PhD Thesis

Determining normal and abnormal lip shapes during movement for use as a surgical outcome measure

CARDIFF UNIVERSITY

PRIFYSGOL CAERDYDD

A thesis submitted in accordance with the conditions governing candidates for the degree of Philosophiae Doctor in Cardiff University

by

Hashmat Popat

August 2012

School of Dentistry

Cardiff University

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …………………………………… (candidate)   Date ……………

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …………………………………… (candidate)   Date ……………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …………………………………… (candidate)   Date ……………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …………………………………… (candidate)   Date ……………

Acknowledgements

The completion of this dissertation would not have been possible without the guidance and help of several individuals who have extended their expertise in various ways through the course of the study.

Firstly I would like to thank my principal supervisor, Professor Stephen Richmond for his direction, encouragement and endless enthusiasm. I am greatly appreciative of his constant support and novel ideas to tackle conventionally managed problems.

I also convey my gratitude to my co-supervisors Professor David Marshall and Dr Paul Rosin for their knowledge and assistance on the computational aspects of the work. This was invaluable throughout the project.

Other members of the School of Computer Science and Informatics I would like to mention include Kirill Sidorov who very candidly allowed the use of his automatic registration tool to be validated. Andrew Aubrey for his explanations of computational terminology. In particular, Gary Tam's help in the automatic extrapolation of landmark coordinates saved countless hours of time.

From the School of Dentistry, I must thank Sue Bryant for her help in landmarking facial images. Arshed Toma, who as a fellow PhD candidate shared ideas and acted as support for the all too often Progress Monitoring Meetings. Alexei Zhurov for his expertise in customising analytic software that was used as part of this project. Rebecca Playle for her statistical guidance in the early stages of the project. Timothy Pickles for his statistical advice and meticulous feedback.

I would also like to thank all the staff, students and patients who agreed to participant in this study. It would not have been possible to generate the data used in this study without their involvement.

Finally, I am deeply indebted to my wife, daughter and parents for their patience, motivation and unconditional support throughout.

Abstract

Craniofacial assessment for diagnosis, treatment planning and outcome has traditionally relied on imaging techniques that provide a static image of the facial structure. Objective measures of facial movement are however becoming increasingly important for clinical interventions where surgical repositioning of facial structures can influence soft tissue mobility. These applications include the management of patients with cleft lip, facial nerve palsy and orthognathic surgery. Although technological advances in medical imaging have now enabled three-dimensional (3D) motion scanners to become commercially available their clinical application to date has been limited. Therefore, the aim of this study is to determine normal and abnormal lip shapes during movement for use as a clinical outcome measure using such a scanner.

Lip movements were captured from an average population using a 3D motion scanner. Consideration was given to the type of facial movement captured (i.e. verbal or non-verbal) and also the method of feature extraction (i.e. manual or semi-automatic landmarking). Statistical models of appearance (Active Shape Models) were used to convert the video motion sequences into linear data and identify reproducible facial movements via pattern recognition. Average templates of lip movement were created based on the most reproducible lip movements using Geometric Morphometrics (GMM) incorporating Generalised Procrustes Analysis (GPA) and Principal Component Analysis (PCA). Finally lip movement data from a patient group undergoing orthognathic surgery was incorporated into the model and Discriminant Analysis (DA) employed in an attempt to statistically distinguish abnormal lip movement.

The results showed that manual landmarking was the preferred method of feature extraction. Verbal facial gestures (i.e. words) were significantly more reproducible/repeatable over time when compared to non-verbal gestures (i.e. facial expressions). It was possible to create average templates of lip movement from the control group, which acted as an outcome measure, and from which abnormalities in movement could be discriminated pre-surgery. These abnormalities were found to normalise post-surgery.

The concepts of this study form the basis of analysing facial movement in the clinical context. The methods are transferrable to other patient groups. Specifically, patients undergoing orthognathic surgery have differences in lip shape/movement when compared to an average population. Correcting the position of the basal bones in this group of patients appears to normalise lip mobility.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AAM	Active Appearance Model
ASM	Active Shape Model
AU	Action Unit
AVSR	Audio Visual Speech Recognition
CBCT	Cone Beam Computed Tomography
CCA	Canonical correlation analysis
CDTW	Continuous Dynamic Time Warping
DA	Discriminant Analysis
DDTW	Derivative Dynamic Time Warping
DSM	Dense Surface Model
DTW	Dynamic Time Warping
DV	Dependent Variable
EMG	Electromyography
FA	Factor Analysis
FACS	Facial Action Coding System
FAR	False Acceptance Rate
FFD	Free-Form Deformation
fps	Frames per second
FRR	False Rejection Rate
GPA	Generalised Procrustes Analysis
GMM	Geometric Morphometrics
ICP	Iterative Closest Point
IOTN	Index of Orthodontic Treatment Need
IV	Independent Variable
LED	Light Emitting Diode
LMI	Linear Measurement Index
LOA	Level of Agreement
MANOVA	Multivariate Analysis of Variance
MDS	Multidimensional Scaling
MRI	Magnetic Resonance Imaging
NHP	Natural Head Posture
NHS	National Health Service
PC	Principal Component
PCA	Principal Component Analysis
PCoA	Principal Coordinate Analysis
PDM	Point Distribution Model
RMS	Root Mean Square
ROC	Receiver Operating Curve
SHP	Standardised Head Position
SSD	Sum of Squared Differences
TPS	Thin Plate Spline
WDDTW	Weighted Derivative Dynamic Time Warping

Introduction

Patients who present with a variety of craniofacial problems can be treated using orthodontics and/or maxillofacial surgery to achieve acceptable dental/facial aesthetics and function. The process for diagnosis, treatment planning and outcome has traditionally relied on imaging techniques that provide a static image of the facial structure. Methods such as photography and two-dimensional (2D) radiography have become convention in the majority of clinical situations. Their low cost, accessibility and ease of interpretation have tended to outweigh the lack of standardisation in head orientation, magnification and lighting. In an attempt to address the limitations of 2D imaging, technological advances have enabled three-dimensional (3D) facial imaging techniques such as laser scanning, optical stereophotogrammetry and cone beam computed tomography (CBCT) to become commercially available. The greater diagnostic yield of 3D imaging when compared to 2D techniques can prove to be valuable for more complex assessment and treatment planning.

Dental/facial aesthetics can be managed adequately using conventional 2D and 3D imaging techniques as they provide a static image of a particular subject at a given point in time. Functional outcome is more difficult to assess, as a technique for quantifying facial movement is required. Early work focused on facial movement was impeded by the problems of devising an adequate method for measuring the face. Techniques varied in methodology ranging from descriptions of specific changes within part of the face based on grading scales to photographic depictions of movement within facial areas. More current approaches have included video-based marker tracking in which multiple landmarks are placed directly onto the face and recorded during a sequence of facial movement. Markers are placed related to corresponding muscle groups and therefore displacements of these markers and the trajectories created can be quantified reflecting the underlying musculature.

Functional assessment is rarely formally assessed in orthodontics and maxillofacial surgery but can be an important consideration as many clinical interventions specifically aim to achieve a satisfactory functional outcome. These include orthognathic surgery, the management of cleft lip and in the rehabilitation of facial nerve paralysis. Although these inventions are primarily being used to correct facial aesthetics they will also have a direct influence on soft tissue mobility. Despite this, there is very little published data on the effect of any these procedures on facial movement. Objective measures of facial movement may allow more accurate diagnosis of muscular deficits, which may in turn lead to better outcome following an intervention.

There are several issues that need to be addressed if objective assessment of facial movement is to be considered. Initially, patient related factors are the most important. One of the priorities is to identify an appropriate measure of facial movement. Many studies have utilised facial expressions and/or verbal facial actions such as words, sentences or numbers. The ideal facial motion should be reproducible over time so that it is performed as near to the same way each occasion with as little variation as possible. In this respect, the effect of a clinical intervention on facial movement can be truly quantified. Other factors related to the measure of movement used include that it should:

- represent the region of the face that is of interest;
- be representative of what would be considered normal movement for the patient/subject;
- be simple/easy to perform.

Once an appropriate measure of facial movement is identified, one can assume that different subjects will vary in their articulation or performance of this measure. Ascertaining the variation of a particular facial motion in a representative sample would outline patterns or similarities in movement

and allow for the construction of references of what could be considered as average facial movement. The process of creating average references is commonplace in orthodontics and craniofacial research. Lateral cephalograms and more recently 3D laser scans from population groups can be age and/or sex matched enabling comparisons to be made between an individual and their respective *normal* template to guide treatment planning and measure outcome.

The acquisition of normal data forms the basis of a control group in clinical studies/trials. This allows comparison of data from an intervention carried out on an experimental group. In the context of facial movement and this thesis, ascertaining the variation in normal facial movement precludes the comparison to potentially abnormally movement. Abnormal movement could be sourced from any one or more of the patient groups identified in the literature review to follow.

Aside from the clinical considerations, quantification of facial movement (particularly in 3D) poses a problem from the technical aspect. Instead of one image/frame for analysis, video sequences may contain several hundred images of one particular facial movement for just one patient. Building profiles of movement from a sample therefore exponentially increases the volume of data to be processed and interpreted. Therefore knowledge of computational approaches is essential to ensure appropriate and efficient data analysis.

Outside of the clinical environment the broader role of facial movement lies within computer animation/vision, psychology, speech therapy and biometrics. Therefore the thesis will draw together multidisciplinary literature to understand current concepts of facial movement, review previous work in the field and present computational/statistical analyses of 3D data.

Contributions

The thesis presents a Geometric Morphometric (GMM) approach to the analysis of lip shape during movement. The main contributions are:

- an assessment of the reproducibility of landmark identification between manual and semi-automatic methods;
- an evaluation of the repeatability of verbal facial gestures when compared to non-verbal;
- the creation of an average model of lip shape during movement based on the most repeatable facial gestures. Experiments demonstrate that the model can discriminate different lip shapes from within the sample;
- data from a patient group pre- and post-surgery was added to the average model. Analyses showed differences in lip shape for the pre-surgical patient group that could be statistically identified from the average model. Post-surgery these differences were not statistically detectable.
- The impact of the study suggests that the average model can act as an outcome measure for lip shape during movement. A similar methodological approach can be adopted to compare matched populations in clinical situations where an objective measure of lip function is required.

Organisation of thesis

The thesis is organised into two distinct parts with Chapters 1 to 4 forming the literature review and Chapters 5 to 9 describing the experimental work undertaken as part of the study. Specifically;

- Chapter 1 describes the fundamental concepts of facial movement commencing with muscular anatomy and neurology. The main measures of facial movement (verbal or non-verbal gestures) are

explored in detail. Finally the clinical application of facial movement analysis is outlined.

- Chapter 2 focuses on the methods of facial movement analysis, presenting the relative advantages and disadvantages of each particular technique.

- Chapter 3 discusses the computational considerations for analysis of facial movement data to include image registration, feature detection and statistical shape analysis.

- Chapter 4 summarises the main issues related to orthognathic surgery as this is the patient group utilised in Chapter 8.

- Chapter 5 states the aim of the study and outlines the challenges and problems posed in achieving this.

- Chapter 6 begins the experimental section of the thesis. Landmark reproducibility is investigated comparing three methods of landmark placement.

- Chapter 7 utilises statistical shape analysis to compare the reproducibility of different facial gestures.

- Chapter 8 extends the findings of Chapter 7 to incorporate the most reproducible facial gestures into an average model of lip shape.

- Chapter 9 adds patient data into the average model in attempt to statistically distinguish lip shape pre- and post-operatively.

Thought has been given not to isolate components and therefore the reader is signposted to previous references throughout the thesis. The final section of the thesis summaries the contributions and findings of the study and describes the potential for future research.

Publications from thesis

Papers

- Popat H, Zhurov AI, Toma AM, Richmond S, Marshall D, Rosin PL. Statistical modelling of lip movement in the clinical context. *Orthodontics and Craniofacial Research* 5 (2) pp. 92-102, 2012.

- Popat H, Richmond S, Marshall D, Rosin PL. Three-dimensional assessment of functional change following Class 3 orthognathic correction - A preliminary report. *Journal of Craniomaxillofacial Surgery* 40 (1) pp.36-42, 2012.

- Popat H, Richmond S, Marshall D, Rosin PL. Facial Movement in 3 Dimensions: Average Templates of Lip Movement in Adults. *Otolaryngology Head and Neck Surgery* 145 (1) pp. 24 – 29, 2011.

- Popat H, Henley E, Benedikt L, Richmond S, Marshall D, Rosin PL. A comparison of the reproducibility of verbal and non-verbal gestures using 3D motion analysis. *Otolaryngology Head and Neck Surgery* 142 (6) pp. 867-72, 2010.

- Popat H, Richmond S, Benedikt L, Marshall D, Rosin PL. Quantitative analysis of facial movement - A review of three-dimensional imaging techniques. *Computerized Medical Imaging and Graphics* 33 (5) pp. 377-383, 2009.

Book chapters

- Richmond S, Al Ali AM, Beldie L, Chong YT, Cronin A, Djordjevic J, Drage NA, Evans DM, Jones D, Lu Y, Marshall D, Middleton J, Parker G, Paternoster L, Playle RA, Popat H, Rosin PL, Sidorov K, Toma AM, Walker B, Wilson C, and Zhurov AI. Detailing Patient Specific Modeling to Aid Clinical Decision-Making. In: *Patient-*

Specific Computational Modeling. Calvo B and Peña E (Eds.) Springer (5) pp. 105-131, 2012.

- Popat H, Richmond S, Marshall D, Rosin P L, Benedikt L. Assessing facial movement. In Kau C H and Richmond S (Eds). *Three-dimensional imaging for Orthodontics and Maxillofacial Surgery.* Wiley Blackwell, pp. 251-266, 2010.

- Benedikt L, Marshall D, Rosin P L, Cosker D, Popat H, Richmond S. Facial Actions for Biometric Applications. In Kau C H and Richmond S (Eds). *Three-dimensional imaging for Orthodontics and Maxillofacial Surgery.* Wiley Blackwell, pp. 267-286, 2010.

Prizes

- Postgraduate Research Day, Cardiff University, Nov 2011
  Peoples' Prize for Oral Presentation   £20

- Speaking of Science, Cardiff University, May 2010
  1st Prize Oral Presentation   £100

Chapter 1

Facial movement concepts

1.1.   Introduction

Lip shape forms an important part of the clinical assessment in specialities such as orthodontics and maxillofacial surgery where patients seek treatment for aesthetic and functional problems. In this context it is static lip shape that is studied. By using photographs and radiographs, lip thicknesses and positions in relation to the jaws and other facial features can be assessed which can aid the treatment planning process.

Dynamic lip shape (i.e. lip shape during movement) also has a significant role in facial aesthetics and function. For example, the presence of visible lip motion can increase word intelligibility in noisy conditions even without the tongue or teeth visible (Summerfield, 1979). Lip shape during movement is also useful as a feature for the understanding of facial expression and this concept is widely used in clinical psychology. In addition other associated areas of behaviour study believe lip shapes are related to personality traits and how individuals are perceived.

Lip shape is altered through neurological innervation which activates muscle groups to achieve verbal and non-verbal communication. Therefore the following paragraphs introduce the reader to the anatomical and functional relationships between the oro-facial muscles and surrounding structures such as the nose, eyes and mouth. Following this consideration is given to the two types of communication forms; verbal and non-verbal communication, and the facial gestures required to evoke these responses. As the thesis focuses on dynamic lip shape, the chapter is concluded by outlining the clinical applications of studying lip shape during movement.

1.2.   Anatomical considerations

The muscles of the face are embedded in the superficial fascia, and majority arise from the bones of the skull and are inserted into the skin (Figure 1-1). The primary function of the facial muscles is to act as sphincters or dilators of the eyelids, nostrils and mouth. This in turn

modifies the expression of the face, protects the eyes as in blinking or acts to articulate the lips in speech.

1.2.1. Muscles of the eyelids

The sphincter muscle of the eyelids is the orbicularis oculi and the dilator muscles are the levator palpebrae superioris and the occipitofrontalis. The orbicularis oculi muscle is composed of a palpebral part that closes the eyelid and an orbital portion that pulls on the skin of the forehead, temple and cheek to "screw up the eye". The corrugator supercilii is a small muscle that lies deep to the orbicularis oculi and produces vertical wrinkles in the forehead as in frowning (Snell, 2008).

1.2.2. Muscles of the nostrils

The sphincter muscle is the compressor naris, which compresses the mobile nasal cartilages. The dilator muscle is the dilator naris, which pulls the ala laterally widening the nasal aperture. In addition, the alar cartilages can be elevated slightly by the procerus and the levator superioris alaeque nasi to wrinkle the skin of the nose (Snell, 2008).

1.2.3. Muscles of the lips and cheeks

The sphincter muscle of the lips is the orbicularis oris, which compresses the lips together. Many of the fibres of this muscle are derived from the buccinator muscle, which lies laterally and compresses both the lips and cheeks against the teeth.

The dilator muscles of the lips radiate outwards and their action is to separate the lips. This movement is usually associated with jaw opening. The muscles arise from the bones and fascia around the mouth and converge to be inserted into the substance of the lips. There are several muscles involved and include levator labii superioris, zygomaticus major, zygomaticus minor, levator anguli oris, risorius, depressor anguli oris, depressor labii inferioris and mentalis (Snell, 2008).
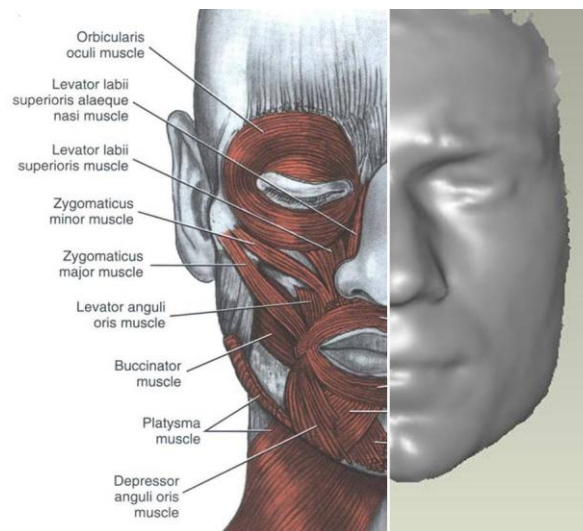
Figure 1-1 *Facial Musculature*

1.3. Neurological basis of facial movement

All muscles of the face are developed from the second pharyngeal arch
and are supplied their motor function by the facial nerve. The facial nerve
also contains a sensory part that supplies taste to the anterior ⅔ of the
tongue, the floor of the mouth and the palate. In addition, parasympathetic
secretomotor fibres are conveyed to the submandibular and sublingual
salivary glands (Snell, 2008). The facial nerve is the 7th cranial nerve and
emanates from the brainstem between the pons and medulla. It receives
impulses from multiple brain areas dependent on the type of movement
(Rinn, 1991). Voluntary movements are controlled by the motor cortex
through the pyramidal tract and represent the majority of the lower facial
muscles. The nerves fibres to the lower facial muscles are primarily contra
lateral in origin so that the right hemisphere of the brain activates motor
neurons of the left facial side, and vice-versa. Voluntary control of the
lower face directs learned behaviour. This provides fine control of the
facial region, which is required for speech articulation and facial
expression.

Involuntary movements are controlled by impulses from subcortical areas
through the extra-pyramidal tract and represent movement of many of the

upper facial muscles. The cerebral cortical projections to the facial motor neurons innervating the upper face are primarily bilateral whereby each cortical hemisphere provides innervation to both sides of the face (Reeves and Swenson, 2008). These movements are those concerned with emotion and feeling

1.4.   Non-verbal facial gestures

Each of the functional muscle units of the face can be innervated with different timing, intensity, and laterality characteristics. These characteristics produce the ability to create thousands of different non-verbal facial gestures or facial expressions. Facial expressions can be associated with different emotional signals or traits (Ekman and Matsumoto, 2008). These include:

- *speech illustration:* raising of the eyebrows when being inquisitive;
- *conversation regulation*: signalling a cue to others when finished talking;
- *emblematic gestures*: movements that symbolically give verbal meaning such as the doubtful look produced by raising the upper lip and pushing the lower lip up;
- *cognitive signalling*: brow furrowing when concentrating or perplexed;
- *functional traits:* using the facial muscles around the mouth area for eating and drinking.

Darwin was first to propose the link between emotions and facial expressions (Darwin, 1872). He claimed that facial expressions are the residual actions of more complete behavioural responses, and occur in combination with other bodily responses such as vocal, postural, gestural, skeletal muscle movements, and physiological responses. Since then, others have further enhanced our knowledge in this field (Ekman et al., 1972).

Table 1-1 *Facial muscle descriptions of Darwin's universal emotions*

Emotion	Darwin's facial descriptors	Non-facial elements
Anger	nostrils raised, mouth compressed, furrowed brow, eyes wide open, head erect	chest expanded, arms rigid by sides, stamp ground, body sways backwards/forwards, tremble
Contempt	lip protrusion, partial closure of eyelids, upper lip raised	snort, body expiration, expiration
Fear	eyes open, mouth open, lips retracted, eye brows raised	crouch, pale, perspiration, hair stands on end, muscles shiver, tremble
Disgust	lower lip turned down, upper lip raised, expiration, mouth open, spitting, blowing out, protruding lips	snort, body expiration, expiration
Sadness	corner mouth depressed, inner corner eyebrows raised	low spirits
Happiness	skin under eyes wrinkled, mouth drawn back at corners	laughter, clapping hands
Surprise	eyebrows raised, mouth open, eyes open, lips protruded	expiration, blowing/hissing, open hands high above head

Table 1-1 outlines facial muscle descriptions of Darwin's universal emotions which allows us to describe facial expressions such as anger by a furrowing of the brow, a tightening of the lips and teeth displayed because these actions are part of an attack response. Disgust is expressed with an open mouth, nose wrinkle, and tongue protrusion as part of a vomiting response (Matsumoto and Ekman, 2008). In summary, facial expressions are elements of a coordinated response involving multiple response systems.

1.4.1.   *Facial Action Coding System*

The Facial Action Coding System (FACS) is a well-known method for measuring and describing facial behaviours. It was developed as a system to categorise facial behaviours by determining how the contraction of each facial muscle (singly and in combination with other muscles)

changes the appearance of the face (Ekman and Friesen, 1978). FACS is generally carried out using videotapes of facial behaviour, which are deciphered by skilled examiners. Behaviours are coded into the muscular contractions that induce an appearance change.

The FACS defines expressions as one of 46 Action Units (AUs). These are either contractions or relaxations of one or more muscles. A selection of peri-oral AUs is described in Table 1-2. A FACS coder *dissects* an observed expression, decomposing it into the specific AUs that produced the movement. The scores for a facial expression consist of the list of AUs that produced it. Duration, intensity, and asymmetry can also be recorded. For example, a happiness expression is considered to be a combination of pulling lip corners (AU 12 + 13), and/or mouth opening (AU 25 + 27) with upper lip raiser (AU 10) and furrow deepening (AU 11).

Table 1-2   *Selected peri-oral AU descriptions (Kanade et al., 2000)*

AU	Facial muscle	Description of movement
10	Levator labii superioris	Upper lip raised; nasolabial furrow deepened producing square-like furrows around nostrils
11	Levator anguli oris	Lower to medial part of the nasolabial furrow deepened
12	Zygomaticus major	Lip corners pulled up and laterally
13	Zygomaticus minor	Angle of the mouth elevated; only muscle in the deep layer of muscles that opens the lips
14	Buccinator	Lip corners tightened. Cheeks compressed against teeth
22	Orbicularis oris	Lips everted
23	Orbicularis oris	Lips tightened
24	Orbicularis oris	Lips pressed together
25	Depressor labii inferioris	Lips parted
26	Masseter	Jaw dropped
27	Pterygoids and digastric	Mouth stretched open

The primary drawback to this comprehensive approach to coding facial behaviour is that it is labour intensive. A trained FACS operator can still take a number of hours to code one minute of video data depending on the complexity and density of facial expressions. In this respect, researchers have attempted to develop automated methods of face display analysis. Facial features have been automatically tracked in digitized image sequences using a hierarchical algorithm for estimating optical flow.   In this context, optical flow is the pattern of motion of the facial surface caused by the relative motion between the observer (i.e. the video camera) and the scene (Fleet and Weiss, 2005). Average agreement with manual FACS coding was 92% or higher for AUs in the brow, eye, and mouth regions. In the cross-validation set, average agreement was 91%, 88% and 81% for AUs in the brow, eye, and mouth regions, respectively (Cohn et al., 1999). It has been stated that an automated system would make facial expression measurement more

widely accessible as a research tool in behavioural science and investigations of the neural substrates of emotion (Bartlett et al., 1999).

1.4.2.   Deliberate versus spontaneous expressions

As the reader will discover in Section 2.4, most of the data that has captured non-verbal facial gestures is done so through subjects deliberately performing a series of facial expressions. As deliberate and spontaneous facial expressions are mediated by separate neural pathways (pyramidal and extra-pyramidal tracks respectively), directed facial actions may differ in appearance and timing from spontaneously occurring behaviour (Rinn, 1984). Indeed, several studies have shown that spontaneous smiles have a slower onset time (Cohn and Schmidt, 2004), are of shorter duration (Heck and Kleck, 1997), and are less asymmetric in lip commissure movements (Skinner and Mullen, 1991) when compared to deliberate smiles. These differences can be important from a psychology perspective as they can affect social judgements of people. For example, smile asymmetry has been shown to evoke negative social judgments and a lower level of trust-worthiness in observers (Brown and Moore, 2002). In addition, a study which temporally altered synthetic smiles showed that by increasing the speed of smile onset decreases their perceived genuineness (Krumhuber and Kappas, 2005).

From a clinical perspective, one may choose to limit the amount of emotion linked to an expression to elicit an objective pose. This may be most relevant when assessing outcome following a surgical intervention. Perceived outcome may influence both the timing and magnitude of an expression. Therefore, making a deliberate individual action will minimize the likelihood of producing an emotional experience (Hager, 1983).

1.5.   Verbal facial gestures

Speech-related facial motions or verbal facial gestures are another dynamic motor activity through which thoughts and emotions can be

expressed. Speech is composed of phonemes (sounds) and visemes (mouth shapes) that are closely related. Throughout the text, visemes are annotated in lower case and underlined if highlighting a particular section of a word. Phonemes are annotated according to language and/or written within slashes.

Table 1-3 shows a phoneme-to-viseme mapping based on General American English (Klatt, 1990). Note than many phonemes are visually synonymous and therefore several sounds can be mapped to a single mouth shape.

Table 1-3 *Phoneme to viseme mapping based on American English*

Viseme	Phoneme	Viseme	Phoneme	Viseme	Phoneme
p	P	k	K	ch	CH
p	B	k	G	ch	JH
p	M	k	N	ch	SH
p	EM	k	L	ch	ZH
f	F	k	NX	ey	EH
f	V	k	HH	ey	EY
t	T	k	Y	ey	AE
t	D	k	EL	ey	AW
t	S	k	EN	ao	AO
t	Z	iy	IY	ao	OY
t	TH	iy	IH	ao	IX
t	CH	ah	AA	ao	OW
t	DX	ah	AH	uh	UH
w	W	ah	AX	uh	UW
w	WH	ah	AY	sp	SIL
w	R	er	ER	sp	SP

For General American English, every phoneme is represented by one or two capital letters. Numerical digits are used as stress indicators (0 = no stress, 1 = primary stress, 2 = secondary stress) and are placed at the end of a syllabic vowel. Word examples are given in Table 1-4.

Table 1-4   *Phonetic word examples based on American English*

Phoneme	Word examples
AO	frost (F R AO1 S T)
UH	should (SH UH1 D)
OY	boy (B OY1)
ER	bird (B ER1 D)
TH	thanks (TH AE1 NG K S)

As the classification of phonemes will vary across languages, the phonetic transcription of British English or Received Pronunciation (Roach, 2004) with word examples is also presented in Table 1-5 and Table 1-6.

Table 1-5   *Phonetic transcription of consonants in Received Pronunciation*

Articulation	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					tʃ dʒ			
Nasal	m			n			ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ			h
Approximant	(w)				r	j	w	
Lateral approximant				l				

Word examples:

p	b	t	d	tʃ	dʒ	k	g
<u>p</u>uppy	<u>b</u>aby	<u>t</u>ree	<u>d</u>og	<u>ch</u>eese	<u>j</u>oke	<u>c</u>oin	<u>g</u>o

f	v	θ	ð	s	z	ʃ	ʒ
<u>f</u>ree	<u>v</u>ideo	<u>th</u>ing	<u>th</u>is	<u>s</u>ee	<u>z</u>oo	<u>sh</u>eep	televi<u>s</u>ion

m	n	ŋ	h	l	r	w	j
<u>m</u>ouse	<u>n</u>ow	thi<u>ng</u>	<u>h</u>ope	<u>l</u>ove	<u>r</u>un	<u>w</u>e	<u>y</u>ou

Table 1-6 *Phonetic transcription of vowels in Received Pronunciation*

Short vowel	Example	Long vowel	Example	Diphthongs	Example
ɪ	sh<u>i</u>p	iː	sh<u>ee</u>p	eɪ	b<u>a</u>by
e	l<u>e</u>ft	ɑː	f<u>ar</u>	aɪ	l<u>i</u>ke
æ	h<u>a</u>t	ɔː	d<u>oor</u>	ɔɪ	c<u>oi</u>n
ʌ	p<u>u</u>ppy	uː	sh<u>oo</u>t	ɪə	h<u>ere</u>
ɒ	b<u>o</u>b	ɜː	h<u>er</u>	eə	h<u>air</u>
ʊ	b<u>oo</u>k			ʊə	t<u>ou</u>rist
ə	<u>a</u>bout			əʊ	r<u>o</u>pe
				aʊ	m<u>ou</u>th

1.5.1.   Articulation and resonance of speech sounds

The production of all speech sounds is the result of modulation of the airflow from the lungs. The speaker must produce a stream of exhaled air and then modulate it in ways that make it audible to the listener. This is done by passing air through the larynx and into the oral or nasal cavities when it is then modified to form different consonant and vowel sounds (Table 1-7). Consonant sounds are classified by voice, place or manner production. Voiced sounds are made when the vocal cords are vibrating, e.g. /b g d z/ (and all vowels). If the vocal cords are apart, voiceless sounds are produced, e.g. /p t k s/ (Wyatt et al., 1996).

Table 1-7 *Classification of different phonemes by articulation (Cawley, 1996)*

Articulation	Speech Sound	Description
Labiodental	/v f/	Lower lip rises and touches upper teeth
Bilabial	/b m p/	Articulation involving both lips
Alveolar	/t d s z/	Tip of tongue to alveolar ridge
Glottal	/h/	Articulation involving the glottis
Velar	/k g/	Back of tongue and velum
Dental	/th/	Front of tongue to upper front teeth
Palatal	/y/	Front of tongue and hard palate
Palatoalveolar	/sh/	Blade of tongue forms a long channel for friction to occur along the posterior alveolar ridge and the hard palate

1.5.2.   Oral cavity

The oral cavity is bounded anteriorly and laterally by the teeth, which are embedded into the alveolar process of the maxilla and mandible. The most important teeth for speech are the incisors. Both the upper and lower central and lateral incisors are used with the lower lip, the tongue, and with each other to create a constriction for sounds such as /f/ and /s/. The hard palate, which forms the roof of the oral cavity, is formed by the palatine process of the maxillary bone in its anterior two-thirds and the palatine bone in its posterior third. The superior alveolar ridge and its palatal rugae generate many speech sounds as a result of actions of the tongue along them.

1.5.3.   Soft palate

The greater part of the soft palate consists of a broad muscle entering the sides of the velum from the temporal bones on each side. These muscles, named the levator palatine muscles elevate the soft palate and close the

entrance to the nasal cavities above. As the levator palatine muscles contract, the soft palate is lifted up and backwards toward the posterior wall of the pharynx. This action, termed velopharyngeal closure is of special relevance, as it occurs to some degree for all of the speech sounds in English, except nasal consonants such as /m/ and /n/ which require nasal resonance. For these exceptions, the port to the nasal cavities is left open by relaxing the levator palatine muscles.

1.5.4.   Tongue

The tongue is capable of moving as a mass in three-dimensions through its extrinsic musculature. The styloglossus muscles are attached to the styloid process of each temporal bone. The muscle fibres run down and forward inserting into the sides of the tongue. Contraction of this muscle pulls the tongue back and up. This movement is important for sounds such as /u/ as in "Sue."

The hyoglossus muscles are attached to the hyoid bone, and the fibres run in a thin sheet up into the lateral base of the tongue. Contraction results in tongue depression. The sound /a/ has a low tongue position.

The genioglossus muscles are attached to the medial surface of the mandible at the superior mental spine. The muscle fibres radiate up and back to insert throughout the length of the tongue, down to and including the hyoid bone. Contraction of these muscles draws the hyoid bone and tongue anteriorly. This is important for speech sounds such as /e/, as in "see."

While the extrinsic muscles determine the gross position of the tongue body, the intrinsic muscles determine the shape of its surface. The superior longitudinal muscle curls the tongue tip up, while the inferior longitudinal muscle depresses the tongue tip. Together, the intrinsic muscles can shape the tongue tip into a variety of configurations.

1.5.5.   Lips

Contraction of orbicularis oris, the major lip muscle can close the lips for bilabial sounds such as /p/ /b/ or /m/. It can also protrude the lips for /u/ or can act as a sound source as in /f/ where air passes through the gap between the upper incisors and lower lip under pressure, causing friction.

1.5.6.   Disorders of speech

Disorders of speech can be classified into either organic or psychogenic in origin (Aronson, 1980) and acquired or developmental. Dysarthrias are a family of motor speech disorders that are organic in origin and comprise the most common acquired disorders of speech and language. They may include neurological and cerebro-vascular conditions such as Parkinsonism, Multiple Sclerosis, Stroke and Cerebral Palsy. Structural abnormalities of particular interest to the orthodontist, or craniofacial clinician that also fall into the heading of organic speech disorders include cleft lip and palate and malocclusions.

1.5.7.   Cleft speech

Resonance and nasality are terms used in speech therapy that describe problems associated with the nasopharynx and nasal cavity. Disorders of resonance, hyper- and hyponasality are distinctive tones of voice. Whilst voice sounds are generated in the larynx; the pharynx, the mouth and the nose shape them.   An adequate functioning velopharyngeal sphincter provides a barrier into the nasal cavity to achieve this. Excess resonance in the nasal cavity results in hypernasal resonance. This characteristic is generally associated with velopharyngeal incompetence as it indicates a lack of separation between the oral and nasal cavities (Barkana, 2009).

Consonant production in cleft palate speech is frequently characterised by nasal airflow (nasal emission and nasal turbulence) with, or instead of oral airflow. It is most readily perceived on production of voiceless consonants /p t k s f sh/ which require maximum pulmonary air pressure. Particular

sentences for elicitation of speech in cleft patients include, "The puppy is playing with a rope" and "Bob is a baby boy" as they target lip and oropharnygeal articulation (Sell et al., 1999).

1.5.8. Malocclusion and speech

A malocclusion results in abnormal relative positions of the teeth, lips and tongue. It has been suggested that there may be three possible mechanisms by which malocclusion and speech may be interrelated (Harvold, 1970). Firstly, there may be an occlusal and/or skeletal problem and a coincidental articulatory problem. There may also be a genetic or metabolic disorder affecting the central nervous system, which will lead to poor motor control and possible distorted morphogenesis. Finally, there may be a true cause and effect where occlusal or structural anomalies affect articulatory skills.

The relationship between increased overjet and speech disorders has been under long dispute. There is a relationship between Class II division 1 malocclusions and the production of the /s/ sound. Target hard and soft tissue positions for correct /s/ sound production include soft palate raised, incisors edge-to-edge, tongue just behind upper incisors and a channel along the dorsum of the tongue to allow air to "squirt" along the tongue and between the teeth.

Orthodontic and speech analysis carried out on children attending an orthodontic department showed that skeletal morphology had little if any influence on the production of interdental stigmatism due to compensation (Blyth, 1956).

In contrast, subjects with an increased overjet have been found to protrude the mandible to a greater extent than a normal occlusion group in the production of the /s/ sound (Benediktsson, 1958). There have also been significant differences found in lip position, incisor position and

tongue position within and between speakers during the production of the /s/ sound, depending on phonetic context (Subtelny and Oya, 1972).

Class III subjects often have a tongue posture that is lower and more retruded than normal (Guay et al., 1978). During phonation of /s/, these subjects further retruded their tongue in an attempt to achieve a normal tongue-tip-to-upper-incisor relationship. Despite these compensatory movements, normal production of the /s/ sound was severely diminished in this sample of patients.

Subjects with anterior openbite are 63% more likely to misarticulate than those subjects with a positive overbite (Fymbo, 1936). There is also a strong association of anterior openbite with lisping (Bernstein, 1954) and the production of significantly defective speech sounds of /s z th/ and /l/ (Pomerantz and Zeller, 1965).

Increased overbite has been associated with lateral lisping (Ingervall and Sarnas, 1962). Occlusal and speech analysis also shows a statistically significant correlation between misarticulation of /s/ sounds and deep overbite (Lubit, 1967).

For certain speech sounds such as bilabials (e.g. /p b m/) where the lips are brought into contact with each other, the ability to achieve lip competence is important. The antero-posterior position of the incisors will have a role to play in either allowing or obstructing the lips to make contact. In addition, the vertical skeletal relationship will be important in determining the distance between the origins of the muscles of the lips, and hence the ease with which lip competence may be achieved.

Despite the evidence cited in the preceding paragraphs suggesting a link between malocclusion and speech, it is rare that the main aim of orthodontic treatment would be the correction of a speech problem. As such it is difficult to draw firm conclusions on the correlation between

malocclusion and speech. The reason for this is that speech is a complex process involving several organs including the brain, neural networks, teeth, lips, tongue and muscular components. The multifactorial nature of speech production may mean than organs mutually compensate to ensure that pronunciation is correct (Johnson and Sandy, 1999).

1.6. Clinical application of facial movement analysis

1.6.1. Orthognathic surgery

Corrective jaw surgery (termed orthognathic surgery) is used to treat a variety of facial and jaw abnormalities in which the upper and lower jaws and the teeth are beyond the scope of orthodontic treatment alone. Common complaints from patients include a concern regarding their facial appearance, an inability to chew their food properly and speech problems (Williams et al., 2005). A combination of orthodontics to place the teeth in the correct position followed by orthognathic surgery to reposition the jaws can provide an acceptable aesthetic outcome (Figure 1-2).

Initial presentation ➡ **3 months post-operative**

Figure 1-2 *Class 3 orthognathic patient*

The stability of this treatment is multi-factorial and includes the age of the patient (Ekstrom, 1982), the amount and type of jaw movement (Costa et al., 2001), the type and material of fixation (Politi et al., 2002), and an adaption of the soft tissues (i.e. lips, cheeks and related muscles) to the new jaw position post-surgery (Van Sickels and Richardson, 1996). Although factors for stability such as the planned jaw movement and fixation material are controllable, others such as the soft tissue reaction

are far less predictable. This is in part due to the lack of data available on how the soft tissues move pre- and post-surgery, and how this is related to stability of the post-surgical jaw position. The role of orthognathic surgery in relation to the context of the thesis is discussed in more detail in Chapter 4.1.

1.6.2. Cleft lip repair

The incidence of cleft lip (and/or palate) is approximately 1 in 700 and although the primary lip repair is carried out at 3 months of age (Figure 1-3), many patients require lip revision in their adolescence.

Figure 1-3   *Cleft lip and repair at 3 months of age*

Cleft lip revision may be carried on aesthetic and/or functional grounds (Talmant, 2006). However, it is hypothesized that increased tissue scarring due to further surgery causes impaired circum-oral movement (Trotman et al., 2010). An objective technique for recording facial movement would be highly beneficial for determining this effect and would form an important factor on balancing aesthetic versus functional gain (Figure 1-4).

Figure 1-4   *Bilateral cleft lip repair in function - lip purse*

1.6.3.   *Facial nerve pathology*

Patients with facial nerve damage exhibit a weakness of the muscles of facial expression. In addition to the functional impairment there may be a direct impact to psychosocial well-being. The most common cause of facial nerve paralysis is Bell's palsy which was previously considered as an idiopathic condition (Peitersen, 1992). It has now been linked to the herpes simplex virus that is thought to induce an inflammatory swelling of the facial nerve in its canal (Murakami et al., 1996). As the canal is long and narrow, swelling can result in damage either by direct pressure or by impairing blood flow in the nerve. Defective taste sensation to the anterior ⅔ of the tongue can also be present. Most cases of uncomplicated Bell's palsy recover well but its severe form long-term facial nerve weakness can occur due to an incomplete process of regeneration. This is more common when a longer course of the nerve is affected (Phillips and Bubash, 2002).

Table 1-8 lists the many other causes of facial nerve palsy. A variety of assessments exist to test function of the facial nerve and these are discussed in detail in Chapter 2. Treatment methods include electrical stimulation, physiotherapy and nerve grafts (Jackson and von Doersten, 1999). Objective assessment of facial movement forms an integral part of

monitoring the recovery and rehabilitation process. A reliable measure of facial movement is required as an indication of what would be considered normal or acceptable movement to aim to achieve during the management of the condition.

Table 1-8   *Causes of facial nerve palsy (Pestronsk, 2012)*

Type	Cause
Congenital	Moebius syndrome (congenital nuclear aplasia) Myotonic dystrophy Hemifacial microsomia
Neurological	Myasthenia gravis Multiple sclerosis
Neoplastic	Facial nerve tumour Parotid tumour
Infectious	Otitis media Bacterial (tuberculosis) Viral (herpes simplex)
Traumatic	Temporal bone fractures Facial lacerations High altitude palsy

1.6.4.   Visual speech recognition

Speech recognition is a technique of understanding conveyed information from individual(s) using visual and auditory signals from the lips, tongue and vocal chords. As each speech sound (phoneme) has a particular lip shape or viseme (Table 1-7), individuals with auditory impairments can utilize these visual characteristics to interpret the movements of the lips allowing them to *speech read*. Indeed, lip movement is known to play an important role in both sign language and communication between the deaf (Pearson, 1981). Adequate visibility of the face clearly benefits speech perception due to the vocal articulators being observable and can help

disambiguate speech sounds that can be confusable from acoustics alone, e.g. the unvoiced consonants /p/ (a bilabial) and /k/ (a velar) (Aleksic et al., 2009). This has motivated research into the automatic recognition of visual speech (AVSR) which has significant advantages over traditional audio-only based automatic recognition which can suffer from background noise (Rajavel and Sathidevi, 2009). Research techniques aim to use reliable imaging of lip movements from large sample groups. Tracking of lip movements allow feature extraction and visual modelling of visemes (Kumar et al., 2007).

1.6.5.   *Biometric face recognition*

In contrast to speech recognition, which aims to decipher different lip sounds from lip shapes, the goal of biometric face recognition is to find similarities in lip or face shapes between individuals. Recent interest in biometric authentication has stemmed from such applications as national ID cards, airport security and surveillance/site assess (Bhattacharyya et al., 2009). Well-known biometrics used in these situations include fingerprints, iris scans and face recognition. Face recognition in 2D suffers weaknesses in identification accuracy when images are taken at different angles or variable lighting conditions are used (Zhao et al., 2003). Therefore, with the ever-increasing demand for heightened security, research has focused on 3D face recognition.

Recent work has aimed to incorporate 3D facial movement into biometrics to raise protection levels further and is dependent upon reliable 3D facial movement capture (Benedikt, 2009). This hypothesises that individuals are able to repeat movements precisely over time and relies on the premise that movements are distinct enough between individuals to act as adequate security measures. Large *face databases* have been created, primarily from the computer science community to test these hypothesises using complex mathematical algorithms and template matching techniques (FACEREC). The databases vary between institutions and can include non-verbal data such as facial expressions, e.g. happiness (smiling),

surprise (eyebrow raising), disgust (mouth opening) and sadness (frowning). Verbal databases tend to be phonetically rich but can include any form of speech with some examples being the uttering of digits '0' to '9', and spoken phrases such as "Joe took father's green shoe bench out" (Pigeon and Vandendorpe, 1997).

1.7. Summary

There are several clinical applications for which facial movement analysis would be of benefit. The applications are dependent on an objective measure of the muscles of facial movement in function.   Facial function can be expressed in the form of verbal or non-verbal gestures. Verbal gestures such as everyday words limit the functional analysis to that of the circum-oral region. Although non-verbal gestures allow muscles of the entire face to be analysed they are highly correlated with emotional responses. The following chapter explores the past and currently employed techniques to capture facial movement in the clinical context.

Chapter 2

Capturing facial movement

2.1. Introduction

Some of the earliest work describing facial movement was carried out in the mid-nineteenth century by Duchenne, a French neurologist (Duchenne, 1862). He triggered muscular contractions by electrically stimulating various areas on the surface of the face and photographed the resulting facial expressions with a camera (Figure 2-1).

Figure 2-1 *Facial expressions triggered by electric stimulation*

From: Mécanisme de la Physionomie Humaine by Guillaume Duchenne, 1862, Fig. 4, p. 277

Duchenne managed to determine that smiles resulting from true happiness not only utilize the zygomaticus major muscles, which raise the corners of the mouth, but also the orbicularis oculi muscles of the eyes. Such genuine smiles are now commonly known as "Duchenne smiles".

Although Duchenne is considered as one of the developers of electro-physiology and electro-therapeutics the large surface electrodes used in his studies would have produced a diffuse electrical current, activating numerous nerve structures and resulting in a loss of control of muscle activations. In addition he often combined muscle stimulation with voluntary movement by his participants and as a result, the specific movement associated with individual muscles could be considered ambiguous. The lack of precision in early work on measuring facial movements rendered many reports as essentially verbal descriptions of

facial expression and did not attempt to characterise the entire face as a whole or to quantify selected facial expressions.

2.2. Grading Scales

It was not until the mid-to-late twentieth century that a move toward grading scales to classify facial movement was proposed. A system for awarding points out of 10 for each of 10 categories (tone, wrinkle forehead, close eyes tightly, blink, wrinkle nose, grin, whistle, blow out cheeks, depress lower lip and tense the neck) was developed (May, 1970). A normal response was graded as 10, a weak but present response as 5, and an absent response or flaccid tone as 0. A composite score for each of the 10 categories was calculated. Although this was an easily applied regional system for the description of facial movement it lacked precise qualitative description as responses can only be categorised as normal, weak or absent.

2.2.1.   House-Brackmann Facial Nerve Grading System

A further review of the subjective grading scales that were of historic importance and clinical use at the time was conducted in 1983 which led to the conclusion that a gross scale which gives a single grade to evaluate facial movement and secondary effects was most appropriate in the subjective description of facial paralysis. A new system of classification was proposed based on this review and the expert opinion of clinicians in the field (House, 1983). The so-called House-Brackmann Facial Nerve Grading System (House and Brackman, 1985), was subsequently adopted by the American Academy of Otolaryngology/Head and Neck Surgery, the American Otological Society and the American Neurotology Society as the standard for reporting facial nerve paralysis (Table 2-1).

Table 2-1 *House-Brackmann Facial Nerve Grading System*

Grade	Description	Characteristics
I	Normal	Normal symmetrical function in all areas.
II	Mild Dysfunction	Slight weakness noticeable only on close inspection. Complete eye closure with minimal effort. Slight asymmetry of smile with maximal effort. Synkinesis barely noticeable, contracture, or spasm absent.
III	Moderate Dysfunction	Obvious weakness, but not disfiguring. Complete eye closure and strong but asymmetrical mouth movement with maximal effort. Obvious, but not disfiguring synkinesis, mass movement or spasm.
IV	Moderately Severe Dysfunction	Obvious disfiguring weakness, inability to lift brow. Incomplete eye closure and asymmetry of mouth with maximal effort. Severe synkinesis, mass movement, spasm.
V	Severe Dysfunction	Motion barely perceptible, incomplete eye closure, slight movement corner mouth. Synkinesis, contracture, and spasm usually absent.
VI	Total Paralysis	No Movement

The scale includes grades that are arbitrary numbers in sequential order from least to most severe deficit and represents a direct and straightforward approach to the clinical description of facial nerve function. Although it is easy to use and offers a simple language for the description of facial function, there is a wide range of facial movements incorporated within a single grade and overlap between grades is highly likely. Another major drawback of this model and other subjective models is the large

inter-observer disagreement in grading so that for the same clinical appearance; different examiners may award different categories. Attempts to make gross scales more sensitive to small changes in facial function by expanding the scale to include intermediate grades has been shown to increase inter-observer disagreement (Coulson et al., 2005).

An objective component was added to the House-Brackmann six-point gross scale in an attempt to increase its validity. By measuring the movement of the eyebrow and corner of the mouth on the side affected by facial nerve palsy as well as the normal side, the grades can be converted into numerical data for statistical analysis. Each reference point scores 1 point for a movement of 0.25cm up to a maximum of 1cm (i.e. 4 increments of 0.25cm for the eyebrow and another 4 increments for the mouth). This gives a total possible score of 8 on each side of the face. These results can be combined with the six-point gross scale and can also be expressed as a percentage of function. As no reference point is defined, the placement of the measuring instrument is very much observer dependent. In addition, the increments of 0.25cm may not be sufficiently sensitive to describe facial movement fully (Linstrom et al., 2000).

2.2.2.   Sunnybrook Facial Nerve Grading System

The Sunnybrook facial nerve grading system (Ross et al., 1996) was developed to address the limitations of the House-Brackmann system. In the main part, that the grading system's range of scoring did not reflect clinically important change. The Sunnybrook system grades the function of the facial nerve using three components: resting symmetry, voluntary movement and synkinesis to give a single percentage from 0 to 100 (Figure 2-2).

Each component can change which will individually contribute to the overall score thereby making the system highly sensitive and specific (Neely et al., 2010a). One of the disadvantages of this system is the requirement for writing down a variety of numerical estimations and then to

perform several calculations to arrive at the result. With the development of electronic records this is no longer a major obstacle but as the scale is subjective, the system is still vulnerable to observer disagreement and bias.

Sunnybrook Facial Grading System		

Figure 2-2 *Sunnybrook Facial Grading System*

2.2.3.   *Facial Nerve Function Index*

The Facial Nerve Function Index (FNFI) is a percentage which uses a single expression, the broad smile, to represent facial function (Fields and Peckitt, 1990). The distance between the lateral canthus and the lateral oral commissure at rest is measured for the normal side and the side affected by facial nerve palsy separately. The same distances are measured with the broadest possible smile. The *at rest* distance minus the smile distance for the affected side is divided by the same calculation on the normal side and multiplied by 100 yielding the FNFI.

Although the system is simple, straightforward and objective it lacks information about velocity, acceleration or synkinesis. It also has built-in observer error as it lacks a fixed reference point which is especially relevant in cases of bilateral facial nerve palsy (Pothiawala and Lateef, 2012).

2.2.4. *Linear Measurement Index*

The Linear Measurement Index (LMI) is a collection of measurements from five standard facial expressions (forehead wrinkle, kiss, nose wrinkle, smile and eyes closed tight). A total of seven inter-landmark distances are specified for the five facial expressions, e.g. superior orbit to inferior orbit for eyes closed tight. The LMI is calculated as a score on the paralysed face as a percentage of the normal side. The step-wise calculation can be seen in Table 2-2 (Burres, 1985a).

Table 2-2 *Calculation of the Linear Measurement Index*

Step	Calculation
1	Calculate the percent displacement (pD) for all of the appropriate measurements for each facial position: pD = change in distance (mm) x 100
2	Average the two pDs from the Nose Wrinkle, Smile and Eyes Closed Tight. Divide the pD from the Forehead Wrinkle in half.
3	Total the five values from each half of the face separately.
4	Corneal Exposure: For each millimetre of corneal exposure, measure by the maximum width of the palpebral fissure, subtract 2% in the Eyes Closed Tight from the score for the paralysed half of the face.
5	Rest Asymmetry: Add the millimetre difference between the two halves of the face for all seven measurements taken at rest. If the total is greater than 20mm, subtract 1% from the score on the paralysed side of the face for each additional millimetre.
6	Add 30 to the score from each half of the face.
7	Calculate the score on the paralysed side of the face as a percentage of the normal side. If the total on the paralysed side after steps 1 to 6 is less than 0, the total function on the paralysed half of the face is considered 0.

To enable statistical analyses, negative scores occurring during the calculation of the linear measurements had to be managed. Here, a negative score of 30 was deemed to be a reasonable lower limit for all LMI scores. Therefore any paralysed half-face with a score less than negative 30 was scored as 0 (Step 6, Table 2-2). Otherwise, a score of 30 was added to the total score for each half of the face before the percent motion on the paralysed side was finally calculated. Although providing a degree of objectivity to the analysis of facial movement, the calculations in achieving the LMI are rather protracted and at times confusing. Clinically, inter-observer variability of the LMI has been recorded between 28 and 40% (Burres and Fisch, 1986). It was suggested that more categories within a grading scale increases the inter-observer variability. This concurs with additions made to the House-Brackmann Scale which was discussed in Section 2.2.1.

2.2.5.   *Nottingham System*

The Nottingham System is comprised of three parts in which the first provides an objective measure of facial movement and the second and third parts provide a record of the presence of absence of secondary defects (Murty et al., 1994).

In the first part, four reference points are directly marked onto the patient's face (supra-orbital point, infra-orbital point, the lateral canthus and the angle of the mouth). Two inter-landmark distances (supra- to infra-orbital distance, and lateral canthus to angle of mouth) are measured for both sides at rest and at maximal effort for eyebrow raising, eyes closed tightly and smiling. The sum of the distances for each side is calculated and the lesser is expressed over the greater as a percentage.

In the second part, the clinician records the presence of absence of secondary defects, e.g. hemifacial spasm, contractures or synkinesis.

Finally, in the third part, the presence of absence of crocodile tears, decreased lacrimation or dysgeusia (determined by a questionnaire). The Nottingham System is reported to be quicker to apply than the LMI and superior to the House-Brackmann Scale due to its objectivity and the provision to record secondary defects (Murty et al., 1994).

2.3. 2D Techniques

2.3.1. Anthropometry

Anthropometry derives its meaning from the Greek to *measure man* and has been widely used to study the variation in human physical form. Medical anthropometry involves the identification of anatomical landmarks through which linear and angular inter-landmark measurements can be obtained. The variation of these measurements within a population can then be studied or indeed compared to other population groups (Farkas et al., 1992). In medical anthropometry, 47 craniofacial landmarks have been described (Farkas, 1994) – a subset of these landmarks is illustrated in Figure 2-3 and defined in Table 2-3.

Figure 2-3 *Subset of anthropometric landmarks*

Table 2-3   *Landmark definitions cross-referenced to Figure 2-3*

No.	Landmark	Definition
1	Vertex (v)	Highest point of the head with the subject in Natural Head Position
2	Trichion (tr)	Midpoint of the hairline
3	Euryon (eu)	Most lateral point on the head
4	Glabella (g)	Most prominent point in the median sagittal plane between the supra-orbital ridges
5	Orbitale superius (os)	Highest point on the margin of the orbit.
6	Exocanthion (ex)	Outer corner of the eye fissure where the eyelids meet
7	Orbitale (or)	Lowest point on the margin of the orbit.
8	Endocanthion (en)	Inner corner of the eye fissure where the eyelids meet
9	Soft Tissue Nasion (n)	Outer point of intersection between the sella-nasion line and the soft tissue profile
10	Pronasale (prn)	Most protruded point of the nasal tip
11	Alare (al)	Most lateral point on the nasal ala
12	Subnasale (sn)	Junction between the lower border of the nasal septum, the partition that divides the nostrils, and the cutaneous portion of the upper lip in the midline
13	Chelion (ch)	Outer corner of the mouth where the edges of the upper and lower vermilions meet
14	Crista philtre (cph)	Point on the crest of the philtrum, the vertical groove in the median portion of the upper lip, just above the vermilion border
15	Labrale superius (ls)	Midpoint of the vermilion border of the upper lip
16	Stomion (sto)	Midpoint of the labial fissure when the lips are closed naturally
17	Labrale inferius (li)	Midpoint of the vermilion border of the lower lip
18	Sublabial (sl)	Midpoint of the labiomental sulcus
19	Pogonion (pg)	Most anterior point in the middle of the soft tissue chin
20	Gonion (go)	Most lateral point at the angle of the mandible
21	Menton (me)	Lowest point in the midline on the lower border of the chin
22	Otobasion inferius (obi)	Lowest point of attachment of the external ear to the head
23	Subaurale (sba)	Lowest point of the ear lobe.
24	Postaurale (pa)	Most posterior point on the helix of the ear
25	Superaurale (sa)	Highest point of the free margin of the ear
26	Tragion (t)	Point above the tragus of the ear

Techniques for anthropometric assessment vary in respect to whether they are direct or indirect (Farkas and Deutsch, 1996). For direct anthropometry, landmarks are drawn on the face and measured directly. Indirect anthropometry requires landmarks to be identified in some manner prior to imaging. The measurements are then recorded on the output of the imaging modality.

The advantage of direct anthropometry is the ability to measure areas of the face covered by hair (e.g. length and height of the head). Disadvantages include soft tissue compression of anatomical areas during measurement distorting results and the prolonged time needed to perform the examination, which is directly influenced by the number of measurements in the assessment. In addition, the reliability of the measurements in direct anthropometry is dependent on precise landmark identification and a reliance on patient cooperation. The technique may not be suitable for young children who can become restless during the assessment (Farkas and Deutsch, 1996).

Indirect methods of craniofacial surface anthropometry have the advantage of a short examination time and less dependence on patient compliance. The disadvantage of indirect anthropometry is largely dependent on the imaging modality used to capture the data. Precise landmark identification is also a consideration here. These issues are discussed in the following sections.

2.3.2.   Photography

Some of the earlier studies on facial motion were based on analyses of 2D images, for example, measurement of the amplitude of facial landmark motions during animations using standardised photographs taken at rest and at maximal facial animation (Johnson et al., 1994). Subjects wore physical markers on selected facial landmarks arranged in a standardised configuration and an adhesive ruler on their faces.

Although the method allowed quantification of all regions of the face simultaneously, manual placement of marking dots and identification of markers' positions in multiple frames of the video sequence increases the opportunity for error in measurement and is a recognized limitation of the assay (Johnson et al., 1994). The method can also be labour intensive and time consuming for both operator and patient. Information from a single camera limits motion information to 2D feature extraction and representation. The use of physical markers may inhibit spontaneous facial motion and limits analysis to preselected features (i.e. facial landmarks). Since no information is given regarding motion of the face between response and maximal excursion, non-linear motion is ignored and there is no representation of the actual path of motion of the facial landmarks (Wachtman et al., 2001).

2.3.3.   Cine Camera

A further study used a 16mm cine camera to quantify the 2D trajectories of various lip landmarks during natural smiling and reported both the amplitude and direction of the landmark motions (Paletz et al., 1994). It has been documented however, that the trajectories represented composite data from multiple smiles because the central and lateral lip landmarks were not sampled simultaneously (Gross et al., 1996).

Another group developed a 2D analysing system without markers using a single camera (Wachtman et al., 2001) but as 2D amplitudes underestimate 3D amplitudes by as much as 43% (Gross et al., 1996), the latter has been recommended for the evaluation of facial motion.

2.4.   3D Techniques

2.4.1.   Direct Measurement

The evaluation and assessment of facial expressions by direct measurement was one of the earlier methods used to analyse and record facial movement (Burres, 1985b). Indeed direct measurements of a

patient's smile taken manually with a handheld ruler compared favourably in accuracy when compared to the same measurements taken from videotaped recordings (Manktelow et al., 2008). Although this can allow values in all three planes to be recorded, patient fatigue and distortion of results due to skin depression are potential problems. In addition, the direction of the movement cannot be objectively assessed. A means of capturing data to enable analysis at a later date was required.

Motion capture systems are an established technique used in the field of gait analysis. They provide kinematic (measurement of the movement of the body in space) and kinetic information (the forces involved in producing the movements). Landmarks are placed on anatomical points as described in Section 2.3.1 and *tracked* during movement using multiple cameras. As the subject moves trajectories of the markers are calculated in 3D – thereby providing *3D motion analysis*. In gait analysis this information usually concerns joint mobility (Davies et al., 1991). More recently, 3D facial motion capture systems have evolved by means of the same principle. There are however different demands and approaches between facial motion capture and body motion capture. With body motion capture, movements of a structured bony framework are analysed which primarily involve generic, fluid movements. Facial motion capture on the other hand has a wider range of movements and expressions, and in addition lifelike nuances and subtle variations. These movements are often less than a few millimetres and therefore require greater resolution and fidelity than usually used in body motion capture. Nonetheless, the main systems which have dominated the field of facial motion capture are the *marker based tracking systems* – similar to body motion capture systems. Using the concept of stereophotogrammetry they consist of groups of one or more cameras that are connected to a personal computer and track multiple landmarks placed on the subject's face. Several of these systems have been described as in-house developments using either *passive* or *active* based markers (Popat et al., 2009).

2.4.2.   *Passive marker based tracking systems*

Passive marker based systems use markers coated with a retro-reflective material to reflect light back that is generated near the cameras lens. The camera's threshold can be adjusted so only the bright reflective markers will be sampled, ignoring skin and fabric. The facial markers can vary in diameter from between 2 to 6mm and can number up to 30 and are placed on specific facial landmarks.

To obtain 3D coordinate data for a marker, two cameras must record the marker position in space. As markers on the face may be carried outside the field of view of the two primary cameras, additional cameras can be used to ensure that data from at least two cameras is always recorded. Prior to recording facial movements, the cameras must be calibrated by way of an object with an array of markers whose positions in space are certified to a known accuracy. For example, C3D (Glasgow Dental Hospital/Faraday Laboratory, University of Glasgow, UK) is a stereophotogrammetric camera system (Johnston et al., 2003) based on two pairs of video cameras placed either side of the subject with integral illumination. Preliminary investigations with the system using a dummy head indicated that it had an accuracy of 0.1mm (Johnston et al., 2001). The main study recruited subjects who were pre-labelled with 20 landmarks on the face. Images were captured while performing a sequence of five facial expressions (rest, natural smile, maximal smile, lip purse and maximal cheek puff) while maintaining natural head posture. They concluded that the extent of facial expression reproducibility was expression specific and that differences exist between males and females.

The Motion Analysis System (Motion Analysis™, Santa Rosa, CA) uses analogue video cameras with a focal length of 25mm to capture data at a sampling rate of 60 frames per second (fps). Spherical retro-reflective markers of between 2 and 4mm diameter were attached to selected facial landmarks. Subjects performed a variety of maximal facial animations including smile, lip purse, cheek puff, eye closure, eye opening and

grimace. Among the findings of these particular studies has been the modelling of facial movement in patients with cleft lip and palate (Trotman et al., 2005), the influence of sex and facial shape in the 3D analysis of facial movement in normal adults (Weeden et al., 2001) and dynamic analysis of differences caused by orthognathic surgery (Nooreyazdan et al., 2004).

Vicon™ (Vicon 250, Vicon Motion Systems, 2000) tracks the motion of 2mm reflective markers placed on the face at 60 fps using 5 infrared video cameras and has been used to quantify spontaneous facial motility in infancy (Green and Wilson, 2006). Four separate reference markers were placed on the forehead to correct for head movement that would otherwise be included in the facial movement signals. Image-processing software identified the marker locations in each 2D infrared camera image to compute its 3D location relative to a calibration plate that was positioned in the data collection room. The system has been documented for its usefulness in research applications (Frey et al., 1994) but has also been reported to be too complicated for daily clinical application (Frey et al., 1999).

FACIAL CLIMA has also been described and uses marker based tracking with infra-red cameras. It has been used on healthy subjects and is proposed in the assessment of the functional outcome of facial paralysis re-animation surgery (Hontanilla and Auba, 2008).

2.4.3. Active marker based tracking systems

Active marker systems use markers that are triggered by an incoming infrared signal and respond by sending out a corresponding signal of their own. This signal is then used to triangulate the location of the marker. The advantage of this system over the passive is that individual markers can work at predefined frequencies and therefore, have their own *identity*. This means that post-processing of marker locations is minimized. One such system is described as the Lip Function Monitor. It uses a facemask

mounted with infrared light emitting diodes (LED's) on one-side and photodiode light detectors on the other. Direct coupling means that when the mouth is closed, the detector cannot *see* the emitters as its path is blocked by the lips. Conversely, when the mouth opens the detector is incident upon the emitter. The photodiode circuits convert the infra-red emission to an output of voltage which can be extrapolated on computer (Dawes and Kelly, 2005). This technique gives a continuous, instantaneous readout of the movement of the LED's, but is only suitable for extra-oral use. Whilst a signal can be picked up for an intra-oral LED, the signal becomes corrupted as it penetrates the soft tissue (Jemt, 1981). Another limitation is that the wires connecting the LED's to the power source have to pass between the lips that may directly interfere with normal speech.

2.4.4.   Marker reproducibility

Although marker based systems have been widely used in dynamic facial analysis, one of the main problems is the error introduced by placement of landmarks between two different sessions. As point landmarks are manually placed by the clinician or observer it is important that the same landmark is identified as closely as possible between images taken at different time points. This concept is termed *landmark reproducibility*.

It has been suggested that a landmark can be considered reproducible if, between placements, the standard deviation from the centroid is 0.5mm or less in all three planes of space (Hajeer et al., 2002). In this study, 25 out of 30 facial landmarks were within this threshold. However another more recent study found only cheilion (left and right), labrale superius and exocanthion (left) to be reproducible if these guidelines are followed (Gwilliam et al., 2006). Other studies suggest that a clinically acceptable level of reproducibility between landmark placement is up to 1mm (Toma et al., 2009). The reproducibility of landmark placement is better with the same (intra-) observer when compared to multiple (inter-) observers (Plooij et al., 2009).

Software programs have been developed which document the use of automatic landmarking tools which would solve the labour-intensive nature of manually landmarking multiple images (Dibeklioglu et al., 2010). Despite an initial image requiring manual landmarking for the algorithm to bootstrap the subsequent images, they could represent an innovation in terms of landmark-based analysis. Despite this, it is unclear whether automatic landmarking is comparable in terms of agreement with manual landmarking.

Clearly systems devoid of markers solve the limitations described above and those that have evolved have been coined as *marker-less.* This is somewhat a misnomer as landmark coordinates need to be identified on the captured image (rather than being directly placed on the subject). Therefore an alternate nomenclature would be to term these systems as *indirect marker based.* Several of these systems have now become commercially available and they represent the latest developments in the field of 3D facial motion capture.

2.4.5.   Indirect marker based/marker-free systems

A relatively early study implemented videotaping to record facial movement without the use of facial markers (Neely et al., 1992). The video recording was digitized and absolute pixel counts were subtracted from sequential frames from the baseline frame to give computer generated x–y plots of relative facial movement. The results recommended the development of computerized systems to analyse facial movement as they enabled objective disease profiles to be constructed. It was not possible to create these profiles when subjective grading scales were used (Neely et al., 1996).

3D Video™ (OGIS Research Institute Co. Ltd., Osaka, Japan) is a 3D computer imaging program which has been used to analyze the motion of the lips (Mishima et al., 2006). The hardware system consists of three

infrared digital video cameras, one colour digital video camera, a synchronizing signal generator, two distributors of the signal, an infrared projector, four digital video recorders and a personal computer. Following calibration of the cameras using a checkerboard cube an infrared pattern was projected onto the subject and images recorded through the infrared cameras at 30 fps while the subject phonated. Lip motion was analysed by dividing the upper and lower vermilions into four areas. The upper vermilion consisted of five manually placed landmarks and one Bezier curved line, with the lower consisting of two landmarks and two Bezier curved lines. Following manually fitting of the Bezier lines, a virtual grid was made over the lips which allowed shift and velocity during lip motion to be calculated in each area (Mishima et al., 2006). By applying a template-matching technique, range images could be produced across the whole images during motion. Transferring their methods, this group have analysed lip motion of cleft patients before and after lip repair (Mishima et al., 2009). Here statistical modelling techniques were employed to analyse the data. The use of these techniques in 3D shape analysis is discussed in Section 3.5.

The SmartEye® Pro system has been used to explore lip movement in patients with muscular dystrophy (Sjogreen et al., 2011). The system can be used with up to six cameras and utilises infrared to capture facial geometry. The authors in this study analysed maximal lip movement during an open smile and lip pucker when compared to the rest position. The data used in the analysis incorporated a measure of overall mouth width distance between the commissures during maximal movement and a measure of asymmetry using the left and right mouth widths measured from the oral commissure to the midline. In addition, horizontal, vertical and anterior-posterior oral commissure displacements from the rest position to maximal movement using the mean distance error were calculated. The results showed that healthy individuals had a fairly symmetric mouth width at rest and when performing an open mouth smile but less so with a lip pucker was in general less symmetric. Mouth width,

mouth width change and mouth width asymmetry for the open mouth smile and lip pucker differed significantly from the healthy adults when compared to the muscular dystrophy group. The authors did not specify the frame capture rate of the system and included a several subgroups of patients as part of the muscular dystrophy cohort, which may limit the interpretability of the results.

Two further commercially available marker-less 3D motion analysis systems; the *3dMDface™ Dynamic System* (3dMD, Atlanta, GA) and the *4D Capture System* (DI3D™, Dimensional Imaging, Glasgow) are discussed in more detail below.

The *3dMDface™ Dynamic System* is based on active stereophotogrammetry using a random infrared speckle projection. Their latest generation system is a six camera, two pod system comprising of six 1.3 mega pixel cameras - four grey scale and two colour (Figure 2-4).

Figure 2-4   *3dMDface™ Dynamic System*

It can capture static images as well as sequential imaging up to 60 fps. The output files are 3D geometry definition files that are generated as one continuous point cloud produced from the two stereo camera viewpoints. The continuous point cloud can theoretically eliminate the data errors

associated with merging/stitching left and right sides from separate cameras. Distortion of the infrared speckle as it projects onto a subject's face allows the generation of 3D coordinates of the face through complex algorithms. Several thousand points make up the cloud for a 3D image of the face. White light is used to capture colour texture images of the face simultaneously with the infrared pattern produced 3D data. Small infrared components of the white light may have the potential to drown some of the infrared speckle pattern and therefore compromise image capture. The use of cold-illumination in the form of two Bowens Tri-lites© (Bowens International Limited, UK) can minimize this effect.

The passive stereophotogrammetric *4D Capture System* from DI3D™ uses a stereo pair of low noise monochrome cameras to capture three-dimensional shape and a single colour camera to capture the colour texture images using only regular white light projection.

Figure 2-5   *DI3D™ 4D Capture System*

The specification for the standard 4D system from DI3D™ is a three camera, single-pod comprising two 2 mega pixel grey scale cameras and one colour camera functioning at 30 fps (Figure 2-5). Their latest high-performance 4D system is also a single-pod with two 4 megapixel grey scale cameras functioning up to 60 fps. Two pod versions of both DI3D™ systems are also available.

2.5. Summary

There are several methods of capturing facial movement, which can range from simple grading scales to commercially available 3D motion scanners. The choice of method employed is multifactorial and may depend on the clinical information required, patient burden, cost, clinical expertise, speed of analysis, as well as the reliability and validity of the system in question. However, as technological advances are made in medical imaging, the availability and use of 3D techniques has increased which has led them to represent the gold standard in terms of objectivity. The following chapter discusses the concepts of analysing and interpreting 3D motion data.

Chapter 3

Modelling shape variation

3.1. Introduction

3D imaging systems output their data to reconstruct facial geometry. As *x, y* and *z* coordinates define a space in which multi-dimensional data are represented the data can be analysed mathematically and statistically which allows shape, structure, size, volume and spatial relationships of anatomical structures to be objectively assessed.

3.2. Image registration

Image registration, in the clinical context has a wide range of potential applications. These include:

- monitoring changes in size, shape or image intensity over time (Morris et al., 1999);
- relating an individual's anatomy to that of a standard atlas (Broadbent, 1939);
- relating preoperative images and surgical plans to the physical patients post-operatively to assess outcome (Noh et al., 2011);
- combining information from multiple imaging modalities, e.g. CBCT, digital study models and stereophotogrammetric surface scans to build a virtual 3D patient (Popat and Richmond, 2010).

The fundamental concept of image registration is that important information can be contained in more than one image. This may be a set of two or more images taken after a specified time period following a treatment intervention. Here, the user may be interested in structural correspondence, i.e. aligning similar anatomical structures to detect the response of the treatment. For the purposes of dynamic imaging which acquires several images from the same patient at one sitting, the user may require functional correspondence, i.e. alignment of functionally equivalent areas of the face to assess movement.

Image registration (also known as *fusion*, *matching* or *warping*) is a key tool for extracting information from these multiple image sets. It can be

defined as the process for determining the correspondence of features between different time points or using different imaging modalities. Thus, given a source and a target image, the goal is to find a transformation, such that the transformed target is similar to the source image. This process is underpinned by mathematics and computer science which utilise semi- or fully-automatic registration algorithms to achieve optimal correspondence between image sets. A registration algorithm can be broadly decomposed into the following components:

- *feature detection* - salient and distinctive points, closed-boundary regions, edges, contours, line intersections, corners, etc. are manually or preferably automatically detected;
- *feature matching* - correspondence between the features detected in the target image and those detected in the source image is established;
- *transformation model* – specifies the way in which the target image can be changed to match the source;
- *optimisation* – varies the transformation model in such a manner to maximise the registration of the target and source images.

Registration algorithms differ in their methodology according to the focus of the application. For example, the field of industrial inspection and tracking requires high volume processing and therefore computing time is paramount. In this respect, the registration algorithm may preference lower quality feature detection to maximise the computing speed. On the other hand, precision feature detection is vital in character recognition biometrics and therefore remote sensing is prioritised. For medical treatment planning, high-resolution 3D images are required for model accuracy and as such image acquisition and computer memory becomes a priority. Given there is no established unified theory for image registration the following sections discuss the contemporary methods and problems encountered within each of the registration components.

3.2.1.   *Feature detection/matching*

Feature detection elements of image registration algorithms can be classified into *feature-based* or *area-based* (also called *intensity-based*). Feature-based methods are further divided into point-based and surface-based registrations (Zitova and Flusser, 2003). They both rely on unique points, corners, edges or curvatures in an image that characterise the geometric layout of an image. The features should be distinct and spread all over the image so as to enable efficient detection across images. It would be expected that the majority of features remain stable over time at fixed positions.

3.2.2.   *Points and Procrustes*

Point-based registration involves identifying corresponding 3D points on the images to be aligned, registering the points and determining the transformation of the images from these points. There are three discrete types of points or landmarks.

Anatomical landmarks may be pins or markers fixed to the patient's skin and visible on the captured images or placed after image capture on the 3D shell (Sections 2.3 and 2.4).   They should represent the same feature in different images and correspond to anatomical structures which can unambiguously defined, for example the inner and outer corners of the eye fissure where the eyelids meet (*endocanthion* and *exocanthion* respectively) or the most anterior point on the tip of the nose (*pronasale*) (Figure 3-1).

Pseudo-landmarks are constructed points on a structure located around the outline or in-between anatomical landmarks. Continuous curves can be approximated using pseudo-landmarks, which make them suitable for approaches such as in the analysis of hand shapes (Grenander et al., 1991).

A further type of landmark is the semi-landmark, which is a point located on a curve and allowed to slip a small distance with respect to another corresponding curve (Bookstein, 1991a). The term semi- is used because the landmark lies in a lower number of dimensions than other types of landmarks, e.g. along a one-dimensional curve in a two-dimensional image. Several criteria have been proposed to slide points along an outline. Two of the most widely used are minimum bending energy (Section 3.3.1) and minimum Procrustes distance (Perez et al., 2006).
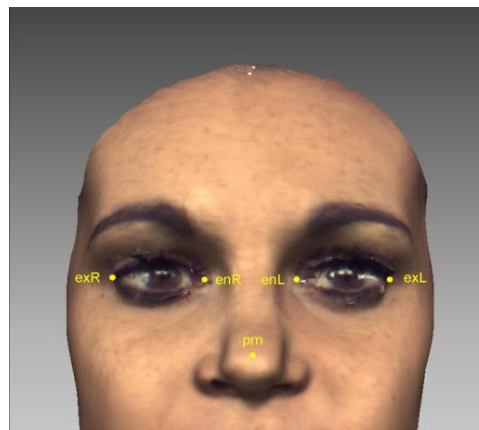
Figure 3-1   *Clearly defined point landmarks*
en L/R endocanthion, ex L/R exocanthion, prn pronasale

Procrustes registration computes the average or *centroid* of each set of points and the difference between the centroids equals the translation that must be applied to one set of points. This point set is then rotated about its new centroid until the sum of the squared distances between the each corresponding point pair is minimised (Dryden and Mardia, 1998). The square root mean of this squared distance is often referred to as the root mean square (RMS). The problem draws its name from the from the Procrustes areas of statistics. Procrustes was a robber in Greek mythology. He would offer travellers hospitality in his house and the opportunity to stay the night in a bed that would perfectly fit each visitor. However it was the visitors that were adjusted to fit the bed by stretching short people and removing the limbs of taller guests. Procrustes analysis is now a well known method of shape analysis and uses the least squares

method to optimally fit two configurations of *N* points in *D* dimensions (Dryden and Mardia, 1998).

For registration of semi-landmarks, the minimum Procrustes distance criterion removes the difference along the curve in semi-landmark positions between the reference form and each specimen by estimating the direction tangential to the curve and removing the component of the difference that lies along this tangent (Sheets et al., 2004).

Procrustes analysis is a *rigid-body transformation*, which, as the name suggests is one that changes position and orientation through translation and rotation without changing size and shape. The issue of rigid-body versus non-rigid registration is discussed in more detail in section 3.3. However, there are several Procrustes variations that allow size change through scaling of which Generalised Procrustes Analysis (GPA) is an example.

3.2.3.   *Surface-based registration*

Boundaries or surfaces, such as the surface of the face can frequently be more distinct than landmarks and therefore if equivalent surfaces of the face can be matched, then a rigid-body registration can be achieved by fitting the two surfaces together. Probably the most well known example of this is iterative closest point (ICP) algorithm which repeatedly selects points on either the target and/or source image and finds the closest points in the other mesh such that the RMS between the two point sets is minimised (Besl and McKay, 1992). Typically the source or reference shell is designated as the frame of the subject at rest and all consecutive target shells are superimposed to this. The main drawback to ICP registration is that the method cannot cope with non-overlapping regions because outliers are never removed. Moreover, when starting from a rough estimation of the translation of one object to another, the convergence is not guaranteed, e.g. anatomically similar areas of the face such as the forehead may register but the facial shells may be orientated

in different directions (Salvi et al., 2007). Commercially available reverse-engineering software can quantify shell-to-shell deviations in the form of a colour map that can easily be interpreted.   Figure 3-2 shows selected colour deviation maps from a smile expression, feature aligned to the rest or source frame. The green colour represents polygon differences between two frames of ± 0.5mm and therefore very little, or no movement. The red spectrum represents strong positive changes that are seen to affect the zygomaticus muscles (up to 5mm) overlying the maxilla as they contract to bring the corners of the lips upwards. In contrast, the blue spectrum corresponds to negative changes as seen in the upper lip and depressor anguli oris muscles which flatten (up to -4.3mm) as the commissures of the lips lift.
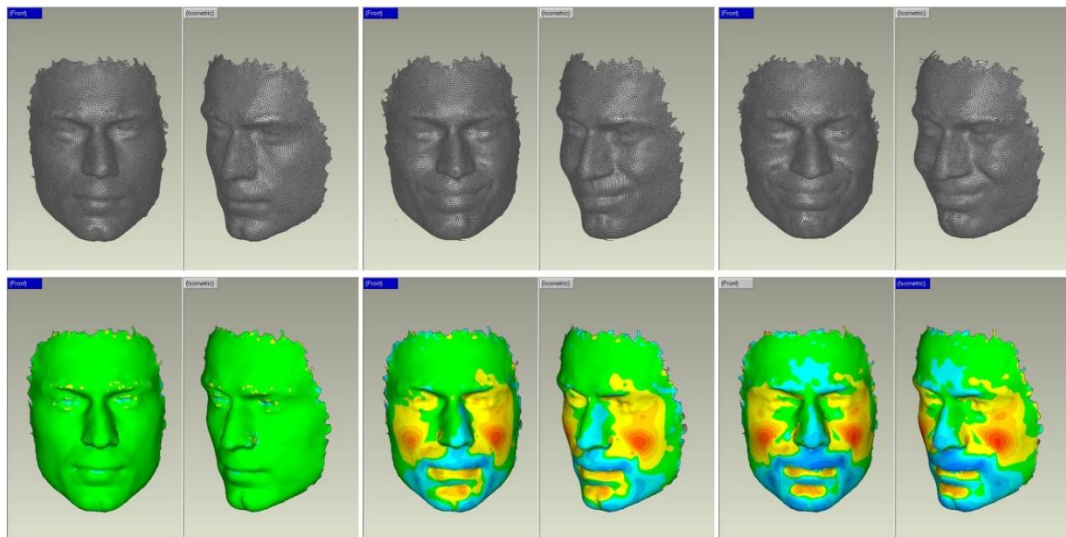
Figure 3-2   *Colour deviation maps from a smile expression sequence T0, T10 and T15 (T=0.02 second)*

The common drawback of feature-based methods is that the respective features may be hard to detect and/or unstable over time. The 3dMDFace™ Dynamic Imaging System has been used in preliminary studies to assess facial movement during a standardised smile expression and the verbal facial gesture /puppy/ via surface matching on the upper face (Popat et al., 2008). It was shown that there were insufficient areas

that remain stable on the face to accurately assess facial changes during a sequence of facial movement using a standardised smile facial expression when compared to using spoken word as small errors in alignment of the upper third of the face can lead to larger errors in the other regions of the face. A solution to this problem which can convey higher accuracy is a feature-based matching approach which separately matches the eyes, the forehead and the nose regions, i.e. the facial regions which are relatively less sensitive to movement during facial movement (Mian et al., 2007).

3.2.4. *Area- based registration*

In contrast to feature-based matching, area-based methods tend to be applied when the images lack prominent details. Here, the distinctive information is provided by grey levels/colours rather than by geometrics such as local shapes and structure (Bunting et al., 2008). They assume that the images will be most similar at the correct registration and tend to be used in voxel based imaging such as magnetic resonance imaging (MRI) and computed tomography (CT) scanning. The simplest measure is by minimising intensity differences, which measures (and minimises) the sum of squared intensity differences (SSD) between images during registration (Crum et al., 2004). Other measures of similarity include correlation coefficient, measures based on optical flow, and information-theoretic measures such as mutual information (Hill et al., 2001). As these imaging techniques are not utilised in the thesis, this section has been only been included for completeness.

3.3. **Transformation models**

The transformation component of a registration algorithm defines how one image can be deformed to match another. Ideally, the function design should transform the features/intensities of the target image to that of the source image such that they should lie as close to correspondences as possible.

Transformations can be either *rigid* or *non-rigid*. A rigid-body transformation in 3D is defined by six parameters (degrees of freedom): three translations and three rotations with the key characteristic that all distances are preserved (i.e. size and shape). Some algorithms increase the number of parameters by allowing for anisotropic scaling (giving nine degrees of freedom) and skews (giving 12 degrees of freedom). A transformation that includes scaling and skews as well as the rigid-body parameters is referred to as *affine*, and has the important characteristics that it can be described in matrix form and that all parallel lines are preserved. A rigid-body transformation can essentially be considered as a subset of affine, in which the scaling values are all unity and the skews all zero. Rigid-body transformations tend to be applied to images from the same modality and same subject. Considering this in the clinical context, individual bones are rigid (at the resolution of radiological imaging modalities), and therefore rigid body registration is widely used in where the structures of interest are either bone or are enclosed in bone.

In some cases, neither rigid-body nor affine transformations will adequately describe the alignment required. Inter-subject registration requires more complex transformations to account for the many anatomical differences between individuals. In addition, non-rigid registration may be necessary to account for differences in posture or organ size and position in serial studies (e.g. pre- and post-surgery).

3.3.1. Thin plate splines

A thin plate spline (TPS) is an interpolation method that finds a minimally bended smooth surface that passes through all given landmarks of two different shape configurations (Bookstein, 1989). They are part of a family of splines that are based on radial basis functions. *Splines* are smooth piecewise polynomial functions (Schonberg, 1945) and are named after devices used by shipbuilders to draw smooth curves (Segner, 1986). The name *thin plate spline* refers to a physical analogy of how an infinitely thin metal plate would behave if it was forced through the same control points

– this is also known as the *bending energy* (Richtsmeier et al., 1992). The TPS model matches the two configurations in a way that minimises the bending energy - if the configurations are identical the bending energy is zero (McIntyre and Mossey, 2003).

TPS registrations are based on the assumption that a set of corresponding points or landmarks can be identified in the source and target images. This is analogous to the use of point landmarks used in the Procrustes method (Section 3.2.2). In spline registration, these corresponding points are called *control points*. TPS transformations interpolate or approximate the displacements which are necessary to map the location of the control point in the target image to its corresponding counterpart in the source image (Bookstein, 1991b). This facilitates the construction and display of a transformation/deformation grid that captures the shape change between the images (Figure 3-3).
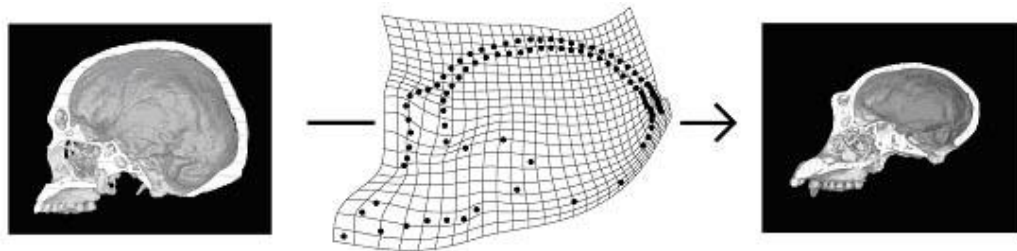
Figure 3-3 *Thin plate spline deformation between a human and chimpanzee skull*

Taken from Virtual Anthropology (Thompson, 1917)

In the clinical field of orthodontics and within the assessment of growth, this morphometric approach has gained increasing popularity. For example, TPS has been used to investigate the shape characteristics of the face and tongue in obstructive sleep apnoea (Pae et al., 1997), the role of the cranial base and the morphological differences of mandibular shape in Class III malocclusions (Singh et al., 1997) as well as in

longitudinal cephalometric comparisons between treated and an untreated Class III patient groups (Baccetti et al., 1999).

Although TPS interpolation finds a smooth transformation between from one set of data to another, it does not find a unique correspondence between two images except at the landmark points (Johnson and Christensen, 2001).

3.3.2. B-splines

B-splines are an alternative to TPS registration for the non-rigid registration of images. They are based on free-form deformations (FFD) and have been widely used for medical imaging as well as computer animation graphics (Rueckert et al., 1999). In TPS registration, each control point has a global influence on the transformation. Therefore, for a large number of control points, the computational function of radial basis function splines becomes complex and time consuming.

The theory behind a FFD is to deform an object by embedding it into a B-spline volume. The shape of the object is changed by moving the B-spline control points (Figure 3-4). This requires a mesh of control points with uniform spacing, in contrast to radial basis function splines which allow arbitrary configurations of the control points (Rueckert, 2001).
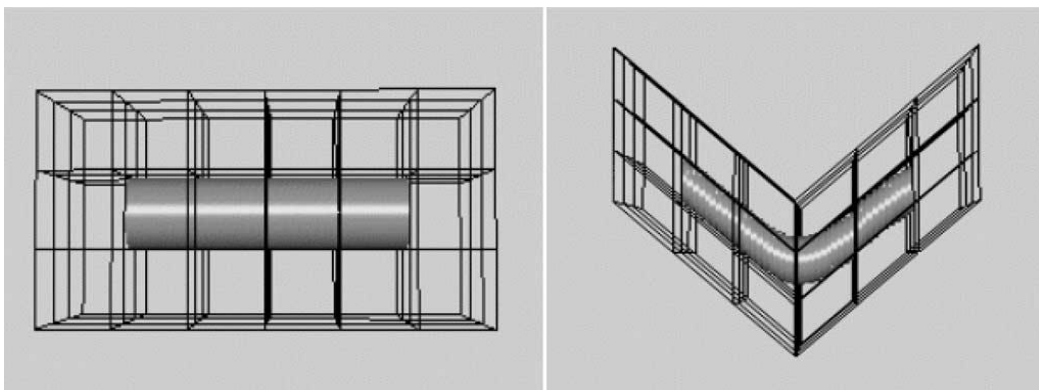
Figure 3-4 *Example of a free-form deformation*

Taken from Figure 1 (Xie and Farin, 2004)

As FFDs are controlled locally, they are computational efficient and changing a control point will only affect the transformation in the local region.

3.4. **Paired/Groupwise registration**

Based on the number of images within the dataset, image registration can be divided into *pairwise* and *groupwise* methods (Zitova and Flusser, 2003). Given two images, called the source image and target image, a registration algorithm (based on any method described in Sections 3.2 and 3.3) uses the pairwise method to calculate the most suitable transformation to bring the two images into correspondence. If multiple images are being considered, consecutive target images can be aligned by repeatedly registering them to source image constituting a repeated pairwise registration. This raises the problem of source image bias as repeated pairwise registration will be affected by the choice of the reference image. If the source image is not characteristic of the sample, alignment will be corrupted and the results are likely to be statistically biased. Groupwise approaches consider the entire group of images within the dataset simultaneously rather than just a pair. In broad terms, groupwise registration utilises as much information as possible from the entire sequence of images to increase the quality and robustness of the registration process (Cootes et al., 2005).

3.5. **Statistical shape analysis**

Statistical methods for representing facial form vary in how closely they model the anatomical structures of the face and how they exploit the shape information. The following sections describe the main geometric morphometric approaches for the multivariate statistical analysis of Cartesian coordinate data, which is usually restricted to landmark point locations.

3.5.1.   Principal component analysis

Principal component analysis (PCA) is a well known statistical technique applied to a set of variables to discover which of those variables in the set form coherent subsets that are relatively independent of one another (Tabachnick and Fidell, 2007). The analysis is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. It is particularly useful for data of a high dimension as this is reduced to a more manageable size for interpretation without loss of information. The goals of PCA have been quoted as (Abdi and Williams, 2010):

- to extract the most important information from the data set;
- to compress the size of the data set by keeping only this important information;
- to simplify the description of the data set;
- to analyse the structure of the observations and the variables.

The first principal component (PC) extracted in the analysis accounts for a maximal amount of total variance in the observed variables. The second component extracted will account for a maximal amount of variance in the data set that was not accounted for by the first component. The second component will be uncorrelated with the first component so that if you were to compute the correlation between components 1 and 2, it would be zero. The remaining components that are extracted in the analysis display the same characteristics; namely each component accounts for a maximal amount of variance in the observed variables that was not accounted for by the preceding components, and is uncorrelated with all of the preceding components. For example, in a biometric study which uses PCA to recognise whether the mouth state is open or closed, PC1 explains over 80% of the total variation in the sample (Bouvier et al., 2008).   This is a variation in mouth opening. Subsequent components correspond to more specific mouth shape variation (e.g. long/wide mouth opening) or differences in commissure/philtrum position. Latter PCs are those

corresponding to extremely small amounts of mouth shape variation. Smaller PCs can be ignored and typically only those components covering 90–95% of the variation should be retained.

PCA is either performed on the co-variance matrix or the correlation matrix of the observed variables. If the measured variables are of different unit scales, the correlation matrix should be used as the input to PCA. This allows standardisation of the dataset such that each variable is transformed to have a mean of zero and a variance of one. The total variance in the data set is the sum of the variances of these observed variables. Therefore, the total variance of PCA using the correlation matrix will be equal to the number of observed variables being analysed, i.e. 6 variables will give a total variance of 6. The co-variance matrix is utilised to preserve variance if the range and scale of variables is similar or in the same units of measure.

A historically important application of PCA within the computer vision industry was in the development of Eigenfaces. (Turk and Pentland, 1991). Eigenfaces was the first successful example of face recognition and utilises PCA to describe a set of features that characterise the variation between different face images. The process is initialized by first acquiring the *training set* which comprises of a number of images of the subjects involved (containing the same number of pixels). Including multiple images per person increases the accuracy of the set due to the increased information available on each known subject. The average face within the training set is then calculated following which the difference of each face from the average is computed. The differences can be used to compute a co-variance matrix for the dataset which reveals how much the two sets of data correlate. The eigenvectors and eigenvalues of the co-variance matrix are then calculated. Each eigenvector has the same dimensionality or number of components as the original images, and therefore can itself be seen as an image. The eigenvectors of the co-variance matrix are therefore called Eigenfaces and they represent the

different ways that the individual images in the training set differ from the mean image.

Eigenfaces is based on 2D analysis and therefore suffers limitations. In particular its performance is dependent on the lighting conditions, changes to head size (scale), orientation and position (i.e. if the face is not centred within the image frame) (Kuhn et al., 1998). Moreover with the current accessibility of 3D imaging techniques, Eigenfaces has largely been superseded.

In 3D analyses, PCA can be seen under several guises dependent on the field of application. In the phenotyping of 3D facial morphology in different population groups, the term dense surface model (DSM) has been used (Hammond et al., 2004). A DSM can be described as a form of point distribution model (PDM) where a large number of densely corresponded surface points are induced or interpolated using a set of manually placed landmarks (Hammond and Suttie, 2012). The resultant model refers to the set of principal components which account for the shape variation in the sample (Hutton et al., 2003). The PCA representations of the faces can be combined with genetic data to study phenotype-genotype correlations.

Active surface models (ASM) are another synonym for using PCA first developed within the field of computer vision (Cootes et al., 1995). They exploit the information provided by PDMs by constructing a training set which is a representation of each subject in a sample as a set of labelled landmark points. The mean positions of the points and the main ways in which the points from each individual vary from the mean are then calculated. The ASM can be used to iteratively search for the modelled structures within the dataset (Cootes et al., 1994). An extension of ASMs is the Active Appearance Model (AAM). Whereas ASMs model shape only, AAMs use PCA to model both shape and texture variations seen in a training set of visual objects (Abboud and Chollet, 2005). Much of the

work on ASMs and AAMs in computer vision has been used for the automatic recognition of facial expression (Fasel and Luettin, 2003).

3.5.2.   Principal coordinate analysis

Principal coordinate analysis (PCoA) is another data reduction method similar to PCA. It is also known as also known as metric multidimensional scaling (MDS). PCoA uses spectral decomposition to approximate a matrix of distances/dissimilarities by the distances between a set of points in two or three dimensions (Gower, 2005).

PCoA searches for similarities between cases by analysing a distance matrix in an attempt to make the distance between any pair of points proportional to the sample distance. The result of a PCoA is a set of coordinates on a number of derived axes such that similar cases are close together. It is not possible to associate these axes with any variables. This is because as PCoA uses a distance matrix the analysis relates only to the cases and therefore information about the original variables is lost. However, this means that it is particularly useful when there are a large number of predictors relative to the number of cases.

3.5.3.   Factor analysis

Factor analysis (FA) is a similar procedure to PCA although the most important distinction is that FA assumes the co-variation in the observed variables is due to the presence of one or more latent variables that exert causal influence on the observed variables. Therefore, mathematically in FA, only the variance that each observed variable shares with the other observed variables is available for analysis. This shared variance is represented by communalities (scores between 0 and 1) that are inserted in the positive diagonal of the correlation matrix. The solution in FA concentrates on variables with high communality values. In contrast, PCA makes no assumption about an underlying causal model and is concerned only with establishing which linear components exist within the data and how a particular variable might contribute to a component.

Interestingly, studies have shown that solutions generated from PCA differ little from those derived from FA (Guadagnoli and Velicer, 1988) although this is largely dependent on the number of variables. For data with 30 or more variables and communalities greater than 0.7 for all variables different solutions are unlikely (Stevens, 2009). This is the converse for less than 20 variables and communalities under 0.4.

3.5.4.   Canonical correlation analysis

The goal of canonical correlation analysis (CCA) is to analyse the relationship between two sets of variables – one set being independent variables (IV), the other dependent (DV). It provides a statistical analysis in which each observation is measured on two sets of variables to see whether the two sets relate to each other.

The first step in canonical analysis is the generation of a correlation matrix – in this case the matrix is subdivided into four parts: the correlations between the DVs, the correlations between the IVs and the two matrices of correlations between the DVs and IVs. The final correlation matrix is a product of all four matrices.

The procedure then calculates the two linear combinations (one from the IV and one from the DV) that have the maximum possible correlation. This is termed the first canonical variate. The procedure then searches for a second pair of linear combinations, uncorrelated with the first such that the correlation between this pair is the next highest – the second canonical variate. Eigenvalues are given to the canonical variates and the corresponding eigenvectors are called canonical correlations - the eigenvalue is equal to the squared canonical correlation of the canonical variate.

Canonical analysis has been described as the most generalised member of the multivariate statistical techniques (Anderson et al., 2009).

Examples of CCA used in clinically related studies are certainly less well found in the published literature than those using PCA. Due to the mathematical nature of the method, the computer vision industry continues to be the purveyor of these analyses with the main application being that of automated facial recognition (Kim et al., 2007).

3.5.5.   *Mahalanobis distance*

The Mahalanobis distance in its simplest form is a measure of the distance between two points in a data set. This includes the distance of a point to the mean of a distribution or the distance between the centroids (means) of two correlated independent distributions (Kapoor et al., 2010). In this respect, the Mahalanobis distance has several applications including the detection of outliers (Rousseeuw and Leroy, 1987), investigating the representativity of data sets (Jouan-Rimbaud et al., 1997) and in discriminatory analyses (Wu et al., 1996). It is calculated using the inverse of the variance–co-variance matrix of the data set of interest (Mahalanobis, 1936). The Mahalanobis distance ($D)$ is a metric value and is measured in terms of standard deviations from the centroid (McLachlan, 1999).

Examples of clinically related research that have utilised Mahalanobis distances include those that have investigated the morphological similarity of human skulls from different populations (Hubbe et al., 2009) and shape analysis for biometric facial recognition (Hill et al., 2011).

One of the quoted drawbacks of the Mahalanobis distance is the equal adding up of the variance normalized squared distances of the groups of interest. In the case of noise free signals this leads to the best possible performance. However, if a feature between the groups is distorted by noise, this feature can have such a high value (due to the squared distances) that it can hide the information provided by other features. Therefore, improvements to the distance measure which give less weight

to the noisy features and more weight to the clean features have been developed (Wölfel and Ekenel 2005).

3.5.6. *Discriminant analysis*

Discriminant analysis (DA) can be used to determine which continuous variables discriminate between two or more naturally occurring groups. Mathematically, a one-way multivariate analysis of variance (MANOVA) and DA are equivalent (Kinnear and Gray, 2010). However, in MANOVA the focus is on the making of comparisons; whereas in DA there is more interest in the prediction of group membership.

DA is usually carried out as part of a two-step process. Firstly, the significance testing of a set of discriminant functions is conducted following which a classification procedure is facilitated. The matrix of total variances and co-variances, and matrix of pooled within-group variances and co-variances are compared via multivariate F tests in order to determine whether or not there are any significant differences (with regard to all variables) between groups. If statistically significant, one proceeds to see which of the variables have significantly different means across the groups. Once group means are found to be statistically significant, classification of variables is undertaken. DA automatically determines an optimal combination of variables so that the first function provides the most overall discrimination between groups; the second provides second most, and so on. Moreover, the functions will be independent or orthogonal (i.e. their contributions to the discrimination between groups will not overlap).

Classification is then possible through calculation of Mahalanobis distances. Each case will have one Mahalanobis distance for each group, and it can be classified as belonging to the group for which its Mahalanobis distance is smallest. Thus, the smaller the Mahalanobis distance, the closer the case is to the group centroid and the more likely it is to be classed as belonging to that group. As previously mentioned, the

Mahalanobis distance is measured in terms of standard deviations from the centroid. Therefore a case which is more than 1.96 Mahalanobis distance units from the centroid has less than 5% chance of belonging to the group represented by the centroid; 3 units would correspond to less than 1% chance.

3.6. Matching time signals

In static facial analysis, faces can be superimposed on stable structures to allow changes over time or after a clinical intervention to be quantified. For example, serial lateral cephalometric radiographs of the same subject taken in Natural Head Position (NHP) can be superimposed on Sella-Nasion to assess growth and treatment changes over time in the vertical and horizontal dimensions. In dynamic facial analysis, a problem arises which is how to measure the differences between sequences that are multidimensional which vary in time, speed and magnitude.

3.6.1. Fréchet distance

Mathematic models that utilise a variety of different algorithms have been implemented to measure similarities between linear time series. Among the simplest is the Fréchet distance which has been used in online signature and hand-writing recognitions to protein structure alignment (Sriraghavendra et al., 2007). The Fréchet distance of two curves can be imagined as a dog walking along one curve C_1 of length p and the dog's owner walking along the other curve C_2 of length q connected by a leash. Both walk independently along their respective curve and can be expressed at any point in time on their curves by $\alpha(t)$ and $\beta(t)$. The Fréchet distance between the two curves is the length of the shortest leash that is sufficient for traversing both curves in this manner (Eiter and Mannila, 1995). As the Fréchet distance is defined over a maximum, small variations in the input can distort the distance function by a large amount (Chouakria-Douzal and Nagabhushan, 2006). As such, the Fréchet

distance has been used as a baseline to assess more sophisticated algorithms (Efrat et al., 2007).

3.6.2. *Dynamic time warping*

Dynamic time warping (DTW) was first introduced in the seventies (Sakoe and Chiba, 1978) for voice recognition. It is a very robust method for measuring the similarity between sequences that vary in time or speed. DTW computes the optimal alignment between two time series by non-linearly warping one signal to the other by stretching or shrinking the signal along its time axis (Figure 3-5).
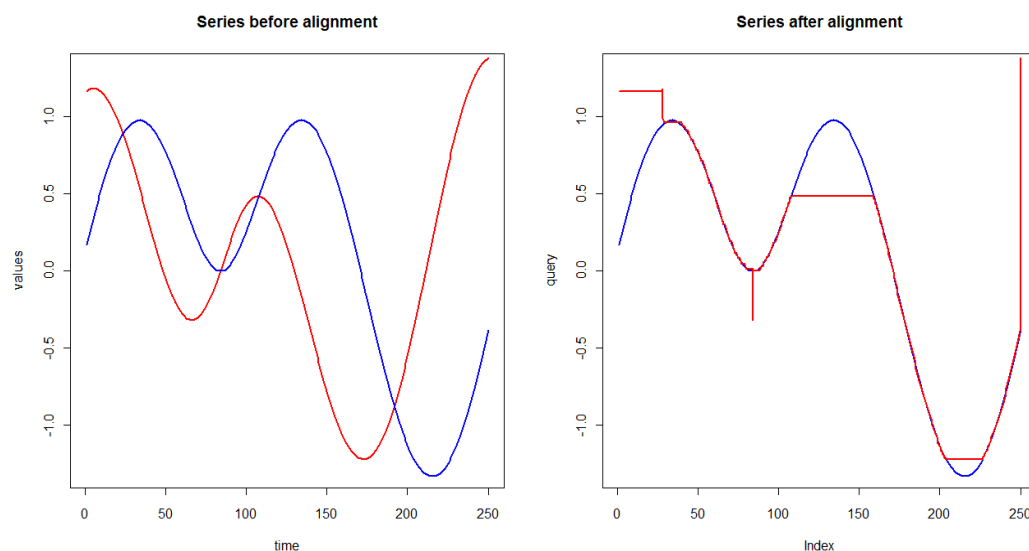
Figure 3-5   *Dynamic Time Warping matching of 2 curves*

From: P. Senin. *Dynamic Time Warping Algorithm Review*. Information and Computer Science Department, University of Hawaii, Honolulu 2008.

The standard DTW approach builds a distance matrix of size m × n. Then an optimal path through the distance matrix (i.e. from (1, 1) to (m, n) is found by moving either one step in horizontal, vertical or diagonal direction (Bailer, 2008). The performance criteria for various DTW algorithms rely on three components (Myers et al., 1980). The first requirement is memory, i.e. the number and size of vectors that need to be stored in

order to compute the accumulated path through the distance matrix. Secondly there should be an emphasis on efficiency or computational speed of the time warping algorithm in computing the optimal path. Lastly and most importantly, the recognition accuracy of the algorithm should be high. For example in speech recognition this relates to the probability (measured as the percentage of occurrences) that the reference word with the smallest distance matches the spoken word in a series of isolated word recognition tests.

Classic DTW suffers from the limitation that it is defined on a sequence of discrete points and therefore is sensitive to the sampling rate. Where the sampling is sparse, there is a lack of resolution in the matching process due to the fact that the algorithm matches only discrete samples rather than continuous curves (Efrat et al., 2007).

Continuous DTW (CDTW) is an extension of classic DTW that maps a sample point in one of the curves to a virtual point between two samples in the other curve (Munich and Perona, 1999). The complexity of CDTW grows exponentially with the sampling rate and therefore computational efficiency can be compromised. In addition, both DTW and CDTW determine their warping paths on two sequences that are similar except for local accelerations and decelerations in the time axis. However the two sequences may also have local differences in the *y*-axis. Consider two data points q_1 and c_1 which have identical values but q_1 is part of a rising trend and c_1 part of a falling trend - DTW will map these points as ideal. Derivative DTW (DDTW) was developed to prevent this problem as a modification of DTW and CDTW. It does not consider the *y*-values of the data points, but rather the higher-level feature of *shape*. Information about shape is obtained by considering the first derivative of the sequences hence DDTW (Keogh and Pazzini, 2001).

A further extension to DDTW called *weighted* DDTW (WDDTW) was proposed recently to include more than the first derivative to the algorithm

(Benedikt et al., 2008). For example, the first derivative may give information on speed, the second on accelerations and decelerations – including further derivatives theoretically provides a more accurate match. However as derivatives are noise sensitive, weighting factors also need to be incorporated so as their inclusion does not detrimentally affect the performance of the algorithm. Indeed, comparisons of curve matching algorithms for automatic face recognition shows that WDDTW outperforms stand alone DDTW, CDTW, DTW and the Fréchet distance (Figure 3-6)
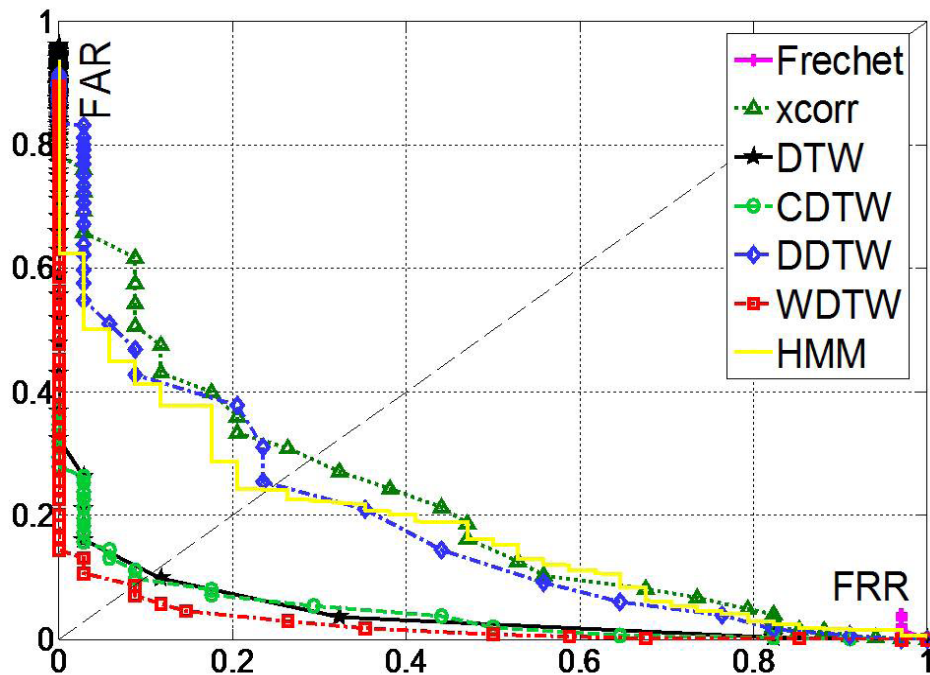
Figure 3-6   *Performance of curve matching algorithms*

false acceptance rate (FAR), false rejection rate (FRR), xcorr = Pearson
correlation coefficient, HMM = Hidden markov model
From: Figure 5 (Benedikt et al., 2008)

3.7.   Summary

The nature of 3D image data necessitates various processes for its analysis and interpretation. The first of which is image registration. Considerations at this stage include whether to use pairwise/groupwise registration or rigid/non-rigid transformation of the data. Non-rigid, groupwise registration would appear to be most appropriate for the analysis of a large inter-subject sample, but this comes at a computational

cost. One should also be mindful that the transformation model does not significantly reduce the shape information available for analysis.

Motion data differs from static image data in that multiple images per subject may be required to be analysed. Here, images can be represented as curves over time and pattern matching techniques used to quantify changes over time between or within subjects.

Chapter 4

Orthognathic surgery

4.1. Introduction

Orthognathic surgery is the art and science of combining orthodontics and maxillofacial surgery to correct dento-facial abnormalities. The incidence of individuals who may benefit from orthognathic surgery is difficult to establish exactly but has been estimated to be up to 2.5% of the United States population (Proffit and White Jr, 2002).

Such deformities may be isolated to one jaw (i.e. the maxilla or mandible) or both. Orthognathic surgery is primarily carried out in adults once growth has ceased and requires a combination of orthodontics and surgical management as skeletal imbalances in the jaws prevent the teeth being moved into an acceptable relationship by orthodontics alone.

The therapeutic goals for orthognathic surgery are summarised in Table 4-1.

Table 4-1   *Therapeutic goals for orthognathic surgery (Wolford, 2007)*

Therapeutic goal	Example
Function	Obtain normal mastication, speech ocular and respiratory function
Aesthetics	Establish facial harmony and balance
Stability	Avoid short- and long-term relapse
Treatment time	Provide efficient and effective treatment time

4.2. Basic concepts

A thorough diagnosis and evaluation is one of the most important aspects of the overall patient management. Evaluation can be divided into patient concerns, clinical and radiographic examination. The establishment of a problem list from the evaluation can help formulate the treatment plan.

Following confirmation of the treatment plan, the orthodontist is initially responsible for moving the teeth into the desired positions within the upper and lower jaws. This will involve elimination of dental compensations, relief of crowding and coordination of upper and lower arch widths which can take up

to 18 months to achieve. Figure 4-1 illustrates a patient with an IOTN of 4m planned for a maxillary advancement and mandibular setback.

Figure 4-1   *Pre-treatment (left) leads to pre-surgical orthodontics (right)*

Rigid skeletal fixation is used to stabilise the jaws in the desired position (Figure 4-2) following which post-surgical orthodontics commences 2 to 4 weeks afterwards.
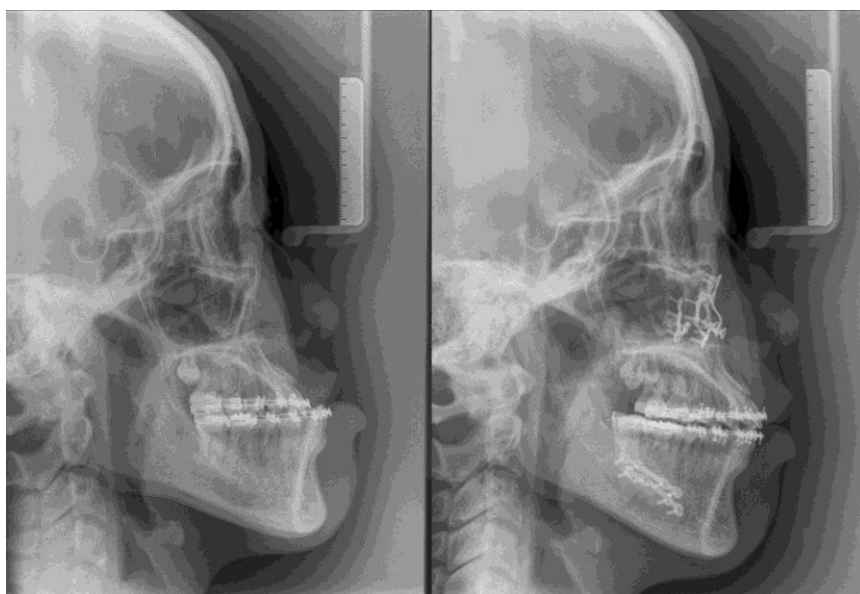
Figure 4-2   *Pre- (left) and post-surgical (right) cephalometric radiograph*

The final positioning of the teeth takes between 3 to 6 months but can occasionally take longer dependent on the orthodontic requirements. The removal of the orthodontic appliances some 6 months after surgery coincides with research which suggests that 90% of the post-surgical swelling has resolved at this stage (Kau et al., 2006a). Figure 4-3 shows the completed case that was introduced in Figure 4-1.

Figure 4-3   *Completed orthognathic case*

4.3.   Soft tissues changes

During the orthodontic planning process, the bony structures (i.e. maxilla and mandible) are used to determine the movements necessary to provide the appropriate soft tissue profile change. The ability to predict the soft tissue response is difficult. Currently, different computer imaging programmes based on mathematical algorithms give the patient and clinician an idea of the expected outcome. Much of this data used in these programmes is derived from information from groups of patients who have previously had surgery (Pektas et al., 2008). Unfortunately much of this data is based on small sample sizes and studies have shown that the predictive capacity of these software programmes to be poor, particularly in the region of the lips (Kaipatur and Flores-Mir, 2009). Figure 4-4 shows the predicted outcome for the patient illustrated in Section 4.2 using Dolphin® 11.0 (Imaging and

Management Solutions, Chatsworth, CA). This is based on upper incisor retraction, lower incisor proclination, maxillary advancement and mandibular setback.
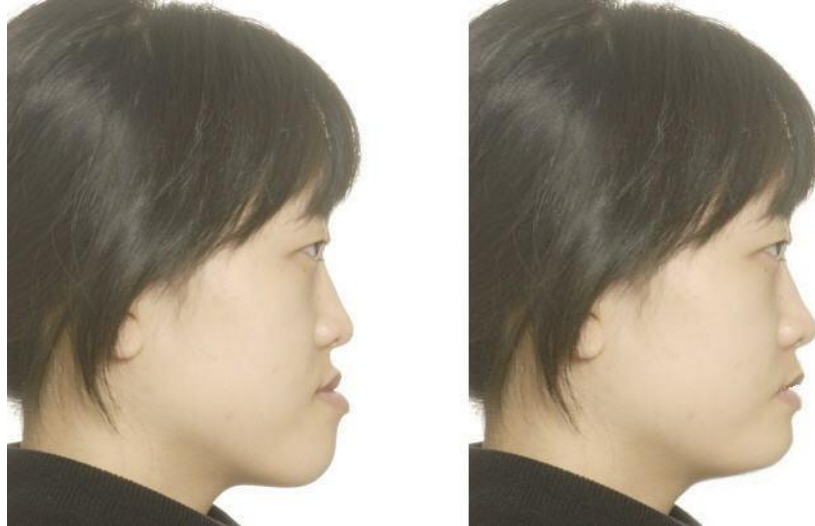
Figure 4-4   *Predicted post-operative software outcome*

(pre-op left, post-op right)

4.3.1.   Incisor retraction

Most orthodontic changes will be reflected in changes in the position and posture of the lips. The literature suggests that with incisor retraction, the upper lip rotates backward around subnasale, reducing the lip prominence (Hershey, 1972). In addition, the thickness of the upper lip increases approximately 1mm for every 3mm of incisor retraction (Rains and Nanda, 1982).

4.3.2.   Maxillary surgery

Different movements of the maxilla have distinct effects on the nasal and labial morphology. Superior repositioning (impaction) of the maxilla causes elevation of the nasal tip, a widening of the alar bases and a decrease in naso-labial angle (Guymon et al., 1988).

Anterior repositioning (advancement) of the maxilla has the greatest effect on the nose and lip resulting in advancement of the upper lip, subnasale and

pronasale, thinning of the upper lip, widening of the alar bases and an increase in the elevation of the nasal tip (O'Ryan and Schendel, 1989).

4.3.3.   Mandibular surgery

The soft tissue changes associated with mandibular advancement include a consequential advancement and lengthening of the lower lip. As the lower labial sulcus and chin adhere to the bony structure of the mandible they following the underlying osseous structures very closely (Dermaut and De Smit, 1989).

Mandibular setback surgery has no effects on subnasale or the tissues higher than this point. Instead, the soft tissues below this point tend to follow the mandible posteriorly, with the chin closest followed by the lower lip (Weinstein et al., 1982). Table 4-2 shows the soft to hard tissue ratios for different types of surgical procedure. Landmark definitions have been outlined in section 2.3.1.

Table 4-2   *Soft to hard tissues ratios for orthognathic procedures*

(H) = horizontal   (V) = vertical

Procedure	Landmark	Ratio	Reference
Upper incisor retraction	ls (H)	-1:1	(Rudee, 1964)
Maxillary impaction	ls (V)	-0.42:1	(Mansour et al., 1983)
	sn (V)	0.29:1	(Hack et al., 1993)
Maxillary advancement	ls (H)	0.9:1	(Freihofer, 1977)
	prn (H)	0.28:1	(Carlotti et al., 1986)
Mandibular advancement	pg (H)	1:1	(Hernandez-Orsini et al., 1989)
	li (H)	85%	as above
Mandibular setback	pg (H)	-1:1	(Robinson et al., 1972)
	li (H)	-0.93:1	(Aaronson, 1967)
	ls (H)	-0.2:1	as above

4.4.   Stability of orthognathic surgery

Stability of the surgical outcome is the ability to maintain skeletal structures in a given position over time. Much of the work related to the stability of

orthognathic surgery can be attributed to the Dentofacial Program at the University of North Carolina, which originally began in 1975. Their *Hierarchy of Stability* which was originally published in 1996 states that the direction of surgical movement followed by the type of fixation, and the surgical technique employed largely influence stability of the outcome. The most stable orthognathic procedure is quoted as maxillary impaction, closely followed by mandibular advancement. Figure 4-5 shows the complete hierarchy and corresponds to changes in the first post-surgical year, which relate directly to the surgical healing, post-treatment orthodontics and short term physiologic adaptation (Proffit et al., 2007).
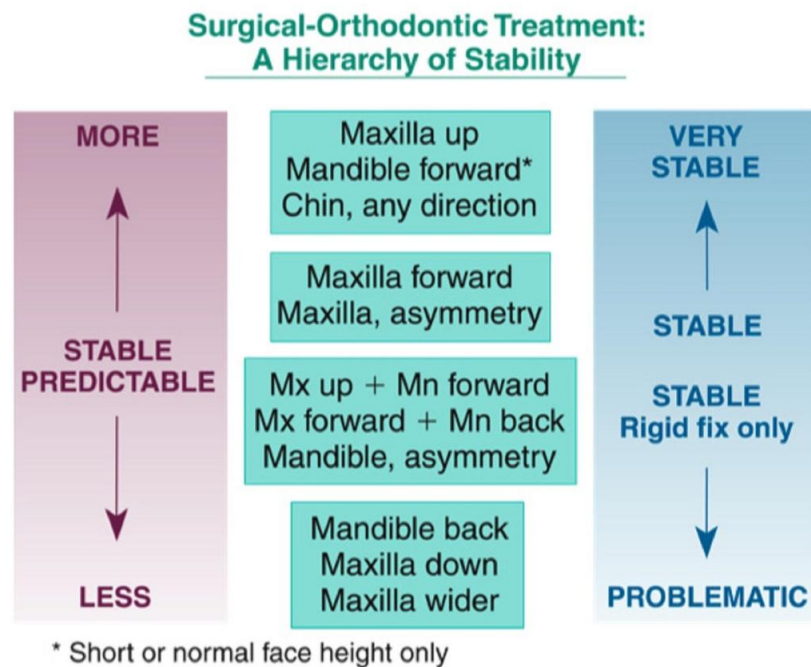
Figure 4-5 *Hierarchy of stability*

From: Proffit, Turvey and Phillips, The hierarchy of stability and predictability in orthognathic surgery with rigid fixation: an update and extension. Head & Face Medicine 2007, 3:21, 1-11

Stability is primarily evaluated from 2D lateral cephalometric radiographs, which are orientated in a standardised manner. A horizontal line along *Sella-Nasion* (and rotated downwards 6º) approximates to NHP and is used as the *x*-axis. A vertical plane perpendicular to this through *sella* is used as the *y*-axis, so that changes in landmark locations can be registered as *x*, *y*

coordinate changes. Due to errors in landmark placement, changes of < 2mm are within the range of error and therefore not considered to be clinically significant (Proffit et al., 2007).

Choice of landmark is dependent on which jaw the surgery was performed on, and in what direction stability is being assessed. For example, superimposing serial lateral cephalometric radiographs can assess stability of mandibular advancement surgery by quantifying the difference in *Pogonion*, *Gonion*, *lower incisor tip* and *B point* in the horizontal plane.

Other studies that have investigated the adaptation of the soft tissue envelope to the new equilibrium have also looked at activities of the masticatory muscles (usually masseter and temporalis) before and after orthognathic surgery using electromyography (EMG). These studies have found that control subjects have better neuromuscular stability (i.e. symmetrical distribution of muscle activity) than patients who are candidates for orthognathic surgery. These were a combination of mandibular prognathic and retrognathic patients who, 6-8 months after surgery exhibited neuromuscular activity closer to the control group (Di Palma et al., 2009, Di Palma et al., 2010). It has therefore been suggested that masticatory muscles in post-surgical orthognathic patients adapt to the new environment achieved (Nakata et al., 2007).

4.5. Facial movement and orthognathic surgery

As outlined in the preceding sections, substantial research has been conducted in studying static facial soft tissue changes following orthognathic surgery. Predictable soft tissue changes can be invaluable for the clinician in formulating a treatment plan. Conversely, very little information is known with regard to how facial movement changes after orthognathic surgery. This is a vital consideration as facial movement and animation has been shown to be an important factor in facial aesthetics. We have seen that the facial muscles originate from the underlying facial bones that are repositioned during orthognathic surgery. Similarly, the muscles are incised and elevated during the surgery and disruption of the muscles may have an effect on the soft

tissue movement of the face. In addition, a further hypothesis is whether impairment in soft tissue movement due to the position of the jaws exists and if so, whether this is normalised post-surgery. This could have an important relationship on the stability of the result. For example, if it was found that facial movement in maxillary advancement surgery did not normalise post-surgery and was associated with relapse, but bi-maxillary surgery did normalise and showed less relapse, then future surgery may indicate a two jaw procedure over a single jaw.

Very few studies have described differences in facial movements before and after orthognathic surgery. One such study was based on 2D data and investigated the changes in an instructed smile before and after orthognathic surgery (Johns et al., 1997). Videotapes were made of 20 subjects pre-surgery and between 3 to 8 months after surgery using 5 landmarks around the upper lip. All patients underwent maxillary surgery only with rigid skeletal fixation. The group was split into those receiving a maxillary advancement or maxillary inferior repositioning. There was no control group. The post-operative facial movement of the group that underwent maxillary inferior repositioning was decreased whereas the maxillary advancement group showed an increase in movement post-operatively. The conclusion drawn was that repositioning the maxilla in an anterior direction would lengthen the underlying facial muscles and therefore result in an increase in the ability of the muscle to elevate and move the skin into which it is inserted. The converse is true for inferior positioning, which will decrease muscle length and therefore cause less movement. Limitations of this study include the 2D data set and the lack of a control or reference group. There was also no mention of the reproducibility of the instructed smile. Given that an emotive non-verbal facial gesture was used, the outcome of the surgery could influence the performance of the gesture.

A more recent study that has published work on facial movement in relation to orthognathic surgery used a passive video-based tracking system and retro-reflective markers (Nooreyazdan et al., 2004). There were 19 subjects included in the study who presented with a range of different jaw

abnormalities. Subjects were asked to perform 7 maximal facial expressions from rest 3 times after a practising session. The position of 34 facial landmarks between pre-surgery, 6 months and 12 months post-surgery were measured. Differences were found between the pre-surgery and 12 month post-surgery visits for the instructed smile, natural smile lip purse, eye closure, grimace and mouth opening movements. This study did not include a control group to reference the movement of the surgical group. In addition the recording time was documented to be 20 minutes suggesting the possibility of facial fatigue as a limitation of the results. In the clinical context, the application is narrow as 34 markers had to be positioned on the subject's face for each recording. The authors of this study also acknowledged the small sample size and heterogeneity of the sample group.

One of the only other studies to have been published in this field used 3D imaging to analyse lip movement in patients with mandibular prognathism (Okudaira et al., 2008). The patient group included 10 subjects with mandibular prognathism due for orthognathic surgery - this was compared with a control group of 20 *normal* subjects. Lip movement was tested through articulation of the vowels /a e i o u/. Four landmarks on the lips that included bilateral commissural points and the midpoints of the upper and lower lips were used to track lip movement. The average landmark distance of the maximum lip displacement from rest was compared between groups. The results showed downward movement of the bilateral commissural points in the prognathic group in the articulation of all vowels. The reason for this, the authors suggest, is that the muscle that depresses the angle of the mouth (i.e. depressor angularis oris) is hyperactive in patients with mandibular prognathism. In addition, the prognathic group showed significant backward movement of the commissural points during articulation of the vowels /u/ and /o/ in which the lips should actually be protrusive. The authors hypothesise that the commissural points are compensating for the anterior-posterior discrepancy between the upper and lower lips.

4.6.   Summary

Orthognathic surgery provides a predictable means of correcting a skeletal deformity. Surgical treatment incises soft tissues and repositions basal bones to achieve this. In the conventional treatment planning process there is little consideration to the soft tissue functional effects of the surgery. Literature suggests that this is an important consideration as adaptation of the soft tissues to the new skeletal position is one the factors related to stability of the outcome. In addition, there is a suggestion that lip movement compensates for the pre-surgical skeletal position of the maxilla and mandible. There is currently no measure to quantify whether this is corrected as a result of the surgery.

The following chapter extends the findings from the current literature and outlines the aims and objectives of the study, in addition to the basic research concepts employed.

Chapter 5

Basic research design

5.1. Research aim

The overall aim of the study is to:

- *determine normal and abnormal lip shapes during movement for use as a clinical outcome measure.*

5.2. Research challenges/problems

The basic considerations in achieving this aim are that the technique should minimise any burden to the patient, the analysis should be quick to apply, simple to interpret and be transferrable in method. Taking these into account, the literature review also highlighted some key issues/challenges that require exploring in detail. These are discussed in the following sections.

5.2.1. Capturing lip movement

Chapter 2 gave a comprehensive overview of the different modalities available to capture not only lip movement but facial movement as a whole. The limitations of the varied approaches suggest that an objective 3D imaging system would represent the gold standard in image capture. As such all subjects in the study are to be scanned using the 3dMDFace™ Dynamic System (3Q Technologies, Atlanta, GA, USA) at 48 fps under standardised conditions. The system is a commercially available ultra-fast 4D surface scanner which was originally commissioned to Cardiff University in 2005 (3dMD, 2005). It captures based on active stereophotogrammetry and uses a random infrared speckle projection to capture both pattern-projected and non-pattern projected white-light images simultaneously. The specifications of the system are detailed under the following headings:

- **Resolution**: six A501kc, 1.3 mega-pixel (MP) area scan cameras with Kodak KAI-1020 Interline Transfer Progressive Scan CCD Sensors (Basler Electric, Highland, IL, USA). Approximately 60 - 100 pixels are required to produce a vertex, and therefore with a 1.3 MP sensor, the output is a facial shell of 15,000 to 20,000 vertices if each vertex is matched in at least 2 images. In instances where the same surface is captured in more than 2 images, the points are measured multiple

times, and a weighted average for that vertex is provided rather than a cumulative score of total vertices. The point cloud is triangulated using an in-house algorithm that leverages bi-product data from the image matching process (see below). The manufacturer was unable to divulge any further specific information regarding this process although the size of the triangles from facial shells ranges from 0.8 to 2mm (Figure 5-1). This is largely dependent on which area of the face the triangles are sampled from.
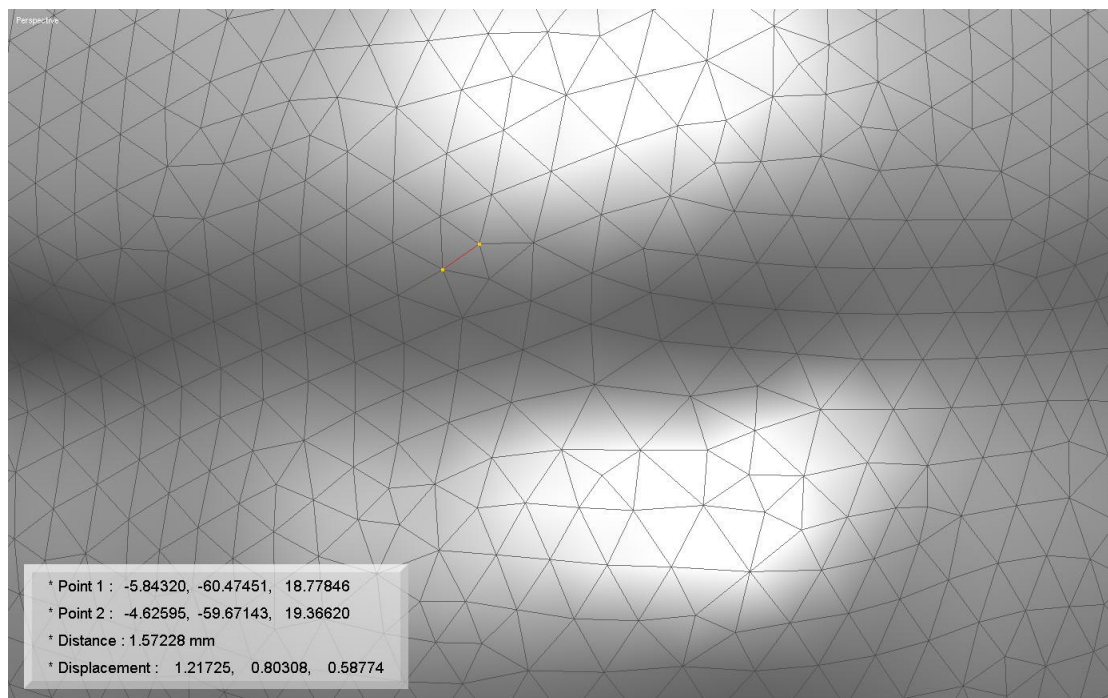
Figure 5-1   *Approximate triangle size taken around the lips*

- **Accuracy**: The point clouds produced from the stereo camera viewpoints are matched via a non-linear least squares algorithm and achieves sub-pixel accuracy (Gruen and Akca, 2005). However, in real-world space the accuracy also depends upon the accuracy of focus for the lenses and the translucency of the skin for the projected wavelength (i.e. skin textures may not all photograph in a universal manner). In this respect, prior to image capture, the camera lenses were checked for the correct focus and subjects were asked to remove make-up, particularly if present around the lips.

- ▪ **Range**: Six Pentax (© Pentax U.K. Limited, Slough, UK) C1641-M
  16mm 1.1:4 TV lenses with a camera-subject distance of
  approximately one metre. In addition, two custom made, infra-red
  projectors with 24V halogen lamps, 50mm lens and fitted filters which
  enable projection of the random infra-red speckle pattern. The
  effective range of image capture is the intersection of the in-focus
  regions of the view frustums between at least one of the projectors and
  two cameras.

Previous work as part of a separate thesis investigated the reliability and
validity of the 3dMD system. Surface registration for a range of facial images
taken consecutively over a one week period showed that 90% of surface
topology was within ±0.5mm between images. Validity was tested by
comparing the same range of facial images captured on a Minolta® Vivid™ VI
900 (Osaka, Japan) 3D optical laser scanning device. Paired surface
registration showed that 76% of the surface topology was within ±0.5mm and
90% to within ±1.0mm using the laser scanner as the gold standard (Popat
2008). These figures were deemed to be within clinically acceptable limits.

5.2.2.   *Feature detection/identification*

Feature detection forms the fundamental basis of the geometric morphometric
approach. As the input data for analysis are Cartesian landmark coordinates
it is crucial for the points used to be valid and reproducible. The conventional
approach for feature detection is via manual placement of landmarks or
markers either directly on the subject's face or following image acquisition.
Automatic landmarking tools could significantly reduce the time required to
process large volumes of data, allowing for larger datasets to be analysed.
However, these tools should be comparable to manual landmarking, which
would be considered as the gold standard. Therefore an automated method
of landmarking 3D facial images is compared to manual examiners (Chapter
6).

5.2.3. *Image registration*

Chapter 3 covered the technical aspects to consider when working with 3D images of which image registration is paramount. Registration is important when comparing structural differences between multiple images over time on the same subject (intra-) or between subjects (inter-). As already discussed in Section 3.3, registration of 3D images can take a rigid or non-rigid approach. Although it has been suggested that non-rigid registration is preferred for inter-subject registration to account for the differences in anatomical variation between individuals, the transformation model should ensure that the maximum amount of shape variation is retained in the registration process. The automated landmarking tool introduced in Chapter 6 utilises a non-rigid approach to image registration whereas the manual method uses a rigid-body transformation. The effect of the type of registration on shape information is discussed in Section 6.3.

5.2.4. *Reproducibility of facial gestures*

When monitoring clinical change over time, the measure used to elicit facial movement should be as repeatable as possible. In this manner, the true effect of an intervention or rehabilitation can be assessed. Should the measure used be less repeatable or prone to a wide variation in its delivery, the performance of the measure could skew the outcome. As Sections 1.4, 1.5 and 2.4 have already discussed, measures of facial movement can be classified into verbal or non-verbal gestures. Verbal gestures include general speech, and non-verbal gestures primarily facial expressions such as smiling, lip purse or cheek puff. The commonality between gesture types is the peri-oral nature of their origin. Therefore it is feasible if using geometric morphometric approaches to utilise a minimum of 6 peri-oral landmarks (Section 2.3.1). Chapter 7 investigates the repeatability of facial movements over time in an attempt to build a hierarchy of reproducibility between different facial movements. There is a recommendation that certain lip shapes during movements are more suitable for use in clinical practice than others.

5.2.5. *Average/normal lip shape*

Chapter 8 utilises the most reproducible facial movements from Chapter 7 to develop a statistical model for average facial movement. In this respect, the model can be used as a reference or outcome measure to compare average facial movement with that of subjects that may have facial movement disorders. In addition to creating an average model, lip shape is also analysed for each movement and differences in lip shape during movement between males and females are investigated.

5.2.6. *Determining abnormal movement*

The final experimental chapter of the thesis aims to investigate whether differences in lip shape during movement of patients with potential movement disorders can be detected from the average model. Here, further statistical shape analysis techniques are employed to determine whether the outcome measure is robust enough to maintain its shape when potentially abnormal lip shapes are included into the sample. Furthermore, whether the methods used are sensitive enough to potentially detect differences in lip shape between the groups (Chapter 9). A further consideration to be addressed is that the sample consists of independent and dependent data (see below) and this will impact of the type of statistical analyses that can be undertaken.

5.2.7. *Subjects/sample*

For the statistical analysis of average facial movement a control group is required. It is proposed that this group will comprise of students and staff the School of Dentistry and Cardiff School of Computer Science and Informatics, Cardiff University. The sample size of the control group will need to be sufficiently large enough to be broadly representative of the population. This should be balanced against the time required for computational processing, image analysis and data storage. Inclusion criteria will be stipulated as age range 18-40 years, no relevant medical history, no previous facial surgery or speech impairment, an average maxillo-mandibular skeletal relationship and British English as the first language spoken. Participants from the control

group will be used to identify reproducible measures of facial movement and construct average templates of facial movements.

To utilise methods of investigating facial movement in clinical groups of patients, a ready sample of subjects with a potential movement disorder is required. The clear example that fits the criteria would be patients with cleft lip. However, due to the low numbers of patients with isolated cleft lip (Fitzsimons et al., 2011) and the large numbers of confounding variables in terms of previous timing, number and type of surgical repair carried out, it was decided against using this particular cohort of patient.

Instead, the patient group asked to participate were those planned for orthognathic surgery. Here, a more homogenous group of individuals with a similar skeletal deformity and clinical presentation could be asked to participant. Approximately 25-50 patients undergo this surgical procedure in Cardiff every year.

Research suggests that post-operative swelling in these patients is reduced by 90% at 6 months (Kau et al., 2006a). Therefore this group of patients could be analysed pre- and post-surgery in the time limits of the study. In addition, the age range of patients undergoing this procedure is similar to the control group allowing for an informal matched comparison. The remaining inclusion criteria are the same as for the control group (i.e. no relevant medical history, no previous facial surgery and British English as the first language). Patients with syndromes and clefts who require orthognathic surgery will be excluded from the study.

5.2.8. *Ethical considerations*

As part of the study intends to include NHS patients, ethical approval from the appropriate NHS Research and Ethics Committee and NHS Research & Development was required. Favourable ethical approval (09/WSE04/44) was granted in December 2009 following which data collection commenced.

Approved participant information sheets and consent forms for both the control and experimental group are included in Appendix A.

5.3.   Summary

The challenges/problems highlighted through the review of the literature have been discussed in this chapter to provide context and foundation for the research undertaken in the study. It is hoped that the sequence of how, what and why in relation to the challenges/problems has set a logical structure for the reader to follow in the experimental sections of the thesis.

Chapter 6

Feature detection/landmark reproducibility

6.1. Introduction

As outlined in Section 3.5, geometric morphometric approaches rely on 3D landmark coordinates for the analysis of shape variation. Therefore, accurate placement of landmarks on 3D facial shells is important to ensure accurate analysis of shape variation.

Several papers have investigated the reproducibility of landmark placement on 3D facial images. Some have compared intra-observer agreement only and others have also included inter-observer agreement. There is a wide variation in the number of the images and repetitions utilised within the methods (Table 6-1).

Table 6-1   *Summary of recent 3D soft tissue facial landmark studies*

Agreement	Raters	Landmarks	Repetitions	Image type	Images	Reference
Intra-	1	24	3	Stereo-photo	6	(Hajeer et al., 2002)
Intra- Inter-	31	24	30	Stereo-photo	10	(Gwilliam et al., 2006)
Intra- Inter-	2	21	2	Laser	30	(Toma et al., 2009)

Recently, automatic methods of image registration have been developed which can allow large volumes of data to be landmarked avoiding the need for time-consuming manual processing (Ruiz and Illingworth, 2008). This would particularly benefit a study of this nature due to the high number of output images obtained when capturing facial movement data. Therefore the following chapter aims to compare the agreement of automatic landmarking when compared to manual landmarking using the manual method as the gold standard.

6.2. Method

3D video data of 30 average subjects (aged 18-40) was captured using the 3dMDFace™ Dynamic System at 48 fps in a standardised manner. Subjects performed a selection of phonetically rich verbal facial gestures,

which included the words *puppy*, *rope*, *baby* and *bob*. This data set comprised of almost 6000 frames of 3D face shapes. To ensure that a variety of subjects and lip shapes were included into the landmarking stage, a random number generator (Haahr, 1998) was used to select the identification numbers of 10 subjects from the sample. Using the same random number generator, a selection of 20 frames from the video sequences of these subjects was made. This allowed 200 face shapes to be selected for analysis representing a 5% proportion of this dataset.

6.2.1. Image processing

The facial shells included for analysis were initially aligned to a Standardised Head Position (SHP) to standardise their orientation in space using commercially available reverse-engineering software (RAPIDFORM™, INUS Technology Inc., Seoul, South Korea). Initially, the landmarks left and right exocanthion, left and right endocanthion, pronasale, soft tissue pogonion were manually identified. Based on these landmarks, the origin of all the facial shells was designated as the mid-intercanthal point. The sagittal plane (Y-Z) was referenced to this point running through the midline of the face (connecting pronasale and soft tissue pogonion), the coronal plane (X-Y) established as average SHP, and the transverse plane (X-Z) intersecting the left and right endocanthion points (Figure 6-1). To ensure consistent SHP was adopted for all facial shells, subjects with significant facial asymmetry were excluded from the study.
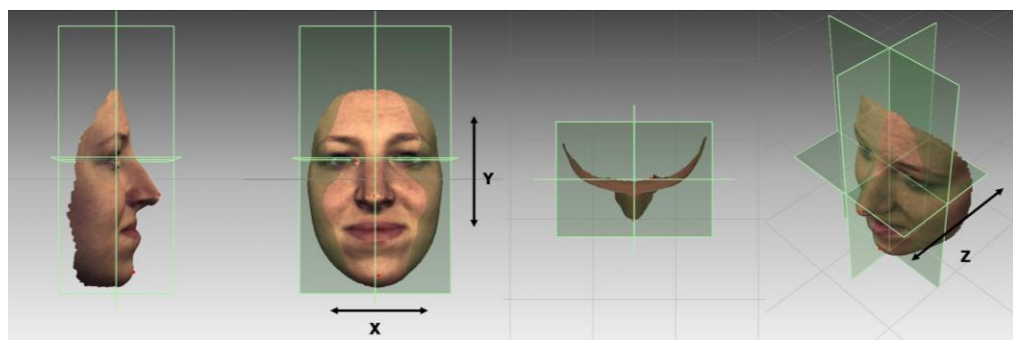
Figure 6-1   *Sample facial shell orientated into Standardised Head Position*

6.2.2. Feature extraction

Following orientation of the facial shells to SHP, two examiners manually placed six anthropometric lip landmarks on each of the normalised facial shells. The landmarks used were labiale superius (*ls*) - the midpoint of the upper vermilion line, labiale inferius (*li*) – the midpoint of the lower vermilion line, crista philtri (*cph* L/R) – the point on the left and right elevated margins of the philtrum above the vermilion line and cheilion (*ch* L/R) – the point located at the left and right labial commissure (Figure 6-2). The colour texture overlying the 3D facial mesh was maintained during landmark identification. In addition, the examiners were able to zoom and rotate the facial shells about the mid-intercanthal (origin) so as to locate the landmarks in the most anatomical correct position (Figure 6-2).
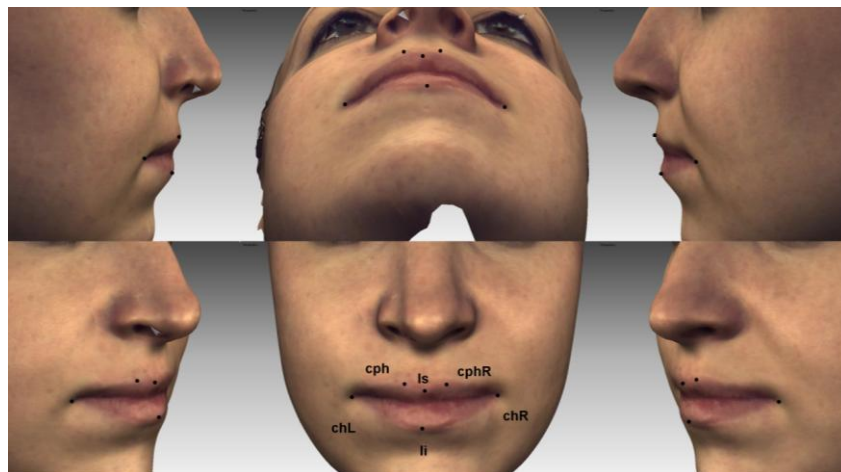
Figure 6-2   *Example shell orientations to aid landmark identification*

The placement of landmarks was repeated by one of the examiners after a 2-week interval. As the origin of all facial shells had been designated as the mid-intercanthal point (x, y, z coordinate = 0, 0, 0) the lip landmark coordinates (x, y, z) relative to this point were recorded. This process was carried out using RAPIDFORM™ software (INUS Technology Inc., Seoul, South Korea).

A recently developed method for semi-automatic, groupwise, non-rigid registration of 3D facial shells was also used to landmark the 200 facial shells (Sidorov, 2010). This was employed for the 20 frames from each subject separately. The method has been described in detail as part of a separate PhD thesis and is utilised here purely as a validation tool. However in brief, the algorithm uses Multidimensional Scaling (MDS) for initial pairwise registration of images to a target/reference image (Sidorov et al., 2011). This is followed by refined groupwise registration using a stochastic optimiser, namely Simultaneous Perturbation Stochastic Approximation (Spall, 2004). After the registration is complete, correspondences between any point on one mesh and any point on any other mesh are known via the reference frame. One facial shell from the automatic registration process was required to be manually landmarked, following which the correspondences in the remaining frames could be interpolated. For this reason the 10 manually landmarked frames (one for each subject) were removed from the analysis, leaving 10 sets of 19 automatically landmarked facial shells. Due to the different processing method and to allow for comparison, one of the examiners also manually landmarked the set of 190 automatically registered facial shells. The x, y, z coordinates for each landmark (18 coordinates in total) were then recorded for both examiners and the automatic registration process.

6.2.3. *Statistical analysis*

The reproducibility of landmark identification at the 2-week interval was assessed for one of the examiners (intra-examiner) and also between the two examiners (inter-examiner). In addition, reproducibility was assessed between one of the examiners and the automatic registration algorithm (manual-auto).

Bland-Altman plots and 95% Limits of Agreement (LOA) of the difference between placements were used to evaluate agreement between intra-examiner, inter-examiner and manual-auto assessments for each landmark coordinate. In addition to assessing agreement, Bland-Altman plots also allow assessment of systematic or random error in landmark placement. The direction of landmark error in NHP is shown in Figure 6-3.
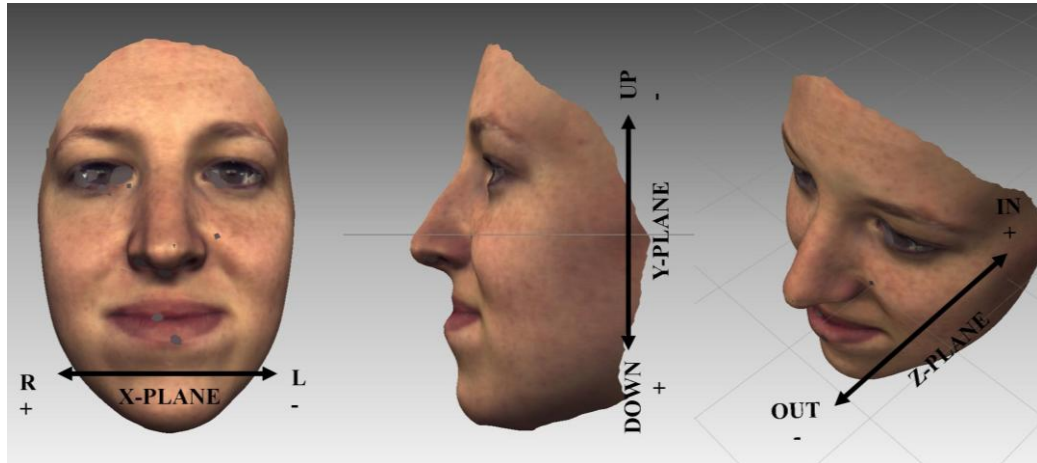
Figure 6-3   *Direction of landmark error*

Landmarks were ranked in order of reproducibility using the mean distance error (Euclidean distance) for between placement times/method using the following equation:

$$D = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$$

Where D is the total distance and Δ is the difference in the respective axes between the two placement times. The reproducibility of each coordinate based on mean distance error was classified according to Table 6-2.

Table 6-2 *Classification of landmark reproducibility*

Mean distance error (mm)	Reproducibility rating
Less than 0.5	Excellent
0.5 to 1.0	Very good
Greater than 1.0 but less than or equal to 1.5	Good
Greater than 1.5 but less than or equal to 2.0	Moderate
Greater than > 2.0	Poor

All statistical data analysis was carried out using SPSS 20.0.0 (SPSS Inc., Chicago, IL).

6.3.   Results

The difference in mean landmark placement between placements for each of the methods is shown in Table 6-3. The data suggests that the highest levels of agreement were seen in the intra-examiner assessment, followed by the inter-examiner. The manual-auto assessment showed the lowest agreement for all landmarks.

Table 6-3   *Individual landmark reproducibility by axis*

Landmark axis	Intra-examiner		Inter-examiner		Manual-auto	
	Mean (mm)	SD	Mean (mm)	SD	Mean (mm)	SD
ls X	0.17	0.22	0.23	1.24	0.58	2.10
ls Y	-0.23	0.24	0.40	2.34	0.81	1.31
ls Z	0.15	0.18	-0.32	1.52	1.02	1.16
li X	-0.01	0.58	0.28	1.08	2.30	1.61
li Y	-0.54	0.69	-0.44	2.58	-2.40	0.58
li Z	0.17	0.11	-1.26	1.55	1.99	1.18
cphL X	0.76	0.78	0.53	0.95	2.26	2.02
cphL Y	0.14	0.44	1.39	2.04	2.03	0.66
cphL Z	-0.05	0.24	0.48	1.41	1.82	1.28
cphR X	0.10	1.17	-0.84	0.96	0.88	1.84
cphR Y	0.14	0.46	1.01	0.29	2.55	1.53
cphR Z	-0.06	0.24	0.59	1.46	0.86	1.16
chL X	0.58	0.97	1.10	1.36	5.70	4.61
chL Y	-0.15	0.67	-0.37	2.12	3.82	2.51
chL Z	-0.10	0.74	0.77	2.77	-2.01	2.95
chR X	-0.87	0.61	0.43	1.76	5.26	2.26
chR Y	0.28	0.60	0.75	2.01	3.60	1.16
chR Z	-0.40	0.42	0.90	1.99	4.70	1.87

Bland-Altman plots were constructed showing the difference in landmark placement (y axis) against the mean landmark placement (x axis) for that particular coordinate. Example plots to illustrate the highest and lowest levels of agreement between placements of landmarks are shown in Figure 6-4, Figure 6-5 and Figure 6-6. A black reference line representing the mean difference in landmark placement is plotted. The upper and lower 95% LOA between placement times/method is shown within an area bounded by two red lines. Random error is observed when readings evenly straddle the y-axis zero. Systematic error results when readings are biased towards landmark placement in a particular direction. Higher levels of agreement should cluster readings close to the zero reference line.

The Bland-Altman plots confirmed that the highest level of agreement was seen in the intra-examiner assessment (Figure 6-4), followed by inter-examiner (Figure 6-5) and manual-auto placement (Figure 6-6). This is supported by the mean distance errors for each landmark (Table 6-4). Here the manual-auto agreement had the largest mean error and standard deviation for each landmark.

For the intra-examiner assessment, landmark reproducibility ranged from *excellent* to *good*. Mean error for *ls* was low implying a high level of agreement (0.47mm, SD = 0.17). The least reproducible landmark was *chL* with a mean distance error between placements of 1.42mm, SD = 0.56. The errors for midline or para-midline landmarks showed a small degree of random error in the x-plane (Table 6-3). In the y-plane, there was a tendency for slight systematic error favoured to higher positioning of these landmarks between placements (Figure 6-4 and Table 6-3). Landmarks further away from the midline showed greater systematic errors between placements in the x-plane albeit very small in magnitude (Table 6-3).

Inter-examiner agreement ranged from *very good* to *good* with the mean distance error between placement for all six lip landmarks being under 1.39mm, SD = 0.57 (Table 6-4). The largest mean distance error was seen in landmark *chR* (Table 6-4). There were higher standard deviations associated with inter-examiner than intra-examiner (Table 6-3). There tended to be bias in the inter-examiner agreement with of the examiners favouring placement of landmarks towards the right-hand side. Here, all landmarks except *cphR* in the x-plane were positive in value between placements implying a consistently right sided difference (Table 6-3 and Figure 6-3). Individually, *cphL* in the y-plane showed the highest error at 1.39mm, SD = 2.04 with contra-lateral landmark *cphR* in the y-plane showing an error of 1.01mm, SD = 0.29 (Table 6-3). These figures are

positive in value which suggests one of the examiners was biased towards slightly inferior (lower) positioning of these landmarks (Figure 6-3).

The mean distance errors for the manual-auto assessment were quite low. All landmarks were considered as poorly reproducible. In particular, the commissures were in high disagreement (Table 6-4). Despite this, mean individual landmark error between placements was low for *ls* in the x- and y-plane and *cphR* in the x-plane and z-plane (Table 6-3).

Overall, landmark *ls* was most reproducible in all three assessments with agreement ranging from *excellent* (intra-examiner and manual-auto) to *very good* (inter-examiner). Landmarks *cph* and *li* were mid-ranked with *ch* most frequently least reproducible. Although this was a generalised similarity across the assessments, *chL* was ranked second behind *ls* in the inter-examiner exercise. The manual-auto mean distance errors were extremely high, and this is discussed in more detail in the following section.

Table 6-4   *Landmark reproducibility by mean distance error (MDE)*

Rank	Intra-examiner			Inter-examiner			Manual-auto		
	Landmark	MDE (mm)	SD	Landmark	MDE (mm)	SD	Landmark	MDE (mm)	SD
1	*ls*	0.47	0.17	*ls*	0.79	0.45	*ls*	2.64	0.92
2	*li*	0.98	0.43	*chL*	1.17	0.47	*li*	2.95	1.45
3	*cphL*	1.09	0.52	*cphR*	1.23	0.55	*cphL*	3.45	1.63
4	*cphR*	1.17	0.54	*cphL*	1.31	0.47	*cphR*	3.64	1.39
5	*chR*	1.29	0.51	*li*	1.32	0.52	*chR*	8.48	3.92
6	*chL*	1.42	0.56	*chR*	1.39	0.57	*chL*	8.30	1.91

Excellent agreement (< 0.5mm)   Good agreement (> 1.0 but ≤ 1.5mm)
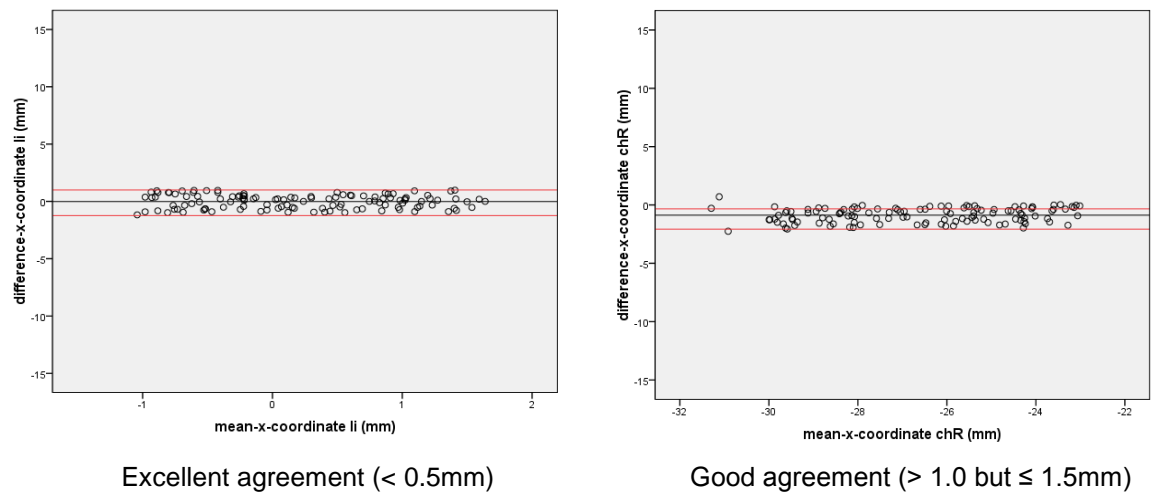
Figure 6-4 *Highest and lowest agreement for intra-observer landmarking*

Very good agreement (< 1mm)   Good agreement (> 1.0 but ≤ 1.5mm)
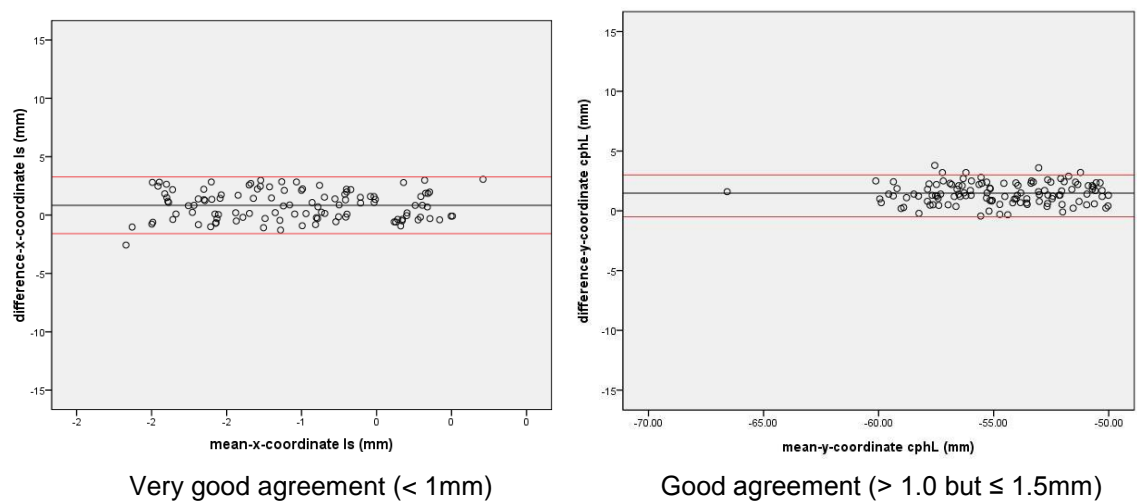
Figure 6-5 *Highest and lowest agreement for inter-observer landmarking*

Very good agreement (< 1mm)   Poor agreement (> 2mm)
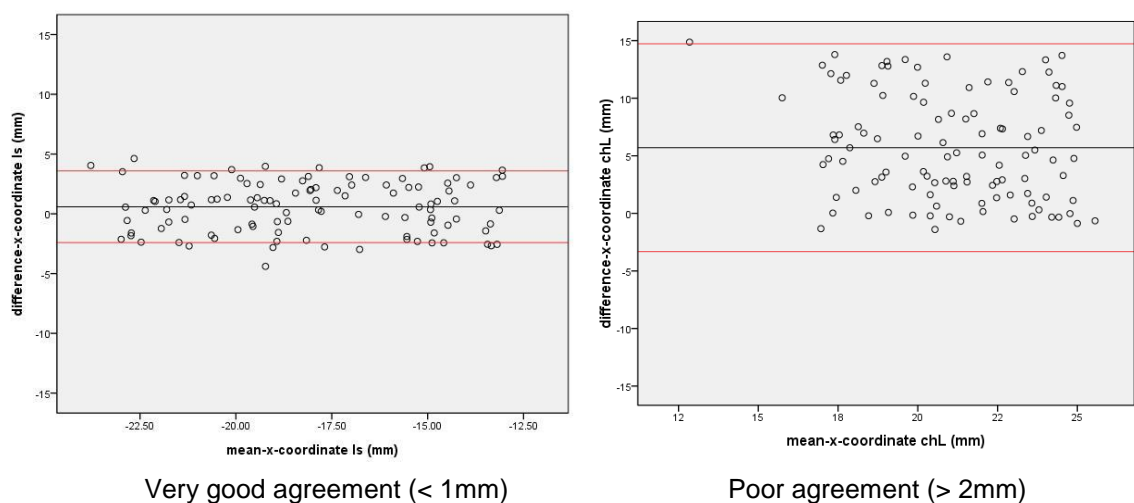
Figure 6-6 *Highest and lowest agreement for manual-auto landmarking*

6.4. Discussion

For the purposes of facial landmark identification, the terms reproducibility and repeatability are used interchangeably throughout the literature. Both refer to the measurement precision or agreement generated on the same specimen using the same test method, under the same conditions on separate occasions. However, repeatability refers to the process being undertaken by the same operator, whereas reproducibility is defined as the closeness of agreement between different operators (Vieira and Corrente, 2011).

The reproducibility of facial soft tissue landmarks is not a novel concept. In orthodontics and cranio-facial surgery, cephalometrics is a well-established technique of using lateral skull x-rays to study hard and soft tissue relationships. Being a 2D analysis, landmarks are first identified following which angles, planes and ratios can be calculated for the vertical and horizontal planes. A meta-analysis of the reproducibility of cephalometric landmarks has suggested that acceptable levels of accuracy between placement times are 0.59mm of total error for the x-plane and 0.56mm for the y-plane (Trpkova et al., 1997). It should be noted that due to several confounding variables, this threshold is not immediately transferrable to 3D soft tissues. Results may vary according to the following:

- image quality: a high quality reproduction of the facial surface in 3D is preferred - this overcomes the limitations of 2D images as discussed in Section 2.3.3 and allows the observer to identify the x, y, and z planes for a particular coordinate with more precision;
- image manipulation: allows the observer to adjust perspective and magnification to allow direct vision during landmark placement;
- facial morphology: landmarks associated with clearly defined facial borders or edges (e.g. *ls* and *cph*) tend to be produce less error in placement when compared to landmarks located on flat or curved surfaces, e.g. glabella (*g*). Gender therefore may also have an influence on landmark error as males are generally considered to have

more defined features than females who tend to have a more gently curving facial anatomy;

* examiner factors: those observers with more experience and knowledge of facial anatomy and the image type used will tend to exhibit less error in landmark identification.

In this section of the thesis agreement in identification of six soft tissue lip landmarks between 2 observers, and 1 observer and an automated method was compared. A high number of images were included to ensure a wide variation in facial morphology. Due to the aim of studying lip movement, the landmarks investigated for reproducibility were limited to six. Their location around the lips along the vermilion border (which is generally well defined) suggests that errors in placement could be low. The landmark with the lowest mean error or most reproducible was *ls*. In addition, a single person's perception of landmark reproducibility (intra-) was better than between different assessors (inter-).

It has been suggested that a landmark can be deemed highly reproducible if the SD from the centroid is 0.5mm or less in all three planes of space (Hajeer et al., 2002). In the original study, 25 out of 30 landmarks were considered to be highly reproducible and these included the landmarks *ls*, *li*, *chL* and *chR* (the landmark *cph* was not used). If these guidelines are followed, then for the intra-examiner data *ls*, *li*, *cphL, cphR* and *chL* can be labelled. However, other authors suggest this is too critical and one should consider the clinical significance as opposed to statistical - in this respect, the SD threshold should be increased to 1.0mm (Gwilliam et al., 2006). All six lip landmarks for both the intra- and inter-examiner placements can now be included if this criterion is adopted.

An implication for the accuracy of manual landmark identification is the original resolution of the image. The images used in this Section of the study were based on triangulated facial shells. As such the triangle size or triangle density for a given area can be regarded as a reflection of the quality of the image/area. Section 5.2.1 outlined that the approximate triangle size for facial

images obtained with the 3dMDFace™ Dynamic System was 0.8 to 2mm. In this respect the manual landmarking errors can be considered as acceptable for the given resolution of the image. Static 3D imaging systems are known to function at significantly higher resolutions than dynamic systems (Sholts et al., 2010) and therefore it is conceivable that landmarking accuracy may well be higher using these systems. However, if dynamic information is paramount, incorporation of method error equal to landmark identification should be employed.

The manual-auto assessment was particularly poor in relation to mean error (excluding *ls*). We attribute this to the image processing associated with the automatic registration algorithm. To allow for computational efficiency during non-rigid registration, the 3D facial surface is decimated to a degree. When investigated further, it was found that the triangle count was being reduced by a count of 7-10 times. Figure 6-7 shows an example of a raw 3D facial shell (left) alongside its automatically registered counterpart (right). The polygon count for the raw shell is 18,505 whereas the decimated shell is 2232. This represents an eight-fold reduction in image resolution. In particular, the most aggressive change can be seen around the lips and mentalis region.
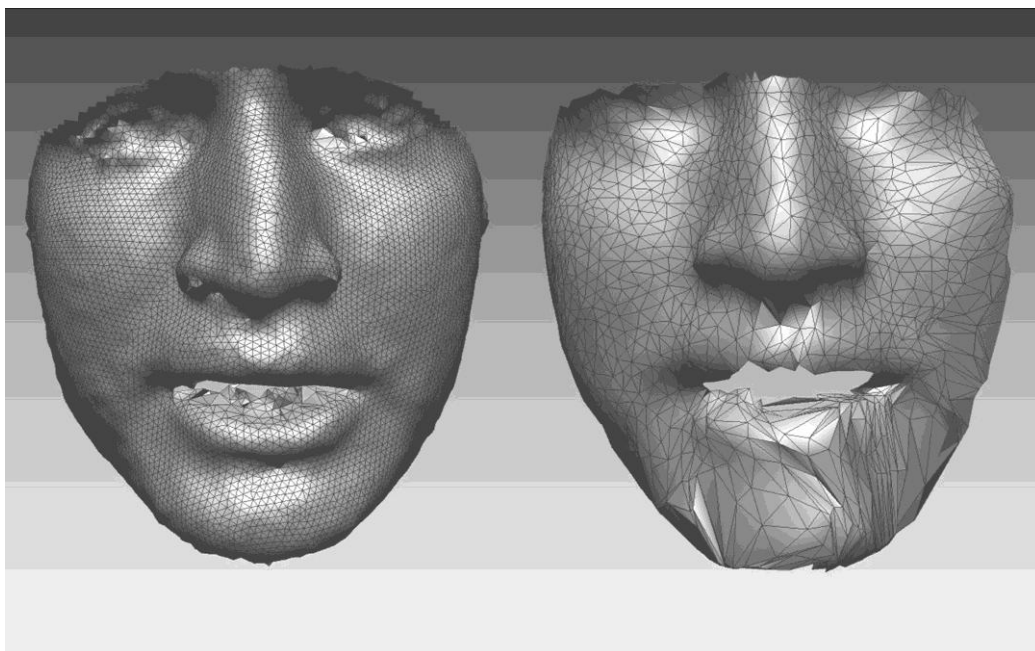
Figure 6-7   *3D facial shell – raw image (left), non-rigid decimation (right)*

The issue of mesh decimation is related to computational time and space. The complexity of a registration problem is proportional to the number of vertices in the facial mesh. To register a set of images, a matrix of (n^3) is formed, where n is the number of vertices in the mesh and therefore also the number of rows in the square distance matrix. Images of high-resolution (high vertex counts) require several hours to be registered making them undesirable to work with, in their raw form. In addition, the computational space requirement for meshes of approximately 20,000 vertices can often exceed the available contiguous memory of desktop computers and therefore efficiency can be achieved if dense meshes are decimated first.

In this example, all of the automatically registered images contain exactly the same number of polygon faces, so that a polygon vertex in one image has a corresponding vertex in the subsequent images. Landmarks placed on areas that show a lower/sparse polygon count will not be as specific as those placed in areas of high-density wire polygons. This will reflect on the degree of reproducibility – in this case, a higher mean error distance. A registration method that facilitates computational efficiency without a significant sacrifice to the final result would be ideal. For example, the decimated mesh could consist of a subset of vertices of the original dense mesh. The difference between the dense and decimated mesh would need to be stored. The necessary registration operations are performed on the decimated mesh with a reduced storage requirement and running time. Finally, using the previously computed difference, the registration results are interpolated to the missing points to give a high resolution final mesh (Sidorov et al., 2007). Therefore, although fully automatic methods of image registration may solve the labour intensive nature of manual image preparation, precautions should be taken in the computerised process to maintain a dense polygon mesh.

6.5.   Conclusion

Both intra-operator and inter-operator methods of landmark identification show clinically acceptable levels of repeatability and reproducibility. The

automated method of image registration for landmark detection utilised in this part of the study cannot currently be recommended over manual identification.

Chapter 7

Reproducibility of facial gestures

7.1. Introduction

We have seen from Sections 1.4 and 1.5 that there are two main measures of facial movement that can be used to assess function. These are either verbal or non-verbal facial gestures. Studies using verbal facial gestures have investigated facial movement during speech (Dawes and Kelly, 2005) and studies involving non-verbal facial gestures have utilised facial expressions (Bartlett et al., 1999).

Deciding upon which facial gesture is most appropriate to employ for a particular application is primarily dependent upon the context of the application. For example in the assessment of facial nerve deficit, the primary objective is to use facial gestures that are reproducible, i.e. repeatable over time. A reproducible measure of facial movement allows valid assessment of changes in facial movement over time and following treatment procedures.

This chapter aims to quantify the repeatability of different facial gestures and establish a hierarchy of similarity between them.

7.2. Methods

Subjects were asked to say four verbal facial gestures; *puppy*, *rope*, *baby* and *bob* in a normal relaxed manner and to perform two non-verbal facial expressions - a *maximal standardised smile* with lips closed based on Action Unit (AU) 12 of the Facial Action Coding System (FACS) and a *maximal normal smile* with lips open based on AU 25. This was carried out at an initial session (T1) and repeated one month later (T2). The sequence was practised prior to image capture to familiarise the subjects with the gestures. To account for variation in the performance of the gestures the subjects were asked to repeat the gestures three times at T1 and T2.

7.2.1.   Image processing

The 3D motion data was processed using an Active Shape Model (ASM) to convert the video data into linear representations. This was carried out for each word/expression for all subjects at T1 and T2. The steps in building the ASM are outlined below:

1. Removal of unnecessary data from the raw image such as poorly defined areas around the ears, hairline and shoulders.

2. Registration of the cropped 3D face shapes in a sequence using Generalised Procrustes Analysis (GPA). Translation, rotation and scaling were incorporated.

3. A point distribution model was created by manually placing twelve landmarks around the lips on all the normalised images. A Thin Plate Spline (TPS) process was then used to *warp* (register) each image to the reference (or resting) frame allowing region-based feature extraction of lip shapes across a sequence.

4. Principal Component Analysis (PCA) was applied to each of the normalised data sets to decompose the vertex displacements into PCs. The Eigenvalues are sorted in descending order so that the first Eigenvector depicts the highest PC. Lower order PCs constitute less of the variation in lip shape and can be considered as noise. The total number of PCs that described 95% of the variation in lip shape for each gesture were retained.

5. A Weighted Derivative Dynamic Time Warping (WDDTW) algorithm (Benedikt, 2009) was applied to the PCA model to match multi-dimensional signals across between T1 and T2 for each subject and each facial gesture. The output is a numeric value representing the similarity in lip movement between two facial gestures across the PCs that were retained (in this case, those PCs

that described 95% of the variation in lip shape). The WDDTW

value ranges from 0 to 1; whereby 0 represents two infinitely

different gestures and a value of 1 represents two identical

gestures.

Step 1 was carried out in RAPIDFORM™ software (INUS Technology Inc.,

Seoul, South Korea). Steps 2-5 were carried out using MATLAB 2009a

(MathWorks, Inc., Natick, MA).

7.2.2.   *Statistical analysis*

In this Section of the thesis, the reproducibility of six facial gestures was

compared. The null hypothesis proposed was therefore as follows; *there*

is no difference in the reproducibility of the six facial gestures tested. A

sample size calculation was carried out using pilot data from 10 subjects

(who were not included in the final sample). To detect a 0.001 (sd 0.002)

difference in similarity between the six gestures using a significance level

of 0.05 and power of 80%, a minimum of 18 subjects was required to

reject the null hypothesis. This was calculated using the software package

G*Power Version 3.1.0 (Appendix B, Table B-2). Twenty-five subjects (11

male and 13 female) with a mean age of 24.3 years were eventually

recruited.

Statistical analysis of the WDDTW data was performed using a (within

subject) repeated measures analysis of variance (ANOVA) in SPSS 16.0

(SPSS Inc., Chicago, IL). Here, the WDDTW scores for the 25 subjects at

T1 for each of the verbal (*puppy*, *rope*, *baby*, *bob*) and non-verbal

gestures *(standardised smile, normal smile)* were statistically compared to

those at T2. The assumption of a repeated measures design, namely that

the same participants participated in all conditions was satisfied.

The other assumption of ANOVA is normal data distribution and therefore

the distribution of the WDDTW scores was visualised using histograms

and quantified with the Shapiro-Wilk test of normality (set at a significance level of 0.05) – there were no violations (Appendix B).

7.3.   Results

All subjects completed the 6 facial gestures at T1 and T2. Figure 7-1 shows the first three PCs extracted for the standardised smile with the corresponding ASM shown adjacent. Here, the *x*-axis represents time and the *y*-axis shows the amplitude or standard deviation of lip movement from the mean lip shape. The first three PCs of this non-verbal facial gesture all explain lip raising. PC1 constitutes the majority of variation in lip movement, and lower order PCs describe less movement and more noise. This is evident as the increasingly irregular lip shape amplitudes from PC1 to the PC3. The gesture lasts 60 frames equivalent to 1.2 seconds.

Figure 7-2 shows PC1 for the four verbal gestures in a similar manner with the corresponding ASMs adjacent. It can be seen here that the words *puppy* and baby are composed of 2 distinct visemes with PC1 explaining lip opening. There is more lip opening associated with the word *baby* than *puppy*. The words *rope* and *bob* are essentially only comprised of one viseme which is lip protrusion. The duration of all four verbal facial gestures is significantly less than the non-verbal smile expressions with *puppy* and *baby* lasting approximately 30 frames (0.6 seconds) and *rope* and *bob* lasting 18 frames (0.36 seconds).
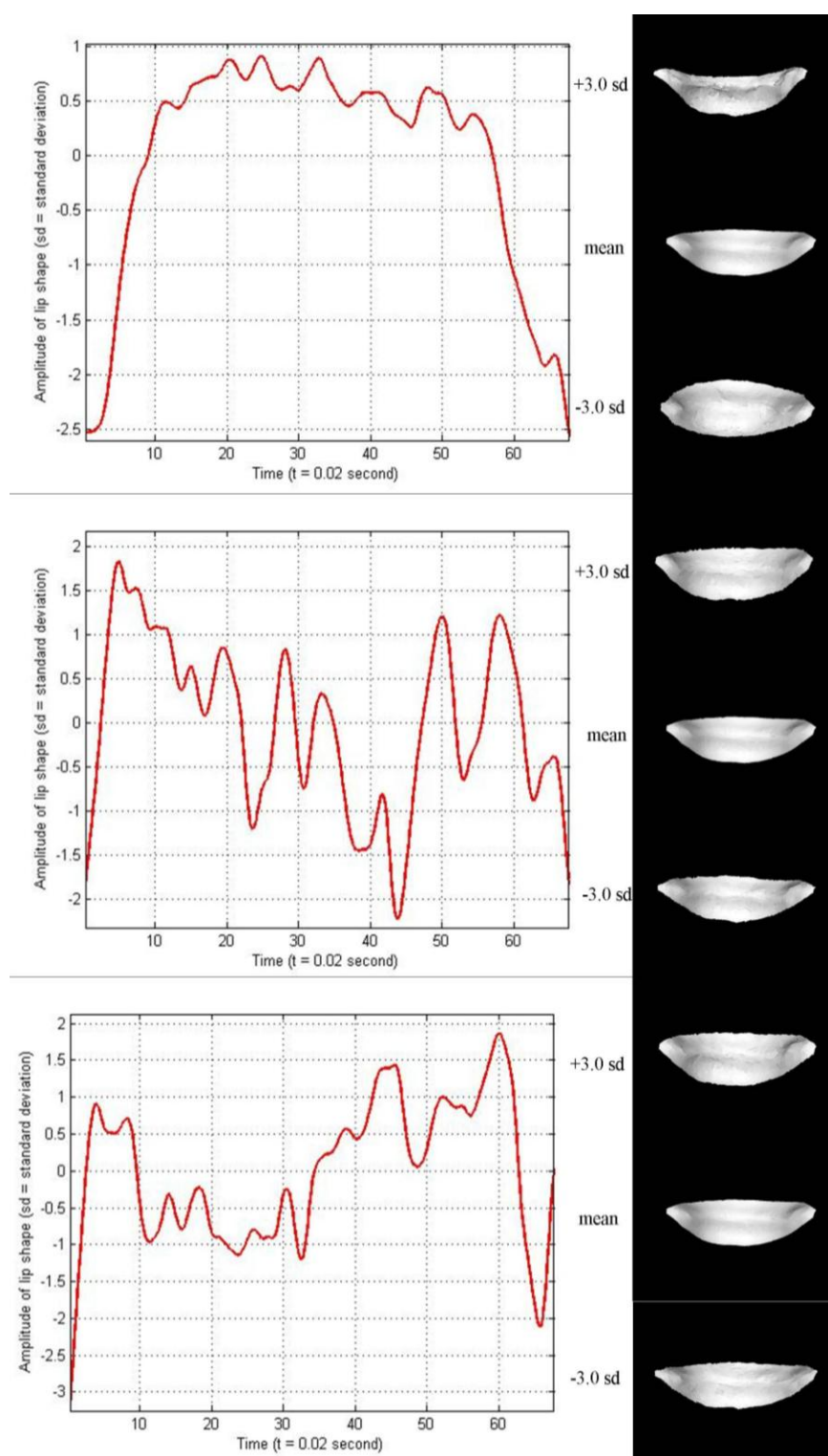
Figure 7-1 *PCA of the standardised smile with associated lip shapes*

PC1 to PC3 (top to bottom)
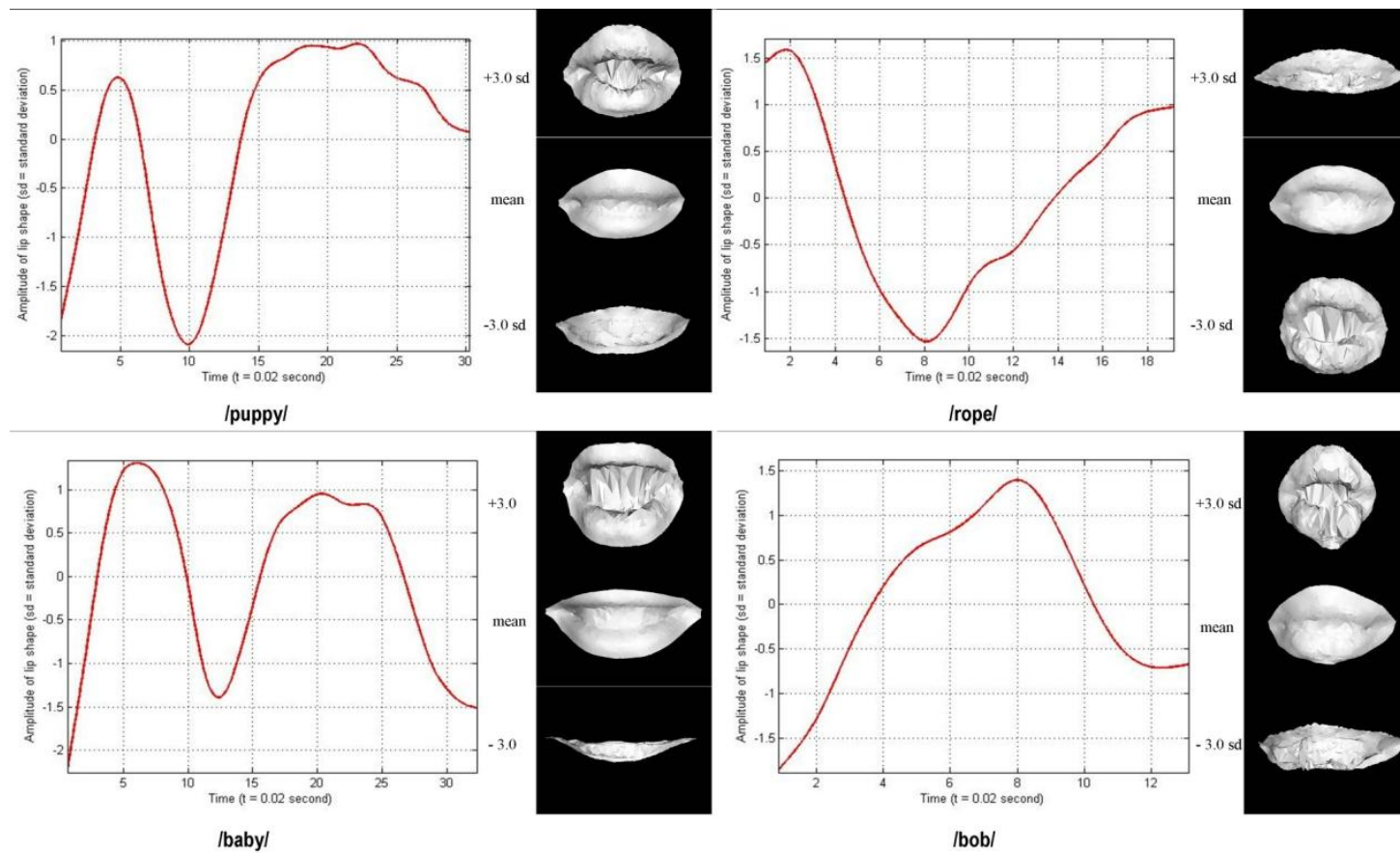
Figure 7-2   *PC1 of the four verbal facial gestures*

The WDDTW algorithm was applied to the PCA signals for each facial gesture between T1 and T2.   Figure 7-3 shows PC1 to PC4 of an individual from the sample saying the word *puppy* at T1 and T2 (WDDTW = 0.9978).
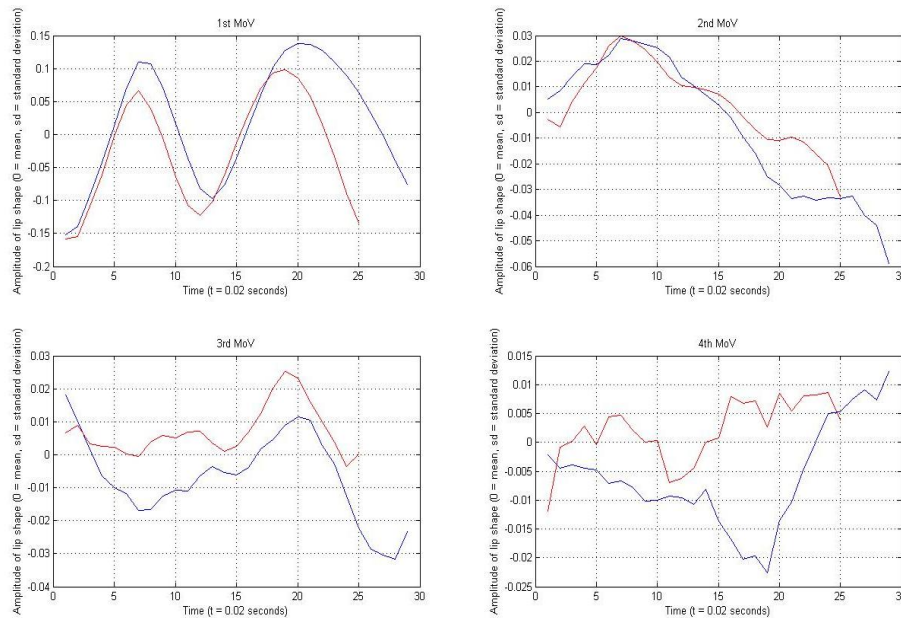
Figure 7-3   *PC1-4 of the word /puppy/ spoken at T1 (blue) and T2 (red)*
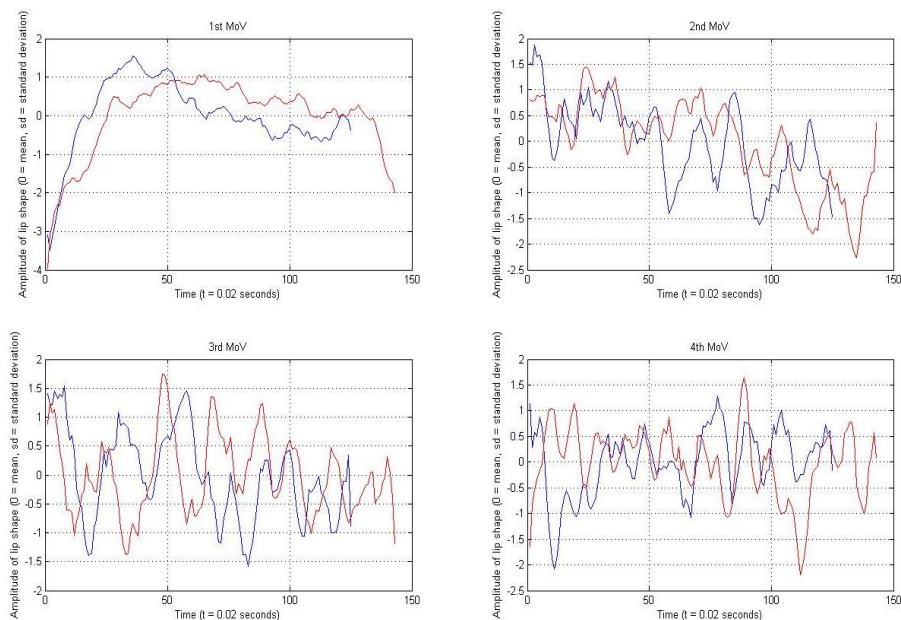
Figure 7-4   *PC1-4 of maximum normal smile expression*
T1 (blue) and T2 (red)

Figure 7-4 shows PC1 to PC4 of a subject performing a maximum normal smile expression (WDDTW = 0.9865). The linear representation of the expression appears noisy and congested for both time points, particularly after PC1. In contrast, PCA for the verbal gesture *puppy* is relatively well defined for T1 and T2 across the first three components. The lower WDDTW value for the normal smile expression when compared to the word *puppy* suggests that the non-verbal gesture is less reproducible.

The variability in reproducibility between T1 and T2 based on the WDDTW values for each facial gesture is shown in Figure 7-5. This implies that all four of the verbal facial gestures were more reproducible than the two non-verbal gestures. The word *baby* was shown to have the highest median reproducibility and smallest range whereas the normal smile was the least reproducible and also the most variable.
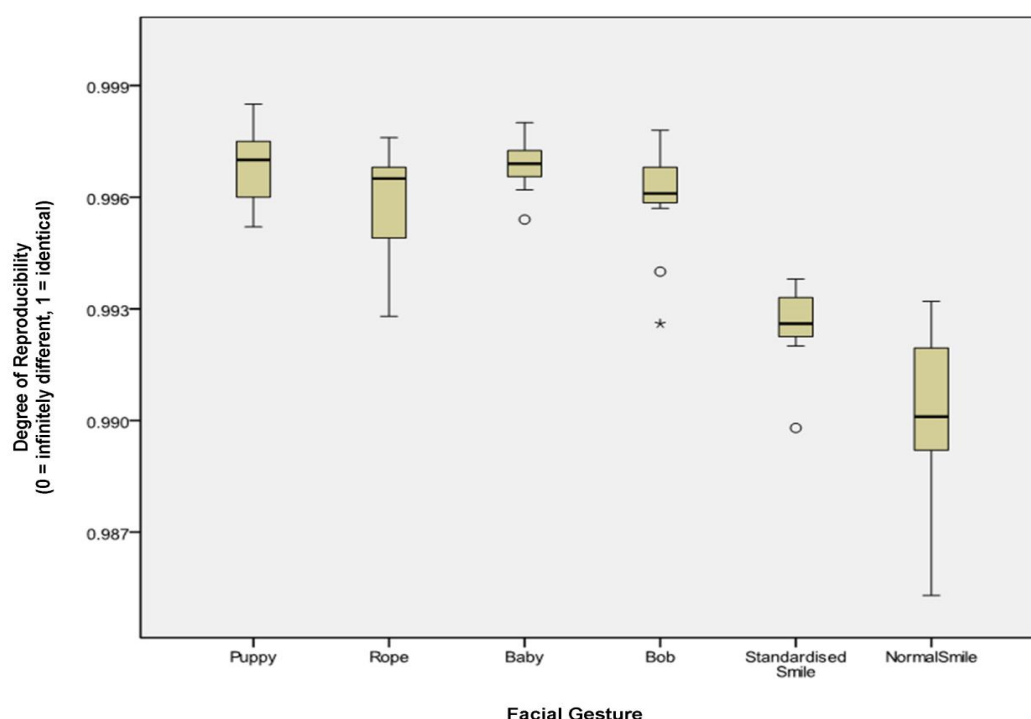
Figure 7-5 *Box plots showing reproducibility of different facial gestures*

º Outlier between 1.5 and 3 interquartile ranges
* Outlier greater than 3 interquartile ranges

The repeated measures ANOVA showed that Mauchly's test of sphericity (which tests for equality of variances of the differences between all possible pairs of groups in the analysis) had been violated, χ^2 (14) = 29.0, p = 0.013, therefore the degrees of freedom were corrected using Greenhouse-Geisser (Greenhouse and Geisser, 1959) estimates of sphericity (ε = 0.474). The results showed that there was a significant difference in the reproducibility between at least one pair of the six facial gestures, F = 45.84, p < 0.001.

Pair-wise comparisons of the 6 facial gestures using the Bonferroni correction (Table 7-1) showed that both *puppy* and *baby* were statistically significantly more reproducible than both *rope* and *bob* (*puppy v rope:* p = 0.014, *puppy v bob: p*= 0.018, *baby v rope:* p = 0.015, *baby v bob:* p = 0.021). There was no difference in the reproducibility between *puppy* and *baby* (p = 0.970) and *rope* and *baby* (p = 0.505). The standardised smile was more reproducible than the normal smile (p = 0.015).

Facial Gesture	Puppy	Rope	Baby	Bob	Standardised Smile	Normal Smile
Puppy		0.014	0.970	0.018	<0.001	<0.001
Rope			0.015	0.505	<0.001	<0.001
Baby				0.021	<0.001	<0.001
Bob					<0.001	<0.001
Standard Smile						0.015
Normal Smile						

Table 7-1   *Pairwise comparisons between different facial gestures showing p-values for reproducibility (Bonferroni correction applied)*

Figure 7-6 shows the number of PCs required to describe 95% of the variation in lip movement for each facial gesture. This shows that both the non-verbal facial gestures are composed of more variations in lip movement than the verbal gestures. This implies that they are complex facial gestures composed of a number of different subtle nuances in lip movement, which may explain their lower reproducibility within the group. Within the verbal gestures, *puppy* and *baby* require more PCs to describe the lip movement than *rope* and *bob*. In comparison, for the verbal facial gestures PC1 represents 75% of the total variation in lip movement, where as for the non-verbal facial gestures used in this study, PC1 only represents 50%.

Based on the WDDTW data, the hierarchy of reproducibility within this subject group from most to least reproducible is:

baby → puppy → rope → bob → standardised smile → normal smile
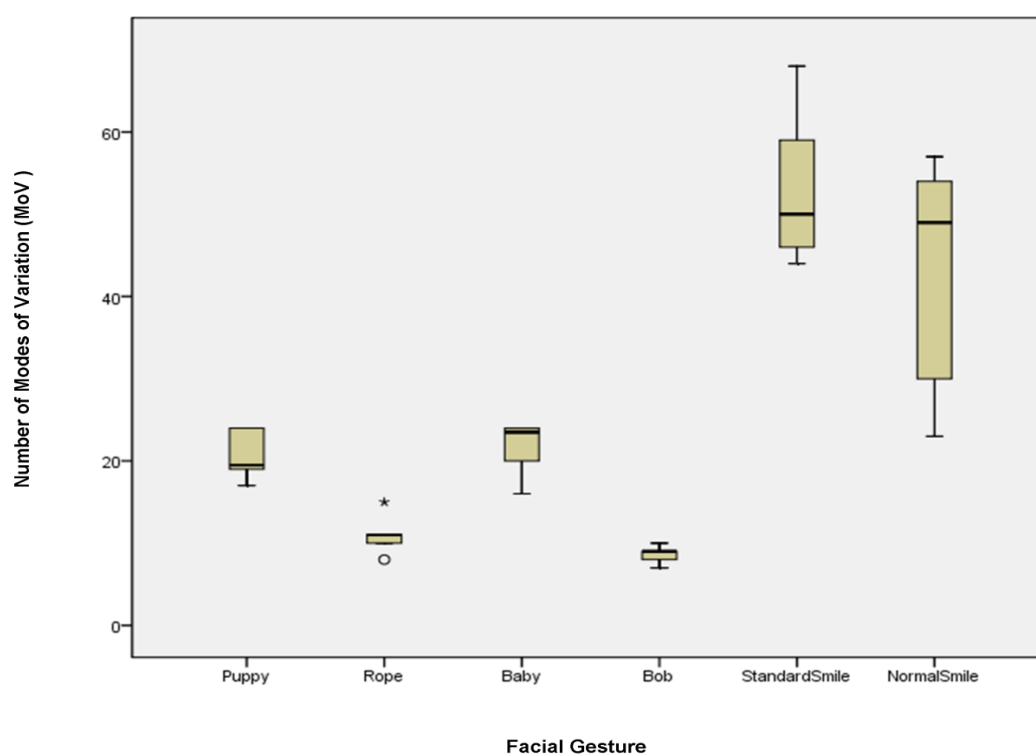
Figure 7-6   *Number of PCs extracted to describe 95% of the lip movement of the different facial gestures*

º Outlier between 1.5 and 3 interquartile ranges
* Outlier greater than 3 interquartile ranges

7.4.   Discussion

Reproducibility has become an important consideration in the objective assessment of facial movement, as monitoring facial movement over time or through intervention requires that the measure of facial movement used is repeatable. If this is not the case, then individual variation or difference in the performed facial gesture between two time points will induce error into the assessment.

It has generally been accepted that facial movement is reproducible within a short space of time (i.e. within 10–20 minutes) (Sawyer et al., 2009) but more important is the assessment of reproducibility over an extended period; as would be the case for investigation of facial movement clinically.

Non-verbal facial gestures form the bulk of the research carried out on reproducibility. Studies have shown a hierarchy of facial expression reproducibility with the rest pose being significantly more reproducible than lip purse, maximal smile, natural smile, or cheek puff (Johnston et al., 2003). Given that facial expressions have been closely linked to emotion, it could be considered that any facial expression could vary dependent on the subject's frame of mind at a particular point in time. This has important implications in assessing facial movement after a clinical intervention as the facial expression could vary dependent on whether the intervention was successful or not.

It is interesting to note that very few studies have investigated the reproducibility of verbal facial gestures. Verbal facial gestures are stimulators of lip movement (Parke and Waters, 1996) and therefore may not be suitable measures of facial movement if areas of the upper face are to be investigated. However for lip articulation in speech and language therapy, functional cleft lip repair evaluation and the assessment of motor deficits in the buccal branch of the facial nerve, verbal facial gestures could be considered appropriate. The four verbal gestures used in this study are words recommended for use in cleft speech assessments (Sell et al., 1999). They are bilabial speech postures which have been shown to give good representation of lip movement and carry little if any emotive connotations (Duffy, 1995). Collectively the four words also represent a range of different lip movements to include lip opening and lip stretch (*puppy* and *baby*) as well as lip purse (*rope* and *bob*).

For the non-verbal facial gestures, a standardized smile expression was used based on AU 12 of the FACS. This involved the subject performing a maximal smile without opening their mouth and without moving any other parts of the face, i.e. no variations other than lower face variations caused by the smile (so no eye widening or eyebrow movement). Subjects were also asked to smile maximally in a *normal* fashion allowing their lips to part – this expression is based on AU 25 of the FACS. These expressions have been commonly used in other studies and represent a balance between *posed* and *normal* facial movement (Houstis and Kiliaridis, 2009).

The sample size calculation was based on the number of participants required to statistically detect a difference in reproducibility of the six facial gestures. In terms of clinical significance, the WDDTW output was taken as a marker of reproducibility of the six facial gestures relative to each other. Therefore it is conceivable that the clinical difference in repeatability between gestures could be negligible but still statistically significant. As such, it is also important to take into account the variability in articulation/expression of a gesture, which is discussed in the following paragraphs.

Results from this part of the thesis found that different gestures require different numbers of variations in lip movement to describe them. Those verbal gestures composed of higher numbers of syllables (i.e. *puppy* and *baby*) require more components to describe the lip movement than words with lesser syllables (*rope* and *bob*). An even greater number of variations in lip movement were required to perform the non-verbal facial gestures suggesting that they are more complex facial movements when compared to spoken words. It may be considered that those gestures that are composed of more variations in lip movements might be less reproducible. This was certainly the case with the non-verbal facial gestures as the normal maximal smile had the highest mean number of variations in lip movement (44 ± 12) and was the least reproducible (mean 0.9901 ± 0.002). However, within the verbal facial gestures, *puppy* and *baby* were more reproducible than *rope* (p = 0.014, p = 0.015) and *bob* (p = 0.018, p = 0.021) but required more components in lip movement to articulate them. This could be explained by virtue that both *rope* and *bob* are plosive in their articulation and therefore this could lead to natural variation in their delivery.

An aspect of using verbal facial gestures in the assessment of facial movement is that during their delivery, there will be a combination of hard and soft tissue movement, i.e. as the mandible opens and the lips articulate. The assessment is essentially a collective and cannot differentiate how much hard or soft tissue movements individually contribute to the movement as a whole. The advantage of using non-verbal facial gestures is that the subject can be instructed to keep their teeth together so that the movement recorded

represents solely soft tissue function. Although the two non-verbal gestures showed a lower level of repeatability, the standardised smile was comparable to the verbal gestures in its (inter-quartile) range of variability between the two time points (Figure 7-5). Due to the low variability in its expression over time, it suggests that use of the standardised smile in clinical outcome evaluation could be adopted. However the lowest reproducibility and highest variability seen in the normal smile suggests that its use clinically showed be taken with caution.

7.5.   Conclusion

The findings would suggest that verbal facial gestures, in particular the words *baby* and *puppy* are more appropriate for use in the assessment of lip movement when compared to smile expressions in a clinical context due to their high level of repeatability over a one-month period.

Chapter 8

Average templates of lip movement

8.1. Introduction

When measuring or recording facial motion, one can assume that different subjects will have distinct movements. Ascertaining the variation of a particular motion in a representative sample of a population would outline trends/similarities in the movement and allow for the construction of references of what could be considered as average facial movement. The technique of creating normal references has been commonly used for static images in orthodontics and craniofacial research. Lateral cephalograms (Hans et al., 1994) and more recently 3D laser scans (Kau et al., 2006b) from population groups can be age and/or sex matched enabling comparisons to be made between an individual and their respective *average* template to guide treatment planning and measure outcome.

In the previous chapter, the most reproducible measures of facial movement were identified as the verbal facial gestures. Therefore the aim of this part of the study is to create average templates of lip movement based on these four words.
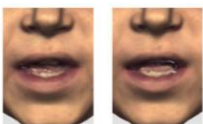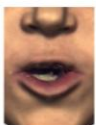
8.2. Methods

115 white subjects (62 male, 53 female) with a mean age of 33.4 years were recruited into this part of the study. Subjects were asked to say four verbal facial gestures (*puppy*, *rope*, *baby*, *bob*) in a normal and relaxed manner during 3D video capture using the 3dMDFace™ Dynamic System.

8.2.1. Image processing

The video sequences were analysed according to the visemes or mouth shapes for each word. The corresponding phonetic descriptions based on British English are shown in Table 8-1. For the four words used in this study there are nine phonemes (including silence). As there is not always one-to-one mapping between phonemes and visemes - seven visemes (rest, p̲uppy, pupp̲y, r̲ope, b̲aby, bab̲y, b o̲b) were analysed in this part of the study (Figure 8-1).

Table 8-1 *Phoneme-to-viseme mapping of the four study words based on British English*

Phoneme	Description	Example	Viseme
Silence	At rest		
/p/ /b/	Plosive consonant	puppy baby	
/ʌ/	Short vowel	puppy	
/eɪ/	Short vowel	baby	
/i/	Long vowel	puppy baby	
/r/	Approximant consonant	rope	
/əʊ/	Diphthong vowel	rope	
/ɒ/	Short vowel	bob	
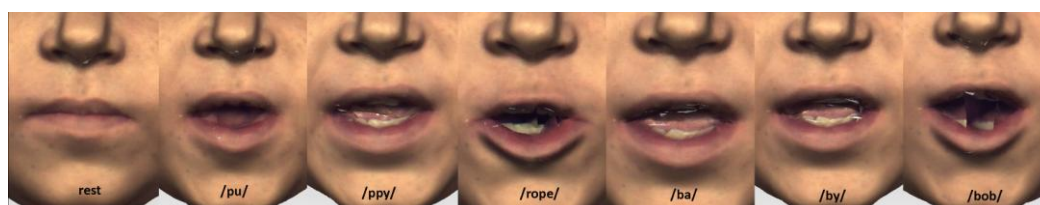
Figure 8-1 *Visual illustration of the seven visemes analysed*

All facial shells were initially aligned to Standardised Head Position (SHP) as described in Section 6.2.1.

Following alignment to SHP, six lip landmarks as previously described in Section 6.2.2 were manually placed by one examiner around the lips for each 3D facial shell. To account for temporal variation in the articulation of the visemes, only the landmark displacement vectors (x, y, z coordinates) for the frame of *maximal lip movement* were recorded. This gave 6 landmark clusters for each lip shape. The frame of *maximal lip movement* represented the point at which the upper and lower lips were most apart in the vertical plane for the visemes p<u>u</u>ppy and <u>b</u>aby, where the commissures were at their widest for the visemes pupp<u>y</u> and bab<u>y</u>, and where the lips were at their most protrusive for the visemes <u>r</u>ope and b<u>o</u>b. This frame was selected by direct observation. If it was not possible to identify the frame of *maximal lip movement* by direct observation, then the images considered to be in contention for selection were all manually landmarked and the image with the largest displacement vectors for a particular viseme was selected. The implication of this methodology is further discussed in Section 8.4.

8.2.2.   Statistical analysis

Generalised Procrustes Analysis (GPA) was used to align the coordinates for all landmarks in the dataset. Translation and rotation were incorporated. Scaling was omitted to observe the effect of size and shape on lip movement within the sample. Following registration, a centroid representing the mean position for each of the 6 clusters of landmarks was derived. Two standard deviations (SD) around each centroid (representing 95% of the variation in x, y, and z from the mean) were calculated for all individuals and plotted as ellipsoids in RAPIDFORM™ software (INUS Technology Inc., Seoul, South Korea). To quantify lip movement the Cartesian displacement vectors from rest to peak amplitude for each viseme were also tabulated.

Principal Component Analysis (PCA) was then applied to the registered coordinates to isolate patterns and relationships between the lip landmarks for each viseme giving 6 separate analyses. This was carried out on the variance-co-variance matrix for each viseme using the software package MorphoJ (Klingenberg, 2011). Kaiser's criterion was implemented which retains only those components with an eigenvalue of 1.0 or more for further investigation (Kaiser, 1960). The eigenvalue of a component represents the total variance explained by that component. Therefore, the first principal component (PC1) will include those landmark coordinates that account for the highest variation in the dataset. Subsequent components account for the next highest variation (PC2, PC3, PC4, and so on) and are independent of the previous components. The PCA data is presented in tabular format with variables ordered into PCs by their loading, which is a representation of the substantive importance of the coordinate to a given PC (Tables 8-8 to 8-13). For a sample size of up to 200, the critical value against which loadings can be considered significant was set at greater than or equal to 0.21 (Stevens, 2002).

Stem plots show the shifts of landmark positions associated with the first three PCs (Figures 8-11 to 8-16). Each stem starts with a dot at the location of the landmark centroid. The length and direction of the stem indicates the movement of the respective landmark from the centroid to the shape change that corresponds to an increase of 0.1 units of Procrustes distance in the direction of the respective PC.

To investigate whether any differences in lip movement exist between males and females, scatter graphs of the PC scores for each individual were plotted labelled by gender. These scores are calculated using regression coefficients and represent estimates of the scores individual subjects would have received on each of the PCs had they been measured directly.

Finally, a Canonical Variate Analysis (CVA) was carried out using the landmark data for all seven visemes (including rest) entered into the model with groupings specified *a priori*. Although PCA is already being utilised to describe the patterns of lip shapes for individual visemes, the purpose of CVA was to identify the shape differences between the visemes. CVA projects multivariate data in a manner that maximises the separation between three or more given groups. It is an extension of discriminant analysis which is utilised in Chapter 9. CVA of N groups (here, 7 groups: rest, puppy, puppy, rope, baby, baby, bob) will produce $N - 1$ axes (here, 6 canonical variates) of diminishing importance. As in PCA, eigenvalues explain the amount of variation in lip shape for a particular canonical variate (CV). In this example, seven groups will produce six CVs. Scatterplots of the CV scores for each individual by viseme allows visualisation of which CVs differentiate the visemes. Significance testing (at a threshold of $p < 0.05$) of the CVs provides a quantitative measure of which CVs statistically differ.

CVA was carried out using SPSS 20.0.0 (SPSS Inc., Chicago, IL).

8.2.3. Data preparation

As PCA was being used descriptively to summarise the relationships in a relatively large set of variables, assumptions regarding the distribution of variables (i.e. normality) were not in force (Tabachnick and Fidell, 2007).

On the other hand, the use of CVA required the assumption of multivariate data normality and homogeneous variance-co-variance matrices between groups to be fulfilled.

Multivariate outliers were identified using the Mahalanobis distance. In this application, the Mahalanobis distance is defined as the distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of all the means of all the variables. The criterion for multivariate outliers is Mahalanobis distance of $p < 0.001$. The

Mahalanobis distance is evaluated as Chi Square (χ^2) with degrees of freedom equal to the number of variables, in this case 18. Therefore any case with a Mahalanobis distance greater than χ^2 (18) = 42.3 would be regarded as a multivariate outlier (Ahrens, 1958). Any cases that were identified as such were removed from the analysis.

Homogeneity of variance-co-variance matrices was assessed through use of Box's M Test. A significance of $p < 0.001$ was set (Anderson, 1984, Box, 1949). Classification on separate-group co-variance matrices was stipulated if heterogeneity was found.

Tests for normality were carried out using SPSS 20.0.0 (SPSS Inc., Chicago, IL, USA).

8.3. Results

8.3.1. Data preparation

There were five multivariate outliers identified outside the threshold of Mahalanobis distance 42.3 (Figure 8-2). These cases were removed prior to CVA. In addition, Box's M Test was violated ($p < 0.001$) and therefore separate-group co-variance matrices were displayed for CVA.
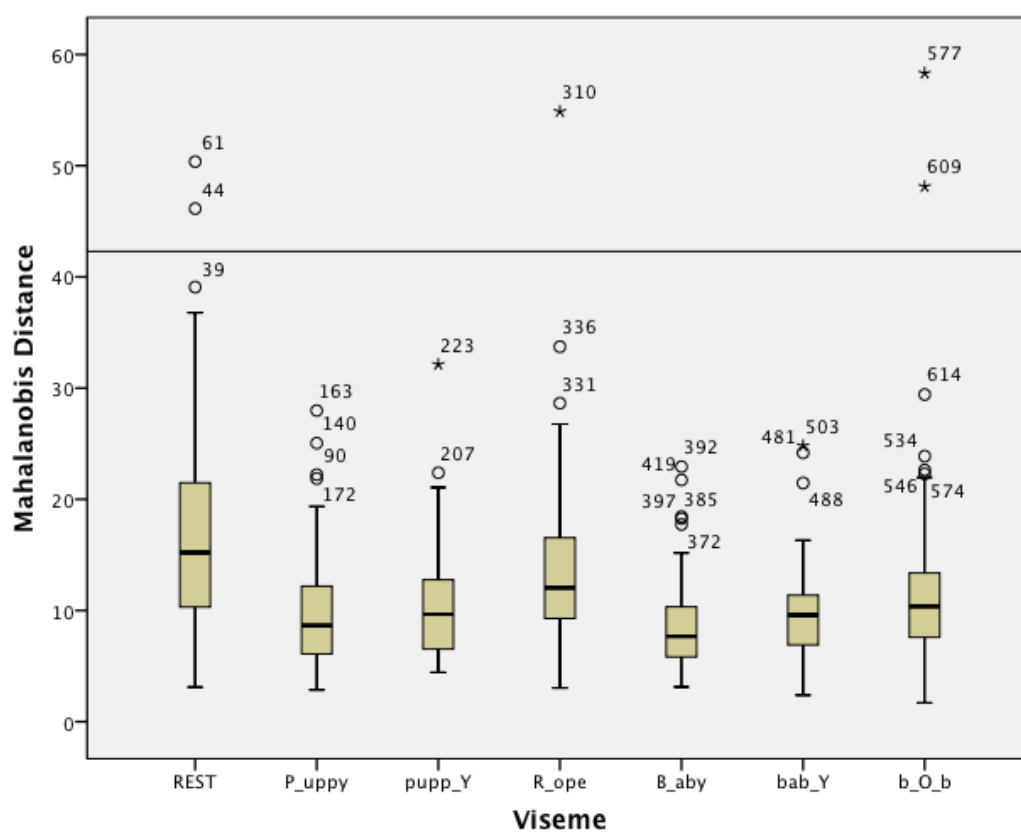
Figure 8-2 *Boxplots showing multivariate outliers for the seven visemes*

8.3.2. GPA

Figure 8-3 shows scatter plots of the raw lip landmarks coordinates from the resting lip shape of the sample and their correspondence after GPA.
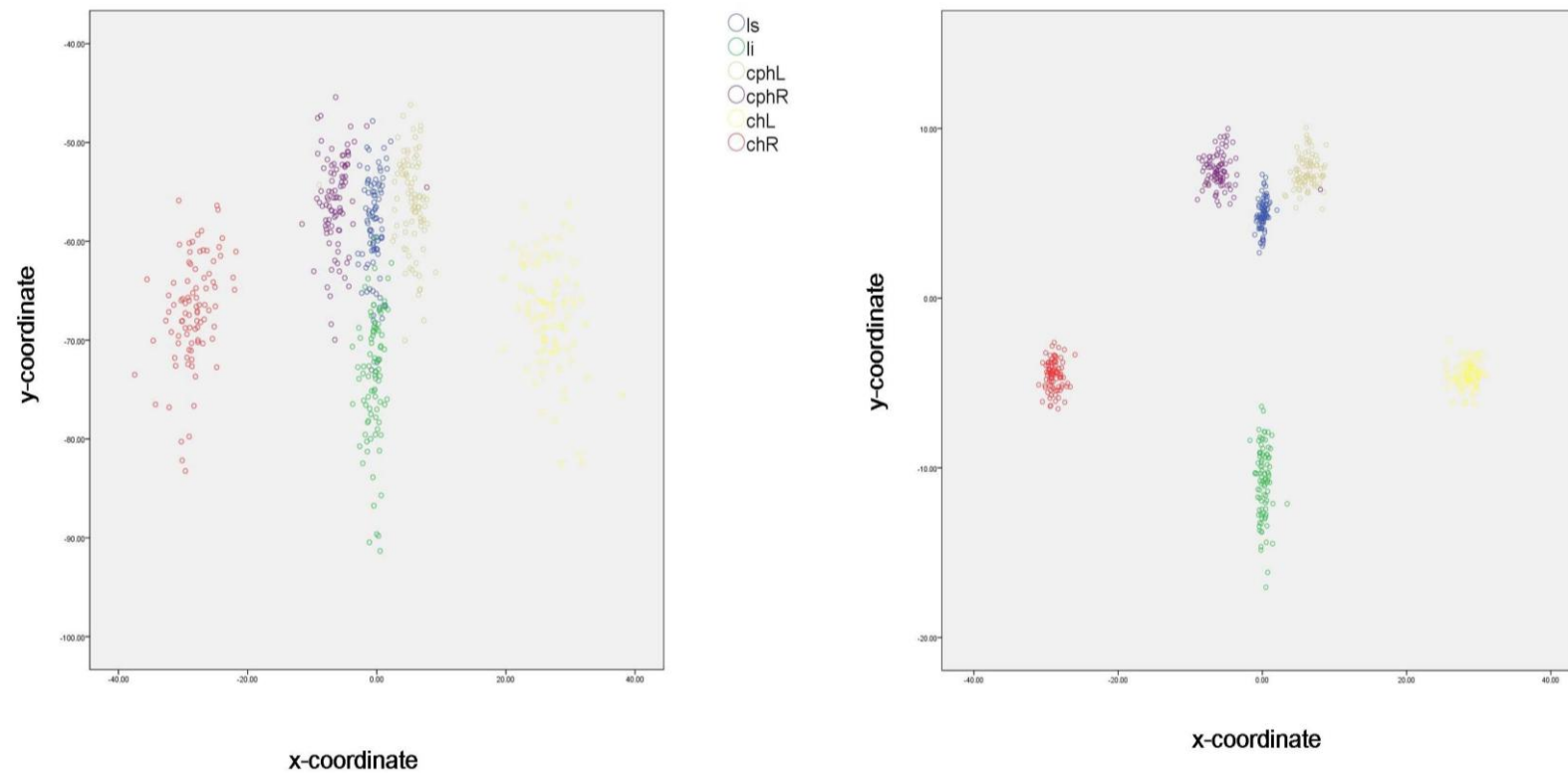
Figure 8-3 *Scatter plots of the x-y coordinates of the resting lip shape*
Raw data (left) and after GPA registration (right)

8.3.3.   Landmark displacements

The variation in resting lip shape between the genders is shown in Figure 8-4. Females (black) had a narrower lip width both between the corners of the mouth and also between the elevated upper vermilion borders when compared to males (white). In addition, females had a shorter overall lip height when compared to males.
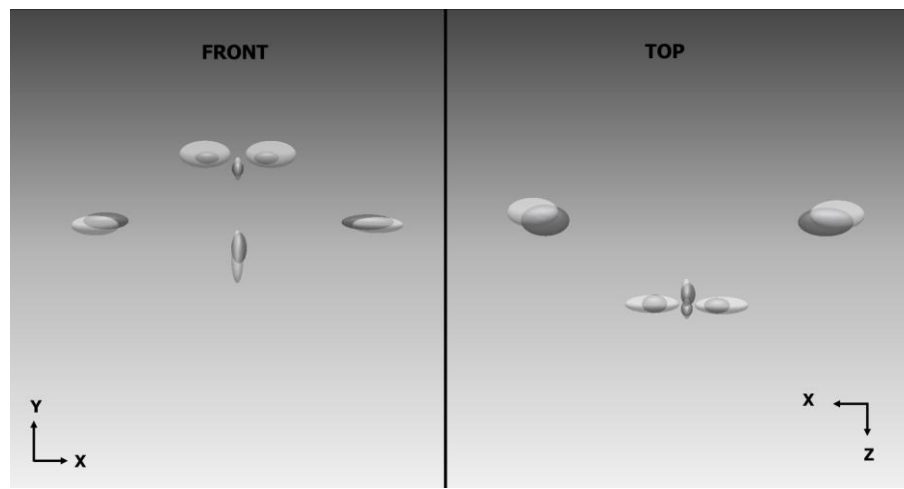
Figure 8-4   *Ellipsoid plot of resting lip shape*

males (white)   females (black)

The subsequent figures show ellipsoid plots of each viseme from the resting lip shape (blue) to peak amplitude. Their associated tables allow quantification of movement (in mm) for each viseme visualised in the figures. In Section 6.3 it was found that the largest mean distance error in landmark reproducibility was recorded at 1.39mm, SD = 0.57 (inter-examiner for *chR*). Therefore, only mean displacements greater than 2.0 mm are considered as contributors to their respective visemes. The reader is referred back to Figure 6-3 for the convention of movement by plane and sign.
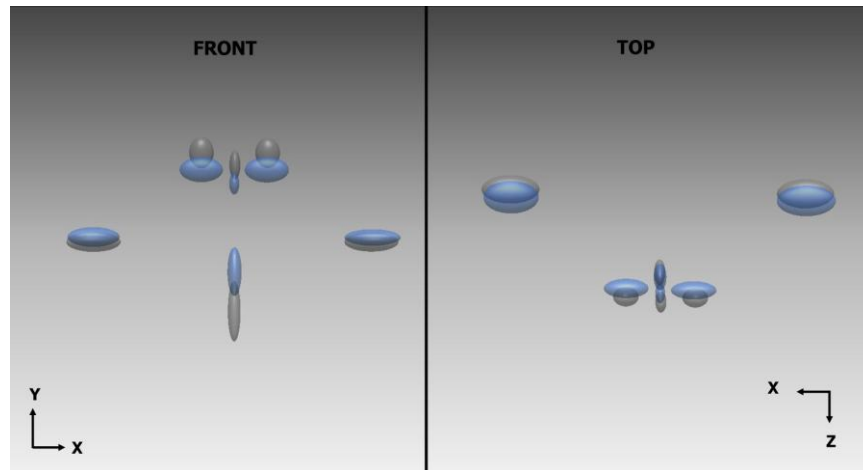
Figure 8-5 *Ellipsoid plot for the viseme puppy*

Rest (blue)   Maximal displacement (grey)

Table 8-2 *Mean (SD) movement for the viseme puppy*

Landmark	Plane of movement (mm)					
	x		**y**		**z**	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.16	0.86	-2.48	2.32	-2.29	2.29
li	-0.28	1.20	9.65	3.59	-0.05	2.75
cphL	-0.22	2.46	-3.05	2.03	-2.15	2.08
cphR	0.01	2.43	-3.46	2.03	-2.17	2.18
chL	-0.34	2.62	-0.24	2.54	0.25	2.87
chR	0.23	2.34	-0.19	2.52	0.90	2.58

(red highlights principal contributors > |2.0mm|)

The viseme puppy can be described as principally a mean downward movement of the lower lip at *li* of up to 10mm. There is an associated mean upward movement of the midline, left and right upper lip at *ls* and *cph* of approximately 3mm. This equates to an overall mouth opening of 13mm.

In addition to the vertical component, there is also a slight mean protrusive movement of the upper lip at *ls* and *cph* of up to 2.5mm.

There is negligible movement in the lateral plane.

The standard deviations for each landmark displacement indicate that there is moderate variation in peak amplitude for this viseme within the sample.
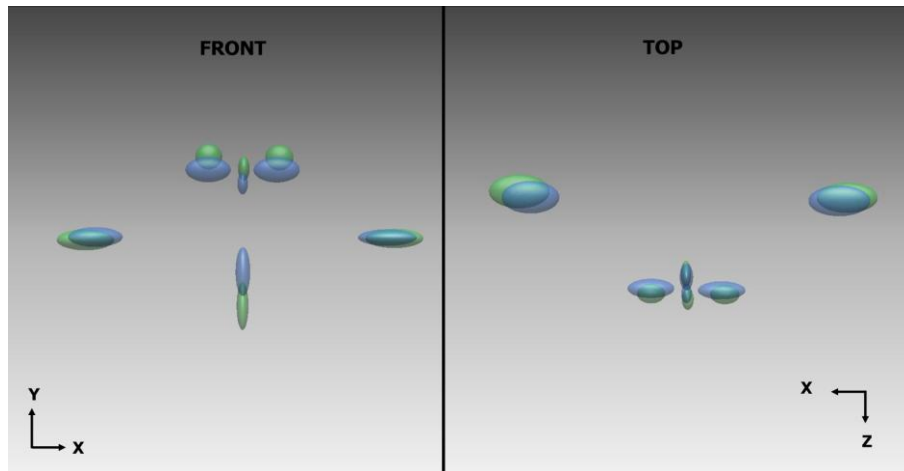
Figure 8-6   *Ellipsoid plot for the viseme puppy*

Rest (blue)   Maximum displacement (green)

Table 8-3   *Mean (SD) movement for the viseme puppy*

Landmark	Plane of movement (mm)					
	x		**y**		**z**	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.26	0.84	1.53	2.24	-1.36	2.19
li	-0.27	1.10	7.63	3.06	0.36	2.81
cphL	-0.47	2.34	1.17	2.20	-1.22	2.11
cphR	-0.08	2.13	1.26	2.20	-1.31	2.17
chL	-1.27	2.95	0.64	2.80	0.23	2.86
chR	1.77	2.77	0.98	2.56	0.85	3.06

(red highlights principal contributors > |2.0mm|)

Except for a downward movement of the lower lip at *li* of 7.63mm, there were no other mean landmark displacements that exceeded 2mm. However, the magnitude of standard deviations suggests that there is moderate variation in peak amplitude for this viseme within the sample.
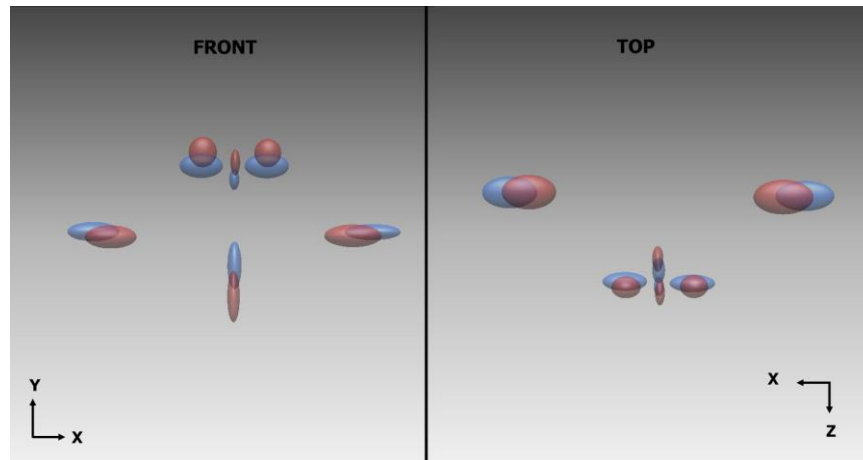
Figure 8-7 *Ellipsoid plot for the viseme ɹope*

Rest (blue)   Maximum displacement (red)

The viseme ɹope is principally composed of a downward movement of the lower lip at *li* with a mean of approximately 7.5mm. There is an associated mean downward movement of the left and right commissures of up to 3mm.

The commissures narrow the mouth aperture through medial movement of *chL* and *chR.*

Table 8-4 *Mean (SD) movement for the viseme ɹope*

Landmark	Plane of movement (mm)					
	x		**y**		**z**	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.19	1.06	-1.54	2.41	-4.96	3.35
li	-0.14	1.20	7.55	3.41	-3.01	3.59
cphL	-0.23	2.47	-1.02	2.29	-4.49	3.20
cphR	-0.11	2.56	-1.04	2.30	-4.64	3.22
chL	3.68	3.03	2.29	2.84	-1.31	4.68
chR	-4.17	2.69	2.80	3.16	-1.47	4.52

(red highlights principal contributors > |2.0mm|)

All landmarks show a mean protrusive element although this was primarily related to the upper lip. The magnitude of the standard deviation particularly in the Z plane suggests that there is a wide variation in protrusive movement for this viseme.
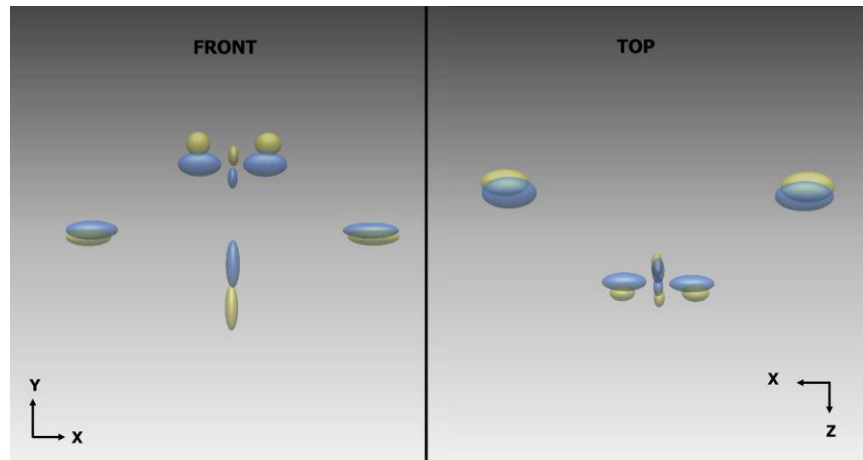
Figure 8-8   *Ellipsoid plot for the viseme baby*

Rest (blue)   Maximum displacement (gold)

Table 8-5   *Mean (SD) movement for the viseme baby*

Landmark	Plane of movement (mm)					
	x		**y**		**z**	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.22	0.95	-1.90	2.26	-1.14	1.86
li	0.01	1.18	9.06	2.99	1.47	2.94
cphL	-0.58	1.92	-1.66	2.14	-0.92	1.86
cphR	0.11	1.82	-1.64	2.14	-1.03	1.88
chL	-1.45	2.37	2.55	2.62	2.36	2.86
chR	0.94	2.46	2.68	2.62	2.54	2.69

(red highlights principal contributors > |2.0mm|)

The principal mean movement for the viseme baby is a downward movement of the lower lip in the order of 9mm. There is also a slight protrusive movement of the elevated margins of the upper lip. However this is only marginally above 2mm in magnitude and the standard deviation suggests a high variation within the sample for these landmarks.
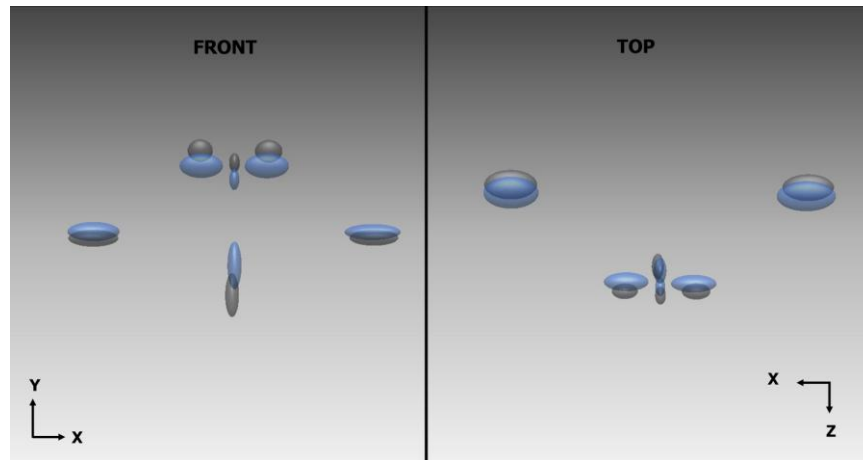
Figure 8-9   *Ellipsoid plot for the viseme baby*

Rest (blue)   Maximum displacement (grey)

Table 8-6   *Mean (SD) movement for the viseme baby*

Landmark	Plane of movement (mm)					
	x		**y**		**z**	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.06	0.95	-2.38	2.16	-0.49	1.91
li	0.24	1.27	5.50	3.01	1.15	2.76
cphL	-0.36	1.96	-2.07	2.14	-0.37	1.86
cphR	0.11	1.78	-2.13	2.13	-0.42	1.90
chL	-1.06	2.56	1.35	2.65	2.35	2.65
chR	0.27	2.37	1.49	2.48	2.34	2.67

(red highlights principal contributors > |2.0mm|)

Mean landmark displacement for the viseme baby involves a combination of downward movement of the lower lip and upward movement of the upper lip. This is in favour of the lower lip in an almost 2:1 ratio. There is also a protrusive element to the corners of the mouth, but in a similar finding to the preceding viseme, the magnitude is only marginally over 2mm and as the standard deviation is relatively high, the variation in the sample in wide.

There is a suggestion of slight asymmetric downward movement of the lip to the right hand side in Figure 8-9 but this was not overly implied by mean landmark displacement of *li* in the x plane.
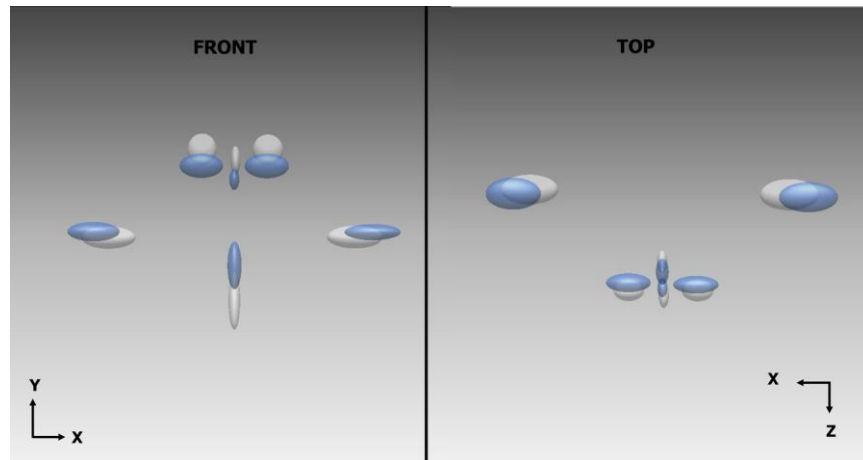
Figure 8-10 *Ellipsoid plot for the viseme bob*

Rest (blue)   Maximum displacement (white)

Table 8-7 *Mean (SD) movement for the viseme bob*

Landmark	Plane of movement (mm)					
	x		y		z	
	Mean	SD	Mean	SD	Mean	SD
ls	-0.26	1.04	-1.18	2.69	-4.24	2.90
li	0.01	1.10	7.93	3.81	-2.28	3.44
cphL	-0.16	2.01	-0.78	2.44	-3.99	2.85
cphR	-0.11	1.99	-0.80	2.46	-3.95	2.74
chL	3.00	2.88	3.19	2.73	-3.31	4.05
chR	-3.01	2.70	3.40	2.63	-2.58	3.82

(red highlights principal contributors > |2.0mm|)

In a similar manner to rope, the viseme bob shows strong protrusive elements for all landmarks in the z plane. In addition, there is contribution from *li* to mouth opening in the order of almost 8mm. The corners of the mouth appear to move towards each other as well as moving downwards. Standard deviations for all principal contributors are relatively high implying a wider degree of variation in movement for the sample.

8.3.4. Principal component analysis

The tables and figures included in this section show PCA for the visemes outlined in the method. For all visemes, five PCs were extracted with an Eigenvalue greater than 1.0. The total variance in lip shape accounted for by these components ranged from 80-92%. The first component controlled mouth opening and mouth width with the coordinates *li* Y, *ls* Y, *cphL* Y, *cph* R Y, *chL* X and *chR* X loaded on PC1 for all visemes. This explained up to 50% of the total variance in lip shape. For p̲uppy, pupp̲y, r̲ope, b̲aby and bab̲y PC1 was visualised as an increase in midline mouth opening and a narrowing of the mouth width associated with the commissures moving towards each other. However this was the converse for b̲o̲b as mouth opening reduced and mouth width increased (Figure 8-16).

PC2 accounted for up to 18% of the variance in lip shape for the visemes. Here, *chL* Y and *chR* Y loaded with *li* Z, showing a superior movement of the commissures in addition to a protrusive movement of the midline point of the lower lip. This was a similar trend across all visemes. PC2 also isolated some lateral movement of the upper elevated vermilion margins. This tended to be a widening as seen for visemes r̲ope, b̲aby, bab̲y and b̲o̲b.

The coordinates *chL* Z and *chR* Z generally loaded on PC3 or PC4 and this was most commonly seen as a retrusive movement associated with the corners of the mouth moving backwards. The total amount of variance explained by this movement was up to 11%.

There were coordinates that had high loadings across several components, which may imply a dominant movement for a particular viseme. However the PCs included were important to note. For example, *li* Y loaded on PC1 in addition to PC2 and PC5 for the viseme pupp̲y. Together these components explained 69% of the total variance for this viseme (47, 15, and 7% respectively). For the same viseme, *chR* Y

loaded across all components except PC1. However together, these components only explained 41% of the variance (15, 10, 9 and 7% respectively).

The coordinates *ls* X and *li* X were consistently loaded on the lower order PCs or often not recorded with a loading greater than 0.21 for the components extracted. Therefore, given that these landmarks were low contributors to the overall variance in the transverse plane, it implies that midline lip movement remained in the midline. In addition, bilateral coordinates were generally loaded in pairs with loading scores of similar magnitude between left and right sides. This suggests lateral symmetry in lip movement too.

Interestingly there was very little protrusive movement of the upper lip with coordinates *ls* Z, *cphL* Z and *cphR* Z not loading high enough to be recorded on the components extracted for any of the visemes.

Table 8-8   *PCA table for the viseme puppy*

Coordinate	PC (% variance)				
	1 (45)	2 (18)	3 (10)	4 (8)	5 (6)
li Y	-0.61	-0.48			
chR X	0.38		0.50		
chL X	-0.36		0.30	0.42	0.65
ls Y	0.31				
cphR Y	0.29				
cphL Y	0.28				
chL Y		0.48	-0.24	-0.26	
chR Y		0.45	0.30		
li Z			0.40	-0.60	-0.22
chL Z			-0.36	0.22	
cphR X			-0.33		-0.69
chR Z				0.40	
cphL X				-0.29	
ls X[a]					
li X[a]					
cphL Z[a]					
cphR Z[a]					
ls Z[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
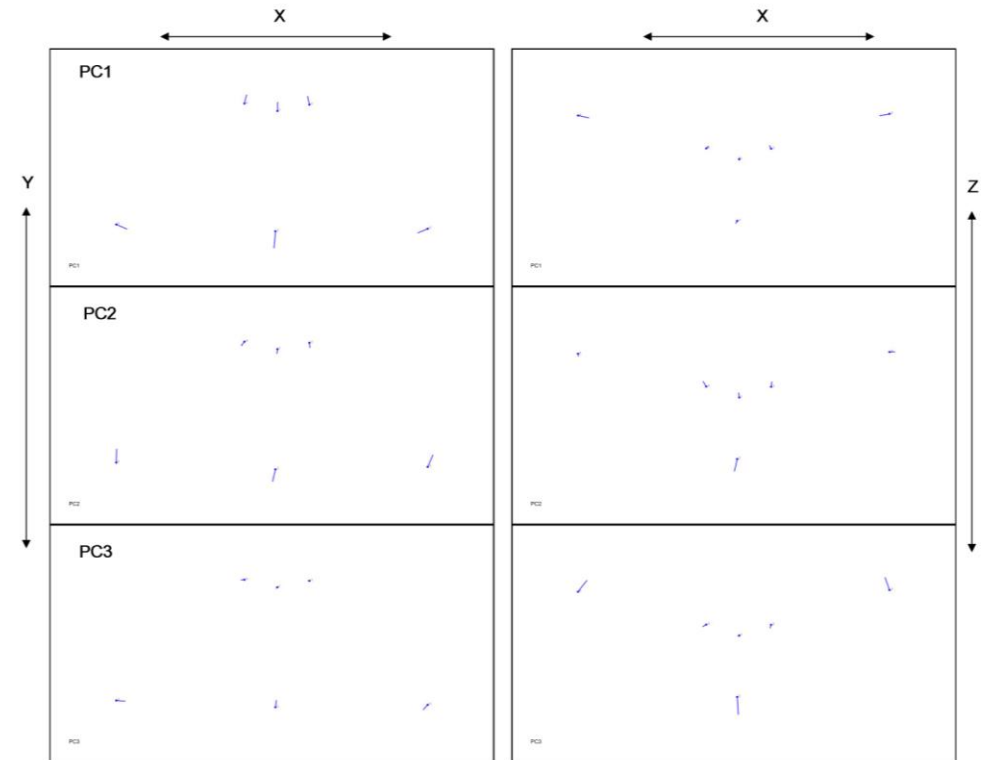
Figure 8-11   *Shape changes associated with PC1-3 puppy*

Table 8-9   *PCA table for the viseme puppy*

Coordinate	PC (% variance)				
	1 (47)	2 (15)	3 (10)	4 (9)	5 (7)
li Y	0.60	-0.31			-0.40
chL X	0.42		-0.23	0.30	0.35
chR X	-0.38			-0.50	
ls Y	-0.29				
cphL Y	-0.27				
cphR Y	-0.24				
chL Y		0.48		-0.35	
li Z		0.46	-0.50		-0.31
chR Y		0.35	-0.22	0.24	0.36
cphL X		0.33	0.54		-0.33
cphR X		-0.27	-0.28	-0.59	
chL Z			0.35		0.33
chR Z			0.21		0.34
ls X				-0.22	
li X[a]					
cphL Z[a]					
cphR Z[a]					
ls Z[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
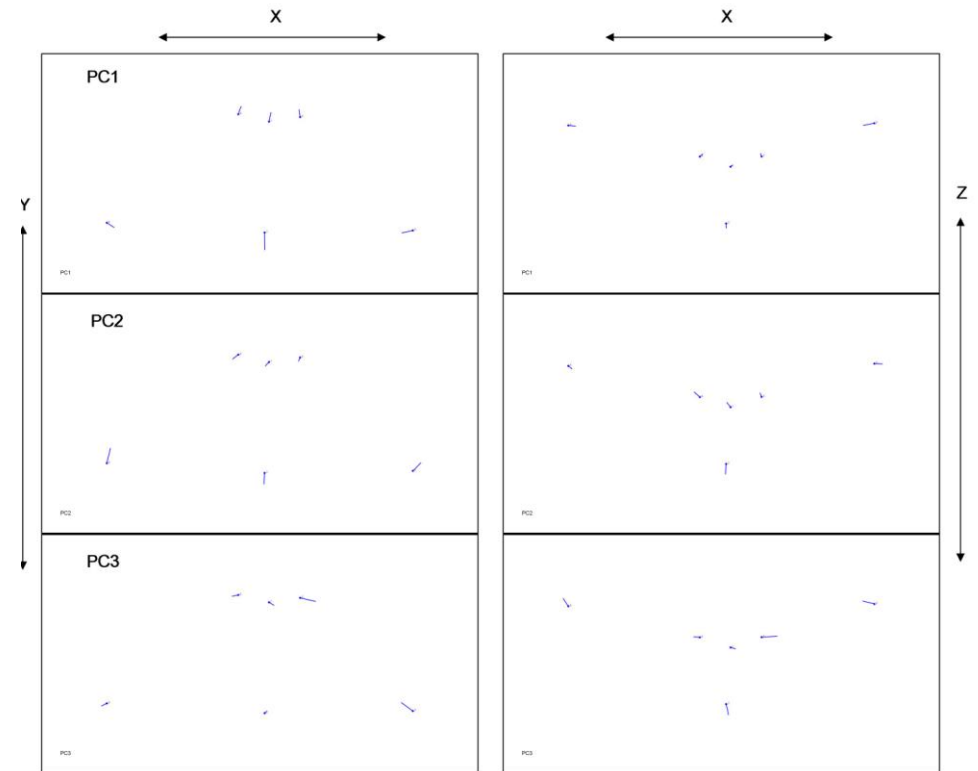
Figure 8-12 *Shape changes associated with PC1-3 for puppy*

Table 8-10 *PCA table for the viseme rope*

Coordinate	PC (% variance)				
	1 (48)	2 (17)	3 (9)	4 (8)	5 (6)
li Y	-0.59	-0.47			
chL X	-0.45			0.48	
chR X	0.31			0.45	-0.39
ls Y	0.29				
cphL Y	0.28				
cphR Y	0.28				
chL Y		0.46			
chR Y		0.42			
cphR X		-0.28	0.36		0.54
cphL X		0.28	-0.39	-0.37	-0.40
li Z		0.27	0.64	0.33	-0.31
chR Z			-0.35		0.21
chL Z			-0.33		0.32
ls X				-0.22	
li X				-0.22	
cphL Z[a]					
cphR Z[a]					
ls Z[a]					

[a] *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
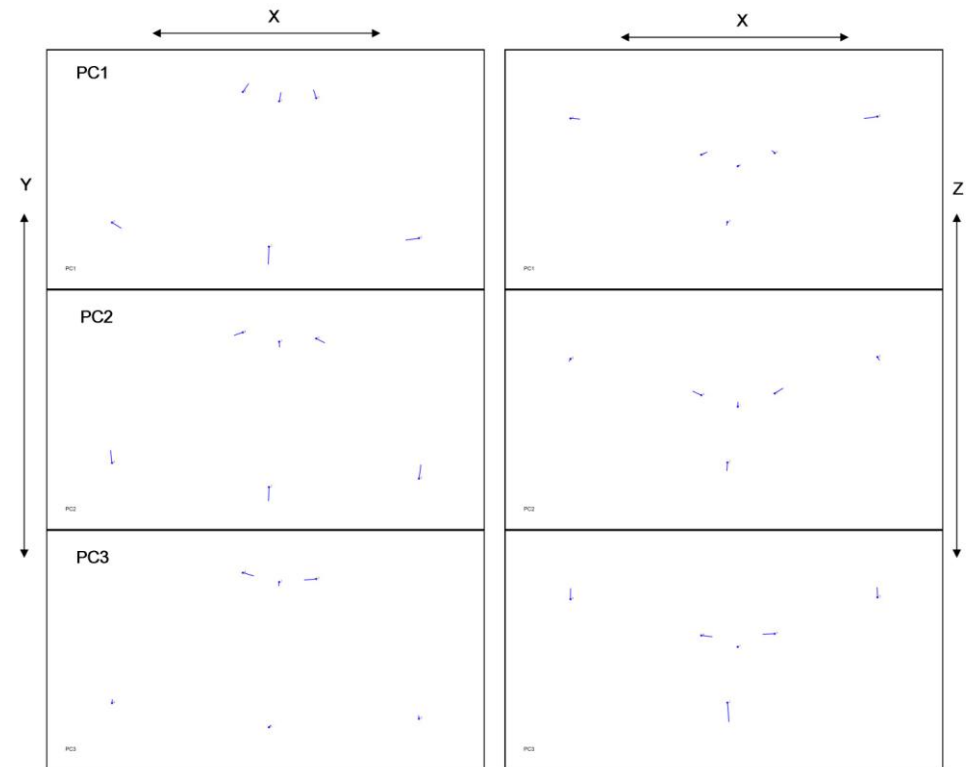
Figure 8-13 *Shape changes associated with PC1-3 for rope*

Table 8-11 *PCA table for the viseme baby*

Coordinate	PC (% variance)					
	1 (45)	2 (16)	3 (11)	4 (9)	5 (6)	6 (5)
li Y	-0.70			0.28		
chL X	-0.40		0.22	-0.47		
chR X	0.26	-0.36	0.34		0.36	
cphL Y	0.25	-0.23				
ls Y	0.24					
cphR Y	0.24					
chL Y		0.55			-0.22	0.26
cphL X		0.40	-0.29	0.25	0.51	
chR Y		0.26	0.40	-0.36		
li Z			0.60	0.43		
chL Z			-0.32	-0.28		-0.22
chR Z				-0.42		0.22
cphR X					-0.71	0.21
li X						-0.80
ls X[a]						
cphL Z[a]						
cphR Z[a]						
ls Z[a]						

[a] *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
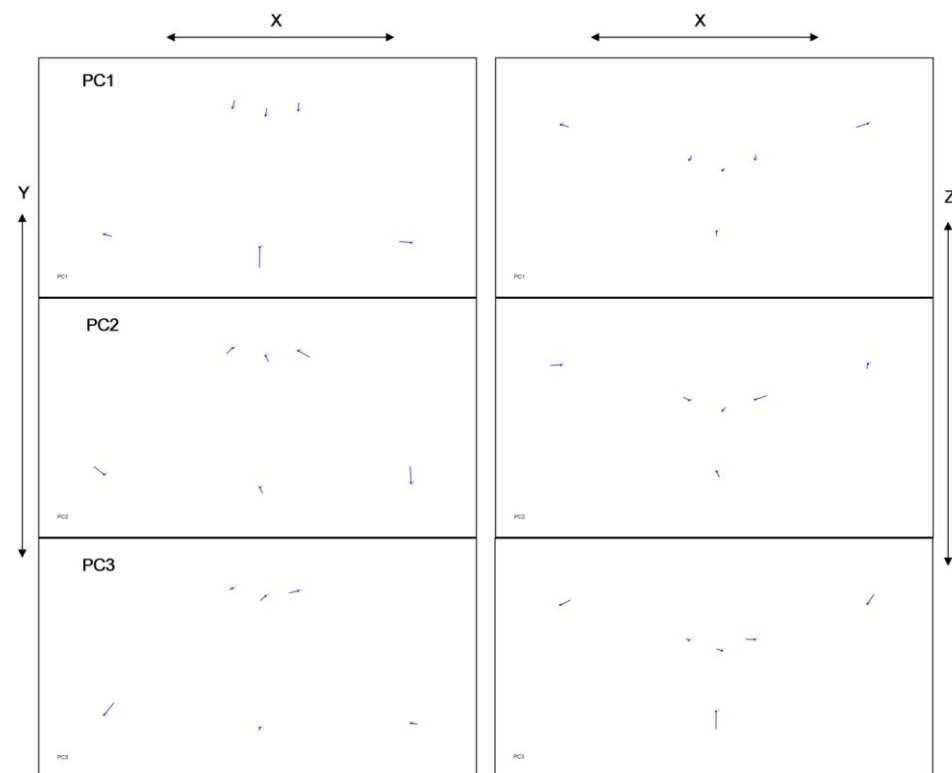
Figure 8-14 *Shape changes associated with PC1-3 for baby*

Table 8-12 *PCA table for the viseme baby*

Coordinate	PC (% variance)				
	1 (41)	2 (17)	3 (10)	4 (10)	5 (7)
li Y	-0.67		-0.30		-0.22
chL X	-0.35		-0.53		
chR X	0.30		-0.32	0.31	-0.32
ls Y	0.29				
cphL Y	0.26				
cphR Y	0.26				
cphL X		0.42	0.29	-0.33	-0.43
cphR X		-0.42			0.59
chR Y		0.40	-0.34		
chL Y		0.40	0.27		0.36
li Z		0.30	0.33	0.66	
chR Z		-0.30	-0.32	-0.31	
ls X			0.24		
chL Z				-0.36	
li X[a]					
cphL Z[a]					
cphR Z[a]					
ls Z[a]					

[a] *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
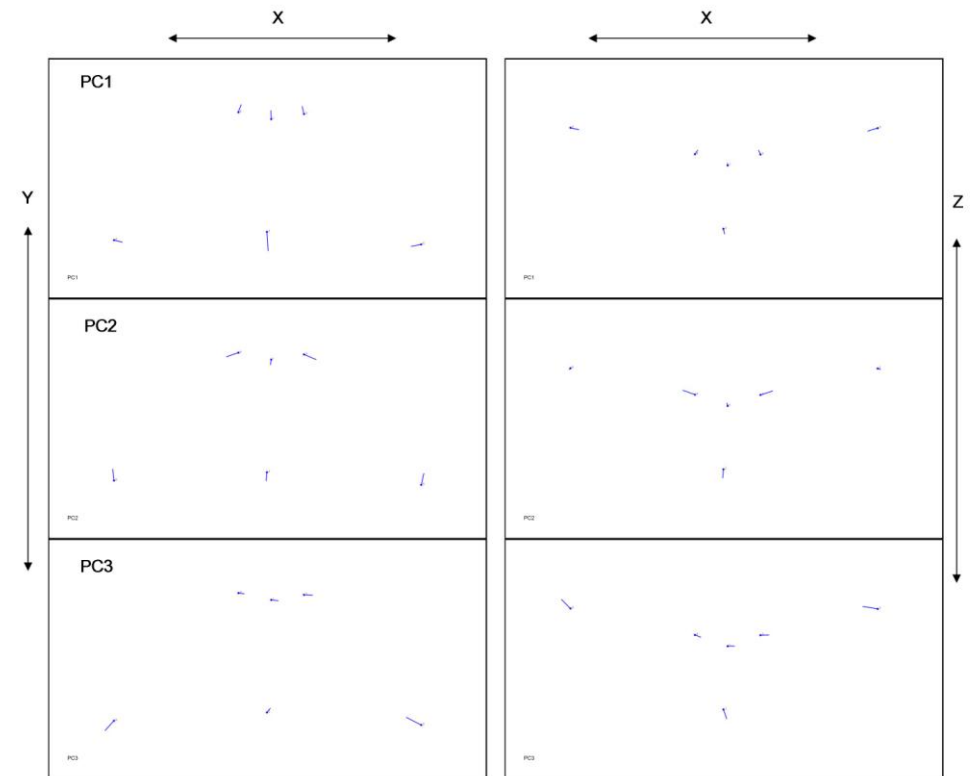
Figure 8-15 *Shape changes associated with PC1-3 for baby*

Table 8-13 *PCA table for the viseme b<u>o</u>b*

Coordinate	PC (% variance)				
	1 (50)	2 (12)	3 (8)	4 (7)	5 (4)
li Y	0.60	-0.31			-0.40
chL X	0.42		-0.24	0.30	0.34
chR X	-0.38			0.50	
ls Y	-0.29				
cphL Y	-0.26				
cphR Y	-0.24				
chL Y		0.47		-0.35	
li Z		0.46	-0.50		-0.32
chR Y		0.35	-0.22	0.23	0.36
cphL X		0.33	0.54		-0.33
cphR X		-0.28	-0.28	-0.59	
chL Z			0.35		0.33
chR Z			0.21		0.34
ls X				-0.22	
li X[a]					
cphL Z[a]					
cphR Z[a]					
ls Z[a]					

[a] *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
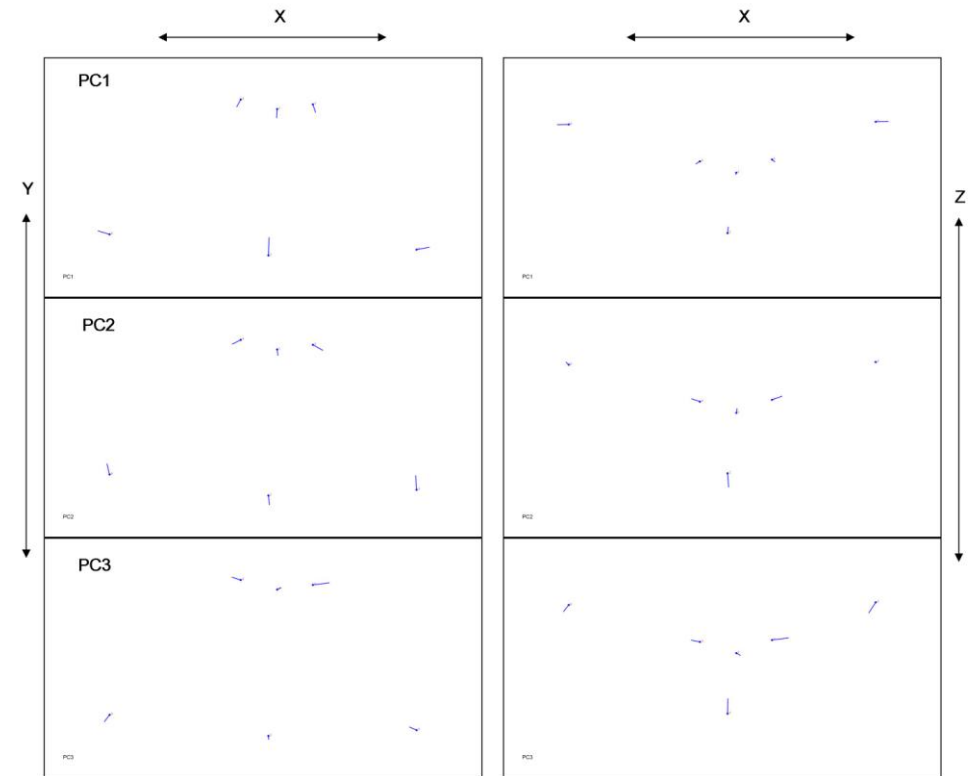
Figure 8-16 *Shape changes associated with PC1-3 for b<u>o</u>b*

Male lip movement was compared to that of female movement by plotting the PC scores for each individual separated by viseme. These scores are calculated using regression coefficients and represent estimates of the scores individual subjects would have received on each of the PCs had they been measured directly. Ellipses encompassing the 95% CI for the PC scores labelled by gender were plotted which allowed visualisation of which PC (if any) discriminated the sexes.

Figure 8-17 shows how the PC scores have been plotted and labelled by gender to compare lip movement between males and females for the viseme p̲uppy. Despite overlaps of the 95% CI ellipses, there is a suggestion of separation along PC1 at the extremes. This may imply a difference between the genders for this viseme in relation to midline lip opening and mouth width (Table 8-8).
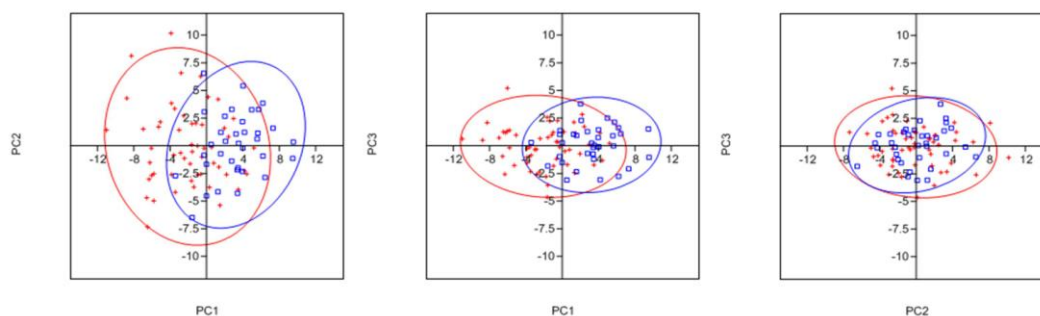
Figure 8-17 *PC1-3 plotted for the viseme p̲uppy by gender*

males (blue) females (red)

Figure 8-18 to Figure 8-22 show plots of the first three PCs for the remaining visemes. The only other viseme to show separation along PCs was b̲o̲b, which in a similar manner to p̲uppy was along PC1. However on this occasion, the separation was in the opposite direction.
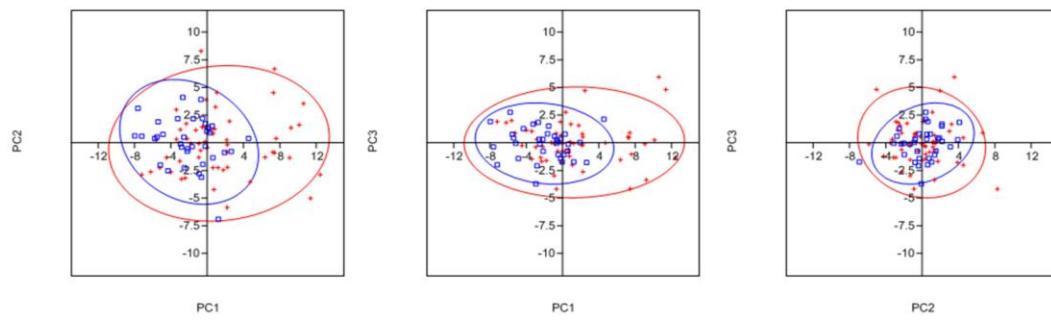
Figure 8-18 *PC1-3 plotted for the viseme pupp*u*y by gender*

males (blue) females (red)
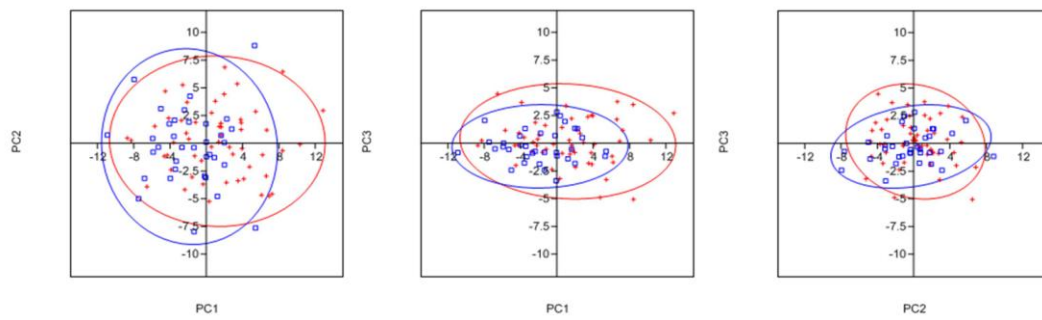
Figure 8-19 *PC1-3 plotted for the viseme *r*ope by gender*

males (blue) females (red)
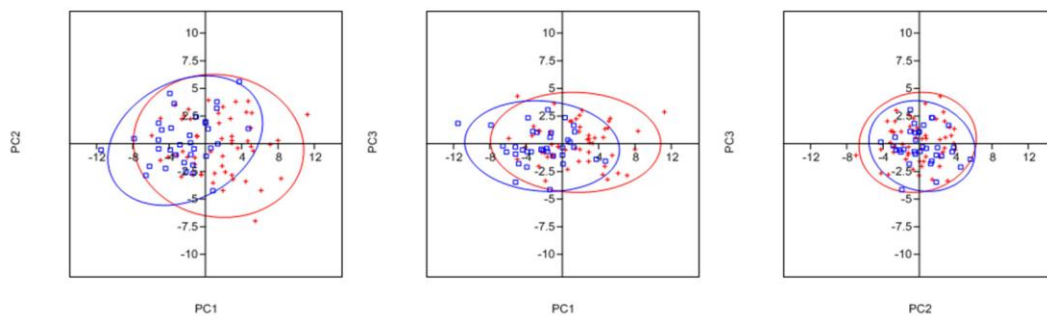
Figure 8-20 *PC1-3 plotted for the viseme *b*aby by gender*

males (blue) females (red)
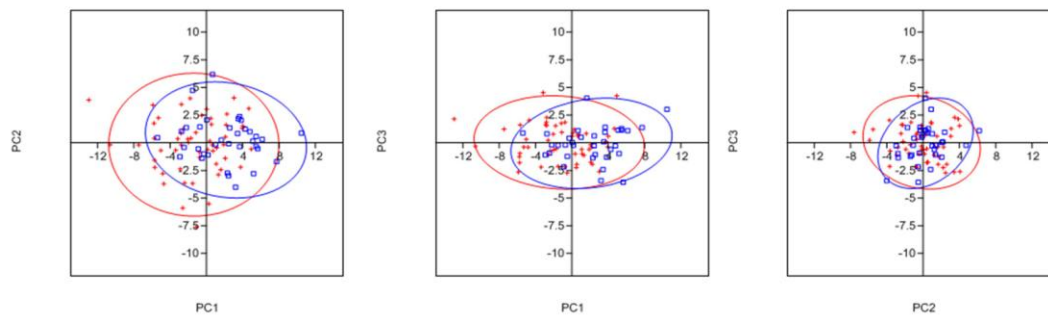
Figure 8-21 *PC1-3 plotted for the viseme ba<u>b</u>y by gender*

males (blue) females (red)
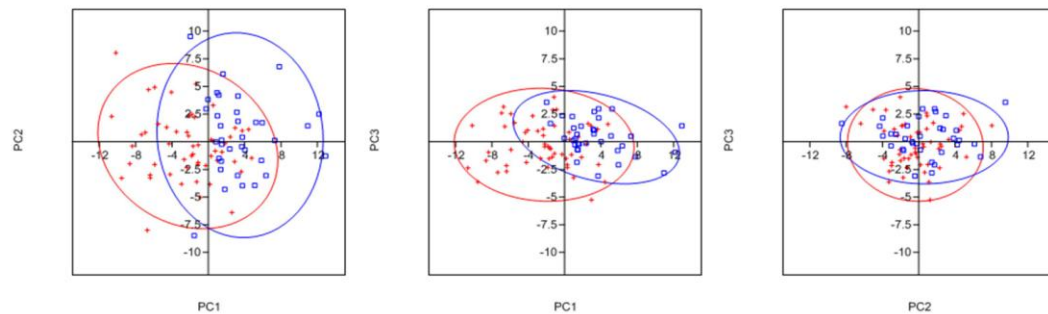
Figure 8-22 *PC1-3 plotted for the viseme b<u>o</u>b by gender*

males (blue) females (red)

Comparable PC scores were seen between males and females for all other visemes. The 95% CI for females was generally wider than for males for all visemes suggesting slightly more variation in lip movement for the female group.

8.3.5.   *Canonical variate analysis*

Six CVs were revealed through CVA with the first explaining 72.8% of the variance, whereas the second explained only 17.2% (Table 8-14). In total, the first two CVs accounted for 90% of the variance with CV3-6 explaining the remaining 10%.

Table 8-14   *Summary of canonical variates*

CV	Eigenvalue	Variance (%)	Cumulative (%)
1	2.21	72.8	72.8
2	0.52	17.2	90.1
3	0.19	6.5	96.6
4	0.09	3.0	99.6
5	0.00	0.2	99.9
6	0.00	0.1	100.0

The significance of the model was tested as a whole, following which each variate was removed in turn to see whether the variates that remained were considered significant (Table 8-15). This showed that in combination, the first four CVs significantly discriminated the seven visemes.

Table 8-15   *Significance testing of canonical variates*

Test of Variates	Significance
1 through 6	0.00
2 through 6	0.00
3 through 6	0.00
4 through 6	0.00
5 through 6	0.98
6	0.95

Despite significance testing highlighting the first four CVs, visualisation of the CV scores for each viseme showed that only CV1 and CV2 clearly differentiated the seven lip shapes (Figure 8-23).
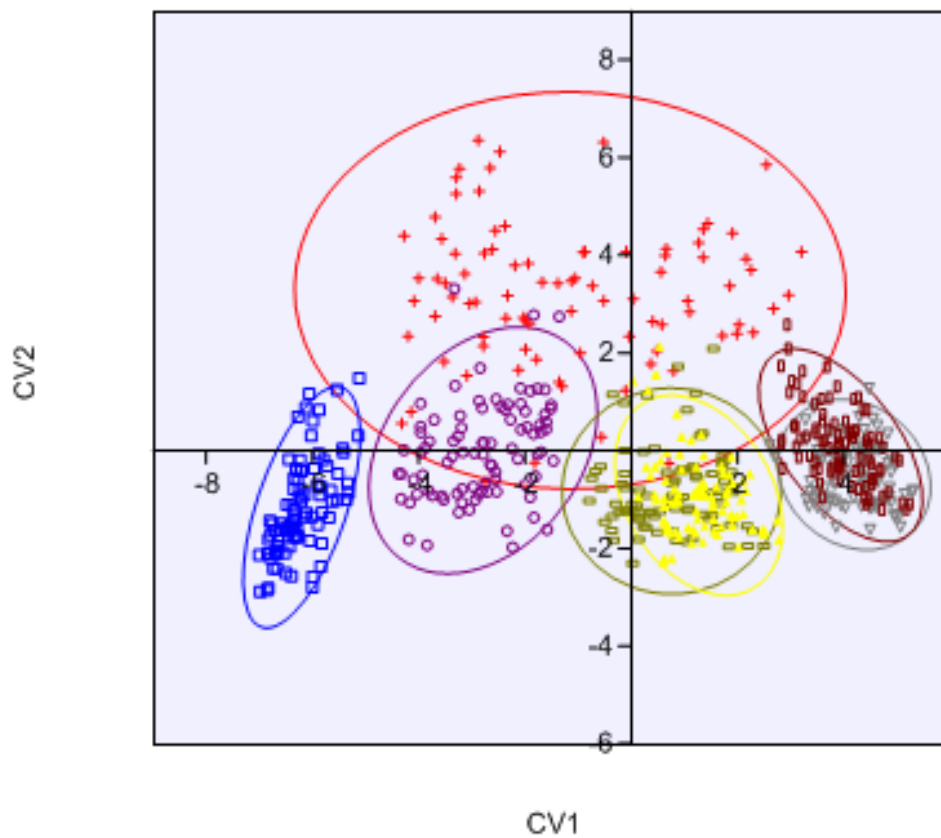
Figure 8-23 *Scatterplot of CV Scores for CV1-2 labelled by viseme*

Red cross = resting   Blue square = p̲uppy
Purple circle = pupp̲y   Yellow triangle = r̲ope
Brown vertical bar = b̲aby   Grey triangle = bab̲y
Khaki line = bo̲b

Encompassed by 95% CIs, the plot shows a wide variation in resting lip shape (red). A shift along the Y-axis (CV2) marked the change from resting lip shape to word articulation. Progression along the X-axis (CV1) differentiated the utterances. P̲uppy (blue) appeared to be the most distinct viseme whereas overlap of the 95% CIs suggested that r̲ope (brown) and bo̲b (yellow), and b̲aby (green) and bab̲y (grey) were extremely similar in peak lip shape.

Table 8-16   *Correlations between landmark coordinates and CVs*

Coordinate	CV					
	1	**2**	**3**	**4**	**5**	**6**
ls Y	0.76	-0.23	-0.28	-0.01	-0.03	-0.21
li Y	-0.70	0.01	0.51	0.30	0.10	0.16
cphL Y	0.67	-0.38	-0.17	-0.26	-0.02	-0.09
cphR Y	0.66	-0.35	-0.22	-0.12	-0.27	0.11
chR X	0.51	-0.50	0.08	-0.24	-0.35	-0.40
chL X	-0.50	0.46	0.38	0.30	-0.12	-0.25
chL Y	-0.25	0.52	-0.35	-0.02	0.21	0.05
ls Z	0.34	-0.50	0.16	0.04	-0.24	0.01
chR Y	-0.29	0.47	-0.15	-0.28	-0.11	-0.23
cphR Z	0.41	-0.45	0.41	0.11	0.21	0.23
li Z	-0.27	0.35	0.32	-0.30	0.04	0.12
cphR X	0.10	0.28	-0.04	-0.10	0.04	0.08
chL Z	-0.11	0.14	-0.78	0.22	0.27	0.08
cphL Z	0.26	-0.30	0.53	-0.24	-0.13	0.17
li X	-0.04	-0.02	-0.35	-0.48	-0.01	0.45
ls X	0.06	-0.04	0.27	-0.06	0.66	0.41
chR Z	-0.16	.016	-0.32	0.22	-0.21	-0.50
cphL X	-0.03	-0.26	-0.24	0.26	0.24	0.29

The pooled within-groups correlations between the landmark coordinates and CVs are shown in Table 8-16. Coordinates are ordered by absolute size of their correlation within a CV. The largest absolute correlations between each coordinate and the first four CVs are highlighted.

CV2, which explained 17.2% of the variance in the sample and was the variate that differentiated resting lip shape from the utterances correlated highly with midline lip protrusion (*ls* Z and *li* Z) and vertical opening at the commissures (*chL* Y and *chR* Y).

CV1, which explained up to 72.8% of the variance, differentiated the visemes correlated with vertical mouth opening (*ls* Y, *li* Y, *cphL* Y and *cphR* Y) and mouth width (*chL* X and *chR* X).

8.4. Discussion

In this section of the thesis, a sample of 115 *average* subjects was used to model ordinary lip movement for different visemes. When reviewing the literature for databases that have used 3D data to construct profiles of average facial movement, a benchmark of approximately 100 subjects is often quoted (Savran et al., 2008, Gupta et al., 2010). In this respect, the numbers of participants recruited to this part of the study would seem acceptable.

The 115 subjects were asked to say four words, which were shown in the previous chapter to be highly reproducible. Each of the words lasted up to one second in recording time and therefore at 48 fps, the total number of 3D facial shells obtained in the dataset was approximately 22,080. It was originally anticipated that automatic 3D facial shell registration and landmark interpolation would allow the entire dataset to be landmarked for analysis. However, the variable mean error associated with the automatic registration tool tested in Chapter 6, precluded manual landmark placement. In this respect, the time required to manually process several thousand 3D images was considered to outweigh the extra information gained from landmarking all the facial shells. Therefore only the maximum displacement vectors for each word were used to create the average templates. Using the frames of maximum displacement also enabled temporal variations between the subjects articulating the words to be accounted for. Clearly the choice of maximal frame could influence the outcome of the results and the reliability of choosing this frame was not investigated. However, as the scanning system functions at a high frame capture rate (48 fps) recommended for use in speech capture (Qifeng and Alwan, 2000), a one-frame discrepancy is unlikely to skew the results although this remains an area for further investigation.

Integration of GPA and PCA was used to analyse lip movement. GPA ensured that all coordinates were aligned in the same 3D space, which compensated for head movements during articulation. Other studies have used head frames to introduce immobile reference points to compensate for head movements (Mishima et al., 2006), but using GPA eliminates this

requirement. The other advantage of the Geometric Morphometric (GMM) approach is that the coordinates of the landmarks are statistically analysed rather than inter-landmark distances. This allows the results of the statistical analyses to be visualised as deformations of landmark configurations thereby increasing the sensitivity as more shape information is analysed (Hennessy and Moss, 2001).

It was shown that by using PCA ordinary lip movements could be characterised. Five PCs described up to 92% of the variance in lip movement for the words. Although PC1 generally described the largest movements in magnitude, i.e. vertical lip opening and mouth width, it was also possible to isolate more subtle movements, e.g. upper elevated vermilion border width. Symmetry of articulation could also be assessed as bilateral landmarks grouped together throughout the analysis. In addition, coordinates *ls* X and *li* X were either consistently in the last component for all visemes or not loaded significantly on higher order PCs implying midline symmetrical movement. Therefore not only can symmetry can be assessed in terms of lip shape using PCA, but when cross-referenced to the equivalent standard deviation envelope plot, lip position/placement can also be quantified, e.g. to see whether the lips are drawn to the stronger side of the face in patients with a unilateral facial nerve palsy.

Although there was a difference in resting lip shape between the genders with males having a larger, less protrusive but more variable lip shape than females, there were no significant differences in lip shape during movement. Plots of the PC scores labelled by gender for all the visemes showed overlap of 95% CI ellipses suggesting that lip shape at maximum articulation was largely similar for males and females. This finding is in keeping with other published literature that has shown the effect of age and sex on facial movement is limited within a normal sample (Sforza et al., 2010, Neely et al., 2010b).

CVA was similar to the results from PCA in that vertical mouth opening, mouth width and lip protrusion were the dominant movements highlighted. However,

CVA was also used to evaluate the between-visemes lip shape differences. There was almost complete overlap of the data from the visemes <u>r</u>ope and b<u>o</u>b, and <u>b</u>aby and bab<u>y</u> implying that visemes were very similar in shape. In terms of the diagnostic information gained, it would seem appropriate therefore to eliminate one of the similar visemes from each pair. This would also allow for a more manageable dataset for analysis. Taking into consideration the findings from Chapter 7, the most reproducible visemes (i.e. <u>r</u>ope and <u>b</u>aby) could be retained.

Although the data generated in this part of the study could be regarded as specific to the geographical area and language it could also act as a template to compare lip shape/movement from a different population. Much like the overlap of 95% CIs for certain visemes implying a similar lip shape, one would assume that there would be overlap for an average population. However if data from a group of subjects with potentially abnormal lip movement was included into the data presented in Figure 8-23, one would hope to reveal PCs/CVs that isolated the disorder and also identified what particular movement was affected. Indeed the following chapter aims to address this.

8.5.　Conclusion

Five principal components describe the majority of lip shape during articulation of the words used in the part of the study. The main lip movements comprised of mouth opening, mouth width and midline protrusion although more subtle movements such as upper lateral commissure width could also be detected. An average reference has been created for lip shape during movement through two canonical variates; one of which distinguishes resting lip shape from the four utterances and the other discriminates between the four utterances.

Chapter 9

Detection of abnormal movement

9.1. Introduction

In the previous chapter lip movement was captured from a sample population to create an average model. In the following sections we describe addition of patient data to the model in an attempt to investigate whether abnormalities in movement can be identified statistically.

9.2. Methods

This part of the study was designed to statistically identify lip movement from a patient group (PG) from that of the average data obtained in Chapter 8 termed the control group (CG). The PG consisted of 30 consecutive subjects (19 male, 11 female) who presented at the Department of Orthodontics, University Dental Hospital, Cardiff with a Class 3 skeletal pattern requiring bimaxillary surgery to correct the dento-facial relationships. Additional inclusion criteria for the PG was a reverse overjet prior to surgery of 5mm or greater – in this respect a combined maxillary-mandibular movement of 7mm (taking into account a 2mm overjet) would allow the effect of surgery on lip movement to potentially materialise. The pre-treatment and pre-surgical cephalometric variables for the PG compared to Caucasian norms (Mills, 1987) are listed in Table 9-1.

Table 9-1   *Cephalometric variables of the PG compared to Caucasian norms*

Variable	PG (pre-treatment)		PG (pre-surgical)		Normal	
	Mean	SD	Mean	SD	Mean	SD
SNA	76°	3.4	75°	3.9	81°	3
SNB	85°	3.5	87°	3.8	79°	3
ANB	-9°	3.5	-12°	3.4	3°	2
Mx-Mnd planes angle	32°	5.2	32°	4.5	27°	4
Upper incisor to Mx plane	119°	6.1	110°	3.2	109°	6
Lower incisor to Mnd plane	78°	5.3	84°	4.2	93°	6

As the CG data had already been obtained, data from the PG was captured in a similar manner with all 30 subjects asked to say four words (puppy, rope, baby,

bob) in a normal relaxed manner and scanned using the 3dMDFace™ Dynamic System at 48 frames per second under standardised conditions. The PG was scanned twice; once prior to orthognathic surgery following orthodontic decompensation (PG_{pre}), and a second time 6 months post-surgery (PG_{post}). The surgical procedures were all undertaken by the same clinician and carried out using rigid-internal fixation.

9.2.1.   *Imaging processing*

In a similar manner to previous sections, six anthropometric lip landmarks were manually identified on the 3D facial shell of maximum lip displacement for each viseme. This included: puppy, puppy, rope, baby, baby and bob. Following their identification, the *x*, *y*, and *z* coordinates of each landmark were recorded (18 coordinates in total).

Generalised Procrustes Analysis (GPA) was applied to align/register the sets of 18 lip landmarks by removing translation and rotation. Scaling was also incorporated here to remove the effect of size on lip movement. Note that scaling was not incorporated in Section 8.2.2 to assess differences in size and shape within the control group (e.g. between males and females). In this part of the study it was postulated that the PG may have size differences in lip shape when compared to the CG, particularly excessive lower lip protrusion. Therefore this effect was scaled so as to remove the influence of lip size on lip shape during movement.

9.2.2.   *Statistical analysis*

To determine whether lip movement between the PG and CG could be statistically distinguished discriminant analysis (DA) was carried out. The coordinates of the lip landmarks for each viseme were used as predictor variables to separate individuals based on their pre-defined groups (i.e. CG, PG_{pre} and PG_{post}). The analysis accounts for the groups being independent and therefore pairwise comparisons of CG v PG_{pre} and CG v PG_{post} were facilitated.

DA creates discriminant scores using a linear discriminant equation such that the ratio of the between groups sums of squares to the within groups sums of squares is as large as possible, i.e. new variables are defined so that they separate the groups as far apart as possible. How well the model performs is reported in terms of the classification efficiency, that is, how many cases would be correctly assigned to their groups using these new variables. To reduce bias in the classification, jack-knifed cross-validation was specified (Lachenbruch, 1967). Here, each subject is classified by the DA variables derived from all cases other than that case. This is repeated for all cases and gives a more realistic estimate of the ability of the variables to separate the groups. Separate discriminant functions were generated for each viseme (i.e. puppy, puppy, rope, baby, baby and bob) and in each discriminant analysis all predictor variables were entered simultaneously for the particular viseme.

Canonical Variate Analysis (CVA) was then applied as an extension to DA. Using the proposed *average* outcome measure presented in Section 8.3.5, the landmark coordinates for the visemes of the PG were added. As the model showed that the visemes rope and bob, and baby and baby were essentially the same, only the most reproducible visemes (i.e. rope and baby) were retained. CVA was carried out for the PG_{pre} and PG_{post} separately.

To report on whether lip shape for each viseme was statistically significant between the groups, Wilk's Lambda (Λ) and the squared Mahalanobis distance (D^2) was used. The proportion of unexplained variability in the analysis is provided by Λ with an associated F test; smaller Λ values indicate greater differences between the two groups in the analysis. The threshold for Wilk's Lambda was set at $p < 0.05$. The D^2 is the distance between the centroids (means) of two groups in multidimensional space as defined by the predictor variables. In the present context, the distance provides a quantitative estimate of how divergent lip shape of the PG is from the control group. For three-dimensional data, a sample size of up to 200, and a significance level of 5%, the critical value for D^2 is quoted as 18.42 (Barnett and Lewis, 1984).

Therefore, a D^2 greater than 18.42 from a given centroid would indicate a less than 5% chance of an individual belonging to that particular group.

Finally, to investigate the statistical findings from the DA/CVA in terms of the actual patterns of lip shape that differed between the CG and PG, PCA was carried out again with the PG data added to the original CG model. Scatter plots of the PC scores for individuals labelled by group were cross-referenced to PCA tables with associated coordinate loadings to isolate differences (if any) between the CG and PG.

DA/CVA and PCA were carried out using carried out using MorphoJ (Klingenberg, 2011).

9.2.3.   *Data preparation*

To ensure the assumptions of DA had not been violated all variables were checked for normality/skewness, univariate/multivariate outliers and equality of variance-co-variance matrices.

Normal distribution of data for each of the 18 landmarks for each viseme separated by group was visually inspected using histograms. Due to the sample size being under 2000, the Shapiro-Wilk statistic was used to assign a p value to the distribution (Park, 2008). A non-significant result ($p > 0.05$) indicated normality.

To identify univariate outliers, boxplots for each of the 18 landmarks for each viseme separated by group were constructed. Cases outside 1.5 box-lengths (box-length = interquartile range) from the interquartile range were highlighted. The 5% trimmed mean was compared to the actual mean to assess the effect of the outliers. If these values were considered to be significantly different, the outlying cases were removed from the analysis.

Multivariate outliers were identified using the Mahalanobis distance in a similar manner to Section 8.2.3. Here, Mahalanobis distances were calculated for each subject, grouped by viseme. To reiterate the description of the method, the Mahalanobis distance is evaluated as Chi Square (χ^2) with degrees of freedom equal to the number of variables, in this case 18. Therefore cases with a Mahalanobis distance greater than $\chi^2(18) = 42.3$ would be regarded as multivariate outliers (Ahrens, 1958). Any cases that were identified as such were removed from the analysis.

Homogeneity of variance-co-variance matrices was assessed through use of Box's M Test comparing CG v PG_{pre} and CG v $PG_{post.}$ A significance of $p < 0.001$ was set. Classification on separate co-variance matrices was stipulated if heterogeneity was found.

Tests for normality were carried out using SPSS 20.0.0 (SPSS Inc., Chicago, IL, USA).

9.3. Results

Appendix C contains the results for the data preparation. The initial section confirms the Shapiro-Wilk statistic for normality of the data distribution. This was carried out for each coordinate, for each group and for each viseme giving a total of 324 analyses. Of these analyses, 15 tested statistically significant implying a deviation from normal. The option at this stage was to consider transformation of all the variables. It was decided to assess the influence of other preparatory tests prior to making this decision.

The effect of univariate outliers was extremely low as shown by the similar values of the 5% trimmed mean compared to the actual mean for all the variables for all visemes. Therefore no cases were removed from the analysis.

Multivariate outliers were assessed using the Mahalanobis distance. Figure C-1 shows boxplots of the Mahalanobis distances for each case by viseme. The

threshold for a multivariate outlier is marked as 42.3. All cases above this threshold were removed from the analysis. This totalled nine cases from the CG for the viseme rope and one case from the CG for the viseme bob.

Table C-13 shows the p values for Box's M Test to assess the homogeneity of variance-co-variance matrices. There were no violations at a threshold of $p < 0.001$.

Due to the multivariate data results of the preparatory tests being either fulfilled or accounted for it was decided against transformation of the data.

9.3.1.   Discriminant analysis

The percentage of actual and predicted group membership using DA for each viseme is presented in Table 9-2. The accuracy of the classification rate using jack-knifed cross-validation is included. The visemes puppy, rope, baby, baby and bob showed clear discrimination in the analysis of CG v PG_{pre} as predicted membership of subjects in PG_{pre} was found to be consistently above 75%. In addition, the classification accuracy for these visemes was relatively high at 90% and above - the visemes puppy and rope showed 100% accuracy in classification. When predicting group membership of CG v PG_{post} there was crossover of the PG_{post} into the CG. This reflected in the classification accuracy of the model with scores below 76.1% for all visemes. These findings imply differences in movement for these visemes for the PG_{pre} when compared to the CG and $PG_{post,}$ and a suggestion of a return to normality post-surgery as classification of lip movement in the PG_{post} reduces in accuracy. The DA for puppy was less clear with the model showing the least predictive accuracy of all the visemes.

Table 9-2 *DA classification accuracy (pairwise comparisons of PG against CG)*

VISEME		Predicted group %			Predicted group %	
puppy		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	81.0	19.0	**CG**	87.9	12.1
	PG$_{pre}$	76.2	23.8	**PG$_{post}$**	80.5	19.5
		65.8% of cases correctly classified			67.1 % of correctly classified	
puppy		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	100.0	0.0	**CG**	91.2	8.8
	PG$_{pre}$	0.0	100.0	**PG$_{post}$**	95.0	5.0
		100% of cases correctly classified			71.6 % of cases correctly classified	
rope		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	100.0	0.0	**CG**	85.3	14.7
	PG$_{pre}$	0.0	100.0	**PG$_{post}$**	75.0	25.0
		100% of cases correctly classified			71.6% of cases correctly classified	
baby		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	95.6	4.4	**CG**	89.7	10.3
	PG$_{pre}$	22.1	77.9	**PG$_{post}$**	70.8	29.2
		90.9% of cases correctly classified			76.1% of cases correctly classified	
baby		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	94.1	5.9	**CG**	85.3	14.7
	PG$_{pre}$	20.3	79.7	**PG$_{post}$**	90.5	9.5
		90.9% of cases correctly classified			67.4% of cases correctly classified	
bob		**CG**	**PG$_{pre}$**		**CG**	**PG$_{post}$**
Actual group %	**CG**	100.0	0.0	**CG**	87.9	13.1
	PG$_{pre}$	15.7	84.3	**PG$_{post}$**	78.9	21.1
		96.6% of cases correctly classified			74.7% of cases correctly classified	

9.3.2. *Canonical variate analysis*

CVA was used to extend the findings of the DA. Nine canonical variates (CVs) were revealed. Visualisation of the CV scores by subject showed that only CV1 and CV2 were able to discriminate between the visemes for both the PG_{pre} and PG_{post} (Table 9-3 and Figure 9-5).

Table 9-3 *Summary of canonical variates*

CV	Eigenvalue	Variance (%)	Cumulative (%)
1	2.39	69.5	69.5
2	0.65	18.9	88.4
3	0.16	4.9	93.3
4	0.11	3.2	96.6
5	0.07	2.1	98.7
6	0.01	0.6	99.3
7	0.01	0.4	99.7
8	0.00	0.2	99.9
9	0.00	0.1	100.0

As in Section 8.3.5, CV2 was the discriminator between resting lip shape and lip shape during movement for all the groups. Lip shape during movement was differentiated by CV1. Table 9-4 shows the within-groups correlations between discriminating variables and CVs with variables ordered by absolute size of correlation within a CV. Here we see that the change in resting lip shape to lip shape during speech (CV2) is correlated with the vertical position of the commissures (*chL* Y and *chR* Y) and lip protrusion (*ls* Z and *li* Z) (Table 9-4). The change in lip shapes for the words (CV1) are differentiated by degrees of lip opening (*ls* Y, *li* Y, *cphL* Y and *cphR* Y) and mouth width (*chL* X and *chR* X) (Table 9-4).

Table 9-4 *Correlations between landmark coordinates and CVs*

(CV1 and CV2 highlighted)

Landmark	CV								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
ls Y	0.84	-0.23	-0.15	0.07	0.00	-0.07	-0.00	0.13	-0.12
li Y	-0.74	-0.10	0.25	0.29	-0.07	0.22	0.20	0.32	-0.01
cphL Y	0.70	-0.35	0.06	0.03	0.11	-0.26	0.01	-0.26	0.14
cphR Y	0.70	-0.33	-0.03	0.01	0.01	-0.29	-0.13	-0.34	0.31
chR X	0.55	-0.45	0.18	-0.34	0.19	0.24	-0.30	-0.26	0.07
chL X	-0.54	0.43	0.17	0.03	-0.12	0.52	-0.09	0.24	0.24
chL Y	-0.27	0.61	-0.41	-0.32	-0.00	-0.02	0.07	-0.31	-0.27
chR Y	-0.30	0.58	-0.05	-0.48	0.04	0.14	-0.40	0.02	-0.04
ls Z	0.39	-0.57	0.31	0.16	-0.30	-0.17	0.27	-0.11	0.21
cphR Z	0.40	-0.50	0.18	0.35	0.46	-0.02	0.26	0.18	0.09
li Z	-0.28	0.41	0.26	-0.36	0.18	0.05	0.05	-0.09	-0.24
cphL Z	0.29	-0.36	0.34	0.20	0.09	-0.07	0.09	0.10	-0.16
chL Z	-0.11	0.13	-0.70	0.16	-0.18	-0.29	-0.04	0.01	-0.00
cphR X	0.07	0.13	0.04	0.59	0.35	-0.22	0.01	-0.43	-0.15
cphL X	-0.01	-0.17	-0.32	-0.29	-0.42	-0.07	0.31	0.21	-0.12
li X	-0.06	0.00	0.07	0.04	-0.01	-0.73	-0.36	0.45	-0.02
ls X	0.04	0.00	-0.32	-0.04	0.11	-0.59	0.65	-0.05	-0.14
chR Z	-0.14	0.17	-0.33	-0.09	-0.19	0.39	-0.42	0.02	0.24

For the ease of visualisation, scatterplots of the CVA scores are first shown by individual viseme (Figure 9-1, Figure 9-2, Figure 9-3 and Figure 9-4) following which a composite scatterplot is presented of all visemes plotted together (Figure 9-5).
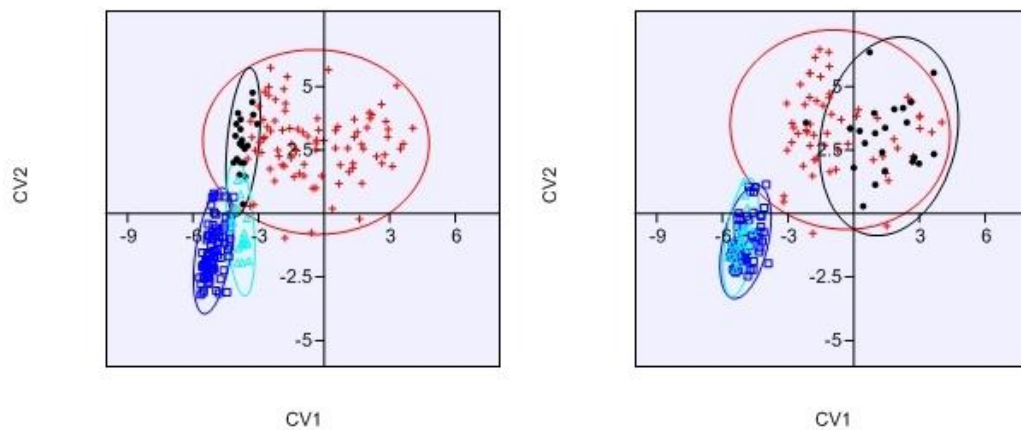
Figure 9-1 *Scatterplot of CV Scores for CV1-2 for puppy*
(PG$_{pre}$ – left   PG$_{post}$ – right)

CG	red cross (resting)	blue square (puppy)
PG	black dot (resting)	cyan triangle (puppy)
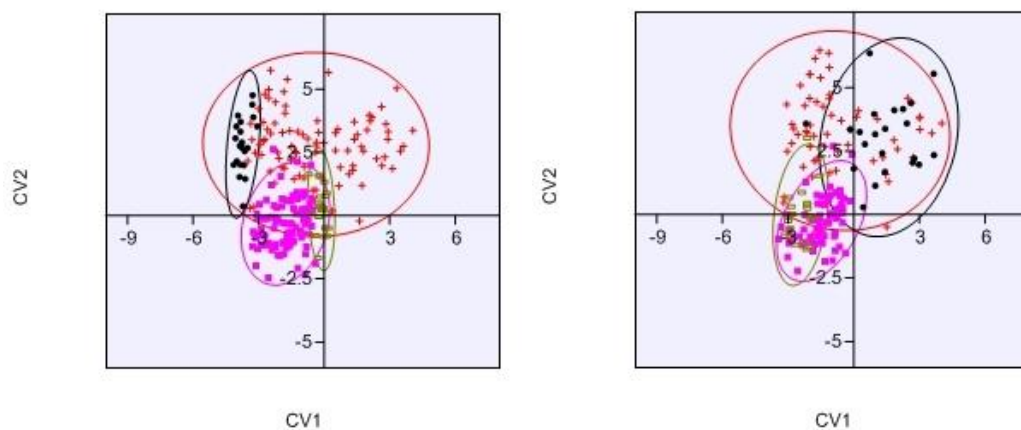
Figure 9-2 *Scatterplot of CV Scores for CV1-2 for puppy*
(PG$_{pre}$ – left   PG$_{post}$ – right)

CG	red cross (resting)	pink block (puppy)
PG	black dot (resting)	khaki line (puppy)
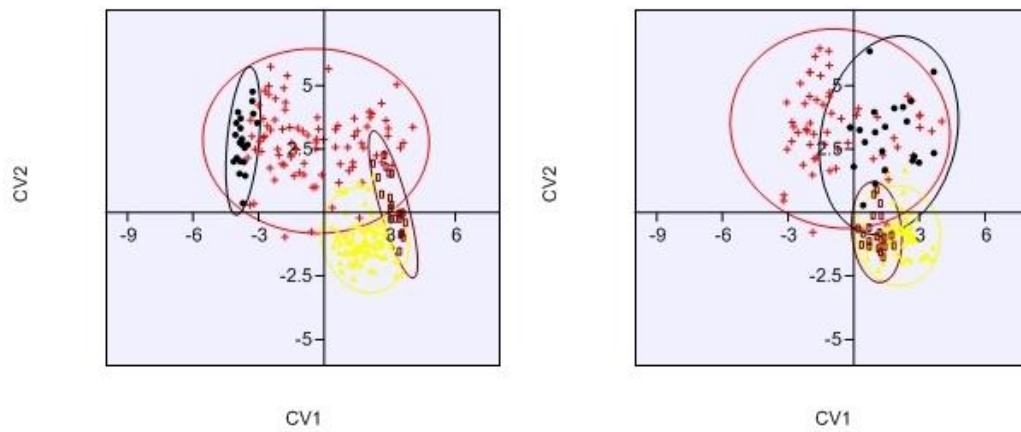
Figure 9-3 *Scatterplot of CV Scores for CV1-2 for rope
(PG$_{pre}$ – left   PG$_{post}$ – right)*

CG	red cross (resting)	yellow triangle (rope)
PG	black dot (resting)	brown vertical bar (rope)
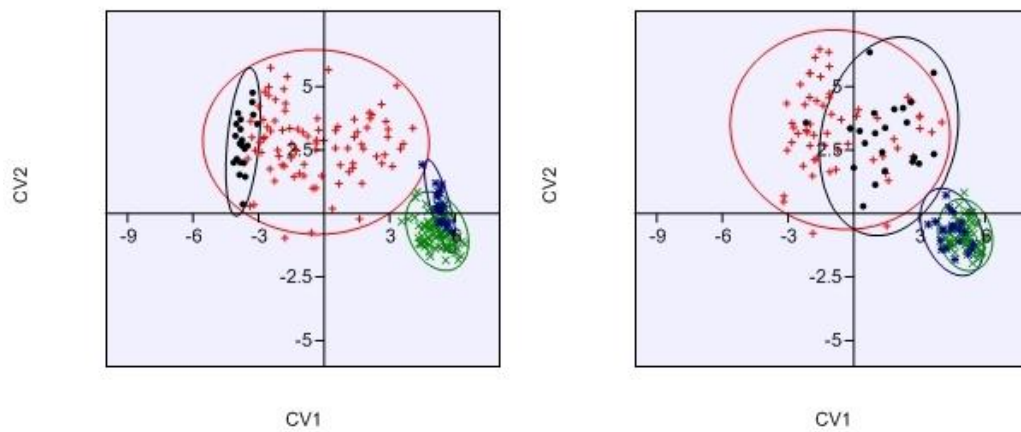
Figure 9-4 *Scatterplot of CV Scores for CV1-2 for baby
(PG$_{pre}$ – left   PG$_{post}$ – right)*

CG	red cross (resting)	green cross (baby)
PG	black dot (resting)	dark blue star (baby)
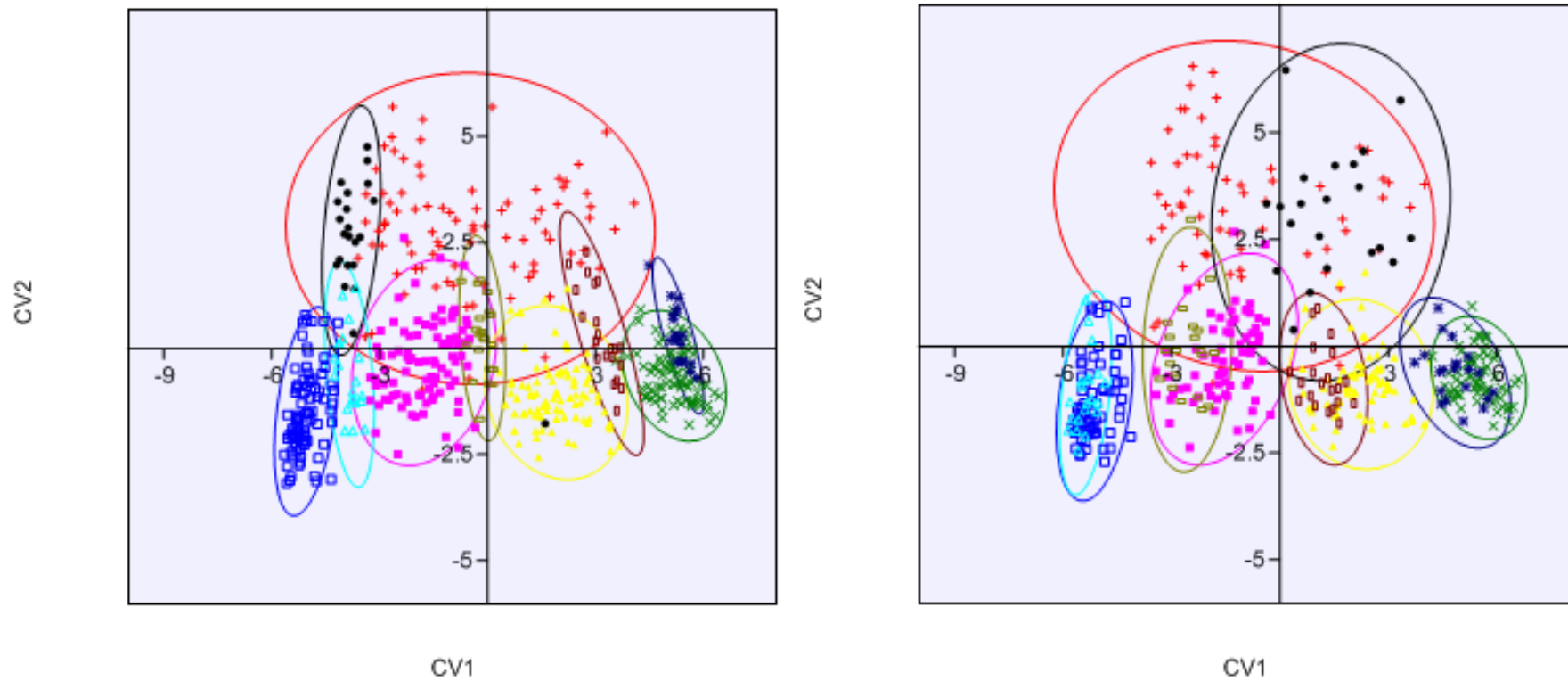
Figure 9-5 *Scatterplot of CV Scores for CV1-2 labelled by viseme*
(PG$_{pre}$ – left   PG$_{post}$ – right)

| CG | red cross (resting) | blue square (puppy) | pink block (puppy) | yellow triangle (rope) | green cross (baby) |
| PG | black dot (resting) | cyan triangle (puppy) | khaki line (puppy) | brown vertical bar (rope) | dark blue star (baby) |

The resting lip shape of the CG and PG$_{pre}$ was essentially similar with the 95% CI ellipse of the PG$_{pre}$ encompassed by the CG, albeit positioned at the left-hand extreme of CV1 (Figure 9-5). Lip shapes during speech for the PG$_{pre}$ appeared to lie interspersed between those of the CG along CV1 whereas for the PG$_{post}$ there was more convergence with the CG (Figure 9-5). In addition, post-surgery, the resting lip of the PG$_{post}$ changed from a relatively narrow variation to one of greater dispersion, which was in keeping with the variation seen in the CG (Figure 9-5).

The findings from the DA and CVA data were supported by Λ and D^2 (at a threshold of p < 0.05 and 18.42 respectively), which showed that there were significant differences for all the visemes between the PG$_{pre}$ and CG (Table 9-5). The largest D^2 was for the viseme baby (D^2 = 32.37). The lowest D^2 was recorded for the visemes rope and bob which registered just outside the D^2 threshold of 18.42 (Table 9-5). There were no differences between the PG$_{post}$ and CG for all the visemes using D^2 although Λ showed that a statistical difference for baby and bob remained post-surgery.

Table 9-5   *Squared Mahalanobis distances and Wilk's Lambda for CG against PG for each viseme*

Comparison	Mahalanobis distance (D^2)	Wilk's Lambda (Λ)	Sig
CG v PG$_{pre}$			
puppy	27.64	0.73	0.04
puppy	21.77	0.05	0.00
rope	19.09	0.03	0.00
baby	32.37	0.41	0.00
baby	26.77	0.37	0.00
bob	19.28	0.12	0.00
CG v PG$_{post}$			
puppy	3.02	0.85	0.45
puppy	3.37	0.85	0.37
rope	4.51	0.74	0.05
baby	4.83	0.72	0.01
baby	3.58	0.80	0.13
bob	4.05	0.75	0.03

9.3.3. Principal component analysis

Figure 9-6 shows the regression scores for PC1-3 plotted by group. The 95% CI ellipses overlap across the first three components with no obvious discrimination between the groups.
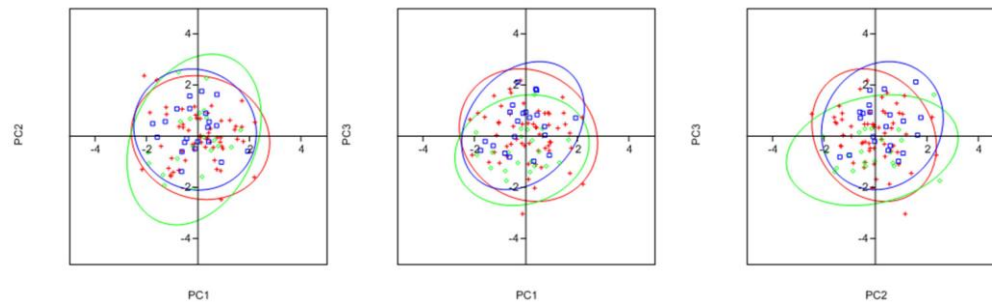
Figure 9-6   *PC1-3 plotted for puppy labelled by group*
Red (CG)   Green (PG$_{pre}$)   Blue (PG$_{post}$)

The wireframe graph in Figure 9-7 allows the changes in lip shape between the groups to be visualised. This highlights a slight difference pre- and post-surgery with the PG$_{pre}$ exhibiting a shorter and wider lip shape when compared to the CG. This appears to conform to the CG post-surgery. There is no statistically significant difference in Λ and D^2 between the pre-, and post-surgical lip shape for the viseme puppy when compared to the CG (p = 0.05 and p= 0.45 respectively).
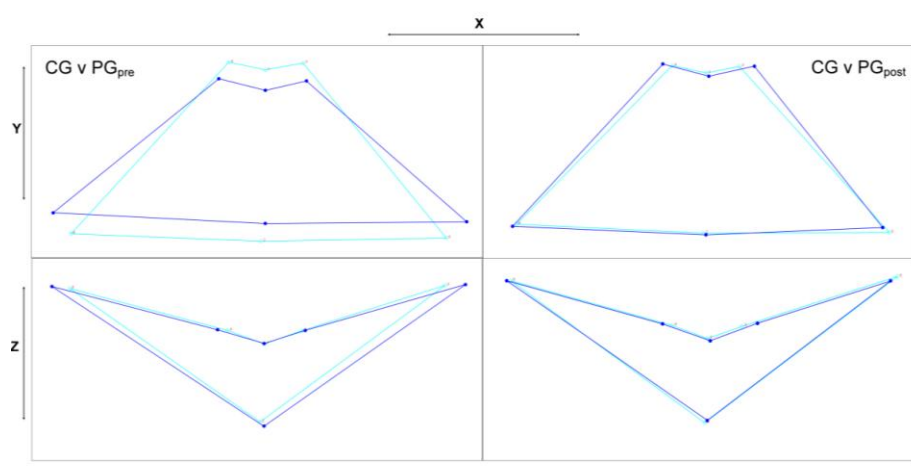
Figure 9-7   *Mean shape changes pre- and post-surgery for puppy*
CG (green)   PG (blue)

In contrast, for the viseme puppy there was evidence of discrimination of PG$_{pre}$ along PC2 which was more apparent when PC2 was plotted against PC3 (Figure 9-8). PCA was employed with the PG data added to the model to isolate which coordinates were loaded onto PC2. This highlighted coordinates *chL* X and *li* Z (Table 9-6). The wireframe graph shown in Figure 9-9 suggests asymmetric lip movement of the PG$_{pre}$ during the viseme puppy with the corners of the mouth drawn to the left-hand side during maximum articulation.
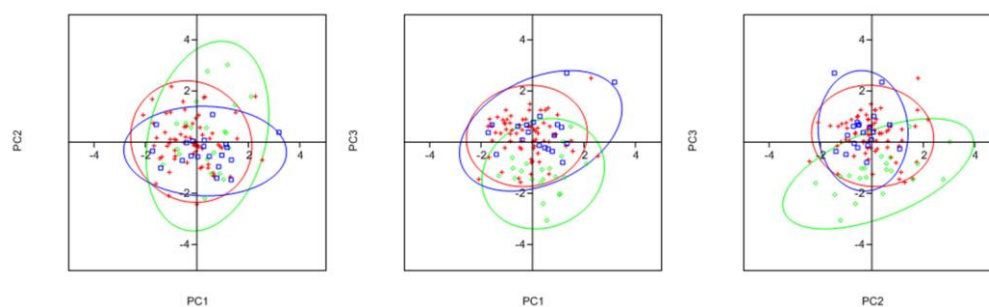
Figure 9-8   *PC1-3 plotted for puppy labelled by group*
Red (CG)   Green (PG$_{pre}$)   Blue (PG$_{post}$)

Table 9-6   *PC table with coordinate loadings for puppy*

Coordinate	PC (% variance)				
	1 (42)	2 (19)	3 (15)	4 (11)	5 (7)
li Y	-0.59		0.30		
chL X	-0.46	-0.42	-0.35		
ls Y	0.32				
cphR Y	0.28				
cphL Y	0.27				
chR X	-0.24				-0.25
li Z		-0.55	0.52	0.25	
chR Z			-0.46		
chL Z			-0.36		
chL Y				0.56	
chR Y				0.42	
cphL X				0.36	
cphR X				-0.24	0.68
ls X[a]					
ls Z[a]					
cphR Z[a]					
cphL Z[a]					
li X[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
Highlighted cells indicate discriminatory loadings

The suggestion of asymmetric movement was strengthened by virtue of *chL* X loaded across the first three components where as *chR* X only featured in PC1. As such, the lip shape difference between the PG_{pre} and CG was considered to be statistically significant (p < 0.001) whereas a return to average (and symmetrical) movement post-surgery was not (p = 0.37) (Figure 9-9).
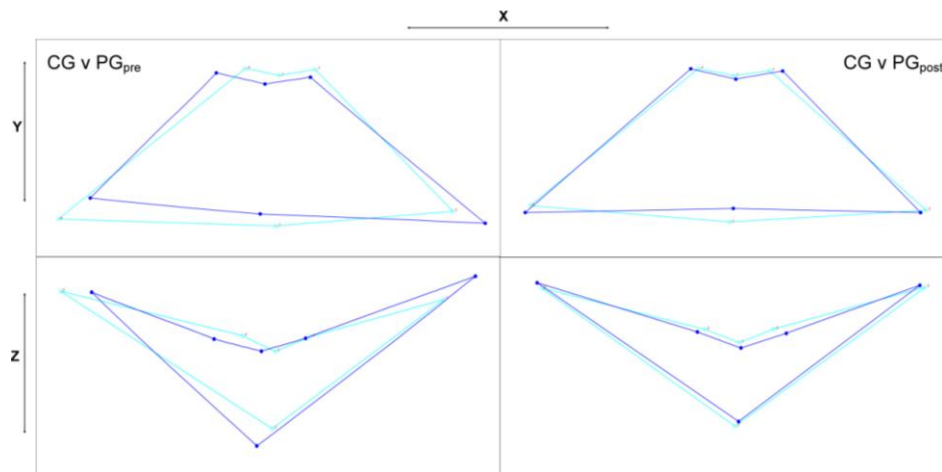
Figure 9-9 *Mean shape changes pre- and post-surgery for puppy*
CG (green)   PG (blue)
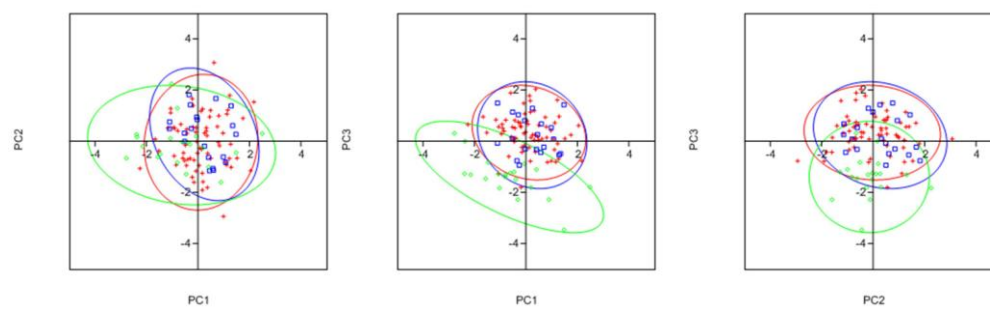
Figure 9-10 *PC1-3 plotted for rope labelled by group*
Red (CG)   Green (PG_{pre})   Blue (PG_{post})

The PC scores for viseme rope showed deviations of the PG_{pre} from the CG
and PG_{post} in respect to PC1 and PC3. The PCA table showed that this was
predominantly associated with vertical and lateral elements of the midline and
para-midline in addition to a protrusive component in the lower lip (Table 9-7).

Table 9-7   *PC table with coordinate loadings for rope*

Coordinate	PC (% variance)				
	1 (49)	2 (15)	3 (10)	4 (9)	5 (7)
li Y	-0.52	-0.42			
chL X	-0.49				
ls Y	0.30				
cphR Y	0.28				
cphL Y	0.27				
chR X	-0.21				-0.23
li Z		-0.48	0.50	0.43	
chR Y			-0.46		
chL Y			-0.36		
chL Z				0.48	
chR Z				0.47	
cphL X				0.26	
cphR X				-0.24	
ls X[a]					-0.21
ls Z[a]					
cphR Z[a]					
cphL Z[a]					
li X[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
Highlighted cells indicate discriminatory loadings

The shape differences for the viseme rope can be visualised in the wireframe graph which shows that lip shape for the PG_{pre} is limited in the vertical component when compared to the CG but more evident in the lateral and protrusive (Figure 9-11). This essentially describes a shorter, wider lip shape for the PG_{pre} for this viseme, which is statistically significantly different to the CG ($p < 0.001$). Post-surgery, the lip shape returns to what is more comparable to average ($p = 0.05$).
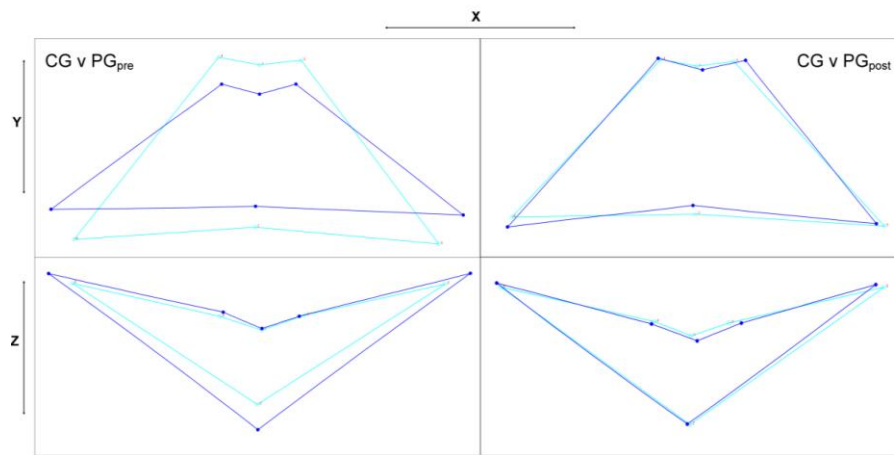
Figure 9-11 *Mean shape changes pre- and post-surgery for ₛope*
CG (green)   PG (blue)

Figure 9-12 shows the PC scores plotted for the groups for the viseme ᵦaby.
Although this did not highlight any one particular PC that discriminated the
groups the mean lip shape changes showed that the PG_{pre} once again
exhibited a shorter and slightly wider lip shape during articulation when
compared to the CG (Figure 9-13).
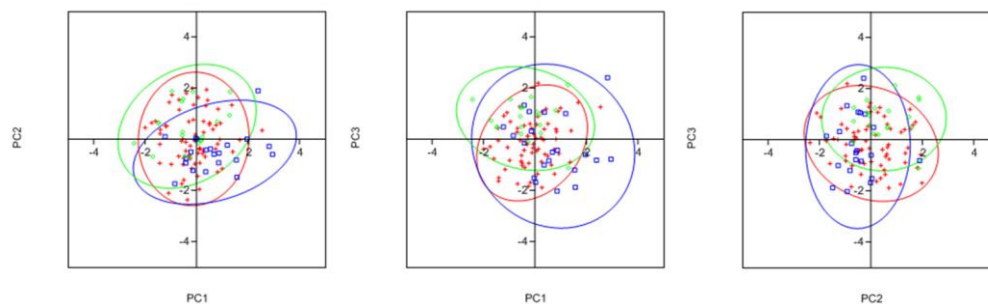
Figure 9-12 *PC1-3 plotted for ᵦaby labelled by group*
Red (CG)   Green (PG_{pre})   Blue (PG_{post})

Despite the mean lip shape of the PG_{post} being more similar to the CG, the
statistics inferred that both the pre- and post-surgical lip shapes were
significantly different to the CG ($p < 0.001$ and $p = 0.01$ respectively).
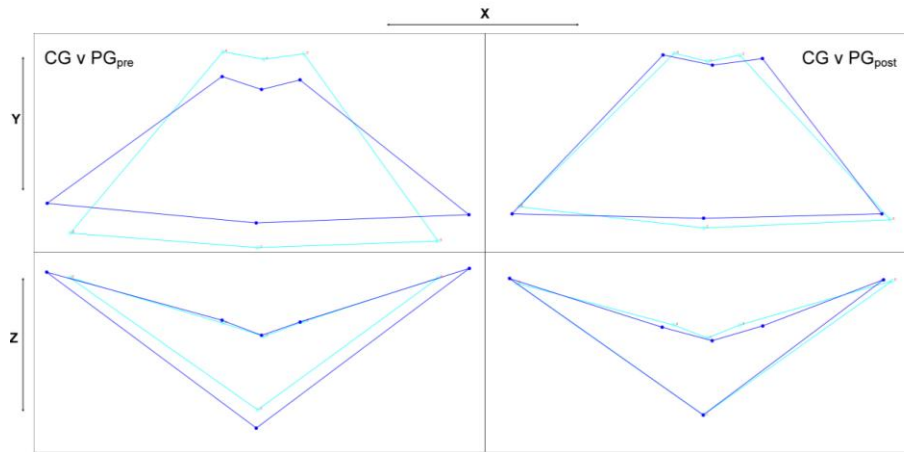
Figure 9-13 *Mean shape changes pre- and post-surgery for baby CG (green)   PG (blue)*

For the viseme baby the PC scores highlighted PC4 as a discriminatory component (Figure 9-14). The coordinates loaded onto PC4 are shown in Table 9-8. There were several coordinates loaded onto PC4, which in a similar manner to the visemes rope and baby described midline and para-midline vertical lip height, commissure width and lower lip protrusion. The mean shape changes can be seen in Figure 9-15.
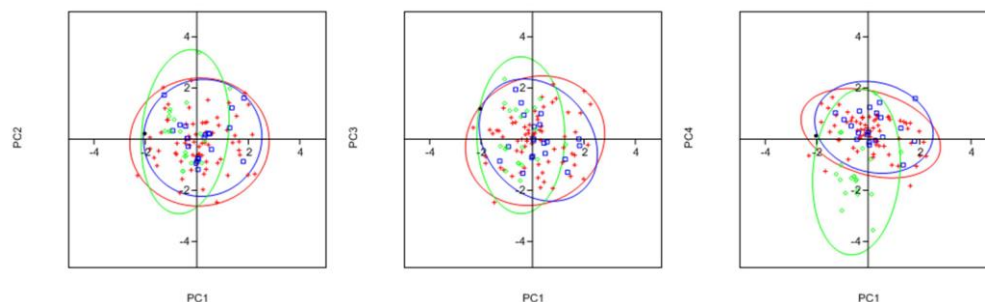
Figure 9-14 *PC1-4 plotted for baby labelled by group*
Red (CG)   Green (PG_{pre})   Blue (PG_{post})

Table 9-8   *PC table with coordinate loadings for baby*

Coordinate	PC (% variance)				
	1 (40)	2 (20)	3 (15)	4 (11)	5 (9)
li Y	-0.55	-0.47		-0.33	-0.23
chL X	-0.38			-0.52	
ls Y	0.34			-0.32	
cphR Y	0.32			-0.23	
cphL Y	0.31			-0.26	
chR X	0.29		0.32	-0.35	-0.32
li Z		0.52	0.54	0.27	
chR Y		0.42		-0.35	
chL Y		0.38		0.27	0.23
cphL X			-0.47	0.29	-0.49
chL Z			-0.34		
chR Z			-0.30		
cphR X			0.25	0.23	0.64
ls X				0.21	
ls Z[a]					
cphR Z[a]					
cphL Z[a]					
li X[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
Highlighted cells indicate discriminatory loadings

Following a similar pattern, the PG_{pre} favoured a shorter lip height but wider lip width when compared to the CG (p < 0.001) – these changes normalised post-surgery (p = 0.17).
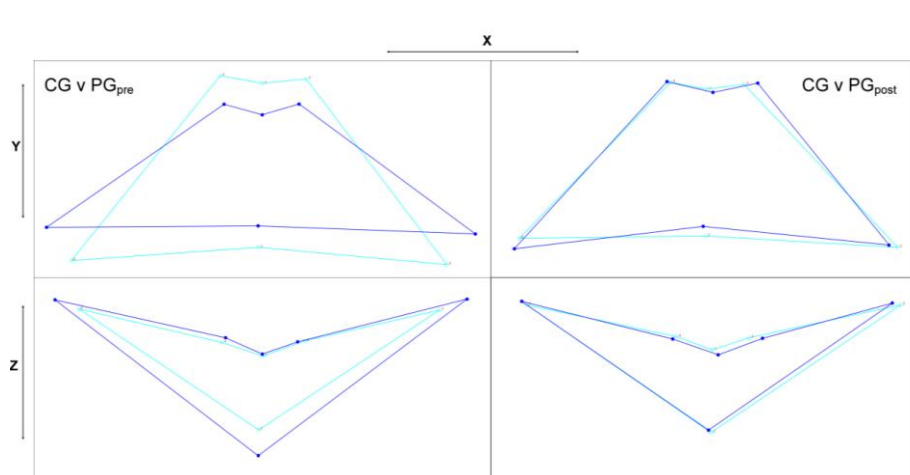
Figure 9-15   *Mean shape changes pre- and post-surgery for baby*
CG (green)   PG (blue)

Scatter plots of the PC scores from the CG and PG for the viseme b<u>o</u>b showed discrimination of the PG_{pre} along PC2 (Figure 9-16). As with the previous plots, the scores from PG_{post} had overlap within the 95% CI ellipse of the CG, which was suggestive of similar lip movement post-surgery.
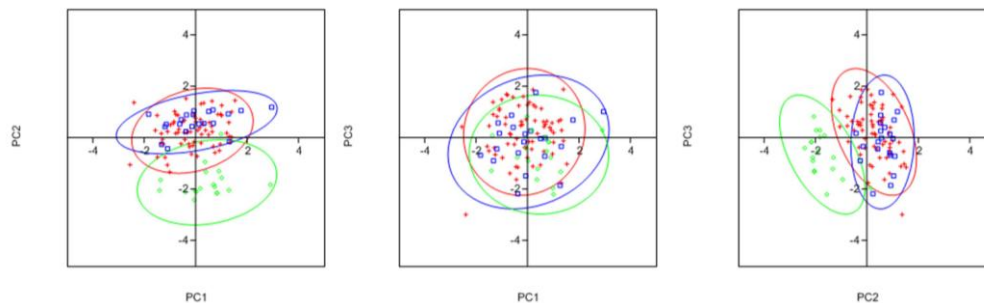
Figure 9-16 *PC1-3 plotted for b<u>o</u>b labelled by group*
Red (CG)   *Green (PG_{pre})*   *Blue (PG_{post})*

The coordinates loaded onto PC2 are highlighted in Table 9-9. This suggests vertical and protrusive differences with *li* Y, *chL* Y, *chR* Y, *cphL* Z and *cphR* Z isolated.

Table 9-9   *PC table with coordinate loadings for the viseme b<u>o</u>b*

Coordinate	PC (% variance)				
	1 (52)	2 (20)	3 (8)	4 (6)	5 (6)
li Y	-0.47	0.48			-0.37
chL X	-0.43		0.23	0.30	
chR X	0.38			0.52	
cphL Y	0.34				
cphR Y	0.32				
ls Y	0.28				
chR Y	-0.24	-0.35		0.25	0.27
chL Y	-0.22	-0.44		-0.31	
li Z		-0.54		0.27	-0.47
cphR Z		0.21			
cphL Z		0.20			
cphR X			0.53	-0.42	
ls Z			0.21		
chL Z				-0.33	0.34
chR Z					0.43
ls X				-0.22	
cphL X[a]					
li X[a]					

[a]   *Not loaded > 0.21 on components with an Eigenvalue of > 1.0*
Highlighted cells indicate discriminatory loadings

The shape changes can be visualised in Figure 9-17, which confirm that pre-surgical lip shape for the viseme b<u>o</u>b tended to be shorter in height but wider across the commissures when compared to the average. In addition, the elevated margins of the upper lip were less protrusive than the CG, but the midpoint more so (p < 0.001). These differences were closer to the average post-surgery despite a statistical inference of significance (p = 0.03).
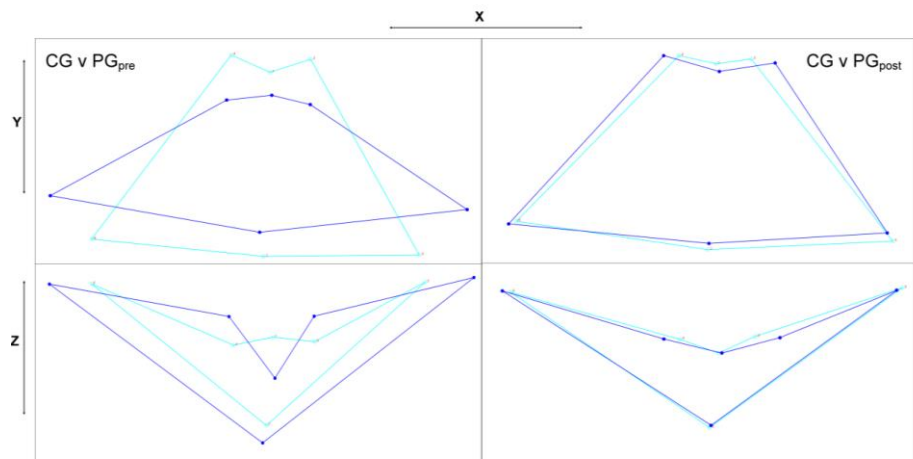
Figure 9-17 *Mean shape changes pre- and post-surgery for b<u>o</u>b CG (green)   PG (blue)*

9.4.   Discussion

The discussion focuses on the methodology behind this section of the study in addition to the implications of the results from a clinical perspective.

Section 4.5 confirmed that research in facial mobility following orthognathic surgery is limited. Of the three studies outlined previously, a further study has been published very recently (Verze et al., 2011). Here, 11 patients with a Class 3 skeletal pattern were scanned once prior to surgery and at 3 further time points post-surgery (1, 6 and 12 months). Subjects were asked to perform 5 basic facial expressions (frowning, eye closure, grimace, smile and lip purse) and scanned using a 3D laser scanner (Cyberware 3030RGB, Cyberware Incorporated, Monterey, CA). A gross outcome measure of mean shell-shell deviation between left and right sides was used. Their findings suggested that orthognathic surgery did not significantly alter facial mobility in the long term as movements 12 months post-surgery were largely similar to those recorded pre-surgically.

Despite being one of the most current publications in the field, the article suffers from several limitations. Patients were scanned using a static laser scanner – it is presumed that patients were scanned at maximum expression although this is not directly stated. The main drawback of using a static scanner to capture dynamic movements (even when only maximum displacement is considered) is that it is difficult for the patient to maintain the pose for an extended period of time. Indeed the Cyberware laser scanner has been quoted as having a total scan time of 18 seconds which may have led to errors in the capture of maximal facial expression (Perez et al., 2006). In addition to the lack of a control group, the sample size of 11 patients was further subdivided into smaller 3 groups, which were compared against each other. The small numbers of subjects question whether the sample is representative of the wider population. There were no formal statistical analyses conducted.

To this extent, the design of this study aimed to address the limitations of previously published research. Thirty patients were recruited to the study – this figure represents a larger sample than in all other previous literature. In addition, the sample size was intended to be large enough to potentially detect a statistical difference in lip movement from the control group, although a formal sample size was not carried out. This cohort represented the total number of Class 3 patients that commenced treatment in the Orthodontic Department during 2010 and were able to be followed up in terms of analysis during the timescale of the project.

The sample was also intended to be as homogenous as possible – the patients all presented with a severe Class 3 skeletal pattern requiring bimaxillary surgery. They were all treated by the same surgeon with rigid internal fixation. The surgical notes confirmed that upper lip closure was carried out using a simple continuous suture technique. No V-Y closures were made as studies suggest that upper lip length and dynamics can change following this procedure (Muradin et al., 2009). Despite this, there were clear within-group variations in lip competency pre-surgically, patient demographics

and antero-posterior movements of the maxilla and mandible. Therefore, the homogeneity of the group is perhaps less than initially suggested.

Both the pre-treatment and pre-surgical cephalometric values suggest that all patients had primarily antero-posterior deficiencies of the maxilla and mandible. The mean maxillary-mandibular planes angle was increased which also implied a vertical component to the aetiology of the malocclusion. Although details of specific treatment plans differed between patients (e.g. whether decompensating extractions were undertaken and the precise maxillary-mandibular movements) the additional inclusion criteria of a reverse overjet of at least 5mm prior to surgery was intended to allow a reasonable anterior-posterior skeletal movement to potentially take its effect on lip movement. Patients were scanned prior to surgery and not the start of treatment for three main reasons. Firstly, the time taken from the initial start of orthodontic treatment to surgery can take up to 18 months. It was considered unacceptable for a time period of this duration between image capture. Orthodontic treatment prior to surgery allows decompensation and arch coordination so that the teeth meet together in a satisfactory manner after surgery. This also allows the true extent of the skeletal pattern to be expressed maximising the influence of the basal bones on lip movement. Finally, it may have been considered that orthodontic appliances would hinder normal lip movement. However given that the patients were scanned some 18 months into treatment we suggest that this would have given ample opportunity for subjects to adapt. In addition, all patients still had their orthodontic appliances in place at the second scanning time 6 months later.

The results of this part of the study showed that there were statistical differences in lip movement of Class 3 orthognathic patients when compared to a control group. These differences in lip movement appeared to normalise towards the control group at 6 months post-surgery. Specifically, the PG_{pre} showed a lack of lip movement in the vertical dimension, preferring a wider, shorter lip shape when compared to the CG. Indeed, the lip shape of the PG_{pre} during movement was more suggestive of a resting lip shape implying an overall lack of movement for all the visemes. It would appear however that

post-surgery, the deficiencies in lip movement were corrected indirectly through the correction of the spatial relationships of the maxilla and mandible. The D^2 values indicated a return to normality with no statistically significant findings between the CG and $PG_{post.}$ Although Λ did find that there were still post-operative differences in the visemes <u>ba</u>by and b<u>o</u>b, the readings were much closer to normality that pre-operatively.

The research methodology utilises pre-existing statistical modelling techniques and applies them to 3D lip movement data showing that that both variation from average and abnormal movement can be detected from two sample populations. DA was used as a classification tool to test statistical significance between the CG and PG. PCA was then applied to the data to isolate those variables that differed between the groups. Alternatives to the use of DA include logistic regression and canonical variate analysis. Logistic regression has been previously adopted to investigate 3D facial morphology (Hammond et al., 2004). Although logistic regression essentially answers the same questions as discriminant analysis, it can be preferred to DA as it is more flexible in its assumptions. Consequently, the predictor variables do not have to be normally distributed, linearly related, or of equal variance within each group (Bookstein, 1991a). However, when the assumptions regarding the distributions of predictors are met, DA may be a more powerful and efficient analytic strategy (Sheets et al., 2004) - therefore given the findings during data preparation, DA was adopted.

Ideally the CG data should be robust enough to withstand significant alteration in CVA space when abnormal data is added to the model. We can see from Figure 9-5 that when comparing the plots of PG_{pre} to PG_{post} there have been minor changes to the CG CVA space although largely the structure remains the same. This would imply that the CG data is fairly robust (when compared to PG sample). The strength of the CG data is influenced by the sample size and the within sample variation in lip shape. In addition the degree to which the abnormal lip shapes vary from the CG may influence the CG structure.

The findings are difficult to compare to other studies already mentioned due to the differing outcome measures and methodology. In this group of patients, orthognathic surgery has changed the anterior-posterior and vertical positions of the maxilla and mandible. Other aspects in relation to verbal articulation are that a maxillary advancement will also lengthen the soft palate and mandibular surgery will alter the position of the tongue. In controlling for as many variables as possible (i.e. image capture, facial gesture reproducibility, landmark identification, subjects, surgeon and surgical procedure) we would hypothesise that correcting the underlying skeletal relationships may ensure that the articulators involved in speech production (lips, tongue and palate) are in a more balanced position in relation to each other. Although phonemes were not investigated as part of this study, it remains a further area for research as to whether phonemic stress exists in association with skeletal deformity and whether this is corrected in line with lip movement post-surgery.

9.5. Conclusion

An average model of lip shape can be used in the diagnosis and outcome of lip shape during movement for patients pre- and post-surgery. In this example, Class 3 orthognathic patients were found to have differences in lip shape during articulation when compared to an average template. Bimaxillary surgery normalised the lip shape to that of the control group.

Final conclusions and summary of contributions

This thesis has primarily focused on methods of capturing, analysing and interpreting facial movement in the clinical context to provide healthcare professionals who manage the relevant patient groups information on diagnosis, treatment planning and outcome assessment. The research was in part motivated by the needs of orthodontists and maxillofacial surgeons who treat patients with potential movement disorders on a daily basis but lack objective measures to quantify facial movement. However, the methods presented in this work are not limited to the clinical setting, but are generic and have wider applicability in speech and language therapy, psychology and computer vision.

The focus for Chapter 6 was on feature detection/landmark identification. The placement of six facial landmarks by two examiners at two time points was compared. In addition, landmark placement by one of the examiners was compared to that of an automated method. The results showed that manual methods of landmark identification were still considered as the gold standard due to inherent issues with image processing via automated registration.

Chapter 7 concentrated on the repeatability of facial gestures comparing verbal with non-verbal gestures. The relevance of this in the clinical context is that the measure of facial movement should be a repeatable as possible – that is, be performed in exactly the same manner over time so that changes over time or after an invention can truly be assessed. Differences in the performance of the gesture over time could potentially give misleading results. Using computer vision techniques it was shown that verbal gestures spoken in a normal, relaxed manner are significantly more repeatable over a one-month period when compared to non-verbal facial expressions.

Extending these findings, Chapter 8 took the most repeatable gestures and used statistical methods to first objectively quantify lip movement/shape in an average sample before using the data to create an average model of lip shape during movement. In Chapter 9, patient data was included into the

average model to test its ability to statistically determine changes in lip shape pre- and post-surgery. The findings showed that the model was sensitive enough to identify these differences prior to surgery and as the differences were not picked up post-surgery it was inferred that corrective surgery normalised lip shape.

In terms of how this methodology can be applied in the clinically setting, data from a single patient or patient groups can be analysed in a similar manner to the Geometric Morphometric approach described and projected onto the average CVA model presented in Section 8.3.5. Abnormal lip movement could be identified and indeed quantified by the distance of the patient data from the 95% confidence interval of the average model thereby acting as a diagnostic tool during clinical examinations and as a functional outcome measure following an intervention/rehabilitation.

Throughout the thesis, the triangulated 3D facial shell was used to extract landmark coordinate data. Triangulation has the advantage of providing a smooth illustration of a complex surface which can aid in the identification of facial landmarks. However, the mesh obtained from a point cloud is generally *improved* during the triangulation process to achieve a smooth surface albeit with a degree of data loss. In this respect there may have been benefits in using raw 3D data when compared to that sampled from a triangulated facial shell. This may not be problematic when using an extremely dense point cloud, for example from a static 3D facial imaging system. Here, slight data reduction may actually benefit computational processing speed with a negligible loss of surface detail. It is known that facial motion capture systems function at a lower resolution that a static 3D system and therefore the issue of raw data analysis versus triangulation may become more of an issue.

In summary, the main contributions of this thesis with respect to their appearance in the text are:

- an assessment of the reproducibility of landmark identification between manual and semi-automatic methods which showed that manual methods remain the gold standard;

- an evaluation of the repeatability of verbal facial gestures when compared to non-verbal which showed that spoken word is significantly more repeatable over time;

- the creation of an average model of lip shape during movement based on the most repeatable facial gestures. Experiments demonstrate that the model can discriminate different lip shapes from within the sample;

- data from a patient group pre- and post-surgery was added to the average model. Analyses showed differences in lip shape for the pre-surgical patient group that could be statistically identified from the average model. Post-surgery these differences were not statistically detectable;

- the impact of the study suggests that the average model can act as an outcome measure for lip shape during movement. A similar methodological approach can be adopted to compare matched populations in clinical situations where an objective measure of lip function is required.

Future research

We have seen that the analysis of facial movement (and in particular lip movement) can have a particular relevance in the management of certain patient groups. Although grading scales have proved to be a quick and simple method of achieving this, a more objective and quantifiable approach would be beneficial in terms of standardising the longitudinal management of patients by different clinicians. The thesis has shown that statistical shape analysis can be successfully applied to 3D motion data in a clinical context. Using a similar methodology, other patient groups such as those undergoing cleft lip revisions could be studied by age matching to a control group.

An extension to the current work would be to look for an association between resting lip shape and lip shape during movement. Phenotypic classification of

lip shape may reveal distinct groups of resting lip shape within the sample. It could be postulated that certain resting lip shapes follow a certain pattern of lip shape during movement. It may however be prudent to collect data for a much larger sample for this to feasibly return a realistic result.

Although the present work has used a landmark based approach for the analysis of lip shape, a study investigating how the surface topology of the lip changes during movement would also be interesting. An extremely high-resolution image during function would be required to create a dense polygon mesh around the peri-oral region. Small muscular deficiencies in the body of the lip could be detected if the image capture was of high enough resolution.

Much like 2D/3D static data has used differences in facial soft-tissue pre- and post-surgery for surgical prediction, the development of a 3D functional database of normal, pre- and post-surgical movements will allow the creation of a similar prediction tool in function. Fusing other imaging modalities such as real-time MRI which carries information on muscular movement will eventually allow the construction of a full biomechanical facial model in function.

References

AARONSON, S. A. 1967. A cephalometric investigation of the surgical correction of mandibular prognathism. *Angle Orthodontist,* 37**,** 251-260.

ABBOUD, B. & CHOLLET, G. Appearance based lip tracking and cloning on speaking faces. *In:* LONČARIĆ, S., BABIĆ, H. & M., B., eds. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005 Zagreb, Croatia. Faculty of Electrical Engineering and Computing, University of Zagreb, 301-305.

ABDI, H. & WILLIAMS, L. J. 2010. Principal Component Analysis. *Computational Statistics,* 2**,** 433-450.

AHRENS, H. 1958. Critical Values of Chi Square. *In:* PEARSON, E. S. & HARTLEY, H. O. (eds.) *Biometrika Tables for Statisticians.* 3rd ed. New York: Cambridge University Press.

ALEKSIC, P. S., POTAMIANOS, G. & KATSAGGELOS, A. K. 2009. Audiovisual Speech Processing. *In:* BOVIK, A. C. (ed.) *The Essential Guide to Video Processing.* London: Academic Press.

ANDERSON, R. E., BLACK, W. C., HAIR, J. F. & TATHAM, R. L. 2009. *Multivariate Data Analysis*, Prentice Hall.

ANDERSON, T. W. 1984. *Introduction to multivariate statistical analysis,* New York, John Wiley & Sons, Inc.

ARONSON, A. 1980. *Clinical Voice Disorders : An Interdisciplinary Approach,* New York, Thieme-Stratton.

BACCETTI, T., FRANCHI, L. & MCNAMARA, J. A., JR. 1999. Thin-plate spline analysis of treatment effects of rapid maxillary expansion and face mask therapy in early Class III malocclusions. *European Journal of Orthodontics,* 21**,** 275-281.

BAILER, W. 2008. A Comparison of Distance Measures for Clustering Video Sequences. *19th International Conference on Database and Expert Systems Application.* Turin

BARKANA, B. D. Detection of the Parameters of Hypernasality. *In:* LATIFI, S., ed. Proceedings of the Sixth International Conference on Information Technology, 2009 Las Vegas, NV. IEEE Computer Society, 1262-1264.

BARNETT, V. & LEWIS, T. 1984. *Outliers in Statistical Data*, John Wiley & Sons Ltd.

BARTLETT, M., HAGER, J., EKMAN, P. & SEJNOWSKI, T. 1999. Measuring facial expressions by computer image analysis. *Psychophysiology,* 36**,** 253-263.

BENEDIKT, L. 2009. *Using 3D facial motion for biometric identification.* PhD, Cardiff University

BENEDIKT, L., KAJIC, V., COSKER, D., ROSIN, P. L. & MARSHALL, D. Facial Dynamics in Biometric Identification. British Machine Vision Conference 1-4 September 2008 Leeds. 1075-1084.

BENEDIKTSSON, E. 1958. Variation in tongue and jaw position in /s/ production in relation to front teeth occlusion. *Acta Odontologica Scandinavica,* 15**,** 275-303.

BERNSTEIN, M. 1954. The relationship of speech defects and malocclusion. *American Journal of Orthodontics,* 40**,** 149-150.

BESL, P. J. & MCKAY, N. D. 1992. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14**,** 239-254.

BHATTACHARYYA, D., RANJAN, R., ALISHEROV, F. A. & CHOI, M. 2009. Biometric Authentication: A Review. *International Journal of u- and e-Service, Science and Technology,* 2**,** 19-28.

BLYTH, P. 1956. The relationship between speech, tongue behaviour, and occlusal abnormalities. *Dental Practitioner,* 20**,** 11-20.

BOOKSTEIN, F. L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 11**,** 567–585.

BOOKSTEIN, F. L. 1991a. *Morphometric Tools for Landmark Data,* New York, Cambridge University Press.

BOOKSTEIN, F. L. Thin-plate splines and the atlas problem for biomedical images. *In:* COLCHESTER, A. C. F. & HAWKES, D. J., eds. Proceedings of the 12th International Conference of Information Processing in Medical Imaging, 1991b London. Springer, 326-342.

BOUVIER, C., BENOIT, A., CAPLIER, A. & COULON, P. 2008. Open or Closed Mouth State Detection: Static Supervised Classification Based

on Log-polar Signature. *Advanced Concepts for Intelligent Vision Systems*, 1093–1102.

BOX, G. E. P. 1949. A general distribution theory for a class of likelihood criteria. *Biometrika* 36, 317-346.

BROADBENT, B. H. 1939. Bolton Standards and Technique in Orthodontic Practice. *Angle Orthodontist,* 7, 209-233.

BROWN, W. M. & MOORE, C. 2002. Smile asymmetries and reputation as reliable indicators of likelihood to cooperate: An evolutionary analysis. *In:* SHOHOV, S. (ed.) *Advances in psychology research.* Huntington, NY: Nova Science Publishers.

BUNTING, P., LUCAS, R. & LABROSSE, F. An Area based Technique for Image-to-Image Registration of Multi-Modal Remote Sensing Data. Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, 7-11 July 2008 2008. 212-215.

BURRES, S. 1985a. Facial biomechanics: the standards of normal. *Laryngoscope,* 95.

BURRES, S. 1985b. Facial biometrics: the standards of normal. *Laryngoscope,* 95, 708-714.

BURRES, S. & FISCH, U. 1986. The comparison of facial grading systems. *Archives of Otolaryngology Head and Neck Surgery,* 112, 755-758.

CARLOTTI, A. E., JR., ASCHAFFENBURG, P. H. & SCHENDEL, S. A. 1986. Facial changes associated with surgical advancement of the lip and maxilla. *Journal of Oral and Maxillofacial Surgery,* 44, 593-596.

CAWLEY, G. C. 1996. *The Application of Neural Networks to Phonetic Modelling.* PhD, University of Essex.

CHOUAKRIA-DOUZAL, A. & NAGABHUSHAN, P. 2006. Improved Frechet Distance for Time Series. *In:* BATAGELJ, V., BOCK, H., FERLIGOJ, A. & ZIBERNA, A. (eds.) *Data Science and Classification.* Springer.

COHN, J. & SCHMIDT, K. L. 2004. The timing of facial motion in posed and spontaneous smiles *International Journal of Wavelets, Multiresolution, and Information Processing,* 2, 1-12.

COHN, J., ZLOCHOWER, A., LIEN, J. & KANADE, T. 1999. Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology,* 36, 35-43.

COOTES, T. F., HILL, A., TAYLOR, C. J. & HASLAM, J. 1994. The Use of Active Shape Models for Locating Structures in Medical Images. *Image and Vision Computing,* 12**,** 355-366.

COOTES, T. F., TAYLOR, C. J., COOPER, D. H. & GRAHAM, J. 1995. Active shape models - their training and application. *Computer Vision and Image Understanding,* 61**,** 38-59.

COOTES, T. F., TWINING, C. J., PETROVIC, V., SCHESTOWITZ, R. & TAYLOR, C. J. Groupwise reconstruction of appearance models using piece-wise affine deformations. *In:* CLOCKSIN, W., FITZGIBBON, A. & TORR, P., eds. British Machine Vision Conference, 2005 Oxford. Alden Group Ltd, 879-888.

COSTA, F., ROBIONY, M., SEMBRONIO, S., POLINI, F. & POLITI, M. 2001. Stability of skeletal Class III malocclusion after combined maxillary and mandibular procedures. *International Journal of Adult Orthodontics & Orthognathic Surgery,* 16**,** 179-192.

COULSON, S. E., CROXSON, G. R., ADAMS, R. D. & O'DWYER, N. J. 2005. Reliability of the "Sydney," "Sunnybrook," and "House Brackmann" facial grading systems to assess voluntary movement and synkinesis after facial nerve paralysis. *Otolaryngology Head & Neck Surgery,* 132**,** 543-549.

CRUM, W. R., HARTKENS, T. & HILL, D. L. G. 2004. Non-rigid image registration: theory and practice. *The British Journal of Radiology,* 77**,** 140-153.

DARWIN, C. 1872. *The expression of the emotions in man and animals,* London, John Murray.

DAVIES, R. B., ÕUNPUU, S., TYBURSKI , D. & GAGE, J. R. 1991. A gait analysis data collection and reduction technique *Human Movement Science,* 10**,** 575-587.

DAWES, K. S. & KELLY, S. W. 2005. An instrument for the non-invasive assessment of lip function during speech. *Medical Engineering & Physics,* 27**,** 523-535.

DERMAUT, L. R. & DE SMIT, A. A. 1989. Effects of sagittal split advancement osteotomy on facial profiles. *European Journal of Orthodontics,* 11**,** 366-374.

DI PALMA, E., GASPARINI, G., PELO, S., TARTAGLIA, G. M. & CHIMENTI, C. 2009. Activities of masticatory muscles in patients after orthognathic surgery. *Journal of Craniomaxillofacial Surgery,* 37**,** 417-420.

DI PALMA, E., GASPARINI, G., PELO, S., TARTAGLIA, G. M. & SFORZA, S. 2010. Activities of masticatory muscles in patients before orthognathic surgery. *Journal of Craniofacial Surgery,* 21**,** 724-726.

DIBEKLIOGLU, H., SALAH, A. A. & GEVERS, T. Automatic landmarking and alignment for facial expression analysis. Signal Processing and Communications Applications Conference (SIU), 2010 IEEE 18th, 22-24 April 2010 2010. 208-211.

DRYDEN, I. L. & MARDIA, K. V. 1998. Generalised Procrustes Methods. *In:* BARNETT, V., BRADLEY, R. A., CRESSIE, N., FISHER, N. I., JOHNSTONE, I., KADANE, J. B., KENDALL, D. G., SCOTT, D. W., BERNARD, B., SMITH, A. F. M. & TEUGELS, J. L. (eds.) *Statistical Shape Analysis.* Chichester: Wiley.

DUCHENNE, B. 1862. *Mecanisme de la physionomie humanie ou analyse electrophysiologique de l'expression des passions.,* Paris, Bailliere.

DUFFY, J. 1995. *Motor speech disorders : substrates, differential diagnosis, and management,* St Louis, Mosby, Inc.

EFRAT, A., FAN, Q. & VENKATASUBRAMANIAN, S. 2007. Curve Matching, Time Warping, and Light Field --- New Algorithms for Computing Similarity Between Curves. *Journal of Mathematical Image and Vision,* 27**,** 203-216.

EITER, T. & MANNILA, H. 1995. Computing the Frechet distance between two polygonal curves. *International Journal of Computational Geometry and Applications,* 5**,** 75-91.

EKMAN, P. & FRIESEN, W. 1978. *The Facial Action Coding System (FACS): A Technique for the Measurement of Facial Action.,* Paolo Alto, California.

EKMAN, P., FRIESEN, W. V. & ELLSWORTH, P. 1972. *Emotion in the human face: guidelines for research and an integration of findings,* New York, Pergamon Press.

EKMAN, P. & MATSUMOTO, D. 2008. Facial expression analysis. *In:* CHEN, K. (ed.) *Scholarpedia.*

EKSTROM, C. 1982. Facial growth rate and its relation to somatic maturation in healthy children. *Swedish Dental Journal Supplement,* 11**,** 1-99.

FACEREC. *Facial Recognition Homepage* [Online]. Zagreb, Croatia. Available: http://www.face-rec.org [Accessed 16th September 2010.

FARKAS, L. G. 1994. *Anthropometry of the Head and Face,* New York, Lippincott-Raven.

FARKAS, L. G. & DEUTSCH, C. K. 1996. Anthropometric determination of craniofacial morphology. *American Journal of Medical Genetics,* 65**,** 1-4.

FARKAS, L. G., POSNICK, J. C. & HRECZKO, T. M. 1992. Growth patterns of the face: a morphometric study. *Cleft Palate Craniofacial Journal,* 29**,** 308-315.

FASEL, B. & LUETTIN, J. 2003. Automatic facial expression analysis: a survey. *Pattern Recognition,* 36**,** 259-275.

FIELDS, M. J. & PECKITT, N. S. 1990. Facial nerve function index: a clinical measurement of facial nerve activity in patients with facial nerve palsies. *Oral Surgery, Oral Medicine, Oral Pathology,* 69**,** 681-682.

FITZSIMONS, K., DEACON, S., VAN DER MEULEN, J. & COPLEY, L. 2011. CRANE Database: Annual Report London: The Royal College of Surgeons of England.

FLEET, D. J. & WEISS, Y. 2005. Optical Flow Estimation. *In:* PARAGIOS, N., CHEN, Y. & FAUGERAS, O. (eds.) *Mathematical Models in Computer Vision.* Springer.

FREIHOFER, H. P., JR. 1977. Changes in nasal profile after maxillary advancement in cleft and non-cleft patients. *Journal of Maxillofacial Surgery,* 5**,** 20-27.

FREY, M., GIOVANOLI, P., GERBER, H., SLAMECZKA, M. & STUSSI, E. 1999. Three-dimensional video analysis of facial movements: a new method to assess the quantity and quality of the smile. *Plastic & Reconstructive Surgery,* 104**,** 2032-2039.

FREY, M., JENNY, A., GIOVANOLI, P. & STUSSI, E. 1994. Development of a new documentation system for facial movements as a basis for the international registry for neuromuscular reconstruction in the face. *Plastic & Reconstructive Surgery,* 93**,** 1334-1349.

FYMBO, L. 1936. The relationship of malocclusion of the teeth to defects of speech. *Archives of speech,* 1**,** 204-216.

GOWER, J. C. 2005. Principal coordinate analysis. *In:* ARMITAGE, P. & COLTON, T. (eds.) *Encyclopedia of Biostatistics.* J. Wiley.

GREEN, J. R. & WILSON, E. 2006. Spontaneous Facial Motility in Infancy. *Developmental Psychobiology,* 48**,** 16-28.

GREENHOUSE, S. W. & GEISSER, S. 1959. On methods in the analysis of profile data. *Psychometrika,* 24**,** 95-112.

GRENANDER, U., CHOW, Y. & KEENAN, D. M. 1991. *Hands: A Pattern Theoretic Study of Biological Shapes,* New York, Springer-Verlag.

GROSS, M. M., TROTMAN, C. A. & MOFFATT, K. S. 1996. A comparison of three-dimensional and two-dimensional analyses of facial motion. *Angle Orthodontist,* 66**,** 189-194.

GRUEN, A. & AKCA, D. 2005. Least squares 3D surface and curve matching. *Journal of Photogrammetry & Remote Sensing,* 59**,** 151-174.

GUADAGNOLI, E. & VELICER, W. F. 1988. Relation of sample size to the stability of component patterns. *Psychological Bulletin,* 103**,** 265-275.

GUAY, A. H., MAXWELL, D. L. & BEECHER, R. 1978. A radiographic study of tongue posture at rest and during the phonation of /s/ in class III malocclusion. *Angle Orthodontist,* 48**,** 10-22.

GUPTA, S., MARKEY, M. K. & BOVIK, C. A. 2010. Anthropometric 3D Face Recognition. *International Journal of Computer Vision,* 90**,** 331-349.

GUYMON, M., CROSBY, D. R. & WOLFORD, L. M. 1988. The alar base cinch suture to control nasal width in maxillary osteotomies. *International Journal of Adult Orthodontics & Orthognathic Surgery,* 3**,** 89-95.

GWILLIAM, J. R., CUNNINGHAM, S. J. & HUTTON, T. 2006. Reproducibility of soft tissue landmarks on three dimensional facial scans. *European Journal of Orthodontics,* 28**,** 408-415.

HAAHR, M. 1998. *Random Integer Generator* [Online]. Dublin: TDSA. Available: http://www.random.org/integers/ [Accessed 27th February 2012].

HACK, G. A., DE MOL VAN OTTERLOO, J. J. & NANDA, R. 1993. Long-term stability and prediction of soft tissue changes after LeFort I surgery.

American Journal of Orthodontics & Dentofacial Orthopaedics, 104**,** 544-555.

HAGER, J. C. 1983. *Asymmetries in Facial Actions.* PhD, University of California.

HAJEER, M. Y., AYOUB, A. F., MILLETT, D. T., BOCK, M. & SIEBERT, J. P. 2002. Three-dimensional imaging in orthognathic surgery: the clinical application of a new method. *International Journal of Adult Orthodontics & Orthognathic Surgery,* 17**,** 318-330.

HAMMOND, P., HUTTON, T. J., ALLANSON, J. E., CAMPBELL, L. E., HENNEKAM, R. C. M., HOLDEN, S., PATTON, M. A., SHAW, A., TEMPLE, I. K., TROTTER, M., MURPHY, K. C. & WINTER, R. M. 2004. 3D analysis of facial morphology. *American Journal of Medical Genetics Part A,* 126A**,** 339-348.

HAMMOND, P. & SUTTIE, M. 2012. Large-scale objective phenotyping of 3D facial morphology. *Human Mutation* [Online]. Available: http://dx.doi.org/10.1002/humu.22054 [Accessed 20th February, 2012].

HANS, M. G., BROADBENT, B. H. J. & NELSON, S. S. 1994. The Broadbent-Bolton Growth Study--past, present, and future. *American Journal of Orthodontics and Dentofacial Orthopedics,* 105**,** 598-603.

HARVOLD, E. 1970. Speech articulation and oral morphology. *American Speech and Hearing Assocation,* Report No. 5**,** 69-75.

HECK, U. & KLECK, R. 1997. Differentiating emotion elicited and deliberate emotional facial expressions. *In:* EKMAN, P. & ROSENBERG, E. (eds.) *What the face reveals.* New York: Oxford University Press.

HENNESSY, R. J. & MOSS, J. P. 2001. Facial growth: separating shape from size. *European Journal of Orthodontics,* 23**,** 275-285.

HERNANDEZ-ORSINI, R., JACOBSON, A., SARVER, D. M. & BARTOLUCCI, A. 1989. Short-term and long-term soft tissue profile changes after mandibular advancements using rigid fixation techniques. *International Journal of Adult Orthodontics & Orthognathic Surgery,* 4**,** 209-218.

HERSHEY, H. G. 1972. Incisor tooth retraction and subsequent profile change in postadolescent female patients. *American Journal of Orthodontics,* 61**,** 45-54.

HILL, D. L. G., BATCHELOR, P. G., HOLDEN, M. & HAWKES, D. J. 2001. Medical Image Registration. *Physics in Medicine and Biology,* 46**,** R1-R45.

HILL, H., CLAES, P., CORCORAN, M., WALTERS, M., JOHNSTON, A. & CLEMENT, J. G. 2011. How Different is Different? Criterion and Sensitivity in Face-Space. *Front Psychol,* 2**,** 41.

HONTANILLA, B. & AUBA, C. 2008. Automatic three-dimensional quantitative analysis for evaluation of facial movement. *Journal of Plastic, Reconstructive and Aesthetic Surgery,* 61**,** 18-30.

HOUSE, J. W. 1983. Facial nerve grading systems. *Laryngoscope,* 93**,** 1056-1069.

HOUSE, J. W. & BRACKMAN, D. E. 1985. Facial nerve grading system. *Otolaryngology, Head and Neck Surgery,* 93**,** 146-147.

HOUSTIS, O. & KILIARIDIS, S. 2009. Gender and age differences in facial expressions. *European Journal of Orthodontics,* 31**,** 459-466.

HUBBE, M., HANIHARA, T. & HARVATI, K. 2009. Climate signatures in the morphological differentiation of worldwide modern human populations. *Anatomical Record (Hoboken),* 292**,** 1720-1733.

HUTTON, T. J., BUXTON, B. F. & HAMMOND, P. 2003. Estimating average growth trajectories in shape-space using kernel smoothing. *IEEE Transactions on Medical Imaging,* 22**,** 747-753.

INGERVALL, B. & SARNAS, K. 1962. Comparison of dentition in lispers and non-lispers. *Odontoiatria Revista,* 13**,** 344-354.

JACKSON, G. & VON DOERSTEN, P. G. 1999. THE FACIAL NERVE: Current Trends in Diagnosis, Treatment, and Rehabilitation *Medical Clinics of North America,* 83**,** 179-195.

JEMT, T. 1981. Chewing patterns in dentate and complete denture wearer - recorded by light-emitting diodes. *Swedish Dental Journal,* 5**,** 199-205.

JOHNS, F. R., P.C., J., BUCKLEY, M., BRAUN, T. W. & CLOSE, J. M. 1997. Changes in facial movement after maxillary osteotomies. *Journal of Oral and Maxillofacial Surgery,* 55**,** 1044-1048.

JOHNSON, H. J. & CHRISTENSEN, G. E. 2001. Landmark and Intensity-Based, Consistent Thin-Plate Spline Image Registration *Information Processing in Medical Imaging,* 2082**,** 329-343.

JOHNSON, N. C. & SANDY, J. R. 1999. Tooth position and speech--is there a relationship? *Angle Orthodontist,* 69**,** 306-310.

JOHNSON, P., BROWN, H., KUZON, W., BALLIET, R., GARRISON, J. & CAMPBELL, J. 1994. Simultaneous quantification of facial movement: the maximum static response assay of facial nerve function. *Annals of Plastic Surgery,* 32**,** 171-179.

JOHNSTON, D., MILLETT, D. T. & AYOUB, A. F. 2003. Are Facial Expressions Reproducible? *Cleft Palate Craniofacial Journal,* 40**,** 291-296.

JOHNSTON, D., MILLETT, D. T., AYOUB, A. F. & BOCK, M. 2001. Validity and reproducibility of 3-dimensional imaging system. *Journal of Dental Research,* 80**,** abstract 228.

JOUAN-RIMBAUD, D., MASSART, D. L., SABY, C. A. & PUEL, C. 1997. Characterisation of the representativity of selected sets of samples in multivariate calibration and pattern recognition. *Analytica Chimica Acta,* 350**,** 149-161.

KAIPATUR, N. R. & FLORES-MIR, C. 2009. Accuracy of computer programs in predicting orthognathic surgery soft tissue response. *J Oral Maxillofac Surg,* 67**,** 751-9.

KAISER, H. F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement,* 20**,** 141-151.

KANADE, T., COHN, J. F. & TIAN, Y. Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000 Grenoble, France. 46-53.

KAPOOR, S., KHANNA, S. & BHATIA, R. 2010. Facial Gesture Recognition using Correlation and Mahalanobis Distance. *International Journal of Computer Science and Information Security,* 7**,** 267-272.

KAU, C. H., CRONIN, A., DURNING, P., ZHUROV, A., SANDHAM, A. & RICHMOND, S. 2006a. A new method for the 3D measurement of postoperative swelling following orthognathic surgery. *Orthodontics & Craniofacial Research,* 9**,** 31-37.

KAU, C. H., ZHUROV, A., RICHMOND, S., CRONIN, A., SAVIO, C. & MALLORIE, C. 2006b. Facial templates: a new perspective in three dimensions. *Orthodontics and Craniofacial Research,* 9**,** 10-17.

KEOGH, E. J. & PAZZINI, M. J. Derivative Dynamic Time Warping. First SIAM International Conference on Data Mining, 2001 Chicago, IL. 1-11.

KIM, T. K., KITTLER, J. & CIPOLLA, R. 2007. Discriminative learning and recognition of image set classes using canonical correlations. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 29**,** 1005-1018.

KINNEAR, P. R. & GRAY, C. D. 2010. Discriminant analysis and logistic regression. *PASW statistics 17 made simple.* Hove: Psychology Press.

KLATT, D. H. 1990. Review of the ARPA speech understanding project. *In:* WAIBEL, A. & LEE, K. (eds.) *Readings in speech recognition.* San Francisco, CA: Morgan Kaufmann Publishers Inc.

KLINGENBERG, C. P. 2011. MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources,* 11**,** 353-357.

KRUMHUBER, E. & KAPPAS, A. 2005. Moving smiles: The role of the dynamic components for the perception of the genuineness smiles. *Journal of Nonverbal Behavior,* 29**,** 3-24.

KUHN, R., NGUYEN, P., JUNQUA, J. C. & GOLDWASSER, L. Eigenfaces and eigenvoices: dimensionality reduction for specialized pattern recognition. Multimedia Signal Processing, 1998 IEEE Second Workshop on, 7-9 Dec 1998 1998. 71-76.

KUMAR, K., CHEN, T. & STERN, R. M. Profile View Lip Reading. *In:* KUH, A. & HUANG, Y., eds. International Conference on Acoustics, Speech and Signal Processing, 2007 Honolulu, HI. 429-432.

LACHENBRUCH, P. A. 1967. An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis. *Biometrics* 23**,** 639–645.

LINSTROM, C. J., SILVERMAN, C. A. & SUSMAN, W. M. 2000. Facial-motion analysis with a video and computer system: a preliminary report. *American Journal of Otology,* 21**,** 123-129.

LUBIT, E. C. 1967. The relationship of malocclusion and faulty speech articulation. *Journal of Oral Medicine,* 22**,** 47-55.

MAHALANOBIS, P. C. 1936. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India,* 2**,** 49-55.

MANKTELOW, R. T., ZUKER, R. M. & TOMAT, L. R. 2008. Facial Paralysis Measurement with a Handheld Ruler. *Plastic & Reconstructive Surgery,* 121**,** 435-442.

MANSOUR, S., BURSTONE, C. & LEGAN, H. 1983. An evaluation of soft-tissue changes resulting from Le Fort I maxillary surgery. *American Journal of Orthodontics,* 84**,** 37-47.

MATSUMOTO, D. & EKMAN, P. 2008. Facial expression analysis. *In:* CHEN, K. (ed.) *Scholarpedia.*

MAY, M. 1970. Facial paralysis, peripheral type: a proposed method of reporting. (Emphasis on diagnosis and prognosis, as well as electrical and chorda tympani nerve testing). *Laryngoscope,* 80**,** 331-390.

MCINTYRE, G. T. & MOSSEY, P. A. 2003. Size and shape measurement in contemporary cephalometrics. *European Journal of Orthodontics,* 25**,** 231-242.

MCLACHLAN, G. J. 1999. Mahalanobis Distance. *RESONANCE Journal of Science Education,* 4**,** 20-26.

MIAN, A. S., BENNAMOUN, A. M. & OWENS, R. 2007. An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 29**,** 1927-1943.

MILLS, J. R. E. 1987. *Principles and Practice of Orthodontics,* Edinburgh, Churchill Livingstone.

MISHIMA, K., YAMADA, T., OHURA, A. & SUGAHARA, T. 2006. Production of a range image for facial motion analysis: A method for analyzing lip motion. *Computerized Medical Imaging and Graphics,* 30**,** 53-59.

MISHIMA, K., YAMADA, T., SUGII, A., MATSUMURA, T. & SUGAHARA, T. 2009. Application of a novel method to analyse lip motion of cleft patients before and after lip repair. *Dento-Maxillo-Facial Radiology,* 38**,** 232-238.

MORRIS, R. J., KENT, J. T., MARDIA, K. V., FIDRICH, M., AYKROYD, R. G. & LINNEY, A. Analysing growth in faces. Proc. Conf. Imaging Science, Systems and Technology 1999, Las Vegas, 1999. 404-410.

MUNICH, M. E. & PERONA, P. Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification *In:* WERNER, B., ed. The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999 Corfu, Greece. 108-115.

MURADIN, M. S., ROSENBERG, A., VAN DER BILT, A., STOELINGA, P. J. & KOOLE, R. 2009. The effect of alar cinch sutures and V-Y closure on soft tissue dynamics after Le Fort I intrusion osteotomies. *J Craniomaxillofac Surg,* 37**,** 334-40.

MURAKAMI, S., MIZOBUCHI, M., NAKASHIRO, Y., DOI, T., HATO, N. & YANAGIHARA, N. 1996. Bell palsy and herpes simplex virus: identification of viral NDA in endoneurial fluid and muscle. *Annals of Internal Medicine,* 124**,** 27-30.

MURTY, G. E., DIVER, J. P., KELLY, P. J., O'DONOGHUE, G. M. & BRADLEY, P. J. 1994. The Nottingham System: objective assessment of facial nerve function in the clinic. *Otolaryngol Head Neck Surg,* 110**,** 156-61.

MYERS, C., RABINER, L. R. & ROSENBURG, A. E. 1980. Performance Tradeoffs in Dynamic Time Warping Algorithms for Isolated Word Recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing,* 26**,** 623-635.

NAKATA, Y., UEDA, H. M., KATO, M., TABE, H., SHIKATA-WAKISAKA, N., MATSUMOTO, E., KOH, M., TANAKA, E. & TANNE, K. 2007. Changes in stomatognathic function induced by orthognathic surgery in patients with mandibular prognathism. *Journal of Oral and Maxillofacial Surgery,* 65**,** 444-451.

NEELY, J., JOAQUIN, A., KOHN, L. & CHEUNG, J. 1996. Quantitative assessment of the variations within grades of facial paralysis. *Laryngoscope,* 106**,** 438-442.

NEELY, J. G., CHERIAN, N. G., DICKERSON, C. B. & NEDZELSKI, J. M. 2010a. Sunnybrook Facial Grading System: Reliability and Criteria for Grading. *Laryngoscope,* 120**,** 1038-1045.

NEELY, J. G., CHEUNG, J., WOOD, M., BYERS, J. & ROGERSON, A. 1992. Computerised quantitative dynamic analysis of facial motion in the paralysed synkinetic face. *American Journal of Otolaryngology,* 13**,** 87-107.

NEELY, J. G., WANG, K. X., SHAPLAND, C. A., SEHIZADEH, A. & WANG, A. 2010b. Computerized objective measurement of facial motion: normal variation and test-retest reliability. *Otology and Neurotology,* 31**,** 1488-1492.

NOH, H., NABHA, W., CHO, J. H. & HWANG, H. S. 2011. Registration accuracy in the integration of laser-scanned dental images into maxillofacial cone-beam computed tomography images. *Am J Orthod Dentofacial Orthop,* 140**,** 585-91.

NOOREYAZDAN, M., TROTMAN, C. A. & FARAWAY, J. J. 2004. Modeling facial movement: II. A dynamic analysis of differences caused by orthognathic surgery. *Journal of Oral and Maxillofacial Surgery,* 62**,** 1380-1386.

O'RYAN, F. & SCHENDEL, S. 1989. Nasal anatomy and maxillary surgery. II. Unfavorable nasolabial esthetics following the Le Fort I osteotomy. *International Journal of Adult Orthodontics & Orthognathic Surgery,* 4**,** 75-84.

OKUDAIRA, M., ONO, T., KAWAMOTO, T. & MORIYAMA, K. 2008. Three-dimensional analysis of lower lip movement during articulation in subjects with mandibular prognathism. *Orthodontic Waves,* 67**,** 93-103.

PAE, E. K., LOWE, A. A. & FLEETHAM, J. A. 1997. A thin-plate spline analysis of the face and tongue in obstructive sleep apnea patients. *Clinical Oral Investigations,* 1**,** 178-184.

PALETZ, J., MANKTELOW, R. & CHABAN, R. 1994. The shape of a normal smile: implications for facial paralysis reconstruction. *Plastic & Reconstructive Surgery,* 93**,** 784-789.

PARK, H. M. W. P. 2008. *Univariate Analysis and Normality Test Using SAS, Stata, and SPSS* [Online]. Bloomington, IN: The University Information

Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University. Available: http://www.indiana.edu/~statmath/stat/all/normality/index.html [Accessed 4th April 2012].

PARKE, F. & WATERS, K. 1996. *Computer Facial Animation*, AK Peters, Ltd.

PEARSON, D. 1981. Visual Communication Systems for the Deaf. *IEEE Transactions on Communications,* 29**,** 1986 - 1992

PEITERSEN, E. 1992. Natural History of Bell's Palsy. *Acta Oto-Laryngologica,* 112**,** 122-124.

PEKTAS, Z. O., KIRCELLI, B. H. & CILASUN, U. 2008. The Use of Software Systems for Visualized Treatment Objectives in Orthognathic Surgery. *In:* BOZOVIC, V. (ed.) *Medical Robotics.* Vienna: I-Tech Education and Publishing.

PEREZ, S. I., BERNAL, V. & GONZALEZ, P. N. 2006. Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy,* 208**,** 769-784.

PESTRONSK, A. 2012. *FACIAL (VII) NERVE DISORDERS* [Online]. St Louis, MO: Washington University. [Accessed 10th October 2012].

PHILLIPS, C. D. & BUBASH, L. A. 2002. The facial nerve: anatomy and common pathology. *Seminars in Ultrasound, CT and MR,* 23**,** 202-217.

PIGEON, S. & VANDENDORPE, L. The M2VTS multimodal face database (Release 1.00). *In:* BIGÜN, J., CHOLLET, G. & BORGEFORS, G., eds. First International Conference on Audio and Video-Based Person Authentication, 1997 Crans-Montana, Switzerland,. Spinger, 403-409.

PLOOIJ, J. M., SWENNEN, G. R. J., RANGEL, F. A., MAAL, T. J. J., SCHUTYSER, F. A. C., BRONKHORST, E. M., KUIJPERS-JAGTMAN, A. M. & BERGE, S. J. 2009. Evaluation of reproducibility and reliability of 3D soft tissue analysis using 3D stereophotogrammetry. *International Journal of Oral and Maxillofacial Surgery,* 38**,** 267-273.

POLITI, M., COSTA, F., ROBIONY, M., SOLDANO, F. & ISOLA, M. 2002. Stability of maxillary advancement for correction of skeletal Class III malocclusion after combined maxillary and mandibular procedures: preliminary results of an active control equivalence trial for semirigid

and rigid fixation of the maxilla. . *International Journal of Adult Orthodontics & Orthognathic Surgery,* 17**,** 98-110.

POMERANTZ, J. & ZELLER, A. 1965. Speech, occlusion and tongue function in elementary school children. *American Journal of Orthodontics,* 51**,** 312-13.

POPAT, H. & RICHMOND, S. 2010. New developments in: three-dimensional planning for orthognathic surgery. *Journal of Orthodontics,* 37**,** 62-71.

POPAT, H., RICHMOND, S., BENEDIKT, L., MARSHALL, D. & ROSIN, P. L. 2009. Quantitative analysis of facial movement-a review of three-dimensional imaging techniques. *Computerized Medical Imaging and Graphics,* 33**,** 377-383.

POPAT, H., RICHMOND, S., PLAYLE, R., MARSHALL, D., ROSIN, P. L. & COSKER, D. 2008. Three-dimensional motion analysis - an exploratory study. Part 1: assessment of facial movement. *Orthodontics and Craniofacial Research,* 11**,** 216 - 223.

POTHIAWALA, S. & LATEEF, F. 2012. Bilateral Facial Nerve Palsy: A Diagnostic Dilemma. *Case Reports in Emergency Medicine***,** 1-3.

PROFFIT, W. R., TURVEY, T. A. & PHILLIPS, C. 2007. The hierarchy of stability and predictability in orthognathic surgery with rigid fixation: an update and extension. *Head Face Medicine,* 30**,** 1-21.

PROFFIT, W. R. & WHITE JR, R. P. 2002. Dentofacial Problems: Prevalence and Treatment Need. *In:* PROFFIT, W. R., WHITE JR, R. P. & SARVER, D. M. (eds.) *Contemporary Treatment of Dentofacial Deformities.* St Louis: Mosby.

QIFENG, Z. & ALWAN, A. On the use of variable frame rate analysis in speech recognition. IEEE International Conference on Acoustics, Speech and Signal Processing, 2000 Istanbul, Turkey. 1783-1786.

RAINS, M. D. & NANDA, R. 1982. Soft-tissue changes associated with maxillary incisor retraction. *American Journal of Orthodontics,* 81**,** 481-488.

RAJAVEL, R. & SATHIDEVI, P. S. Static and Dynamic Features for Improved HMM based Visual Speech Recognition *In:* TIWARY, U. S., SIDDIQUI, T. J., RADHAKHRISHNA, M. & TIWARI, M. D., eds. Proceedings of the

First International Conference on Intelligent Human Computer Interaction, 2009 Allahabad, India. Springer, 184-194.

REEVES, A. G. & SWENSON, R. S. 2008. Cranial Nerves - Facial Movement and Sensation. *In:* SWENSON, R. S. (ed.) *Disorders of the Nervous System.* Hanover, NH: Dartmouth Medical School.

RICHTSMEIER, J. T., CHEVERUD, J. M. & LELE, S. 1992. Advances in anthropological morphometrics. *Annals Review Anthropology,* 21**,** 283–305.

RINN, W. E. 1984. The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin,* 95**,** 52-77.

RINN, W. E. 1991. Neuropsychology of facial expression. *In:* FELDMAN, R. & RIME, B. (eds.) *Fundamentals of nonverbal behaviour.* New York: Cambridge University Press.

ROACH, P. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association,* 34**,** 239-245

ROBINSON, S. W., SPEIDEL, T. M., ISAACSON, R. J. & WORMS, F. W. 1972. Soft tissue profile change produced by reduction of mandibular prognathism. *Angle Orthodontist,* 42**,** 227-235.

ROSS, B. G., FRADET, G. & NEDZELSKI, J. M. 1996. Development of a sensitive clinical facial grading system. *Otolarnygology Head and Neck Surgery,* 114**,** 380-386.

ROUSSEEUW, P. J. & LEROY, A. M. 1987. *Robust Regression and Outlier Detection,* New York, Wiley.

RUDEE, D. 1964. Proportional profile changes concurent with orthodontic therapy. *American Journal of Orthodontics,* 43**,** 324.

RUECKERT, D. 2001. Non-rigid Registration: Concepts, Algorithms and Applications. *In:* HAJNAL, J. V., HILL, D. L. G. & HAWKES, D. J. (eds.) *Medical Image Registration.* Florida: CRC Press LLC.

RUECKERT, D., SONODA, L. I., HAYES, C., HILL, D. L., LEACH, M. O. & HAWKES, D. J. 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging,* 18**,** 712-721.

RUIZ, M. C. & ILLINGWORTH, J. 2008. Automatic landmarking of faces in 3D-ALF [sup 3D]. *In:* IZQUIERDO, E. & LIU, G. (eds.) *5th International Conference on Visual Information Engineering.* Xi'an, China: IEE.

SAKOE, H. & CHIBA, S. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing,* 26**,** 43-49.

SALVI, J., MATABOSCH, C., FOFI, D. & FOREST, J. 2007. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing,* 25**,** 578–596.

SAVRAN, A., ALYÄUZ, N., DIBEKLIOGLU, H., CELIKTUTAN, O., GÄOKBERK, B., AKARUN, L. & SANKUR, B. 2008. Bosphorus database for 3D face analysis. *Biometrics and Identity Management,* 5372**,** 47–56.

SAWYER, A. R., SEE, M. & NDUKA, C. 2009. Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry. *Otolarnygology Head and Neck Surgery,* 140**,** 76-81.

SCHONBERG, I. J. 1945. Contributions to the problem of approximation of equidistant data by analytic functions. *Quarterly Applied Mathematics,* 4**,** 45-99 and 112-141.

SEGNER, D. 1986. The shape of the human face recorded by use of contour photography and spline function interpolation. *European Journal of Orthodontics,* 8**,** 112-117.

SELL, D., HARDING, A. & GRUNWELL, P. 1999. GOS.SP.ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). *International Journal of Language and Communication Disorders,* 34**,** 17-33.

SFORZA, S., GALENTE, D., SHIRAI, Y. F. & FERRARIO, V. F. 2010. A three-dimensional study of facial mimicry in healthy young adults. *Journal of Cranio-Maxillo-Facial Surgery,* 38**,** 409-415.

SHEETS, H. D., KEONHO, K. & MITCHELL, C. E. 2004. A combined landmark and outline-based approach to ontogenic shape change in the Ordovician Trilobite *Triarthrus becki. In:* ELEWA, A. (ed.) *Applications of Morphometrics in Paleontology and Biology.* New York: Springer.

SHOLTS, S. B., WÄRMLÄNDER, S. K. T. S., FLORES, L. M., MILLER, K. W. P. & WALKER, P. L. 2010. Variation in the Measurement of Cranial Volume and Surface Area Using 3D Laser Scanning Technology. *Journal of Forensic Science,* 55**,** 871-876.

SIDOROV, K. 2010. *Groupwise non-rigid registration for automatic construction of appearance models of the human craniofacial complex for analysis, synthesis and simulation.* PhD, Cardiff Univeristy.

SIDOROV, K., MARSHALL, D., ROSIN, P. L. & RICHMOND, S. Towards Efficient 3D Facial Appearance Models. *In:* METAX, D. & POPOVIC, J., eds. Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation, 2007 San Diego, CA. Eurographics Association Aire-la-Ville, 287.

SIDOROV, K., RICHMOND, S. & MARSHALL, D. Efficient Groupwise Non-rigid Registration of Textured Surfaces. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2011 Colorado Springs, USA. IEEE Computer Society, 2401-2408.

SINGH, G. D., MCNAMARA, J. A., JR. & LOZANOFF, S. 1997. Thin-plate spline analysis of the cranial base in subjects with Class III malocclusion. *European Journal of Orthodontics,* 19**,** 341-353.

SJOGREEN, L., LOHMANDER, A. & KILIARIDIS, S. 2011. Exploring quantitative methods for evaluation of lip function. *Journal of Oral Rehabilitation,* 38**,** 410-422.

SKINNER, M. & MULLEN, B. 1991. Facial asymmetry in emotional expression: A meta-analysis of research. *British Journal of Social Psychology,* 30**,** 113–124.

SNELL, R. S. 2008. The Head and Neck. *In:* MONTALBANO, J. (ed.) *Clinical Anatomy by Regions.* Baltimore: Lippincott Williams & Wilkins.

SPALL, J. C. 2004. Stochastic Optimization. *In:* GENTLE, J., HÄRDLE, W. & MORI, Y. (eds.) *Handbook of Computational Statistics.* New York: Springer-Verlag.

SRIRAGHAVENDRA, E., KARTHIK, K. & BHATTACHARYYA, C. Frechet distance based approach for searching online handwritten documents. 9th International Conference on Document Analysis and Recognition, 2007 Critiba, Brazil. IEEE Computer Society, 461-465.

STEVENS, J. 2009. Exploratory and Confirmatory Factor Analysis. *Applied Multivariate Statistics for the Social Sciences.* 5th ed.: Routledge Academic.

STEVENS, J. P. 2002. *Applied multivariate statistics for the social sciences,* Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

SUBTELNY, J. & OYA, N. 1972. Cineradiographic study of siblants. *Folia Phoniatrica et Logopedica,* 24**,** 30-50.

SUMMERFIELD, Q. A. 1979. Use of visual information for phonetic perception. *Phonetica,* 36**,** 314-331.

TABACHNICK, B. G. & FIDELL, L. S. 2007. Principal components and factor analysis. *In:* HARTMAN, S. (ed.) *Multivariate statistics.* 5th ed. Boston: Allyn & Bacon.

TALMANT, J. C. 2006. Evolution of the functional repair concept for cleft lip and palate patients. *Indian Journal of Plastic Surgery,* 39**,** 196-209.

THOMPSON, D. W. 1917. On growth and form. London: Cambridge University Press.

TOMA, A. M., ZHUROV, A., PLAYLE, R., ONG, E. & RICHMOND, S. 2009. Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. *Orthodontics and Craniofacial Research,* 12**,** 33-42.

TROTMAN, C. A., FARAWAY, J. J. & PHILLIPS, C. 2005. Visual and statistical modeling of facial movement in patients with cleft lip and palate. *Cleft Palate-Craniofacial Journal,* 42**,** 245-254.

TROTMAN, C. A., FARAWAY, J. J., PHILLIPS, C. D. & VAN AALST, J. 2010. Effects of lip revision surgery in cleft lip/palate patients. *Journal of Dental Research,* 89**,** 728-732.

TRPKOVA, B., MAJOR, P., PRASAD, N. & NEBBE, B. 1997. Cephalometric landmarks identification and reproducibility: A Meta analysis. *American Journal of Orthodontics and Dentofacial Orthopedics,* 112**,** 165-170.

TURK, M. A. & PENTLAND, A. P. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3**,** 71-86.

VAN SICKELS, J. E. & RICHARDSON, D. A. 1996. Stability of orthognathic surgery: a review of rigid fixation. *British Journal of Oral and Maxillofacial Surgery,* 34**,** 279-285.

VERZE, L., BIANCHI, F. A., DELL'ACQUA, A., PRINI, V. & RAMIERI, G. A. 2011. Facial mobility after bimaxillary surgery in class III patients: a three-dimensional study. *J Craniofac Surg,* 22**,** 2304-7.

VIEIRA, S. & CORRENTE, J. E. 2011. Statistical methods for assessing agreement between double readings of clinical measurements. *Journal of Applied Oral Science,* 19**,** 488-492.

WACHTMAN, G., COHN, J., VANSWEARINGEN, J. & MANDERS, E. 2001. Automated Tracking of Facial Features in Patients with Facial Neuromuscular Dysfunction. *Plastic & Reconstructive Surgery,* 107**,** 1124-1133.

WEEDEN, J. C., TROTMAN, C. A. & FARAWAY, J. J. 2001. Three dimensional analysis of facial movement in normal adults: influence of sex and facial shape. *Angle Orthodontist,* 71**,** 132-140.

WEINSTEIN, S., HARRIS, E. F. & ARCHER, S. Y. 1982. Lip morphology and area changes associated with surgical correction of mandibular prognathism. *Journal of Oral Rehabilitation,* 9**,** 335-354.

WILLIAMS, A. C., SHAH, H., SANDY, J. R. & TRAVESS, H. C. 2005. Patients' motivations for treatment and their experiences of orthodontic preparation for orthognathic surgery. *Journal of Orthodontics,* 32**,** 191-202.

WÖLFEL, M. & EKENEL , H. K. Feature Weighted Mahalanobis Distance: Improved Robustness for Gaussian Classifiers. 13th European Signal Processing Conference EUSIPCO, 2005. Citeseer, 1-4.

WOLFORD, L. M. 2007. Surgical Planning in Orthognathic Surgery. *In:* BOOTH, P. W., SCHENDEL, S. A. & HAUSAMEN, J. (eds.) *Maxillofacial Surgery.* St Louis: Churchill Livingstone.

WU, W., MALLET, Y., WALCZAK, B., PENNINCKX, W., MASSART, D. L., HEURDING, S. & ERINI, F. 1996. Comparison of regularised discriminant analysis, linear discriminant analysis and quadratic discriminant analysis, applied to NIR data. *Analytica Chimica Acta,* 329**,** 257-265.

WYATT, R., SELL, D., RUSSELL, J., HARLAND, K., HARDING, A. & ALBANY, E. 1996. Cleft palatal speech dissected : a review of current

knowledge and analysis. *British Journal of Plastic Surgery,* 49**,** 143-149.

XIE, Z. & FARIN, G. E. 2004. Image registration using hierarchical B-splines. *IEEE Transactions on Visualization and Computer Graphics,* 10**,** 85-94.

ZHAO, W., CHELLAPPA, R., PHILLIPS, P. J. & ROSENFIELD, A. 2003. Face Recognition: A Literature Survey. *ACM Computing Surveys,* 35**,** 399-458.

ZITOVA, B. & FLUSSER, J. 2003. Image registration methods: A survey. *Image and Vision Computing,* 21**,** 977-1000.

3DMD. 2005. *3dMD Launches Its Production 4D Technology-Based 3dMDface Dynamic System For New UK Customer, Cardiff University* [Online]. Available: http://www.3dmd.com/NewsEvents/news_article.asp?ID=128 [Accessed 19th January 2011].

Appendix A

Participant Information Sheets

Consent Forms

Cardiff University
Heath Park
Cardiff
CF14 4XY

University Dental Hospital
Heath Park
Cardiff
CF14 4XY

PARTICIPANT INFORMATION LEAFLET

3D Facial Changes following Jaw Surgery (Control)

We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen to you if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Part 1

What is the purpose of this study?
For most people talking is a normal facial movement but as different people say things differently, their faces will also move differently. The purpose of this part of the study is to use 3D images to see how people's faces move when they talk. This will help us to find out what average movement of the face is like. We will compare this average facial movement with the facial movement of people who are going to have jaw surgery.

During the study, we will be testing a new imaging system and working with 3D images for the first time.

Why have I been invited?
Adults who have never had, or who are not planning to have jaw surgery are being asked whether they wish to take part in this section of the study.

Do I have to take part?
It is up to you to decide. We will describe the study and go through this information sheet. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, without giving a reason. This would not affect the standard of any care you receive in the future.

What will happen to me if I take part?
The 3D imaging system consists of 6 cameras (3 either side of you).

You will be required to sit on a stool about 1.5 metres away from the cameras and say the words /puppy/ /rope/ /baby/ and /bob/ while we take the 3D image.

This will allow us to see how the face moves when saying different words.

The 3D image will only take 10 seconds to take and if you decide to take part, we will take an image of you at the start, one week and one month later. So will we have taken three images in total.

What are the possible risks of taking part?
There is no health risk associated with having the 3D images taken.

What are the possible benefits of taking part?
There is no intended benefit to you directly from taking part in the study but the information we obtain from this study will help us to better understand how different peoples' faces move when saying different words.

Will my taking part in this study be kept confidential?
Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

This completes Part 1

3D Facial Changes following Jaw Surgery (Control), V2, 17.11.09

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What will happen if I don't want to carry on with the study?
If you withdraw from the study we will need to use the data collected up to your withdrawal.

What if there is a problem?
If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions (02920 746734). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital.

Will my taking part in this study be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential. The 3D images taken of you will be stored securely and coded. Your name will be kept anonymous. This is in line with the Data Protection Act (1998). Only authorised persons such as the researchers and regulatory authorities will be allowed to view identifiable data.

The data will be held for 15 years after the end of the project which is line with Cardiff University's Data Protection Policy following which it will be destroyed.

What will happen to the results of the research?
The information will be published in 2013 as part of a PhD Thesis. You will not be identified in any report or publication. A copy of the published results can be sent to you if a request is made in writing.

Who is organising and funding the research?
Cardiff University are organising and funding this research study.

Who has reviewed the study?
All research is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This

study has been reviewed for conduct by the South East Wales Research Ethics Committee.

Further information and contact details

Hashmat Popat
Clinical Lecturer in Orthodontics
University Dental Hospital
Cardiff
CF14 4XY

Telephone: 02920 746734
Email: popath@cardiff.ac.uk

Thank you for your time.

Cardiff University   University Dental Hospital
Heath Park   Heath Park
Cardiff   Cardiff
CF14 4XY   CF14 4XY

PARTICIPANT INFORMATION LEAFLET

3D Facial Changes following Jaw Surgery

We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen to you if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Part 1

What is the purpose of this study?
At the moment there is very little information about what changes will happen to the way the face moves after jaw surgery. The purpose of this study is to investigate these changes by taking a three-dimensional (3D) image of the way the face moves before and after jaw surgery.

We will be testing a new imaging system and working with 3D images for the first time.

Why have I been invited?
Adults who are having braces and jaw surgery are being asked whether they wish to take part in this study.

Do I have to take part?
It is up to you to decide. We will describe the study and go through this information sheet. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, without giving a reason. This would not affect the standard of any care you receive in the future.

What will happen to me if I take part?
The 3D imaging system consists of 6 cameras (3 either side of you).

You will be required to sit on a stool about 1.5 metres away from the cameras and say the words /puppy/ /rope/ /baby/ and /bob/ while we take the 3D image.

This will allow us to see how the face moves when saying different words.

The 3D image will only take 10 seconds to take and if you decide to take part, we will take an image of you before you start your brace treatment, and again just before you have your jaw surgery and a final image 6 months after your surgery. So we will have taken 3 images in total.

What are the possible risks of taking part?
There is no health risk associated with having the 3D images taken.

What are the possible benefits?
There is no intended benefit to you directly from taking part in the study but the information we obtain from this study will help us to better understand what changes happen to the way the face moves after various types of jaw surgery.

Will my taking part in this study be kept confidential?
Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

This completes Part 1

3D Facial Changes following Jaw Surgery (Control), V2, 17.11.09

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What will happen if I don't want to carry on with the study?
If you withdraw from the study we will need to use the data collected up to your withdrawal.

What if there is a problem?
If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions (02920 746734). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital.

Will my taking part in this study be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential. The 3D images taken of you will be stored securely and coded. Your name will be kept anonymous. This is in line with the Data Protection Act (1998). Only authorised persons such as the researchers and regulatory authorities will be allowed to view identifiable data.

The data will be held for 15 years after the end of the project which is line with Cardiff University's Data Protection Policy following which it will be destroyed.

What will happen to the results of the research?
The information gained will be published in 2013 as part of a PhD Thesis. You will not be identified in any report or publication. A copy of the published results can be sent to you if a request is made in writing.

Who is organising and funding the research?
Cardiff University are organising and funding this research study.

Who has reviewed the study?
All research is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity.

3D Facial Changes following Jaw Surgery (Control), V2, 17.11.09

This study has been reviewed for conduct by the South East Wales Research Ethics Committee.
Further information and contact details

Hashmat Popat
Clinical Lecturer in Orthodontics
University Dental Hospital
Cardiff
CF14 4XY

Telephone: 02920 746734
Email: popath@cardiff.ac.uk

Thank you for your time.

CONDIFF
UNIVERSITY

PRIFYSGOL
CAERDYD

NHS
WALES
GIG
CYMRU

Cardiff University   University Dental Hospital
Heath Park   Heath Park
Cardiff   Cardiff
CF14 4XY   CF14 4XY

CONSENT FORM

Title of Project: 3D Facial Changes following Jaw Surgery (Control)

Name of Researcher: Hashmat Popat
                       Clinical Lecturer in Orthodontics
                       University Dental Hospital
                       Cardiff
                       CF14 4XY

Please initial box

1.   I confirm that I have read and understand the participant information sheet
      (control) V2, 17.11.09 for the above study and have had the opportunity to ask
      questions and have had them answered satisfactorily.

2.   I understand that my participation is voluntary and that I am free to withdraw at
      any time, without giving any reason, without my medical care or legal rights being
      affected.

3.   I understand and relevant sections of my medical notes and data collected during
      the study may be looked at by responsible individuals from the research team,
      regulatory authorities or from Cardiff and Vale NHS Trust, where it is relevant to
      me taking part in this research. I give permission for these individuals to have
      access to my records.

4.   I agree to take part in the above study

_____ _____ _____
Name of Participant   Date   Signature

_____ _____ _____
Name of Person taking consent Date   Signature
(if different from researcher)

_____ _____ _____
Researcher   Date   Signature

3D Facial Changes following Jaw Surgery (Control), V2, 17.11.09

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYD

NHS WALES GIG CYMRU

Cardiff University
Heath Park
Cardiff
CF14 4XY

University Dental Hospital
Heath Park
Cardiff
CF14 4XY

CONSENT FORM

Title of Project: 3D Facial Changes following Jaw Surgery

Name of Researcher: Hashmat Popat
Clinical Lecturer in Orthodontics
University Dental Hospital
Cardiff
CF14 4XY

Please initial box

1.   I confirm that I have read and understand the participant information sheet V2, 17.11.09 for the above study and have had the opportunity to ask questions and have had them answered satisfactorily.

2.   I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

3.   I understand and relevant sections of my medical notes and data collected during the study may be looked at by responsible individuals from the research team, regulatory authorities or from Cardiff and Vale NHS Trust, where it is relevant to me taking part in this research. I give permission for these individuals to have access to my records.

4.   I agree to take part in the above study

_____   _____   _____
Name of Participant   Date   Signature

_____   _____   _____
Name of Person taking consent Date   Signature
(if different from researcher)

_____   _____   _____
Researcher   Date   Signature

3D Facial Changes following Jaw Surgery (Control), V2, 17.11.09

Appendix B

Tests of normality on WDDTW data (Section 7.2.2)
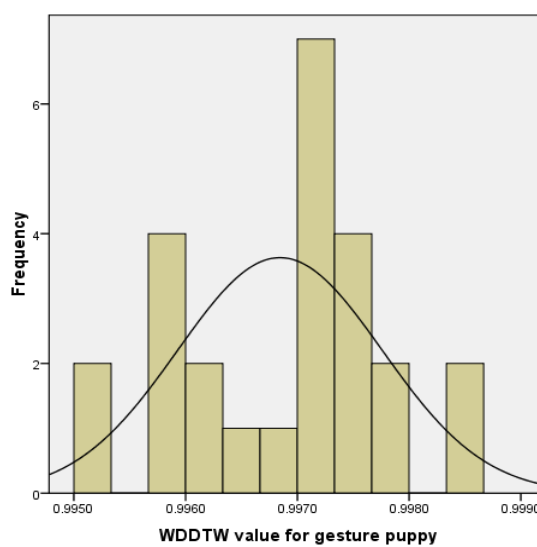
Figure B-1   *Frequency histogram of WDDTW scores for puppy*
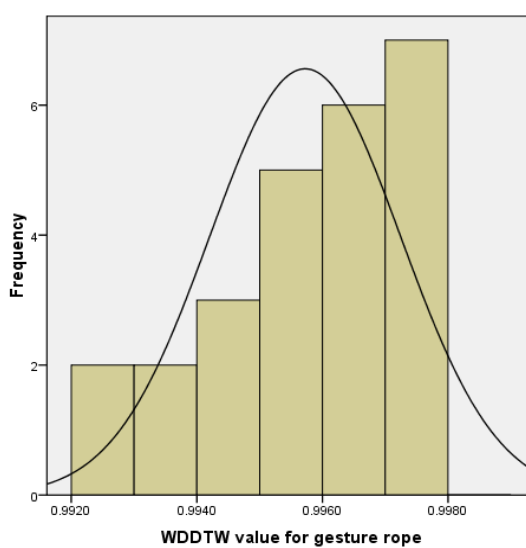
Figure B-2   *Frequency histogram of WDDTW scores for rope*
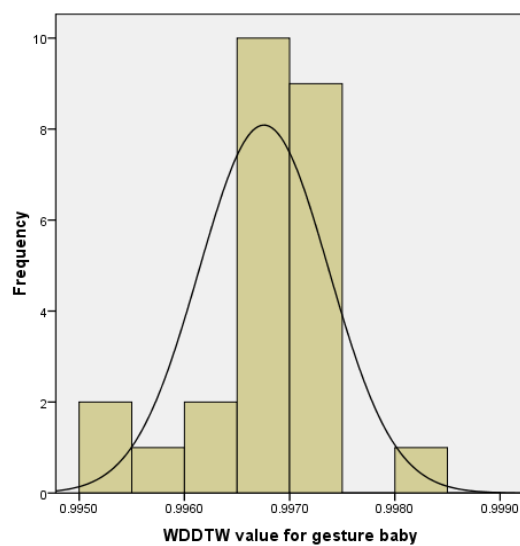
Figure B-3   *Frequency histogram of WDDTW scores for baby*
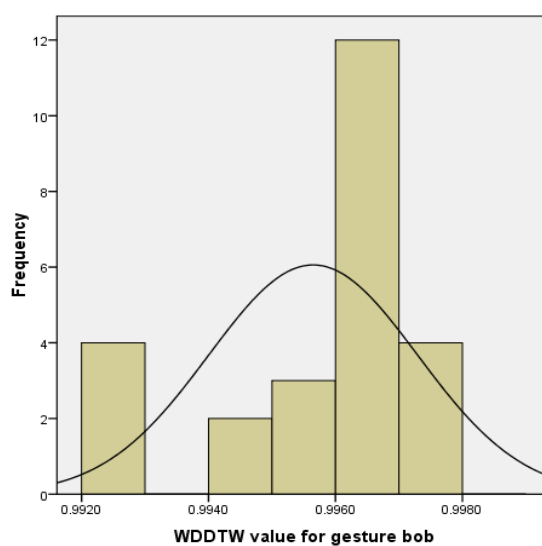
Figure B-4   *Frequency histogram of WDDTW scores for bob*
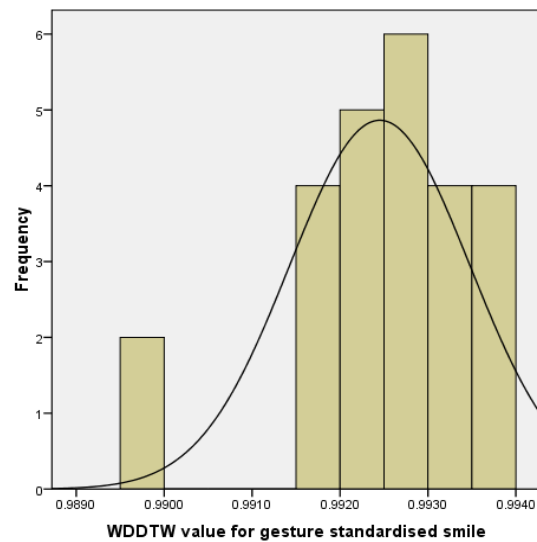
Figure B-5 *Frequency histogram of WDDTW scores for standardised smile*
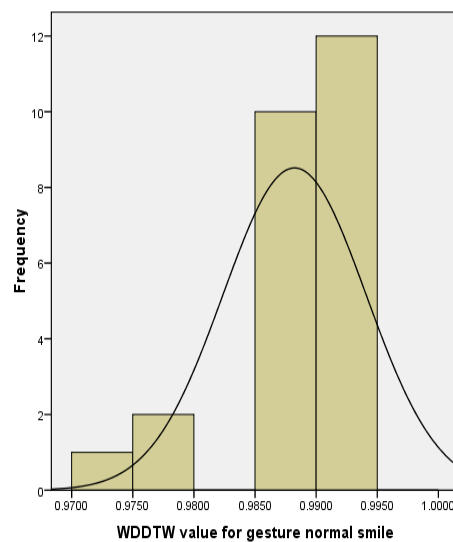
Figure B-6 *Frequency histogram of WDDTW scores for normal smile*

Table B-1 *Shapiro-Wilk test of normality for WDDTW data*

Gesture (WDDTW value)	Shapiro-Wilk
Puppy	0.79
Rope	0.10
Baby	0.84
Bob	0.07
Standardised Smile	0.05
Normal Smile	0.70

Table B-2   *Repeated measures ANOVA power calculation*

Parameter	Value
Statistically significant difference in WDDTW	0.001
Standard deviation	0.002
α error probability	0.05
Power (1 - β error probability)	0.80
Number of groups	6
Repetitions	2
Critical F	3.11
Total sample size	18
Actual Power	0.97

Appendix C

Data preparation for DFA

Table C-1   *Normality tests for viseme landmarks puppy*

Coordinate	CG			PG_{pre}			PG_{post}		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	-0.25	0.81	0.65	0.22	-1.50	0.05	-0.11	-0.21	0.99
ls Y	0.02	0.32	0.88	-0.08	0.64	0.33	0.69	0.46	0.48
ls Z	0.14	-0.42	0.37	0.18	-0.59	0.75	0.52	-0.51	0.39
li X	-0.04	-0.13	0.78	0.11	0.34	0.98	0.22	-0.30	0.52
li Y	-0.53	0.39	0.16	0.23	0.61	0.58	-0.07	-0.93	0.70
li Z	0.40	-0.18	0.36	0.46	-1.04	0.07	-0.12	-0.92	0.49
cphL X	0.12	0.15	0.65	-0.43	-0.25	0.87	0.14	-0.95	0.87
cphL Y	-0.17	0.26	0.98	-0.17	-1.38	0.12	0.42	-1.24	0.05
cphL Z	-0.06	-0.51	0.03*	-0.26	0.52	0.57	0.24	-1.22	0.13
cphR X	-0.18	0.25	0.87	-0.66	0.68	0.37	0.14	-1.17	0.36
cphR Y	0.40	-0.16	0.32	-0.43	-0.37	0.68	0.19	-1.21	0.10
cphR Z	-0.06	-0.17	0.99	0.04	-0.85	0.77	0.64	-0.45	0.15
chL X	-0.02	1.49	0.39	0.06	-1.16	0.22	-0.60	-0.84	0.04*
chL Y	-0.20	0.07	0.47	-0.19	-0.58	0.70	-0.52	-0.74	0.23
chL Z	0.36	-0.78	0.02*	-0.11	0.05	0.34	1.37	2.60	0.02*
chR X	0.55	1.03	0.08	0.07	0.70	0.64	0.17	0.18	0.94
chR Y	0.18	-0.53	0.45	0.25	0.15	0.80	-0.37	-1.12	0.15
chR Z	0.40	0.28	0.61	0.92	2.24	0.25	0.31	-0.34	0.82

*   *Statistically significant (p < 0.05)*

Table C-2   *Normality tests for viseme landmarks puppy*

Coordinate	CG			PG_pre			PG_post		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	0.36	2.51	0.06	1.14	0.68	0.01*	1.06	1.49	0.05
ls Y	0.06	-0.12	0.64	0.83	0.29	0.07	-0.22	0.02	0.58
ls Z	-0.14	-0.69	0.33	0.10	-1.28	0.11	0.29	-0.26	0.97
li X	-0.09	0.13	0.78	-0.23	0.73	0.83	0.58	-1.02	0.06
li Y	0.04	1.98	0.07	0.06	1.33	0.71	0.50	-0.32	0.59
li Z	0.47	0.23	0.22	0.38	-0.33	0.23	-0.57	0.09	0.74
cphL X	0.55	0.72	0.19	0.51	-0.56	0.12	0.02	-0.31	0.63
cphL Y	0.08	0.66	0.76	-0.03	-0.28	0.96	-0.42	0.26	0.36
cphL Z	0.18	0.77	0.26	0.24	-0.44	0.96	-0.35	0.53	0.86
cphR X	-0.64	0.13	0.01	-0.39	-0.09	0.89	-0.07	-0.87	0.72
cphR Y	0.01	0.48	0.93	0.81	-0.15	0.08	-0.13	-0.58	0.61
cphR Z	0.15	-0.20	0.83	-0.23	-.046	0.70	0.89	2.94	0.18
chL X	0.12	0.68	0.34	0.22	0.74	0.76	-0.14	1.59	0.05
chL Y	-0.36	-0.73	0.07	-0.05	-0.32	0.99	-0.25	-0.93	0.55
chL Z	0.36	0.85	0.12	0.12	-0.55	0.99	1.36	2.53	0.01*
chR X	1.47	6.30	0.01*	0.24	0.47	0.75	-0.71	0.84	0.50
chR Y	-0.01	0.14	0.99	0.25	0.55	0.86	-1.23	1.16	0.06
chR Z	0.31	-0.03	0.28	-0.72	-0.26	0.14	-1.22	2.42	0.07

*   *Statistically significant (p < 0.05)*

Table C-3   *Normality tests for viseme landmarks ̱rope*

Coordinate	CG			PG$_{pre}$			PG$_{post}$		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	-0.29	0.59	0.54	0.37	-0.84	0.23	0.08	-0.61	0.68
ls Y	-0.18	0.02	0.92	-0.19	-0.71	0.47	-0.31	-0.58	0.63
ls Z	-0.34	1.06	0.30	0.30	-1.03	0.18	-0.02	-1.18	0.40
li X	0.59	0.87	0.05	-1.48	4.00	0.01*	0.01	-0.40	0.67
li Y	0.57	1.79	0.03	0.25	1.05	0.54	-0.34	1.16	0.68
li Z	0.29	0.02	0.56	-0.28	0.38	0.74	-0.06	-0.81	0.72
cphL X	-0.13	0.03	0.89	0.64	0.09	0.45	0.55	-0.65	0.16
cphL Y	-0.44	0.24	0.30	-0.79	0.49	0.18	-0.26	2.84	0.62
cphL Z	0.18	0.69	0.65	0.11	-0.70	0.37	0.11	-0.70	0.15
cphR X	0.39	0.07	0.43	0.48	-0.06	0.78	-0.17	-0.31	0.20
cphR Y	-0.39	-0.15	0.28	-0.59	0.46	0.25	0.59	0.80	0.75
cphR Z	-0.09	-0.44	0.16	0.38	-0.69	0.33	4.47	2.00	0.54
chL X	0.03	0.03	0.94	-0.01	0.38	0.53	0.18	-0.51	0.07
chL Y	0.24	0.47	0.48	-0.43	0.43	0.89	-0.36	0.12	0.98
chL Z	0.15	-0.46	0.70	0.07	-1.04	0.64	0.94	0.78	0.94
chR X	0.52	0.30	0.24	-0.46	-0.55	0.42	0.62	0.58	0.65
chR Y	0.32	0.88	0.47	-0.68	0.37	0.55	-0.01	-0.34	0.98
chR Z	-0.12	-0.75	0.42	0.21	-0.47	0.98	-0.26	-1.26	0.14

*   Statistically significant (p < 0.05)*

Table C-4   *Normality tests for viseme landmarks b̲aby*

Coordinate	CG			PG$_{pre}$			PG$_{post}$		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	0.15	0.84	0.43	0.63	0.64	0.23	-0.78	0.48	0.34
ls Y	-0.34	0.09	0.42	-0.58	-0.89	0.06	-0.41	-0.16	0.82
ls Z	-0.79	1.62	0.72	0.01	-1.34	0.25	0.31	0.62	0.15
li X	0.05	-0.92	0.05	0.88	0.03	0.06	0.24	-0.03	0.83
li Y	-0.01	-0.63	0.18	0.35	0.52	0.62	-0.27	-0.31	0.75
li Z	0.18	-0.49	0.55	0.27	-0.78	0.73	-0.49	-0.93	0.15
cphL X	-0.10	-0.44	0.67	0.48	0.94	0.59	0.61	-0.06	0.37
cphL Y	-0.34	-0.58	0.18	0.12	-1.30	0.16	0.11	-1.21	0.34
cphL Z	0.66	1.43	0.14	-0.25	-0.41	0.87	-0.36	0.43	0.28
cphR X	-0.02	-0.59	0.55	-0.34	-10.07	0.10	-5.38	0.00	0.54
cphR Y	0.21	0.49	0.55	0.30	-0.92	0.46	0.41	-0.27	0.86
cphR Z	0.25	0.12	0.75	0.64	0.33	0.37	0.05	-0.01	0.77
chL X	0.07	-0.45	0.50	-1.32	2.12	0.05	0.17	-1.06	0.37
chL Y	-0.23	-0.36	0.56	-0.08	-0.43	0.92	-0.61	-0.20	0.79
chL Z	0.38	-0.35	0.22	0.56	0.49	0.64	1.37	3.39	0.02*
chR X	0.78	1.90	0.21	0.21	0.61	0.55	-0.57	0.44	0.59
chR Y	-0.20	0.40	0.52	-0.06	0.14	0.37	-0.92	-0.27	0.34
chR Z	-0.09	-0.45	0.33	0.10	-0.64	0.75	-0.57	-0.28	0.34

*   *Statistically significant (p < 0.05)*

Table C-5 *Normality tests for viseme landmarks baby*

Coordinate	CG			PG$_{pre}$			PG$_{post}$		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	-0.40	0.46	0.52	-0.01	0.71	0.69	-0.30	0.62	0.60
ls Y	-0.66	0.67	0.08	0.38	-1.27	0.06	0.51	0.36	0.70
ls Z	-0.33	0.16	0.58	0.17	-0.86	0.49	-0.13	-0.42	0.56
li X	0.03	-0.05	0.59	-0.45	-0.90	0.17	-1.05	1.79	0.02*
li Y	0.12	-0.22	0.58	-1.06	1.09	0.09	-0.71	-0.20	0.10
li Z	0.30	-0.67	0.05	0.26	-0.76	0.52	-0.14	-1.33	0.16
cphL X	-0.04	0.01	0.68	0.34	-0.28	0.42	0.73	-0.09	0.09
cphL Y	0.13	0.55	0.76	0.63	0.36	0.33	0.54	0.86	0.36
cphL Z	0.27	-0.01	0.81	-0.20	-1.29	0.14	-0.28	-0.55	0.84
cphR X	0.07	-0.02	0.77	-1.46	2.56	0.01*	-0.58	-1.10	0.05
cphR Y	0.28	1.71	0.06	1.35	2.38	0.05	0.57	0.34	0.64
cphR Z	-0.24	0.91	0.20	0.32	-1.12	0.05	-0.09	-1.32	0.18
chL X	0.04	0.41	0.10	-1.52	1.80	0.00*	-1.03	1.19	0.12
chL Y	0.03	-0.52	0.50	0.29	-0.66	0.80	0.18	-0.71	0.90
chL Z	0.22	-0.47	0.45	0.37	0.34	0.25	0.29	-0.64	0.68
chR X	0.16	1.16	0.06	1.28	2.59	0.06	-1.05	1.50	0.12
chR Y	0.38	-0.47	0.11	0.23	1.06	0.71	-0.12	-0.48	0.88
chR Z	-0.03	-0.33	0.59	-0.23	-1.67	0.01*	0.04	-0.36	0.66

* *Statistically significant (p < 0.05)*

Table C-6   *Normality tests for viseme landmarks b<u>o</u>b*

Coordinate	CG			PG_{pre}			PG_{post}		
	Skewness	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**	**Skewness**	**Kurtosis**	**Sig.**
ls X	-0.01	-0.36	0.91	-0.55	1.27	0.64	0.35	-0.73	0.66
ls Y	-0.16	-0.04	0.72	-0.59	0.62	0.37	-0.73	-0.32	0.09
ls Z	-1.20	4.23	0.00*	-0.54	0.43	0.47	0.17	-0.98	0.52
li X	0.31	0.37	0.50	-0.59	1.04	0.52	-0.75	3.49	0.02
li Y	0.51	0.47	0.25	-0.74	-0.40	0.10	-0.63	0.49	0.43
li Z	0.16	0.16	0.67	-0.06	-0.29	0.79	0.20	0.50	0.99
cphL X	-0.10	0.52	0.31	0.56	1.16	0.26	0.28	-0.72	0.72
cphL Y	-0.01	-0.51	0.85	0.78	0.18	0.29	0.47	2.10	0.42
cphL Z	0.18	-0.39	0.75	2.16	3.02	0.00*	-0.07	-1.34	0.12
cphR X	-0.36	0.20	0.45	-1.12	1.37	0.05	0.44	-0.31	0.66
cphR Y	-0.15	-0.54	0.55	0.41	-0.75	0.34	-0.07	2.33	0.20
cphR Z	-0.02	-0.19	0.82	2.16	3.08	0.00*	0.20	0.13	0.89
chL X	-0.28	0.31	0.67	-0.48	-0.64	0.21	-1.30	3.23	0.09
chL Y	-0.01	0.66	0.85	0.30	-0.09	0.56	0.26	-0.07	0.96
chL Z	0.54	0.07	0.10	1.21	1.73	0.06	0.31	-0.74	0.67
chR X	0.44	0.38	0.46	-0.30	-0.39	0.89	-0.12	0.12	0.66
chR Y	0.28	0.01	0.81	-0.96	1.79	0.21	-0.22	0.06	0.92
chR Z	0.12	-0.44	0.75	0.20	-1.37	0.08	0.06	1.69	0.05

*   *Statistically significant (p < 0.05)*

Table C-7 *Check for univariate outliers viseme landmarks puppy*

Landmark	CG Mean	CG 5% Trimmed Mean	PG_{pre} Mean	PG_{pre} 5% Trimmed Mean	PG_{post} Mean	PG_{post} 5% Trimmed Mean
ls X	0.17	0.18	0.18	0.18	0.29	0.29
ls Y	7.78	7.78	7.02	7.01	7.60	7.54
ls Z	7.06	7.05	6.50	6.49	7.07	7.05
li X	-0.09	0.04	0.22	0.21	0.15	0.15
li Y	-16.75	-16.66	-16.18	-16.25	-16.76	-16.75
li Z	3.82	3.79	4.49	4.46	3.71	3.70
cphL X	5.65	6.64	5.82	5.84	6.38	6.38
cphL Y	9.79	9.80	9.10	9.11	9.87	9.84
cphL Z	6.23	6.22	5.76	5.76	6.15	6.13
cphR X	-5.43	-5.42	-5.90	-5.87	-5.89	-.589
cphR Y	9.86	9.83	9.24	9.26	9.96	9.96
cphR Z	6.05	6.05	5.57	5.57	6.02	6.01
chL X	27.11	27.12	28.09	28.08	26.79	26.85
chL Y	-5.34	-5.32	-4.67	-4.66	-5.25	-5.24
chL Z	-11.36	-11.37	-10.98	-10.95	-11.10	-11.16
chR X	-27.56	-27.61	-28.41	-28.42	-27.73	-27.74
chR Y	-5.34	-5.34	-4.50	-4.50	-5.41	-5.40
chR Z	-11.80	-11.83	-11.36	-11.40	-11.86	-11.88

Table C-8 *Check for univariate outliers viseme landmarks puppy*

Landmark	CG Mean	CG 5% Trimmed Mean	PG_{pre} Mean	PG_{pre} 5% Trimmed Mean	PG_{post} Mean	PG_{post} 5% Trimmed Mean
ls X	0.28	0.27	0.09	0.06	0.35	0.32
ls Y	7.20	7.19	6.72	6.69	6.93	6.93
ls Z	6.54	6.55	6.14	6.14	6.63	6.62
li X	0.05	0.06	-0.03	-0.01	0.21	0.21
li Y	-15.98	-15.97	-15.63	-15.64	-15.93	-15.35
li Z	3.87	3.83	4.70	4.68	4.08	4.13
cphL X	5.92	5.89	6.09	6.07	6.55	6.53
cphL Y	9.04	9.03	8.55	8.56	8.93	8.69
cphL Z	5.61	5.61	5.46	5.45	5.76	5.78
cphR X	-5.32	-5.29	-6.11	-6.08	-5.70	-5.70
cphR Y	9.19	9.18	8.70	8.66	9.11	9.08
cphR Z	5.54	5.54	5.26	5.36	5.62	5.58
chL X	28.06	28.05	27.21	27.01	27.80	27.89
chL Y	-4.58	-4.56	-4.21	-4.20	-4.63	-4.61
chL Z	-10.38	-10.40	-10.91	-10.88	-10.51	-10.54
chR X	-29.25	-29.05	-29.89	-29.32	-29.23	-29.13
chR Y	-5.05	-4.87	-4.14	-4.16	-5.03	-4.97
chR Z	-11.19	-11.21	-10.91	-10.88	-11.59	-11.52

Table C-9 *Check for univariate outliers viseme landmarks <u>rope</u>*

Landmark	CG Mean	CG 5% Trimmed Mean	PG_{pre} Mean	PG_{pre} 5% Trimmed Mean	PG_{post} Mean	PG_{post} 5% Trimmed Mean
ls X	0.31	0.32	0.32	0.30	0.37	0.36
ls Y	8.51	8.52	8.26	8.25	7.00	7.01
ls Z	7.45	7.47	7.52	7.49	6.00	5.98
li X	-0.02	-0.05	0.04	0.09	-0.05	0.04
li Y	-17.69	-17.76	-17.20	-17.22	-16.32	-16.26
li Z	2.51	2.47	2.58	2.63	4.50	4.52
cphL X	6.41	6.42	7.45	7.40	7.04	6.99
cphL Y	10.41	10.45	10.32	10.34	9.05	9.06
cphL Z	6.23	6.22	6.22	6.20	5.42	5.45
cphR X	-5.78	-5.82	-6.73	-6.78	-6.88	-6.85
cphR Y	10.57	10.61	10.44	10.47	9.30	9.18
cphR Z	6.49	6.48	6.26	6.25	2.12	2.00
chL X	25.85	25.58	25.88	25.86	27.69	27.67
chL Y	-5.75	-5.77	-5.75	-5.70	-4.41	-4.38
chL Z	-10.84	-10.85	-10.83	-10.85	-10.79	-10.84
chR X	-26.77	-26.82	-26.95	-26.91	-28.28	-28.34
chR Y	-6.05	-6.06	-6.07	-6.01	-4.52	-4.51
chR Z	-11.65	-11.64	-11.72	-11.73	-11.32	-11.31

**Table C-10** *Check for univariate outliers viseme landmarks <u>baby</u>*

Landmark	CG Mean	CG 5% Trimmed Mean	PG_{pre} Mean	PG_{pre} 5% Trimmed Mean	PG_{post} Mean	PG_{post} 5% Trimmed Mean
ls X	0.19	0.19	0.14	0.13	0.42	0.25
ls Y	8.13	8.15	6.98	7.01	7.99	8.01
ls Z	7.19	7.20	6.40	6.40	7.22	7.22
li X	-0.06	-0.03	-0.18	-0.21	-0.078	-0.07
li Y	-17.21	-17.21	-16.50	-16.57	-17.00	-17.11
li Z	3.37	3.36	4.63	4.60	3.66	3.68
cphL X	5.74	5.75	5.71	5.69	6.78	6.76
cphL Y	10.07	10.09	9.03	9.01	10.00	9.90
cphL Z	6.22	6.22	5.58	5.58	6.21	6.23
cphR X	-5.49	-5.44	-5.56	-5.54	-5.98	-5.95
cphR Y	10.11	10.12	9.19	9.18	10.05	10.02
cphR Z	6.22	6.21	5.43	5.42	6.29	6.30
chL X	26.91	26.90	28.36	28.53	26.63	26.31
chL Y	-5.53	5.52	-4.30	-4.30	-5.29	-5.27
chL Z	-11.29	-11.31	-10.88	-10.93	-11.23	-11.32
chR X	-27.29	-27.32	-28.46	-28.47	-27.47	-27.42
chR Y	-5.58	-5.57	-4.34	-4.35	-5.74	-5.70
chR Z	-11.72	-11.72	-11.17	-11.16	-12.16	-12.14

Table C-11   *Check for univariate outliers viseme landmarks baby*

Landmark	CG Mean	CG 5% Trimmed Mean	PG$_{pre}$ Mean	PG$_{pre}$ 5% Trimmed Mean	PG$_{post}$ Mean	PG$_{post}$ 5% Trimmed Mean
ls X	0.07	0.83	0.07	0.07	0.36	0.38
ls Y	7.61	7.65	6.45	6.43	7.35	7.33
ls Z	7.20	7.20	6.21	6.20	7.30	7.32
li X	-0.22	-0.22	-0.25	-0.23	-0.36	-0.34
li Y	-15.54	-15.53	-15.42	-15.30	-15.11	-15.02
li Z	3.61	3.60	5.11	5.10	3.80	3.80
cphL X	5.87	5.88	5.52	5.51	6.55	6.51
cphL Y	9.57	9.56	8.56	8.55	9.61	9.58
cphL Z	6.30	6.29	5.59	5.60	6.37	6.37
cphR X	-5.73	-5.74	-5.50	-5.40	-5.91	-5.89
cphR Y	9.70	9.69	8.73	8.67	9.67	9.63
cphR Z	6.30	6.34	5.30	5.30	640	6.41
chL X	27.67	27.66	28.90	28.96	27.27	27.35
chL Y	-5.65	-5.66	-4.14	-4.16	-5.53	-5.53
chL Z	-11.60	-11.61	-11.06	-11.12	-11.62	-11.65
chR X	-27.66	-27.72	-28.73	-28.82	-27.91	-27.85
chR Y	-5.69	-5.71	-4.18	-4.19	-5.99	-5.97
chR Z	-11.81	-11.82	-11.16	-11.16	-12.26	-12.28

Table C-12   *Check for univariate outliers viseme landmarks bob*

Landmark	CG Mean	CG 5% Trimmed Mean	PG$_{pre}$ Mean	PG$_{pre}$ 5% Trimmed Mean	PG$_{post}$ Mean	PG$_{post}$ 5% Trimmed Mean
ls X	0.29	0.29	0.43	0.44	0.47	0.46
ls Y	8.57	8.59	7.23	7.26	8.32	8.34
ls Z	7.57	7.60	6.78	6.80	7.42	7.41
li X	0.15	0.14	-0.12	-0.09	0.18	0.18
li Y	-17.84	-17.84	-16.72	-16.63	-17.62	-17.55
li Z	2.85	2.83	4.22	4.22	2.93	2.92
cphL X	6.08	6.09	6.16	6.01	7.23	7.21
cphL Y	10.63	10.63	9.45	9.42	10.53	10.51
cphL Z	6.41	6.40	2.52	2.37	6.43	6.42
cphR X	-5.68	-5.66	-5.68	-5.61	-6.40	-6.43
cphR Y	10.63	10.65	9.55	9.53	10.66	10.67
cphR Z	6.50	6.50	2.51	2.35	6.49	6.48
chL X	25.55	25.58	27.49	27.55	25.38	25.48
chL Y	-5.96	-5.97	-4.58	-4.60	-5.75	-5.75
chL Z	-11.31	-11.35	-10.87	-10.94	-11.09	-11.11
chR X	-26.40	-26.46	-28.28	-28.24	-26.86	-26.84
chR Y	-6.03	-6.04	-4.93	-4.86	-6.13	-6.12
chR Z	-12.02	-12.03	-11.55	-11.56	-12.19	-12.19
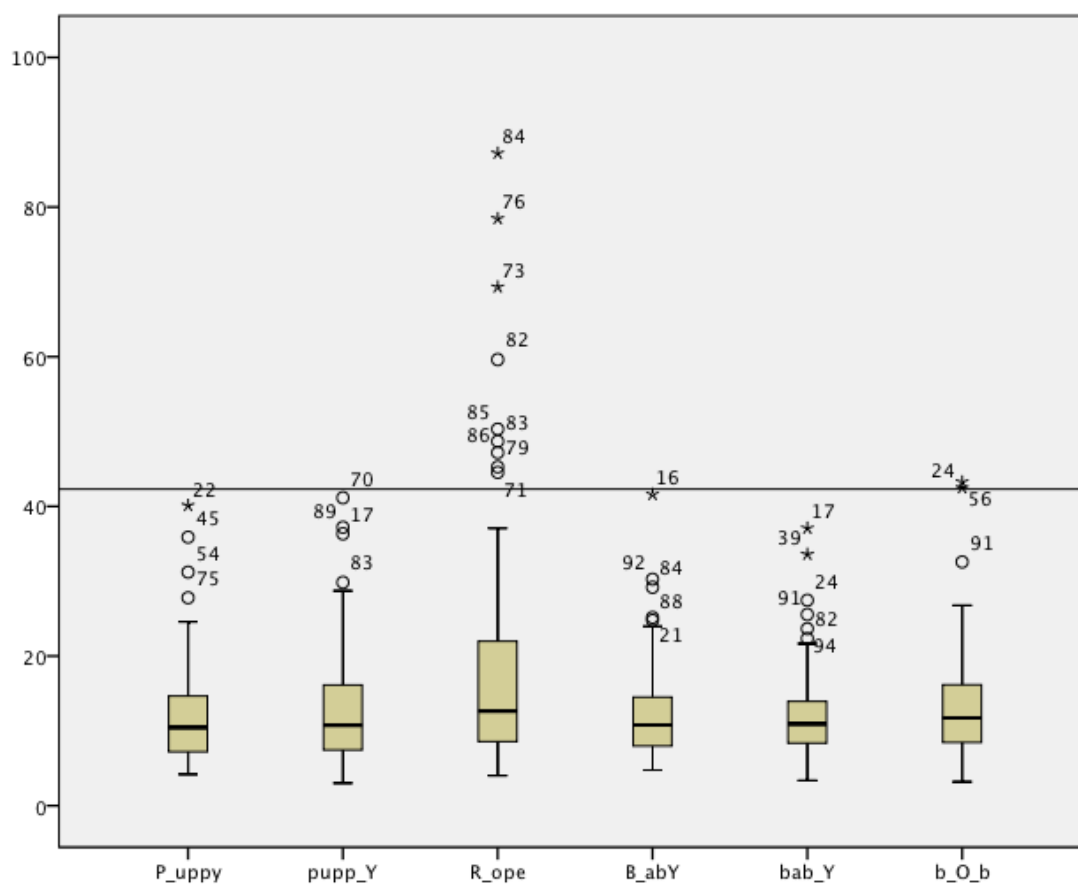
Figure C-1 *Boxplots of multivariate outliers by viseme (threshold 42.3)*

Table C-13 *Significance of Box's M Test*

Viseme	Box's M Sig.	
	CG v PG_{pre}	CG v PG_{post}
p̲uppy	0.078	0.762
pupp̲y	0.035	0.029
r̲ope	0.026	0.045
b̲aby	0.003	0.032
bab̲y	0.013	0.678
b̲o̲b	0.002	0.035

Threshold $p < 0.001$