background colour         */

LT_INTENSITY    1.0                 /* light source intensity    */
LT_ORIGIN       0.0 0.0 -1.0        /* light source position     */
LT_COLOR        1.0 1.0 1.0         /* light source colour       */

MATERIAL        skin
COLOR           1.0 1.0 1.0         /* skin tone                 */

MATERIAL        hair
COLOR           0.29 0.0 0.0        /* hair colour               */

MATERIAL        eye_col
COLOR           0.35 0.0 0.0        /* eye colour                */

MATERIAL        lip_col
COLOR           1.0 0.4 0.3         /* lip colour                */

PRIMITIVE       face                /* object name               */
NAME            MESH                /* type of primitive         */
PTS_FILE        new_face.pts        /* associated data-file      */

DISPLAY         face
MADE_OF         skin

DISPLAY         eyebrows
MADE_OF         hair

DISPLAY         eyes
MADE_OF         eye_col

DISPLAY         lips
MADE_OF         lip_col
```

Animation Scripts

Animation sequences are stored in textual files for ease of comprehension and modification. An example of a facial animation sequence is followed by an example of a conformation animation sequence.

Facial Animation Sequence

```
An example of a facial animation sequence
EXPR            4         surprise
START_FRAME     10
END_FRAME       11
START_TENSION   0.3
END_TENSION     0.9
LAW             accelerate

EXPR            5         fear
START_FRAME     14
END_FRAME       17
START_TENSION   0.2
END_TENSION     0.5
LAW             linear

AU              5         Upper_Lid_Raiser
START_FRAME     18
END_FRAME       22
START_TENSION   0.0
END_TENSION     0.5
SIDE            both
LAW             linear

EXPR            6         wink
START_FRAME     18
END_FRAME       22
START_TENSION   0.0
END_TENSION     1.0
LAW             linear

EXPR            0         happiness
START_FRAME     23
END_FRAME       27
START_TENSION   0.1
END_TENSION     1.0
LAW             decelerate
```

Conformation Animation Sequence

```
An example of a conformation animation sequence
FROM           current
TO             man
START_FRAME    1
END_FRAME      4
LAW            accelerate

FROM           man
TO             woman
START_FRAME    5
END_FRAME      8
LAW            linear
```

Appendix D

Video Sequences

Animation Sequences on Video

Several animation sequences have been recorded on a VHS video cassette to demonstrate the facilities and capability within FACES. The sequences were recorded directly off a graphics display device and hence show some jittering due to incompatibilities between the device and the video recorder. Nevertheless, the sequences do provide an indication of the type of effects that can be achieved using FACES.

SEQUENCE	TIME	LENGTH	REVERSE	PAUSE
various-effects	40 secs	121 frames	no	0.05
various-effects	20 secs	121 frames	yes	0.0
man-to-woman	20 secs	27 frames	no	1.0
man-to-woman	20 secs	27 frames	yes	1.0
blush	15 secs	10 frames	no	0.0
pale	10 secs	10 frames	no	2.0
wink	10 secs	15 frames	no	0.0
vertical eye-roll	10 secs	10 frames	no	0.0
horizontal eye-roll	10 secs	10 frames	no	0.0
FACES system	8 mins, 10 secs			

Table D: Video Sequences

Examples of animation sequences to demonstrate use of the facilities that are available in FACES. The sequences include conformation animation as well as facial animation.

Appendix E

Structure of FACES (Fold-Out)

Appendix F

Refereed Paper

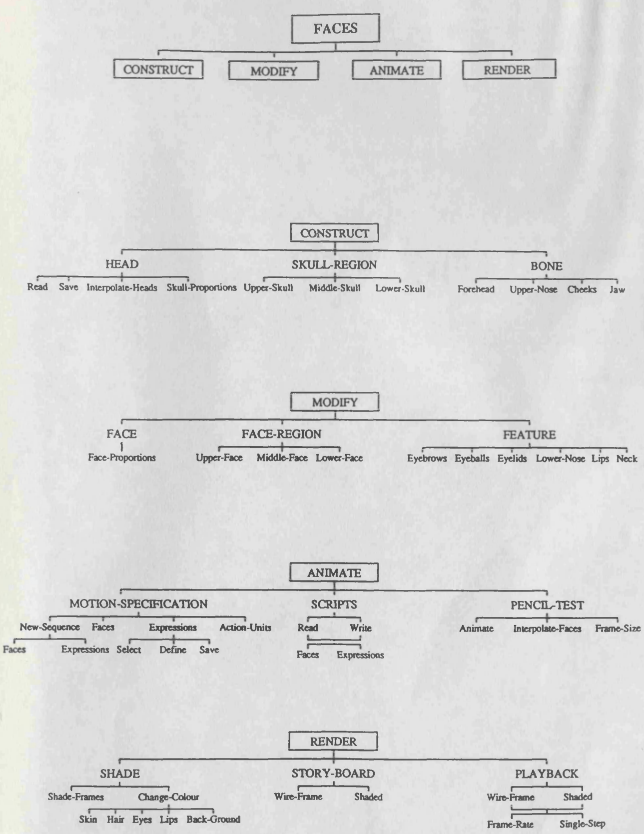

Figure E.1: Structure of FACES

FACES is an acronym for the *Facial Animation, Construction and Editing System.* The hierarchical organisation of the system is illustrated in this tree diagram which shows the four sub-systems that comprise FACES.

Appendix F

Refereed Paper

FACES: Facial Animation, Construction and Editing System [1]

Manjula Patel and Philip J. Willis

Computer Graphics Group
School of Mathematical Sciences
University of Bath
BATH
U.K.

February 1991

Abstract

The aim of the *Facial Animation, Construction and Editing System* (FACES) is to provide a software simulation of the human face. Attention has focused on the face as an important means of non-verbal communication. The interactive composition and modification of the human head and its subsequent animation, have been identified as being of particular interest.

The novelty of FACES is that it integrates the modelling and animation of faces using a three-layer anatomical model. FACES consists of four sub-systems: CONSTRUCT, MODIFY, ANIMATE and RENDER. The CONSTRUCT and MODIFY sub-systems enable changes to be made to the structure of the head, at both global and local levels, enabling specific faces to be created. The ANIMATE sub-system caters for motion specification and control, so that real and exaggerated facial expressions can be animated. The RENDER part of the system facilitates the generation of realistic images and their real-time playback.

In this paper we consider the system from the user's point of view, examining the facilities which are provided, their appropriateness and practicality.

---

[1]Presented at EUROGRAPHICS '91, Vienna, September 1991

Introduction

Two of the most important functions of the human face are identification and communication. Reconstruction of realistic faces from skeletal remains is of immense interest in forensic medicine. At the moment facial reconstruction requires extensive effort and very skilled crafting to develop a clay model. Consequently, there is only one person in the UK and very few in the world, who have the required expertise [17]. Archaeologists currently employ similar techniques which are equally time-consuming.

A computer model has an important advantage over a physical one, namely that it can be made to move, change expression and even appear to talk!

Recognition and identification of faces is also an important aspect of human psychology, particularly in the field of criminology. The manual Identikit and Photofit systems have been found to be inadequate [4]. Although the more recent *Electronic-Fit* or *E-Fit* is a computerised system [8], the images produced are static and 2D. In addition, the facial composites are totally devoid of expression. A system for the modelling of faces and expressions could prove to be of great use in the identification process —this at present remains untested.

Within the entertainment industry, computer animation needs to facilitate *character animation* [9, 10, 16] which is based on the caricaturing of human personalities and emotions. Both aspects can be communicated through facial expression and bodily gestures. This type of non-verbal communication has been exploited extensively in traditional cartoon animation [18].

Computer animation of synthetic faces has received considerable attention recently, it is considered a challenge for two main reasons: the face is an irregular structure, which varies from person to person. This problem is compounded by the motion of the face which involves complex interactions and deformations of both the features and skin. Also, faces are very familiar to us; we have a well developed sense for distinguishing which expressions are natural for a face. Consequently, we are likely to notice the smallest deviation from our perception of how a face should appear and behave.

A majority of facial movements result from either speech or the display of emotion; each of these is complex. Speech synthesis is a major area of research in its own right [1, 2] and is therefore not addressed in this project. We instead concentrate on the face as an important means of non-verbal communication.

Besides speech synthesis, two major problems can be identified: the creation and modification of faces, and control over facial animation. We present FACES —*The Facial Animation, Construction and Editing System* as a means of investigating such issues. The hierarchical nature of this interactive system is shown in Figure 1. It comprises four sub-systems: CONSTRUCT, MODIFY, ANIMATE and RENDER.

The CONSTRUCT and MODIFY sub-systems enable changes to be made to the structure of the head and face at global and local levels to facilitate the creation of specific faces. The ANIMATE sub-system caters for motion specification and control, so that sequences of both real and exaggerated facial expressions can be created. The RENDER part of the system enables the generation of realistic images, their examination and real-time playback.

In this paper we provide an overview of FACES to familiarise the reader with the function of each of the sub-systems. Further details and issues relevant to various parts of the system are discussed in [15].

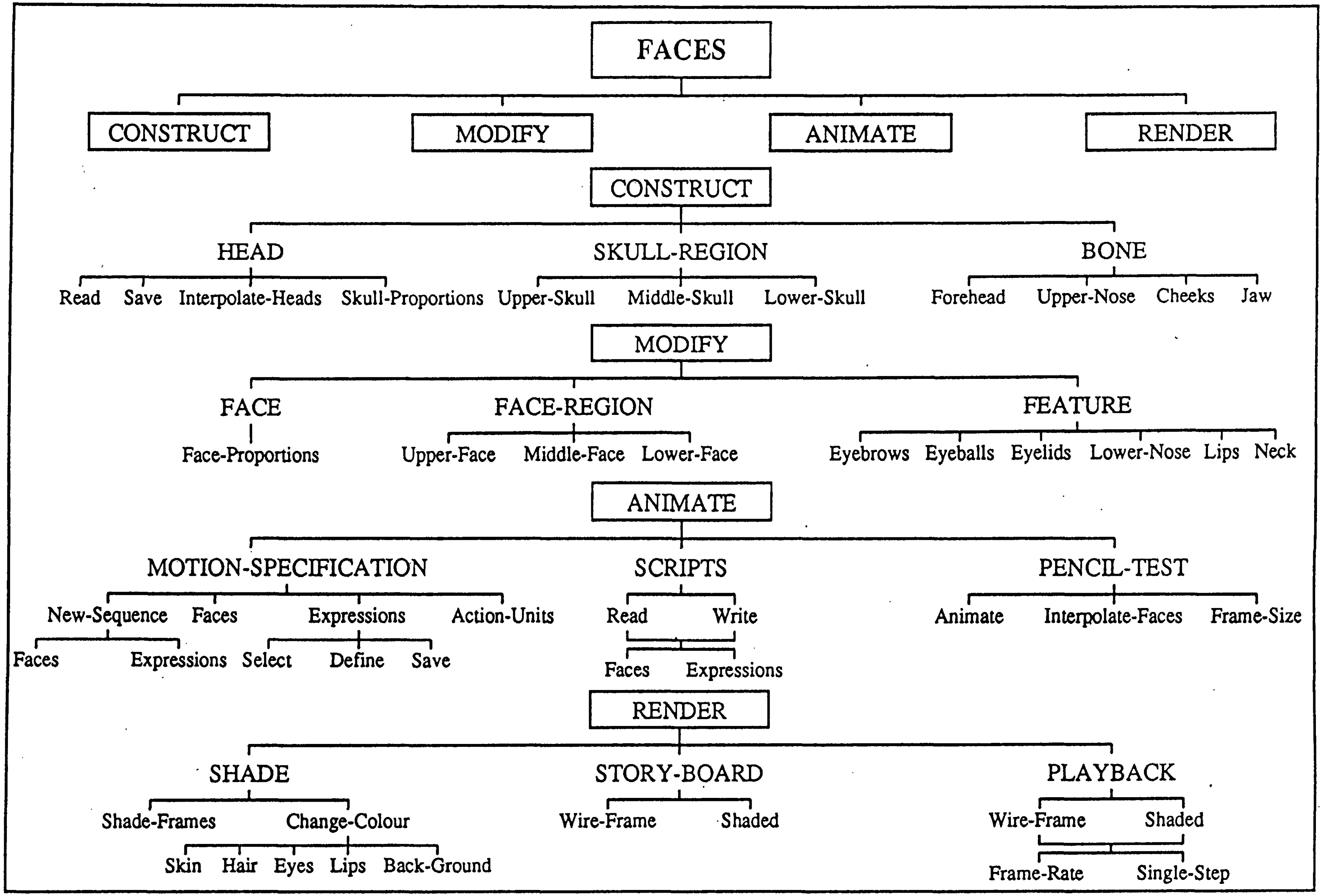

Figure 1: FACES: Facial Animation, Construction and Editing System

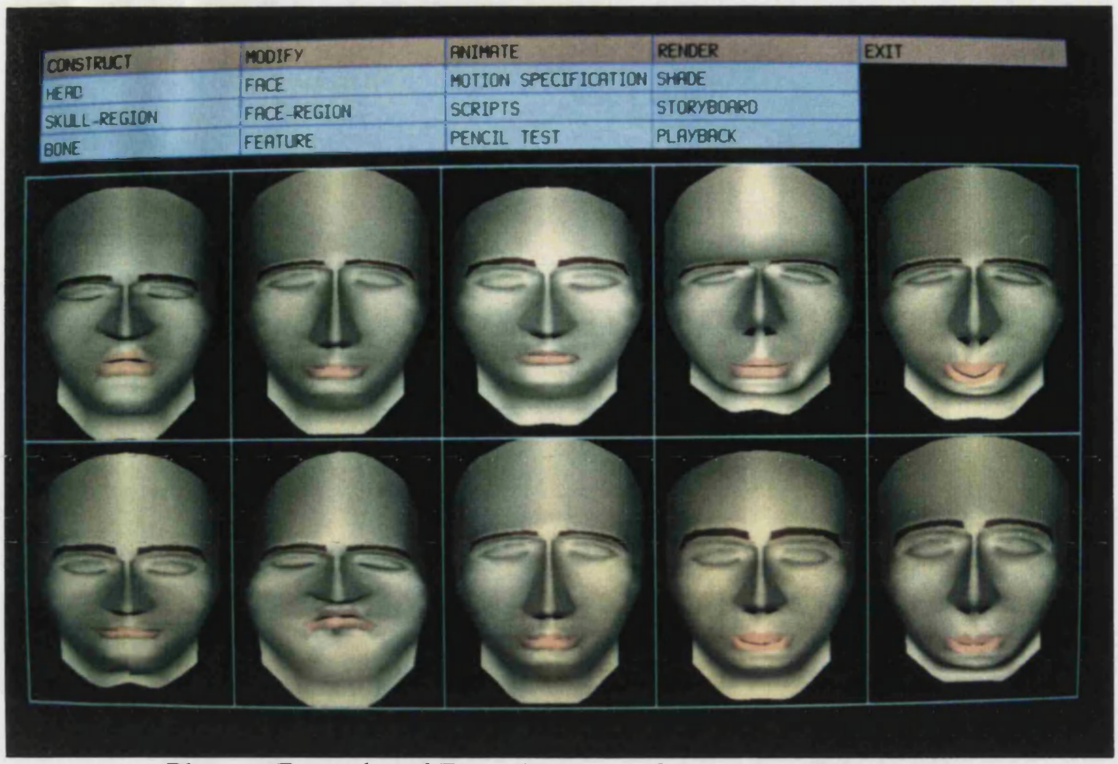

Plate 1: Examples of Faces Constructed From One Facial Mesh

1 Constructing Faces

One of the greatest problems with investigating facial animation is that of acquiring a 3D computer model of the face. Many researchers have used digitised data [5, 10, 13, 14, 19]. This strategy has revealed several drawbacks; first and foremost a physical representation is necessary. Second, the digitising process can be extremely tedious and time-consuming. In addition, the data acquired is specific to a particular face.

The CONSTRUCT and MODIFY parts of FACES attempt to overcome these difficulties. Despite individual variability, it is apparent that all faces have a similar general form, suggesting that generic modelling may be appropriate. Together, the two sub-systems cater for the modelling of specific faces such as those illustrated in Plate 1.

A model of the head consists of three layers: the surface skin; a representation of the underlying skull; and a muscle layer connecting the bone and skin which is represented using muscle vectors [19]. Within the CONSTRUCT part of the system, the user works with the skull since this determines the overall proportions of the head and face [11, 12]. All changes made to the skull also influence the shape and proportions of the surface skin, the face.

The structure of the CONSTRUCT sub-system is shown in Figure 1. It comprises three major sections to enable both global and local modifications to be made: HEAD, SKULL-REGION and BONE.

The HEAD option allows new head data to be brought into FACES, with the aid of the following menu-items:

Read
Save
Interpolate Heads
Skull-Proportions

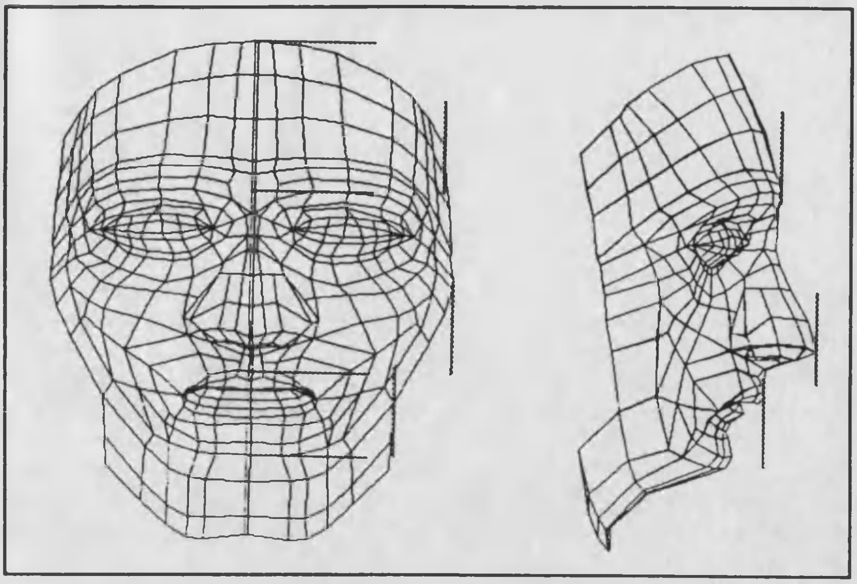

Figure 2: Regional Modifications to the Skull

The *Read* and *Save* options provide obvious functions, they restore and save head structures to and from disk files. *Interpolate Heads* provides a method of generating new heads, through interpolation between two existing ones. This process has been made flexible by allowing the user a choice of interpolation techniques. Each of the intermediate heads is displayed, and is then available to become the "current" head in FACES. Selection of the *Skull-Proportions* option caters for global changes to the skull; these are performed with the use of a 3D scaling function.

The SKULL-REGION option in CONSTRUCT, is used to make regional modifications to the structure of the skull, which has been partitioned into three sections: *upper-skull, lower-skull* and *middle-skull.* The sections can each be stretched or compressed in both the x-y and y-z planes, as shown in Figure 2.

Further down the hierarchy, selection of the BONE option enables amendments to be made to more specific parts of the skull including: the *frontal bone* (forehead); *nasal bone* (upper-nose); *zygomatic bone* (cheeks) and the *mandible* (jaw). The bones can be altered with the use of geometric deformations. Some of the bones are treated as pairs while others are single. For paired bones the user may specify one of *Left, Right* or *Both* and thus gain control over a particular side of the head.

2 Modifying Faces

While it is the skull that determines the overall proportions of the head and face, there are variations in the surface characteristics of the face which cannot be accurately determined from the underlying bone structure. These include muscle thickness and *soft* features such as the eyebrows, eyelids, eyeballs, lower-nose, lips and neck. Such characteristics vary greatly from individual to individual. This is the major reason why faces are so different from each other and therefore recognisable. The MODIFY sub-system addresses these aspects of face creation; amendments made using this part of FACES affect only the surface skin and muscle thickness. Modifications can be made at three different levels: FACE, FACE-REGION and FEATURE.

Within the FACE option the *Face-Proportions* facility allows global amendments to be made to the surface skin, again by the use of a 3D scaling function. Subsequent to seeing the effect of changes in relative proportions, both in CONSTRUCT and MODIFY, the user may either *Keep* or *Ignore* the changes.

Selecting the FACE-REGION option caters for alterations in the relative proportions of facial skin sections, of which there are three: the *upper-face, middle-face* and *lower-face.* The upper-face consists of the eyebrows and forehead; the lower-face is the area from the chin to the upper lip; and the middle region comprises the area from the upper lip to the eyes. In the context of regional modifications to either the skull or the face, a stack-based *undo* operation is available to facilitate experimentation.

The final part of the modelling process is refinement, involving scaling and positioning of the soft features of the face. Geometric transformations are used to modify such elements. Some of the features are treated as pairs while others are single features. For example eyelids, eyebrows and eyeballs are paired features and independent control is available over the left and right sides. During FEATURE and BONE modifications, the undo facility operates only on the last change made.

3 Animating Faces

So far we have only considered the modelling of *conformation* aspects of a static head. Conformation is concerned with the structure of a head model, while animation addresses the motion characteristics of the face. The menu-hierarchy associated with the ANIMATE sub-system is illustrated in Figure 1. This part of FACES enables a user to build up an animated sequence of the face once it has been modelled using CONSTRUCT and MODIFY. The three major options within ANIMATE are: MOTION SPECIFICATION, SCRIPTS and PENCIL-TEST.

3.1 Creating and Controlling Facial Movement

MOTION SPECIFICATION allows a user to create two types of animated sequence. The first comprises a sequence of facial expressions applied to a face and is illustrated in Plate 2. The other is useful for special effects such as transforming one character into another as demonstrated in Figure 3, courtesy of Caplin & Jeremy [7]. In FACES the two types of sequence are separate and distinct. The options available under MOTION SPECIFICATION are:

New-Sequence
Faces
Expressions
Action Units

The *New-Sequence* option initialises the system for the specification of a completely new animation sequence. It is useful whenever an animator needs to start afresh. Animation sequences are persistent, so that they continue to exist until FACES is exited, or *New-Sequence* is chosen, or a script is read in. By default, once a sequence exists in the system any further selections are added onto that sequence.

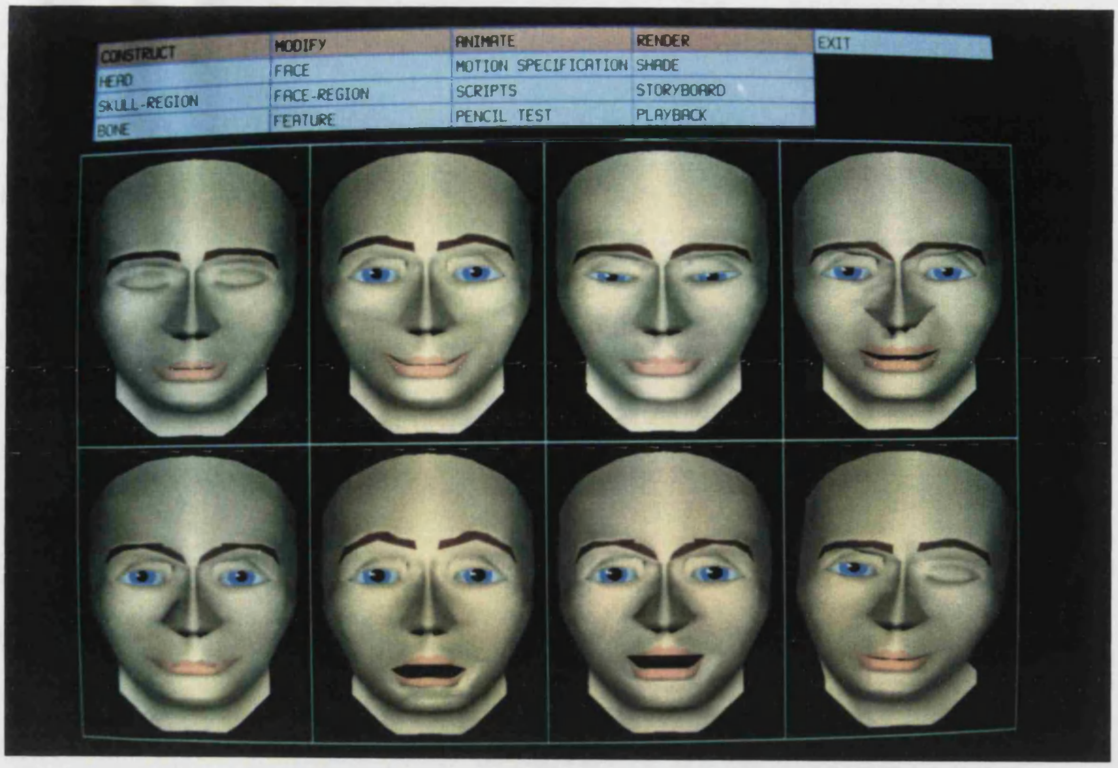

Plate 2: Facial Expressions Generated Using FACES

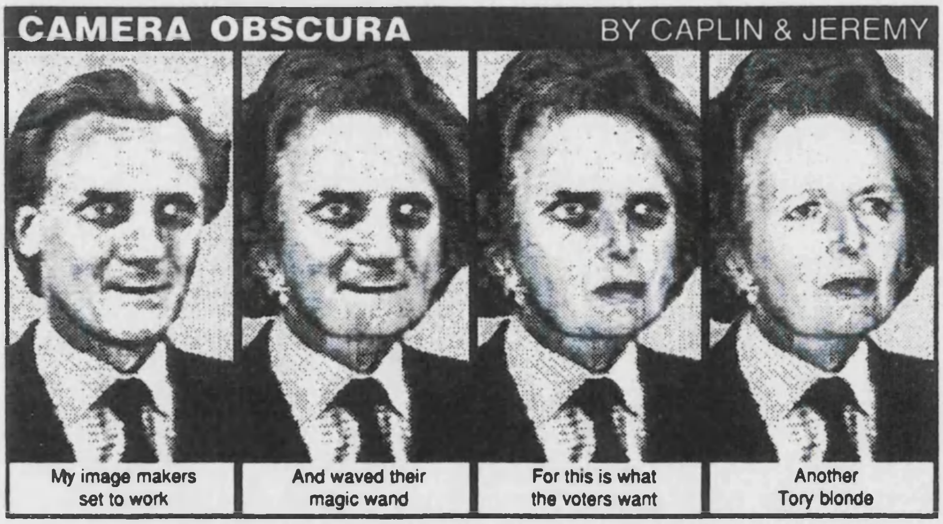

Figure 3: Interpolation Between Two Well-Known Faces

3.1.1 Conformation Animation

Selection of the *Faces* option allows a user to build up a sequence of face interpolations. For each interpolation the user is prompted for the names of the data-files containing the start and end heads; the start and end frames; and the law of interpolation to be used:

linear
accelerate
decelerate
accelerate-decelerate

For either the start or end head it is possible to specify "current" to represent the head being used in FACES at a particular time.

3.1.2 Expression Animation

FACES differs from the Human Factory System [10] in that faces and expressions are independent, so that expression sequences may be applied to any face or conformation sequence.

To build up an expression sequence over several frames, a user selects from either pre-defined *Expressions* or from the lower-level *Action Units* (AUs) listed in Figure 4. Selections from these two sets may be freely intermixed; they may also overlap in time, so that they are *active* over the same frames. Users have further flexibility because expressions and AUs do not need to be selected in time or frame order. Hence new expressions or AUs may be inserted into an existing sequence to refine it.

Motion in FACES is based on parameterised keyframe animation. The deformations result from the simulation of muscle contractions in the face. The user controls the start and end frames for each expression or AU as well as the law of interpolation to be used.

Generation of facial expressions is controlled by a three level hierarchy: muscles, AUs and expressions. The lowest level is based on three types of muscle: *linear, sphincter* and *sheet* [19].

Developing a motion description technique based on the anatomy of the face requires the user to learn the muscular basis. To avoid such inconvenience, the muscular level is hidden from the user, who is presented with access to the AUs shown in Figure 4. Each AU describes the contraction of one facial muscle or a group of related muscles. They are derived from the *Facial Action Coding System*(FACS) [3]. FACS is based on a highly detailed analysis of real facial expressions. It results from a major body of work and identifies 58 AUs which separately, or in various combinations are capable of characterizing any human expression; this comprehensive approach is capable of dealing with more than 7000 expressions [6].

In addition to the creation of facial expressions and their control, there are further AUs to facilitate controlled motions of the entire head, to allow: turning, tilting and nodding.

Some AUs are capable of operating on either side of the face independently of the other; such control makes it possible to create asymmetric effects which are important for some expressions such as winking and various other idiosyncracies.

As well as the facial surface conveying emotion, the eyes are one of the most important features

Exit	Backup
Neutral-Face	Inner-Brow-Raiser
Outer-Brow-Raiser	Brow-Lowerer
Upper-Lid-Raiser	Cheek-Raiser
Lid-Tightener	Lips-Towards-Each-Other
Nose-Wrinkler	Upper-Lip-Raiser
Lip-Corner-Puller	Lip-Corner-Depressor
Lower-Lip-Depressor	Chin-Raiser
Lip-Stretcherer	Lip-Pressor
Lips-Part	Jaw-Drop
Cheek-Puff	Cheek-Suck
Eyes-Closed	Turn-Left
Turn-Right	Head-Up
Head-Down	Tilt-Left
Tilt-Right	Eyes-Track
Pupil-Dilation	

Figure 4: Action Units Available in FACES

in creating expressions [18]. They also play a significant rôle in maintaining attention during conversation. We therefore include the capability to track a given focal point in 3D-space, and also cater for horizontal and vertical eyeball rolling. In fact, each eyeball may focus on a separate point for special effects such as modelling cross-eyes. Blinking and pupil dilation can also be employed for further subtlety. Such fine control enables the animator to model both realistic and unrealistic effects.

Although AUs allow for great subtlety in creating facial expressions, they were felt to work at too low a level for certain purposes. The animator may therefore work at a higher emotional level with predefined expressions such as happiness, sadness and disgust. The intensity of any expression or AU is controllable to achieve the required degree of emotion.

3.2 Creating and Saving Expressions

As well as being able to create sequences of facial expressions, users may also construct and save away static expressions, and thus build up libraries of pre-defined expressions. A user has the capability to combine primitive AU operations into higher level emotions corresponding to familiar expressions such as anger, fear and surprise. These can then be selected for use in animation sequences. The *Expression* option in fact allows access to several others:

Select Expression
Define Expression
Save Expression

Define Expression enables a user to interactively sculpt an expression from the same range of AUs as are used in generating animation sequences. A stack-based *undo* facility is available to enable experimentation while creating expressions. When the user saves an expression away, it immediately becomes available for selection as part of the pre-defined expression database.

3.3 Storing and Retrieving Animation Sequences

Having built up an animation sequence it is useful to save it away and read it back into FACES at some later time. The SCRIPTS option in ANIMATE enables a user to do precisely this; it allows access to two further options:

Read
Write

Read and *Write* cater for the retrieval and storage of animation sequences from and to disk files.

Expression scripts are stored in terms of expressions and AUs, exactly the entities that the user works with. They indicate the expression or AU; the frames over which it is active; the start and end intensities and the law of motion to be used for interpolation. Conformation animation scripts comprise: the start and end heads; start and end frames; and the interpolation law to be used for inbetweening. The use of a textual recording technique makes the scripts easy to read, understand and modify using a standard text editor.

When scripts are read back into FACES, two options are available to the user. If an animation sequence is in existence, it may either be overwritten by the new script or alternatively the new sequence may be incorporated into the old.

3.4 Previewing and Testing Animated Motion

The testing of motion is an essential aspect of generating sequences of animation. The pop-up menu associated with the PENCIL-TEST option allows a choice of the following operations:

Animate
Interpolate Faces
Frame-Size

Animate generates a wireframe preview of an expresion animation sequence applied to the current face. If a user is unhappy with the face it is possible to return to the CONSTRUCT and MODIFY sub-systems. At present dissatisfaction with the expression sequence means that the user must either generate a completely new sequence, or exit FACES and edit the script corresponding to the sequence. A means of interactively modifying a sequence within FACES would be of use here.

The *Interpolate Faces* option caters for special effects such as transforming one character into another as demonstrated in Figure 3 [7]. It enables the generation of an animated sequence consisting of an interpolation between two faces. At this stage one of two effects is possible: if an animation sequence has been defined it is used to generate expressions on the inbetween face; if no animation sequence exists an interpolation is performed without facial expressions.

Testing the smoothness of the generated motion is accomplished within the PLAYBACK facility in the RENDER sub-system, which is described in section 4.

The *Frame-Size* option allows interactive control over the size of the frames to be generated.

4 Rendering Faces

Once users are happy with the model of the head and its motion, they require a means of generating more realistic sequences. The structure of the RENDER sub-system is shown in Figure 1. Using the options SHADE, STORYBOARD and PLAYBACK, it allows fully shaded, realistic images of an animated face to be generated.

SHADE caters for both the shading of the frames and the interactive modification of the colours of skin, hair, eyes, lips and the background. At present only the eyebrows are affected by hair colour. The user also has control over the first and last frames to be shaded so that experimentation with colours is possible before a whole sequence of frames is submitted to be shaded. The shading process itself is optimised and generates each colour image using Gouraud shading within a few seconds for rapid visual feedback.

The STORYBOARD option allows examination of consecutive frames of a sequence; it is illustrated in Plates 1 and 2. This facility can be of invaluable help to an animator during the development of a sequence. Again the user controls the first and last frames to be displayed.

PLAYBACK caters for the real-time display of short sequences under user control. The user may select either *Frame-Rate* or *Single-Step* display mode in order to test animated motion. Within the *Frame-Rate* option the user has additional control such as over the first and last frames to be displayed, and over the pause rate between frame updates. In *Single-Step* mode the user has total control and frames are updated with the press of a puck button. In both cases frames continue to be displayed cyclically until the user decides to exit. STORYBOARD and PLAYBACK are capable of operating on either wireframe or shaded sequences.

5 User Interface

The user-interface integrates the four components of the system, and provides a uniform manner in which the animator is allowed access to them. Since screen space is limited, two types of menu are in use: permanent and pop-up. At the top of the screen, a permanent horizontal menu-bar acts as the gateway which allows interactive access to the four sub-systems:

CONSTRUCT	MODIFY	ANIMATE	RENDER	EXIT

Whenever a top-level menu-item is selected the relevant menu-box is highlighted and a list of the major menu options in the sub-system appears underneath it. The list remains on the screen until its parent menu-item is re-selected.

Parts of the system are modeless while others are guided. This gives users a considerable amount of flexibililty in the way that they work, while at the same time ensuring that FACES has all the relevant information in order to perform a particular task. As far as possible the system checks to ensure that preconditions are met for correct operation. For example, selection of the *Animate* option in PENCIL-TEST is meaningless if an expression animation sequence does not yet exist within the system.

The hierarchical organisation of the system necessitates the nesting of menus to a depth of several levels. The tedium of backtracking to higher levels has been overcome by the provision of an *Exit* option in pop-up menus; it helps a user to escape from a menu hierarchy. Furthermore, whenever it is available the *Exit* option consistently appears as the first menu-item.

6 Discussion and Conclusions

FACES is an interactive system aimed at rapid feedback to increase productivity. The underlying philosophy is that of a system for rapid prototyping. This approach has the potential to save the user a considerable amount of both time and effort.

Many subtle and flexible controls exist over the modelling and animation aspects of the face. In addition, the use of interactive modifications together with an undo facility is a powerful method of catering for experimentation.

Nevertheless, the system does need to be developed further in order to facilitate realistic applications. The major area that requires additional attention is the development of a cohesive heterogeneous structure for the bone, muscle and skin. This structure will allow the thickness of individual muscles to be varied and volume changes in muscle mass to be accounted for during animation. We are also investigating the potential of a 3D Identikit type of approach for the construction of new faces.

The head model currently in use has been adequate for experimental purposes. For more usable applications, the model needs features such as tongue, teeth, ears and hair. There is a further requirement for superficial elements such as spectacles, beards, moustaches and scars.

In addition, some problems with maintaining the consistency of the mesh model have become apparent. One possible resolution is to convert to a patch model based on either Bézier or B-spline curves; this is however likely to make the system compute-intensive and therefore slower.

We provide limited opportunities for shading selected parts of the face: only eyebrows, lips, eyes and skin may be varied. This facility needs to be extended so that the colour of selected portions of the face can be changed to cater for emotional visual cues such as blushing due to embarrassment, or pallor due to fear or anger. At present the most realistic method of representing skin is by texture-mapping [20], which is currently lacking from the system.

There is also room for development of the user interface in providing more appropriate valuators and names for particular operations. For example BONE and FEATURE modifications could be more aptly named.

In FACES conformation and animation are treated as separate and distinct. Composition and modification of faces is addressed by the CONSTRUCT and MODIFY sub-systems, while the motion aspects are dealt with in the ANIMATE sub-system. This separation of functionality allows differing strategies to be used for modelling and animation within a single system.

FACES encapsulates a considerable amount of information regarding facial shape, movement and expression. The system has tremendous scope; the applications are numerous and vary greatly. The implications of such a system are profound and liable to cause a few eyebrows to be raised!

Acknowledgements

The authors would like to thank several people who have taken the time and trouble to provide information in their own areas of expertise: *Adrian Bowyer, Michael Brooke, John Halas, Tony Kitson, John Lasseter, Johnathan Musgrave* and *Richard Neave.*

Thanks are also due to *Stephen Caplin* for permission to use the cartoon strip "Camera Obscura" and to *Andrew Cook* for proof reading this article and providing many constructive comments.

We are grateful for the financial support provided by the *Science and Engineering Research Council* to one of the authors from October 1988 to September 1990.

References

[1] N.M. Brooke. Visible speech signals: Investigating their analysis, synthesis and perception. In *The Structure of Multimodal Dialogue.* Elsevier Science Publishers (North-Holland), 1989.

[2] N.M. Brooke and Q. Summerfield. Analysis, synthesis and perception of visible articulatory movements. *Journal of Phonetics*, 11:63–76, 1982.

[3] P. Ekman and W. Friesen. *Manual for the Facial Action Coding System.* Consulting Psychologist Press, Palo Alto, California, 1978.

[4] H.D. Ellis. An investigation of the use of the Photofit technique for recalling faces. *British Journal of Psychology*, 66:29–37, 1975.

[5] A. Emmet. Digital portfolio: Tony de Peltrie. *Computer Graphics World*, 8(10):72–77, October 1985.

[6] D. Goleman. The 7000 faces of Dr. Ekman. *Psychology Today*, 15:42–49, 1981.

[7] Caplin & Jeremy. Camera obscura. Weekend Guardian, 17th November 1990.

[8] T. Kitson. Personal communication, February 1989. Scientific Research and Development Centre, Home Office, London, UK.

[9] J. Lasseter. Principles of traditional animation applied to 3D computer animation. *Proceedings ACM SIGGRAPH, Computer Graphics*, 21(4):35–44, July 1987.

[10] N. Magnenat-Thalmann and D. Thalmann. *Synthetic Actors in Computer Generated 3D Films.* Springer-Verlag, 1990.

[11] J. Musgrave. Personal communication, June 1990. Lecturer in Anatomy and Anthropology, University of Bristol, UK.

[12] R.A.H. Neave. Personal communication, June 1990. Director of Medical Illustration, University of Manchester Medical School, UK.

[13] F.I. Parke. Computer generated animation of faces. *Proceedings ACM National Conference*, 1:451–457, August 1972.

[14] F.I. Parke. A parametric model for human faces. PhD Dissertation, University of Utah, 1974.

[15] M. Patel. *Making FACES.* PhD thesis, In Preparation, School of Mathematical Sciences, University of Bath, UK, 1991.

[16] S. Porter. Made for the stage. *Computer Graphics World*, 13(8):60–69, August 1990.

[17] A. Roy. Making the face fit. Sunday Times Magazine, May 1990.

[18] F. Thomas and O. Johnston. *Disney Animation—The Illusion of Life.* Abbeville Press, New York, 1981.

[19] K. Waters. A muscle model for animating three-dimensional facial expressions. *Proceedings ACM SIGGRAPH, Computer Graphics*, 21(4):17–24, 1987.

[20] J.F.S. Yau and N.D. Duffy. A texture mapping approach to 3D facial image synthesis. *Proceedings Eurographics UK '88*, pages 17–30, April 1988. Brighton.