# CREATING CROWD CHARACTERS THROUGH PROCEDURAL DEFORMATION

A Thesis

by

NATHAN M. LINDIG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

| | |
|---|---|
| Chair of Committee, | Louis G. Tassinary |
| Committee Members, | Tim McLaughlin |
| | John C. Keyser |
| Head of Department, | Wenping Wang |

May 2021

Major Subject: Visualization

ABSTRACT

Crowd simulations are a popular narrative tool in the entertainment industry and rely heavily on crowd characters for their effectiveness. The traditional processes employed for modeling characters are expensive and impractical for creating crowd character models, leading to the development and use of combinatorial methods. While combinatorial methods are effective at creating large numbers of unique character variations, they lack the flexibility to create characters outside the range of possible variations. This limitation is particularly troublesome in production circumstances that entail frequent fluctuations in character asset requests and in real-time interactive experiences. This is primarily because it is difficult to predetermine the various states in which crowd characters will be rendered. This thesis explores a procedural, parametrically driven system for creating crowd characters that allows for greater flexibility in changing scenarios. This system uses craniofacial anthropometric data - sourced from existing literature - to create a multidimensional parametric structure from which procedurally randomized values can be generated and selected. Once new parametric values are generated, a base craniofacial model is deformed accordingly, resulting in a unique model that represents the parameter values assigned. The results of this method demonstrate the ability of a procedurally driven, craniofacial anthropometric facial deformation system to rapidly and efficiently create a large set of unique crowd character facial models that share a single UV map and polygonal topology. The results of this research indicate such a procedurally driven crowd character creation system would provide artists with the flexibility to quickly iterate and produce desirable results in scenarios such as interactive games.

# DEDICATION

To my loving wife Ashley and beautiful daughter Thea whose support and patience empowered me to persevere to the completion of this thesis.

# CONTRIBUTORS AND FUNDING SOURCES

## Contributors

This work was supported by a thesis committee consisting of Dr. Louis G Tassinary (advisor) and Professor Tim McLaughlin of the Department of Visualization and Dr. John C. Keyser of the Department of Computer Science and Engineering.

The data analyzed for Chapter 3 was publicly available.

All work conducted for the thesis was completed by the student independently under the direct supervision of Dr. Tassinary.

## Funding Sources

No outside source of funding was provided.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

# 1. INTRODUCTION AND LITERATURE REVIEW

## 1.1 Introduction

Human crowd simulations have become a commonly used narrative element for many computer generated stories and experiences. Because crowds are a common shared experience for many individuals, they serve as an effective tool for enhancing the immersiveness of a story. Crowds provide a reference point from which people can more effectively engage with a story's narrative.

Efficient and believable crowd simulations are difficult and remain an active research in computer graphics both commercially and academically. Crowd simulations can serve many purposes and each purpose entails different challenges that range from technical to artistic in nature. For example, a crowd simulation for the purposes of evacuation and disaster scenario planning entails technical challenges specific to accurately recreating the complex behaviors and movement of individual agents within a crowd in relation to specific environmental factors (Sticco, Frank, & Dorso, 2020; Ju et al., 2010). A crowd simulation for the purposes of entertainment and storytelling entails a combination of technical and artistic challenges specific to the creation of a crowd that is both aesthetically pleasing and believable (El-Ali et al., 2016; Yen, Gustafson, Lo, Northrup, & Sun, 2018).

Some of the research on this topic focuses on the methods and systems used to simulate the movement and behaviors of individual agents within the crowd (Bandini, Manzoni, & Vizzari, 2004, 2006; Bandini, Federici, & Vizzari, 2007; Braun, Musse, de Oliveira, & Bodmann, 2003; Durupinar, Allbeck, Pelechano, & Badler, 2008; Thalmann, 2007; Ulicny, de Heras Ciechomski, & Thalmann, 2004; Zhou et al., 2010), while other research in the field focuses on enhancing and optimizing the production methods for creating aesthetically pleasing and believable crowds for film and video games (Mourino et al., 2017).

The present thesis proposes a method for reducing the production costs involved in the creation of crowd simulations for various CG entertainment fields such as film, games, and virtual reality. Crowd simulations created for entertainment necessarily focus on various aesthetic qualities, resulting in unique challenges that alter and often increase production costs.

The process of creating character models for crowd simulations typically require a significant degree of pre-planning. The primary objective of this pre-planning is to achieve the most effective result with the least amount of resource expenditure (A. Sidenblad, personal communication, September 27, 2020). In determining how best to achieve an effective result in a cost effective manner, the desired visual fidelity is the first objective that must be specified. Several factors are considered when determining the desirable fidelity for a crowd simulation. These are the representational style (e.g., realistic vs non-realistic), family resemblance (e.g., degree of variety), screen time and screen space accumulation.

### 1.1.1 Family Resemblance

One of the most significant factors related to the effectiveness of the resulting crowd simulation is the level of perceived variety among the crowd's members. The level of variety deemed desirable will vary based on the narrative requirements.. For example, a level of variety in which the crowd members appear to be clones of each other may be off-putting to the audience in one context yet perfectly acceptable in a different context (e.g., Star Wars: Attack of the Clones). In addition, the desirable level of variety may often be dictated by where the narrative directs the viewer's attention. In industries where each decision is evaluated in terms of its effect on the final visual result, it is sometimes difficult to allocate resources to a task for which the mark of success is for the results of the task to go unnoticed.

### 1.1.2 Screen Time and Screen Space Contribution

The desired level of fidelity is also dependent on the factors of screen time and screen space contribution. The factor of screen time describes how often and for how long the asset (i.e., any 3D digital object) is seen by the viewer. The factor of screen space accumulation describes what percentage of the screen space does the asset fill. As an asset moves closer to the camera, or vice versa, the more screen space the asset accumulates. These factors impact the viewer's ability to discern and assess the level of visual variety in the crowd. The level of detail discernible to the viewer increases the longer an asset is on screen and the closer the asset is to the camera.

In traditional film production, the artists of the project have complete control over the final visual result viewable to an audience whereas factors such as screen space accumulation and total screen time are determined by the production team of the film and cannot typically be altered by the viewer. In contrast, artists involved in interactive CG experiences do not have complete control over the final visual result presented to the audience. The nature of the interactive experience affords the audience the opportunity to alter the visual result presented to them. In three dimensional video games and virtual reality experiences the user or player is frequently given the ability to, in real-time, translate and rotate the render camera so as to explore a virtual world. By changing the camera, the user can change what objects in the world are rendered, how much of an object is rendered, and for how long an object is rendered. The screen space accumulation and total screen time of each object, therefore, is unpredictable and presents challenges to the artists of the production when determining the appropriate and desirable level of visual variety for a crowd simulation.

One common approach in these circumstances is to focus on a limited set of parameters related to the interaction between players and crowd characters. Look up tables of pre-rendered options are then developed and used to display the most successful result for the players current state. Such "lookup table" approaches can work effectively yet inevitably limit the range, precision and nuance of crowd simulations. To make more robust forms of storytelling possible a new system or tool is needed to accurately, expeditiously and intuitively create unique and diverse production quality crowd characters that remain believable at dynamically changing viewing distances.

As described, the planning process is vital to determining an effective method for producing crowd characters. Understanding the target visual quality of the crowd character models significantly affects whether the method selected for modeling them will be sufficient. To further illustrate the importance of planning to determine an effective method for modeling crowd characters, it is important to understand how modeling approaches differ between hero characters and crowd characters.

### 1.1.3 Hero Characters vs. Crowd Characters

It is common practice in the CG entertainment industry to distinguish between what are known as “hero characters" and crowd characters. Hero characters play a significant role within the events of the narrative in an entertainment piece, while crowd characters provide background and context to the events. The number of crowd characters almost always far exceeds the number of hero characters in a production and, as a result, the available modeling resources available for each hero character far exceeds that available for each crowd character. To create every crowd character using the same traditional modeling methods employed in the creation of hero characters is almost always resource prohibitive.

### 1.1.4 Two Continuums

Creating a computer generated crowd simulation is a difficult task for a modeling department. One option is to model a single character and then duplicate it as many times as there are members in the desired crowd. The other option is to model each and every member of the crowd uniquely. The former is the most time efficient option possible, but every member of the resulting crowd will be exactly the same. The latter is the least time efficient option, but every member of the resulting crowd would be entirely unique, effectively creating a crowd of hero characters. Duplicating a single character model multiple times would work well when creating an army of clone soldiers. Many stories, however, are not limited to crowds of completely identical characters. Most of the crowds we encounter exhibit at least some level of variability amongst the individuals within them. The option of modeling each character of the crowd might be cost effective for very small crowds but would quickly become cost prohibitive as the crowd grew in size.

### 1.1.5 Combinatorial Modeling Method

The common result of this planning process is the creation of some form of a combinatorial approach to generate the models of crowd characters (A. Sidenblad, personal communication, September 27, 2020). There are multiple examples of combinatorial approaches to crowd character model generation in the computer graphics industry.

During the production of the animated feature film Wall-E, combinatorics were used to expeditiously create the crowds of robots that operated the starliner Axiom (*Introduction to combinatorics*, n.d.). By creating variations of interchangeable parts, they were able to create a large number of visually unique robots. The total time spent modeling was dependent solely on the total number of variations for each robot body part rather than the total number of crowd characters. For example, if one thousand robots were needed to make a crowd scene, modelers could create ten different robot heads, ten different robot bodies, and ten different pairs of robot arms. The different pieces could then be combined to create up to one thousand different robots if each robot had some combination of a head, body, and a pair of arms.

Dirksen and colleagues detail how a combinatorial system was used to generate a crowd of characters in the film Madagascar 3: Europe's Most Wanted. The "ManA's Multidimensional Rig outputs any of 3 bodies, 6 heads, and various wardrobe items for a total of 119,750,400 unique combinations," (Dirksen, Fischer, Kim, Vassey, & Vogt, 2012).

In both examples, the application of combinatorics to the task of creating crowd models allowed for the artists at Pixar and DreamWorks Animation respectively, to create diverse crowds without having to individually model each character in the crowd.

In the context of a traditional film production, a combinatorial method for creating crowd characters is typically successful at achieving the desired level of visual variety determined in the art direction for the production. This is primarily due to the predictability of the medium. Without this predictability modelers would either have to plan their approach based on assumptions that could be detrimental or consider the worst of all possible scenarios and plan their approach accordingly. The potential costs of these alternatives include inefficiency, time lag, and quality. Fortunately this is not often the case for films considering that the entirety of the film is painstaking planned during pre-production to avoid such events.

Most interactive experiences, however, provide the user with control of the virtual camera making the prediction of the factors described above - where the crowd is viewed from, how long the crowd is viewed, and what of the crowd is viewed - nearly impossible to completely predetermine.

### 1.1.6 Crowds in Interactive Experiences

Imagine an interactive story experience in which the user is told, at various points along the story, to interact with virtual persons within the experience in order to further progress in the story. In most contexts the user is given some form of visual guide, description, or identification that aids him or her in locating a specific virtual character that is predetermined by the developer(s) of the experience for the required interaction with the user. An example of this can be seen in the game Horizon: Zero Dawn (2017) in which the main character Aloy - controlled by the player - is frequently tasked with interacting with specific preprogrammed virtual characters scattered across the virtual world of the game (Guerrilla Games, 2017). In most of these cases the character is located amongst a crowd of other characters and is identifiable to the user by a hovering icon associated with the player's current objective. For each interaction in this story, the player can only interact with the predetermined virtual character to complete the objective required to further progress in the game's story. The player is not able to interact with a different character to fulfill the same objective.

For most scripted interactive experiences this form of predetermined character interaction is a sufficient and successful storytelling tool. When considered from the perspective of character modeling, this provides a variety of advantages. Because the characters that a player will interact with are predetermined, modelers can focus their effort on not only these characters but also on the portions of these characters that will be under the most scrutiny by the player. The other crowd characters in the game will not be under the same visual scrutiny by the player and thus do not require as much effort from the modelers.

This is an intentional limitation used by the game designers, developers, and modelers to make the process of producing the game possible with the allotted time and resources for the project. This is how the development of many interactive experiences accounts for the factors described above in order to optimize the process of modeling crowd characters. These limitations would not be suitable for the developers of other interactive experiences that desire to take a different approach to character interaction.

For example, imagine an interactive experience in which the user is directed to interact with a virtual character within a crowd that is not predetermined by the developers. In this scenario the user may choose to interact with any member of the crowd to fulfill the objective required to further progress through the story. This form of character interaction would provide a variety of unique storytelling possibilities that are not currently available to developers because of the challenges inherent to creating the crowd characters in this context. Because the characters required for interaction are not predetermined, the modelers do not have the advantage of focusing their efforts on specific crowd characters. The modelers must account for any possible choices by the user when determining how to go about modeling the characters of this crowd. In order for the process to remain within the budgeted time and resources, the modelers might choose to compromise the desired quality of the crowd character models in order to make the effort and time required for the process affordable. They might also choose to compromise the individuality of the different crowd characters so that less time and effort could be spent making each one unique from the others. Each of these compromises would negatively affect the quality of the overall crowd and likely also impact the immersiveness of the experience.

This challenge is particularly important to consider for the facial models of crowd characters because the gaze and attention of the viewer is naturally and instinctively drawn towards faces (Kesner et al., 2018; Massaro et al., 2012; Savazzi et al., 2014; Villani et al., 2015). As a result of the attention garnered by faces, the level of uniqueness and variation between all the faces of crowd characters plays a vitally significant role in the overall perception of individuality and variability in a crowd or lack thereof. If the faces of the characters in a crowd are discernibly similar to each other, the crowd will likely appear unrealistic to the viewer. Conversely, if each face in the crowd appears discernibly unique from all other faces, the crowd will likely appear realistic and provide a more immersive experience to the viewer.

### 1.1.7 Problem

As discussed above, there are many scenarios for which the combinatorial modeling method is cost effective and produces a desirable visual result. They key shared factor for each of these scenarios being the predictability of the factors with which the result is deemed successful or sufficient. The importance of the predictability of these factors to the success of a combinatorial modeling method is influenced by the nature of crowd simulations involving several characters. The magnitude of character models in a crowd simulation limits the ability for modeling iterations. When the factors used to determine the desirability of the level of visual variety in the crowd change in such a way that the level of visual variety of the existing crowd characters is no longer desirable, iteration is required on the crowd character models in order for them to meet the increased visual demands. In these scenarios, the combinatorial modeling method lacks the flexibility to serve as an effective tool to iterate on this number of models. In order to generate new crowd characters that do not exist within the set that can be produced by combining the existing sets of individual parts, new parts have to be produced to extend the capabilities of what the combinatorial system can produce. In situations where additional fidelity is needed in the new crowd character models, this is a particularly costly task.

This limitation in the ability of a combinatorial modeling method to produce new crowd facial models in productions with changing needs is illustrated abundantly in the example of interactive experiences simulating real people. In these experiences, the aim of the crowd character modeling process is to portray the human characters in the most photo-realistic manner possible. This process often begins with 3D scanning the heads of actors, models, or staff. These 3D facial scans are then processed to conform to a universal topology. Once a database of heads is collected those faces can be mixed with body models, hair models, clothing models as well as different skin, hair, and clothing textures to generate unique combinations, thereby producing unique crowd members. This combinatorial method is limited to the number of existing parts contained in the database. If additional crowd facial models are requested, beyond the capabilities of what can be through the combinatorial system supplied by the database, the only option within the combinatorial system is to expand the database which would most likely require the added expense of additional 3D facial scanning and processing.

### 1.1.8 Solution

For this scenario, and the other scenarios previously discussed for which the factors determining the level visual variety among the characters of a crowd simulation that is discernible to the viewer are unpredictable, an alternative to the combinatorial method is needed. In contrast to the capabilities of the combinatorial method, a method that is parametrically and procedurally driven to randomly generate heads would offer the flexibility to quickly generate new heads for these scenarios in a visually successful and cost-effective manner. Therefore, this thesis will explore a parametrically automated and artistically directable method for creating a multitude of 3D digital facial models that exhibit significant and realistic variety to serve as the faces of crowd characters in crowd simulations. Additionally, this thesis will explore the use of anthropometric measurements in guiding such a parametric system, so as to limit the types of awkward or unrealistic results that a fully randomized system is capable of producing (A. Sidenblad, personal communication, September 27, 2020).

## 1.2 Literature Review

### 1.2.1 CG Production Methods

#### *1.2.1.1 Character Development Process*

There are various fields in computer graphics that use computer-generated three-dimensional characters to tell stories, whether it be for entertainment, advertisement, or education. The process for developing these characters has been a subject of significant study for many individuals and many of the foundational concepts and principles of the process were discovered before the advent of computer graphics (Thomas, Johnston, & Thomas, 1995; Woods, 2002)).

The character development process primarily takes place during a stage of production known as pre-production. Pre-production involves many processes that lay the foundation for the production as a whole. In the film and game industries, pre-production is vital to the creation of a compelling, cohesive, and efficiently produced story. Characters are an extremely important storytelling tool, and as a result, the development of these characters is a vital process in pre-production. In most film and game studios, various artists are tasked with designing both the visual characteristics of the characters, as well as, the personalities, motivations, backgrounds, and other such qualities of the characters. The effectiveness and immersiveness of a story are often dependent on the quality and complexity of its characters. This is because a story's success is dependent on its reception by the audience, and an audience's ability to relate and invest in the characters of a story will often determine the audience's reception of the story. With this understanding, it is vital for character designers to give the characters of the story the level of relatability that will enable an audience to invest interest and attention to the experiences, actions, and desires of these characters.

#### *1.2.1.2 Modeling Process*

Once the characters have been designed and the visual appearance of the characters have been defined the production of the characters moves to the modeling artists. The task of the modelers is to create three-dimensional models of the characters that match or invoke the two-dimensional drawings and concept art that were created by artists during pre-production. The modeling process is by no means trivial and requires both technical and artistic skill. Similar to a sculptor, a modeler requires an adroit understanding of the artistic principles of form, shape, and composition. While the tools and medium used by a modeler differ from that of a traditional sculptor, the requisite technical mastery of their respective resources is a shared trait. In the case of the modeler, there are various tools that can be used to create three dimensional surfaces and a variety of types of surfaces with which the modeler can mold and shape. The most commonly used types are Polygonal models and Non-Uniform Rational B-Splines models, also known as NURBS.

#### *1.2.1.3 The Modelers Medium*

Polygonal models are composed of a combination of vertices, edges, and faces. The level of detail of a polygonal mesh is determined by the polycount of the mesh. The polycount refers to the total number of polygonal components a particular mesh is composed of. In other words, the level of detail of a model is limited by its polycount. A model with a low polycount will not be able to support the same level of detail as a model with a high polycount. Much like a low-resolution raster image can not support as much detail as a high-resolution raster image. NURBS models are mathematically defined surfaces that are determined by the location and relation of control points in three-dimensional space. NURBS models are very effective when used to create round and curved surfaces. When the rendered results of a low-poly polygonal sphere and a NURBS sphere - in which the number of control points of the NURBS sphere are similar to the number of vertices of the polygonal sphere - are compared, it is evident that the contour of the sphere, in the rendered results of the NURBS sphere, appears much smoother than that of the polygonal sphere. In the current technological environment that modelers work in, the most commonly used type of model is polygonal meshes that are rendered using a form of sub-divisional approximation (Parke & Waters, 2008). There are multiple reasons for this:

1. Polygonal models are typically the more intuitive model type to use when creating significantly complex forms

2. A sub-divisional approximation of the model can be rendered, which increases the detail and resolution of the model so that it appears smoother. The quality of these results is not discernibly less than that of NURBS models.

3. The computational requirements to render complex polygonal models using a sub-divisional approximation scheme are fairly affordable for artists as a result of significant advancements in the computational power of modern computers.

*1.2.1.4 The Modelers Tools*

There are many different modeling techniques and software tools available to modelers in the creation of character models and each is unique in its advantages and disadvantages. Ultimately, they all entail some combination of the following operations on the various components of a geometric surface description: addition, subtraction, transformation, and reorganization. As a modeler adjusts and changes a mesh through the use of these techniques, they continually refine the form towards the desired result. The modeler is finished when the geometric surface description or mesh accurately matches the visual concept of the desired character or object.

As is often the case for tasks that require a skilled artistic hand, the process of creating a complex three dimensional form of significant aesthetic quality is highly time intensive. This is particularly true when the complex three dimensional form is that of a human character.

The amount of time required for a skilled modeler modeler to create a finished character model is determined by a variety of factors. These factors include: how skilled the modeler is; how effective the modeling tools used are; how complex the desired character is. Ultimately the process can take anywhere from multiple hours to multiple days for a complex character mesh, even for a highly skilled modeler using industry-standard software.

When considering the time required for the character modeling process in the context of real-world film and game production (i.e., limited budgets) it is important to determine the quantity of time and resources that affordably can be invested in the different characters of a story. There are a variety of factors that influence these determinations, such as a character's role in the progression of a story, the total amount of time the character will be viewed by the audience, the proximity from which the character will be viewed by the audience, and the complexity of the character's design. These considerations typically result in a binary classification of characters; namely, a character is either a hero or a crowd member.

### 1.2.2 Characters

Typically the category of main characters contains a small number of individuals within a particular story. These characters, often referred to as heros, play significant roles in the development and continuation of the story and are viewed by the audience multiple times and from different perspectives. As a result, these characters are given the greatest priority in terms of resources given to their development.

The category of crowd characters often contains a large number of individuals within a particular story. These characters as individuals do not play frequent or significant roles in the development and continuation of the story. That being said, the crowd, being a collection of the individuals, may be considered a character that can perform an important role at specific points in a story. It is evident that a disparity exists when considering the differences between hero and crowd characters. In most scenarios, there are significantly less hero characters than that of crowd characters, but the respective levels of narrative significance dictates more resources be invested in each hero character and less be invested in each crowd character.

This disparity is at the core of the interests of this research. A method that enhances from the traditional modeling process in terms of its expediency is needed to ameiliorate this deficit, so that crowd characters of a visual quality similar to that of hero characters can be created without requiring budgetary modifications.

In order to effectively assess or develop methods for the production of quality crowd character models, it is important to understand the artistic subject matter. As with any artistic depiction, whether it be in the form of abstraction or realistic portrayal, an important factor in the level of representational successfulness is the artist's understanding of the subject matter. For this reason I will now briefly discuss the larger phenomenon of crowds as well as the historical development of crowd simulation methods and software.

### 1.2.3 Crowd Characteristics

Many individuals interact with crowds daily. "In its ordinary sense the word 'crowd' means a gathering of individuals of whatever nationality, profession, or sex, and chance that may have brought them together" (Bon, 2009, p. 20). A crowd can therefore be defined simply as a group of individuals in which the actions of the individuals are either indiscernible from one another, are less important than the actions of the collective, or shared by multiple individuals. For example, a collection of individuals walking through a public setting, such as an airport or mall, would be considered a crowd by this definition. The actions of the individuals are neither uniquely discernible nor necessarily shared by many. Each individual is acting upon his or her own motivations and ideas rather than that of a collective. Some motives may be shared between the individuals but it is not the result of the crowd's influence. To many, this would still be considered a crowd even though it does not meet the criteria of a psychological crowd. According to Bon (2009, pp. 19-28), a psychological crowd is formed when each individual adopts the mental state of the collective. This is the point in which the individuals have a shared motivation and actively work towards the collective goal. For many researchers, the phenomena of crowds is primarily of interest for this reason. A psychological crowd is "a new being which displays characteristics very different from those possessed by each of the cells singularly," (Bon, 2009, p. 23). As a unique entity a crowd consists of discernible characteristics. In the most general sense, the characterizations of a crowd fall into one of two primary categories: the unique characteristics of the individuals in a crowd and the characteristics unique to the collective. Both categories of characterization are vital in a crowd's description.

Therefore to understand and accurately simulate a crowd, one must understand both how individuals collaborate in a crowd, as well as how the uniqueness of each individual contributes to defining the crowd as a whole. The individual thus plays a pivotal role in crowd simulation. In addition to the distinction of individual versus collective characteristics, other characteristics include those pertaining to the visual aesthetics, quality and form of movement, behavior, psychology, sound, and interaction. The influence of each characteristic per individual on the resulting whole cannot be understated.

### 1.2.4 Crowd Simulation

A significant tool for the study of crowds and their characteristics is the digital simulation of crowds. Some researchers employ crowd simulations in an effort to study potential outcomes of scenarios in which crowds attempt to escape dangerous events in various environments. Other researchers employ crowd simulations to study how changes in external and internal factors affect the characteristics of a crowd. In terms of the entertainment industry, the interest and use of crowd simulations is primarily as a storytelling device.

The history of the modeling and simulation of crowds began with the work of Reynolds (1987). Reynolds sought to simulate the complex collective motion seen in animal groupings such as flocks of birds or schools of fish. To Reynolds, the movement observed in the natural phenomena of flocks was reminiscent of a particle simulation. The birds of a flock could be simulated as particles with additional rules that determine their individual motion. The interaction between each particle would determine the unique behavior of the collective.

> Each simulated bird is implemented as an independent actor that navigates according to its local perception of the dynamic environment, the laws of simulated physics that rule its motion, and a set of behaviors programmed into it by the 'animator.' The aggregate motion of the simulated flock is the result of the dense interaction of the relatively simple behaviors of the simulated birds. (Reynolds, 1987, p. 25)

Reynolds refers to each member of a flocking simulation as a "boid." The rules that determine a boid's movement relative to other boids were remarkably simple; viz., avoid collisions, fly with your neighbors, and stay with the flock. When the simulation is run, various calculations are performed to determine what the current state of each boid is in relation to the neighboring boids in the system, as well as, how the boid should respond in accordance with the rules of collision avoidance, velocity matching, and flock centering. Collision avoidance and flock centering cause boids to remain together without getting so close as to collide or intersect with one another. Velocity matching causes boids to fly similarly to the neighboring boids both in speed and direction. The combination of these rules on the individual boids creates a unique collective motion similar to that of a flock.

Flocks of birds and schools of fish share many similar characteristics to that of human crowds. They are moving organisms with discernible characteristics formed of a collection of similar yet distinct entities moving and working in relation to each other. Each entity within the collective possesses the ability to make individual decisions yet act in a manner determined by the desires and actions of the collective. Crowds, however, are significantly more complex than that of flocks or schools. It is for this reason that the boids model is not sufficient as a simulation method for crowds.

Much research has since been performed in an effort to develop more suitable methods for crowd simulation. Durupinar et al. (2008) developed a system that uses personality traits as parameters for creating variation within a crowd simulation. In the research Modeling Individual

Behaviors in Crowd Simulation, a system is developed to test "the impact of individual agents characteristics in emergent groups" (Braun et al., 2003, p. 1). Much of the research that has been done on the topic of crowd simulations address primarily the accuracy of simulating a crowd's movement (Bandini, Federici, and Vizzari, 2007; Bandini, Manzoni, and Vizzari, 2006; Bandini, Manzoni, and Vizzari, 2004).

### 1.2.5 Crowd Implementation

In addition to the academic study in the advancement of crowd simulation methods there are a variety of commercial crowd simulators currently available that have improved the quality of crowd simulation development. These simulators vary in regard to the method of simulation, the types of crowds they are most suited to simulate, the possible visual quality of the simulation, and the intended field for which the software is designed. In their book, Simulating Crowds in Egress Scenarios, Cassol, Musse, Jung, and Badler (2017) discuss several important commercial crowd simulators.

Such commercial crowd simulators have been designed primarily for the purposes of study and research rather than entertainment or advertising. In the context of the entertainment and advertising industries, however, the visual quality of crowd simulations is of singular importance. In these fields, the simulation and depiction of crowds must accurately reflect both a crowd's motion as well as it's visual appearance. This means that an entertainment-centric crowd simulator must be capable of or compatible with industry-standard rendering software (e.g., Massive) The most important fact regarding "real" crowds is that they are composed of unique individuals.. Such bedrock diversity is an important factor to address in the depiction of crowds as it plays a significant role in their realism and believability.

### 1.2.6 Facial Recognition

As discussed previously, the importance of individuality within a crowd cannot be overstated. The distinct interpersonal characteristics of each individual within a crowd have a significant influence on the resulting crowd, and contribute greatly to the complexities observable in each crowd.

The importance of individuality applies to both the non-physical and physical aspects of the individuals within a crowd. The non-physical aspects being the attributes of an individual such as his or her personality, background, intelligence, emotional state, motivations, and general psychology. The physical aspects referring primarily to the visual attributes of each individual.

When considering the possible variations of the physical appearance of the individuals, or agents, within a crowd, there are a number of different aspects to consider such as clothing, hairstyle, skin color, height, and weight. There is one aspect of an individual's physical appearance that contributes to the determination of one's visual uniqueness and ability to identify or recognize a specific individual more than any other - the human face.

Not only is the range of variation inherent to human faces limitless, but in addition, humans are expertly trained and immensely capable in the act of facial recognition. We are experts at facial recognition, as evidenced by the fact that we are "are familiar with thousands of faces and are able to recognize individuals despite changes in hairstyle, hair color, facial hair, presence of eyeglasses, and so on," (Diamond & Carey, 1986, p. 108). As inherently social creatures, facial recognition plays a vital role in our ability to interact with others and maintain social bonds (Benton, 1980; Landau, 1989).

### 1.2.7 Facial Modeling, Rigging, Animation

There are many recent examples of computer-generated faces in the entertainment industry that one can point to as evidence of just how complex, physically accurate, and ultimately realistic the state of the art has become. Completely digital characters such as Thanos and The Hulk in Marvel's Avengers End Game and Grand Moff Tarkin in Star Wars Rogue One are evidence that the visual quality currently attainable in the technological climate of the entertainment industry is astonishing. While the public is often quick to criticize these fully digital characters the truth of the matter is artists today working with digital tools are capable of creating extremely photo-realistic computer-generated faces.

In order to appreciate how far the art of created computer-generated faces has come, we must look back to the inception of computer-generated faces. The first computer-generated images of a three dimensional digital face were created by Frederic I. Parke in 1971 (Parke & Waters, 2008). Parke (1972, p. 453) created a polygonal facial surface by defining one of half the face consisting of "124 polygons defined by 202 vertices" and then mirroring it. By 1974, Parke successfully developed a method for animating and deforming a polygonal facial mesh by parameterizing his facial model. Parke's parameterized face model employed "parameters which control the interpolation, translation, rotation or scaling of the various facial features," and were "divided into two main categories, those controlling expression manipulation and those controlling facial conformation," (Parke, 1974, p. 10).

The computational processing capabilities available to Parke during his research were limited, making the processing of each polygonal component of a face for the purposes of animation or changing a face's conformation impractical. Using a parametric model reduced the amount of information that needed to be stored and processed, allowing for facial animation and conformation modification to be possible under the computational limitations of the time. As computer processing capabilities and graphics hardware have improved, the need for a parametric face model has diminished and other methods for animating a face and modifying a facial conformation have increased in popularity as a result of the increased capabilities for artistic expression and high-fidelity visual results they offer artists.

But just as technology has advanced significantly since the development of Parke's parameterized face model, the scale of projects in the entertainment industry and the forms of computer generated content have increased just as significantly — as is made evident by the state of the art digital character examples that were previously mentioned. Artists in the entertainment industry are often pushing the boundaries of what is possible under the budgetary constraints of the projects they work on and the computational limitations of their time. Often the methods used by artists for hero characters are sufficient, effective, and practical for those characters but become insufficient as the scope of character related tasks grows. More specifically, these same procedures cannot be equitably applied to the creation of crowd characters. The concept of requiring unique methods for the treatment of crowd characters as compared to hero characters in the context of the production pipeline is well illustrated in the work Improving crowd quality through interdepartmental collaboration on Madagascar 3: Europe's Most Wanted (Dirksen et al., 2012).

In this work a method is presented for bringing crowd characters into the foreground to act temporarily as hero characters. The authors refer to this process as hero promotion. "High quality deformations allow crowd characters to be pushed closer to camera, where their acting performance is highlighted. Right from the beginning of production, all of our generic characters were designed so that their facial expression would read the same way when the same animation was applied, allowing the Crowds department to cast generic characters with any head variation with no additional overhead. With crowd characters much closer to camera, they also receive more Director feedback. To allow direction of these performances, we developed a system called Hero Promotion, which makes hero assets from crowd characters," (Dirksen et al., 2012).

In this technical paper, Dirksen et al. (2012) present how the method used for the modeling of the crowd characters in Madagascar 3: Europe's Most Wanted (2012) is that of a combinatorial approach in which a collection of face, body, clothing, and accoutrement models are combined together to create unique variations. While this method is somewhat effective in producing some level of diversity in the crowd characters' appearances, it is ineffective at capturing the level of diversity observed in real crowds. Additionally the modeling method presented in this paper does not provide a similar level of flexibility and adaptability as the method presented for enhancing the performance of crowd characters when promoted to hero characters. The method, presented here, for enhancing the performance of the crowd characters during the hero promotion process, is very effective and affordable for enhancing the quality of crowd characters to match that of hero characters, at least in terms of their movement. This is a particularly valuable development for improving the depiction of crowds in computer graphics through affordable means. This also gives evidence to the need for a similarly adaptive and affordable method for enhancing the visual aesthetics of the models of crowd characters during the hero promotion process. As discussed previously, an important visual quality to address in the depiction and portrayal of crowds is that of the diversity and visual variety amongst a crowd's members. A crowd's visual accuracy and realism is greatest when each crowd character is a unique conformation of individual visual features.

As stated by Parke (1982, p. 68), “parameterized models are powerful tools for facial image synthesis and animation.” A parameteized approach in this instance would allow for facial conformations to be procedurally generated in a systematic manner. The ability to procedurally generate unique facial conformations would certainly provide flexibility considering the quantity of characters that crowds entail. As it relates to parameterizing faces in terms of their conformation, Parke (1974, p. 42) concluded “it is not clear exactly what conformation parameters are desirable” and “ there are a number of additional conformation controls that might be tried.” When Parke (1982, p.62-63) set out to develop his parameter set he decided to base them “on structural understanding, wherever possible,” and supplement them “as necessary by parameters based on observation.” As it pertains to his set of conformation parameters, Parke (1982, p. 63) acknowledged the following:

> The development of truly complete conformation parameter sets appears very difficult. Little in the way of theory exists to support their development, and the variations in facial structure from one individual to another are far less understood than the ways in which a given structure varies from one expression to another.”

Because the effectiveness of any parameterized procedural system is dependent ultimately on the effectiveness of the chosen set of parameters, it is essential that an accurate reference source or description method for the form and shape of real human faces is identified. Without this information there is a significant lack of structure with which to inform and dictate the process of morphing one facial shape into another. This understanding led to the search for an extensive and robust system of real human facial description containing quantitative metric data that have logical three dimensional positional or vector associations. The culmination of this search resulted in the review of relevant literature in the field of craniofacial anthropometry.

### 1.2.8 Anthropometric Measurement Systems

“Anthropometry is a direct means of facial measurement that uses standard landmarks and instrumentation to compare populations,” (J. P. Porter, 2004, p. 78). These measurement systems have been used to gain a deeper understanding of the diversity of facial conformation seen across human populations. In addition to the pursuit of greater anthropological understanding, many other fields have shown significant interest in anthropometric data. Surgeons have used data collected with anthropometric measurements to make informed decisions about surgical procedures to successfully treat facial disfigurements (Farkas, Katic, & Forrest, 2005; Ferrario, Sforza, Schmitz, & Santoro, 1999; Ibrahim et al., 2016; Vegter & Hage, 2000, 2001; Wong et al., 2008). Forensic scientists and law enforcement have historically used anthropometric measurements as an aid to criminal identification (Davis, Valentine, & Davis, 2010; Evison & Bruegge, 2010; Kleinberg, Vanezis, & Burton, 2007; Kleinberg, 2008; Mancusi, 2010; G. Porter & Doran, 2000; Taylor, 2000; Tome, Vera-Rodriguez, Fierrez, & Ortega-Garcia, 2015).

Leslie G. Farkas may be described as the pioneer of modern craniofacial anthropometry. He challenged the canons of facial proportions and devoted most of his professional work to craniofacial anthropometry. In writing Anthropometry of the Head and Face, Farkas (1994) expanded upon the anthropometric measurements identified in Anthropologie by (Knußmann, 1988) by adding other landmarks that would aid in the measurement of the various facial deformities he studied throughout his career. Over time further landmarks have been added to better measure non-NAW (North American White) faces. These landmarks and anthropometric measurements have become standard for researchers interested in craniofacial anthropometry.

#### *1.2.8.1 Collection of Anthropometric Measurements*

The common practice for anthropometric measurement involves the use of measurement tools such as sliding calipers, angle meters, and soft measuring tape. Using these measuring tools, researchers collect a variety of linear, angular, and arc measurements pertaining to the specific conformation of a human face (Farkas, 1994; Venkatadri, Farkas, & Kooiman, 1992). Each craniofacial anthropometric measurement is an accurate descriptor of a specific face. That being said the use of craniofacial anthropometric measurements as visual descriptors are not effective in most common facial identification scenarios. For example, suppose an individual was tasked with observing members of a crowd for the purpose of identifying an individual within a crowd that corresponds to a provided value of a specific, single craniofacial measurement. This task would be extremely difficult. One reason for this is a single craniofacial measurement value can be shared by multiple people.

Craniofacial anthropometric measurements have been used in the artistic visual recreation of human heads throughout history. Many modern artists are taught the work of Leonardo Da Vinci when studying how to effectively create depictions and portraits of humans. “According to Da Vinci, in a well-proportioned face, the size of the mouth equals the distance between the parting of the lips and the edge of the chin, whereas the distance from chin to nostrils, from nostrils to eyebrows, and from eyebrows to hairline are all equal, and the height of the ear equals the length of the nose,” (Vegter & Hage, 2000, p. 1091). Even before Da Vinci’s work in the Renaissance, Egyptians and Greeks had studied and used proportions to enhance the depiction of humans. “A common element of anthropometry of all times is that man has tried to catch physical proportions into values,” (Vegter & Hage, 2000, p. 1090).

#### *1.2.8.2 Faces Database*

One particular modern advancement in craniofacial anthropometric has allowed researchers to collect measurements digitally. The 3D Facial Norms Database was “designed to provide the research and clinical community with access to high-quality craniofacial anthropometric normative data. Unlike traditional craniofacial normative datasets that are limited to measures obtained with handheld calipers and tape measures, the anthropometric data provided here are based on digital stereophotogrammetry, a method of 3D surface imaging ideally suited for capturing human facial surface morphology,” (Brinkley et al., 2016a).

The researchers that developed the 3D Facial Norms Database (Brinkley et al., 2016a) utilize 3dMD digital stereophotogrammetry imaging systems (Atlanta, GA; www.3dmd.com) to capture three-dimensional surface scans of human subjects. Trained individuals then located in three dimensional space a total of 24 specified landmarks on each 3D facial model. Because the landmarks exist in three-dimensional space, various inter-landmark linear measurements could be easily calculated.

### 1.2.9 Forensic Face Categorization and Composite Sketch Systems

#### *1.2.9.1 History of Anthropometry in Forensics*

The fields of forensic science, criminal justice, and law enforcement have also shown interest in craniofacial anthropometry as an aid to criminal identification and conviction. Alphonse Bertillon was one of the first individuals to develop a technique in which facial analysis was used for the purposes of criminal identification. Alphonse Bertillon was "a prominent French anthropologist, who in 1882 was made chief of an identification bureau then established in connection with the Prefecture of Police in Paris," (Bertillon & McClaughry, 1896, p. vii).

In his book Signaletic instructions including the theory and practice of anthropometrical identification, Bertillon identifies and describes a system that uses "anthropometry as a method for identification", (Bertillon & McClaughry, 1896, p. 14). Bertillon's system is comprised of three major elements (Bertillon & McClaughry, 1896):

- Anthropometric Signalment: which consists in measuring with the utmost precision, under prescribed conditions, some of the most characteristic dimensions of the bony structure of the human body
- Morphological Signalment: which is the observation of the bodily shape and movements, and even the most characteristic mental and moral qualities
- Pathological Signalment: the observation of the peculiarities of the surface of the body, resulting from disease, accident, deformity or artificial disfigurement, such as moles, warts, scars, tattooings, etc.

The Bertillon system was intended as a solution for the expeditious categorization and identification of criminals. At the time, Parisian police had used photography as a means of comparison for identifying individuals. According to Bertillon, "the collection of judicial portraits thus brought together soon became so numerous that it became physically impossible to find, to discover, among them the likeness of an individual who concealed his name," (Bertillon & McClaughry, 1896, p. 12). Bertillon determined the task of identifying a suspect by comparing the individual's appearance to a collection "100,000 photographs" in the absence of his or her name as highly impractical, especially in consideration of the total number of individuals the Parisian police were required to accurately identify on a daily basis. Bertillon developed his signaletic system as a method of categorization so that the Parisian police's collection of photos could be organized by signalment rather than name. Bertillon believed that signalment through anthropometric measurement could not be falsified in the way one could alter his or her name or appearance and as a result, prove infallible in one's identification. Over time the Bertillon system began to be used additionally as an aid in the identification of suspects of whom the appearance is known but the name is not. Such as in the case of crimes in which a witness sees the perpetrator and can offer a description of his or her physical appearance (Taylor, 2000). The desire to leverage eye witness testimony in the development of forensic art grew continually following the work of Bertillon. Forensic art being, "any art that aids in the identification, apprehension, or conviction of criminal offenders, or that aids in the location of victims or identification of unknown deceased persons," (Taylor, 2000, p. 3).

Karen T. Taylor, who worked as a forensic artist for 18 years and was "a forensic art instructor for many years at the FBI Academy and other law enforcement academies, universities, and medical schools," (Taylor, 2000) discusses the history and development of forensic art in her book Forensic Art and Illustration. According to Taylor (2000, p. 15), "The Bertillon System of Identification: Signeletic Instructions, including the Theory and Practice of Anthropometrical Identification with illustrations by Duprec has provided a basis for future composite kits, catalogues, and even computer-generated recall systems." Taylor describes the field of forensic art as having four primary categories. The most commonly known of which being "Composite Imagery". Taylor (2000, p. 197) defines composite imagery as "graphic images made up from the combination of individually described component parts". One of the most commonly known forms of composite imagery is that of the composite sketch. Stephen Mancusi, who "was the senior forensic artist and a first grade detective for the New York City Police Department for almost 27 years," (Mancusi, 2010, p. xix) describes a composite sketch as "a drawing of a victim's or witness's perception of a perpetrator at the time he or she was observed," (Mancusi, 2010, p. 6). A composite sketch is created by a composite sketch artist whose task, according to Mancusi, is to "successfully gather, interpret, and illustrate the information obtained from the victim's memory," (Mancusi, 2010, p. 6). As a result, composite sketch artists typically are both a skilled portraitist and investigator.

Throughout the 1900s, the practice of forensic art and more specifically composite sketches was utilized with greater regularity (Taylor, 2000, pp. 15-42). The increased utilization of forensic art and continual technological advancement of the 1900s resulted in the following major developments.

#### *1.2.9.2 Forensic Database History*

Two of the most well-known techniques that developed were that of the Identikit and Photofit systems. According to (Frowd et al., 2005, p. 34):

> The Identikit system, favoured in the USA, uses facial features printed on acetate transparencies. The kit originally contained line drawings, but later, photographic elements (referred to as Identikit II). The Photofit system was adopted primarily in the UK and is similar to Identikit II. However, rather than acetates, facial features in Photofit are printed on jigsaw-like pieces that slot into a template.

As computers became more commonly used by law enforcement, an opportunity for advancement in composite sketch techniques appeared. Software, such as FACES by IQ Biometrix were created to make composite sketch techniques more easily accessible, available, and operable for law enforcement. IQ Biometrix's FACES 4.0 is their latest and most advanced product and is used by "thousands of police agencies worldwide – including the CIA, FBI and the US Military…" (*FACES 4.0 - IQ Biometrix - Faces Software*, n.d.). The FACES 4.0 software has an "expanded database of 4,400 facial features," (*FACES 4.0 - IQ Biometrix - Faces Software*, n.d.). The software allows the user to look through various versions of each feature and add them to the workspace where the user can see the composite sketch he or she is creating. Once a specific feature is selected it is placed in the workspace and the user can then adjust various aspects of the position of the features in the workspace. By going through each feature and selecting an option, a user can create a complete and plausible face sketch. Another unique feature in the software allows the user to create a randomly generated face. What is particularly interesting about this feature is that each randomly generated face looks distinctly unique from the last. One limitation of the random face generation feature is that many of the results appear disjointed, meaning the faces are often not believable. One possible reason for this is that the randomization process is not anthropometrically informed.

In the research work, An Anthropometric Face Model Using Various Techniques, DeCarlo, Metaxas, and Stone (1998) explore an anthropometrically informed method for randomly generating faces. DeCarlo, Metaxas, and Stone found anthropometric measurements to be an effective tool for the creation of diverse and unique faces. In comparison to the randomly generated faces produced by the FACES 4.0 software, the results produced by DeCarlo, Metaxas and Stone are always anthropometrically plausible. However, the results of DeCarlo, Metaxas and Stone are less unique from one another in comparison to the results of the FACES 4.0 software.

One reason for this may be that the results exhibit a familial similarity to one another and as a result do not create the appearance of diversity within a crowd. This research work used the Farkas system of anthropometric measurement previously discussed to drive the anthropometric face model. This allowed them to create plausible faces that were anthropometrically different from one another. While these anthropometric measurements are successful at manipulating the face wholistacly, they are not successful at capturing the uniqueness of each facial feature with great detail. This results in the face models exhibiting familial diversity from one another - meaning the faces appear diverse in the manner seen among family members. The level of diversity observed in real crowds is significantly greater than the level of diversity achieved by DeCarlo, Metaxas, and Stone.

### 1.2.10 Conclusion

In my review of the existing literature and relevant work regarding crowd simulation, I have concluded an opportunity exists to advance the methods used to approximate and represent crowds digitally. Much of the existing literature addresses the movement of crowds and the members within them. An opportunity for advancement lies in how the individual members of the crowd are depicted. The task of creating crowd character models that are distinctive and unique is important to accurately and realistically depicting a crowd digitally. Upon analysis of how characters are processed in the CG entertainment industry, it is evident that the traditional methods used for creating hero character models are cost-prohibitive when used for creating crowd character models as a result of their quantity. While there are existing alternative methods for creating crowd character models expeditiously, their effectiveness is linked closely to the circumstances the characters are modeled for. As established in the work of Dirksen et al. (2012), there are times when the circumstances that the crowd characters are designed under change (e.g., the process of "hero promotion" discussed by Dirksen et al. (2012)).

To effectively generate unique crowd character faces expeditiously, the method should be capable of plausible randomization and procedural. This would require a parametric structure for understanding and describing the range of unique form and shape of human faces. As established in this literature review, craniofacial anthropometrics have a historical precedent for use as a system for identifying, recording, and depicting the unique conformations of human faces. Specifically, the work of DeCarlo et al. (1998) establishes a precedent for the use of the Farkas system of craniofacial anthropometric measurement in driving a CG facial model deformation system.

# 2. METHODOLOGY

The process of creating high quality facial models for CG characters is a laborious and time intensive process. Because crowd simulations entail a large quantity of characters, methods such as the combinatorial method previously discussed have been developed as alternatives to sculpting each crowd characters' facial model uniquely through traditional modeling techniques. In the development of real-time experiences where the factors determining the level of visible variety among the characters of a crowd simulation are unpredictable, the existing alternatives to the traditional crowd character modeling process lack the flexibility to expeditiously and effectively address these unpredictable circumstances. This thesis explores an alternative method for creating facial models for crowd characters in an effort to identify a method that provides a substantial improvement in terms of cost. In the pursuit of optimization, it is common practice in the computer graphics industry to employ procedural systems to either mitigate or entirely eliminate the more laborious and repetitive aspects of a process. The use of procedural systems allows artists to achieve a desired result within a shorter time frame, allowing the artist to either move on to another task or use the remaining time to iterate on the result and potentially improve its quality. Both of these outcomes provide improvements to the benefit to cost ratio of the process.

In keeping with this standard approach, this thesis explores a method for creating unique 3D digital facial models for crowd characters that is both parametrically automated and artistically directable. The primary objective of this method is to expediently produce multiple character facial models that exhibit significant and realistic variety. In addition, the method explored in this thesis is developed with the aim of being suitable for a professional production environment. To that end, this method is designed with the following characteristics in mind:

- Able to create production quality models
- Capable of creating realistically diverse and unique models
- Anatomically consistent and correct

- Able to edit each crowd member at varying levels of detail
- Parametrically automated
- Capable of believable randomization
- Capable of exporting models for use in later stages of production

The method explored in this thesis utilizes a framework of parameters, based on craniofacial anthropometric measurements, for the procedural manipulation of a facial conformation. Specifically, the Farkas system of craniofacial anthropometric measurement was determined as a suitable framework from which parameters for this method could be derived. In addition to providing a structure for relative anatomical accuracy, the quantitative nature of the data allows for the simple extraction of the three dimensional transform information that is necessary for the technical performance of this procedural method. This decision also follows the precedent set by DeCarlo et al. (1998).

The specific craniofacial anthropometric data used for the procedural system of this method is a synthesis of the data collected by Farkas (1994), Farkas et al. (2005), and (Marazita, Weinberg, & Raffensperger, 2017). In developing the procedural system, a parameter is created for each craniofacial anthropometric measurement listed in the synthesized data. The full set of parameters can be found in Figures A.1-A.6 in Appendix A. These parameters change the three-dimensional coordinate positions of craniofacial landmarks that act as drivers for a unique deformation system designed to change the conformation of a base facial geometry. The facial geometry used in the evaluation of the method presented in this thesis was designed and modeled using industry standard reference, techniques, and widely accepted practices.

Upon successful implementation of this method, the results will be visually assessed and presented to illustrate the effectiveness of the method in generating believable diversity.

# 3. IMPLEMENTATION AND RESULTS

## 3.1 Implementation

To implement the method described, a series of steps were performed including: data consolidation, production of the base facial mesh, landmark localization and placement, script development, and creation of the deformation system. As an overview, the first of these steps (i.e., data consolidation) provides the information with which the procedural system generates values for the parameters driving the deformation system. The base facial mesh was then modeled using reference of industry standard methods and the base facial meshes used in various sculpting software. Once the base facial mesh was complete, landmarks were placed on its surface in correspondence with the instructions provided by Farkas (1994) and those provided by the researchers that produced the 3D Facial Norms Database (Marazita et al., 2017). At this stage the procedural deformation system could be developed. A script to perform the necessary functions of system was written in Python and executed in Maya. The script is designed to procedurally generate parameter values for a deformation system driven by controls corresponding to craniofacial anthropometric landmarks.

### 3.1.1 Data Used In the Implementation

As described in the Methodology section, the data identified for the implementation of the method outlined in this thesis includes a combination of data from three sources:

- Anthropometry of the Head and Face (Farkas, 1994)
- 3D Facial Norms Database (Marazita et al., 2017)
- International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races (Farkas et al., 2005)

There is a wide range of available data within these sources. Upon thorough analysis of the available data, the following specific information was identified as particularly useful to implementing the method of this thesis:

- Standardized craniofacial anthropometric metrics
- The two landmarks associated with each metric
- The mean value for each metric
- The standard deviation found for each metric

This information provides a structure and network of parameters that a procedurally driven deformation system could be effectively designed around. Once the data to be used in the implementation had been identified, it was transferred and organized in such a way as to be easily parsed and used in the python code that would control the procedural generation of each face. The data was organized into the following three text files.

- Landmark Data: A text file containing information specifically related to the different facial landmarks used
- Metric Data: A text file containing information specifically related to the different anthropometric metrics used
- International Anthropometric Data (IAD): A text file containing the mean and standard deviation information for different measurements for each ethnic group studied in International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races (Farkas et al., 2005).

#### *3.1.1.1 Landmark Data*

The information pertaining to craniofacial landmarks that was included in the landmark text file includes the anatomical name of the landmark, the abbreviation associated with the landmark, whether the landmark is found on both sides of the face, left and right, or along the vertical line of symmetry of the face, and the initial three-dimensional coordinate position of the landmark.

#### *3.1.1.2 Metric Data*

The information pertaining to craniofacial anthropometric measurements that was included in the metric text file includes (1) the anatomical name of the anthropometric measurement, (2) the two landmarks associated with the measurement in the form of their abbreviations, (3) whether the measurement is a linear one dimensional vector along a single cardinal axis - and if so which cardinal axis (X, Y, or Z) - or a linear three dimensional vector not along a single axis, and (4) the central point, if any, used to maintain the ratio between the two landmarks and central point.

#### *3.1.1.3 International Anthropometric Data (IAD)*

The data contained within the IAD text file includes craniofacial anthropometric information for each gender of a variety of different ethnic groups. The mean and standard deviation are included for the following list of craniofacial anthropometric measurements.

- Trichion-Nasion Height
- Facial Height
- Morphological Facial Height
- Lower Facial Height
- Maximum Facial Width
- Mandibular Width
- Intercanthal Width
- Palpebral Fissure Length
- Outercanthal Width
- Nasal Height
- Nasal Width

- Labial Fissure Width
- Aurale Height

### 3.1.2 Base Facial Mesh

The method implemented in this thesis requires a base facial mesh that the procedural system can deform. The base facial mesh created for this implementation, as depicted in Figure 3.1, is an original design and was developed with reference to industry-standard methods and the base facial meshes used in various sculpting software. Figure 3.1 depicts the base facial mesh from various camera angles (i.e. $0°$ ,$30°$, and $90°$) and different render methods (i.e. a faceted normals render and a smoothed normals render).

### 3.1.3 Localization and Placement of Landmarks

Once a base facial mesh was complete, the process of identifying the three-dimensional positions for each landmark could be performed. These landmarks correspond to the craniofacial anthropometric measures that will serve as the parameters for this procedural deformation system. Each parameter is a vector between two points (i.e., landmarks). Traditionally in craniofacial anthropometric research, the process for collecting measurements begins with identifying the position of each landmark on the surface of a subject's face. To create consistency in the collection of craniofacial anthropometric measurements, guidelines were created to address the process of identifying the location of each landmark on the surface of a face. Farkas (1994) provides his own guidelines for this process in the book Anthropometry of the Head and Face, and the researchers of the 3D Facial Norms Database provide a list of their guidelines as well (Brinkley et al., 2016b). These guidelines were used to identify the position of each landmark on the base facial mesh as depicted in Figure 3.2. A more complete listing of each landmark and its location on the base facial mesh can be found in Figures C.4, C.1, C.2, C.3 in Appendix C.

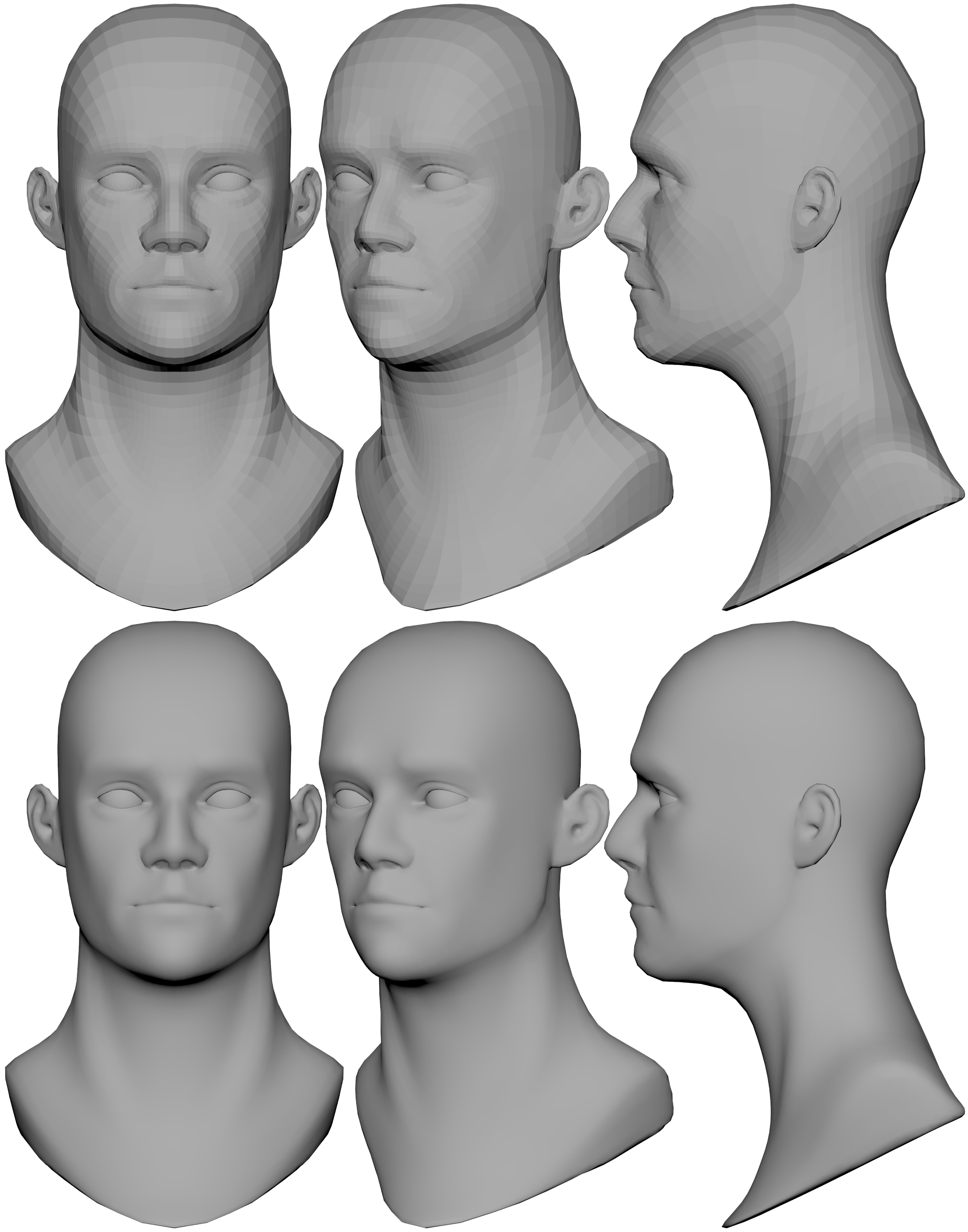


Figure 3.1: The base facial mesh produced for this implementation. The upper image depicts a faceted render of the polygonal surface. The lower image depicts a smoothed render of the polygonal surface.

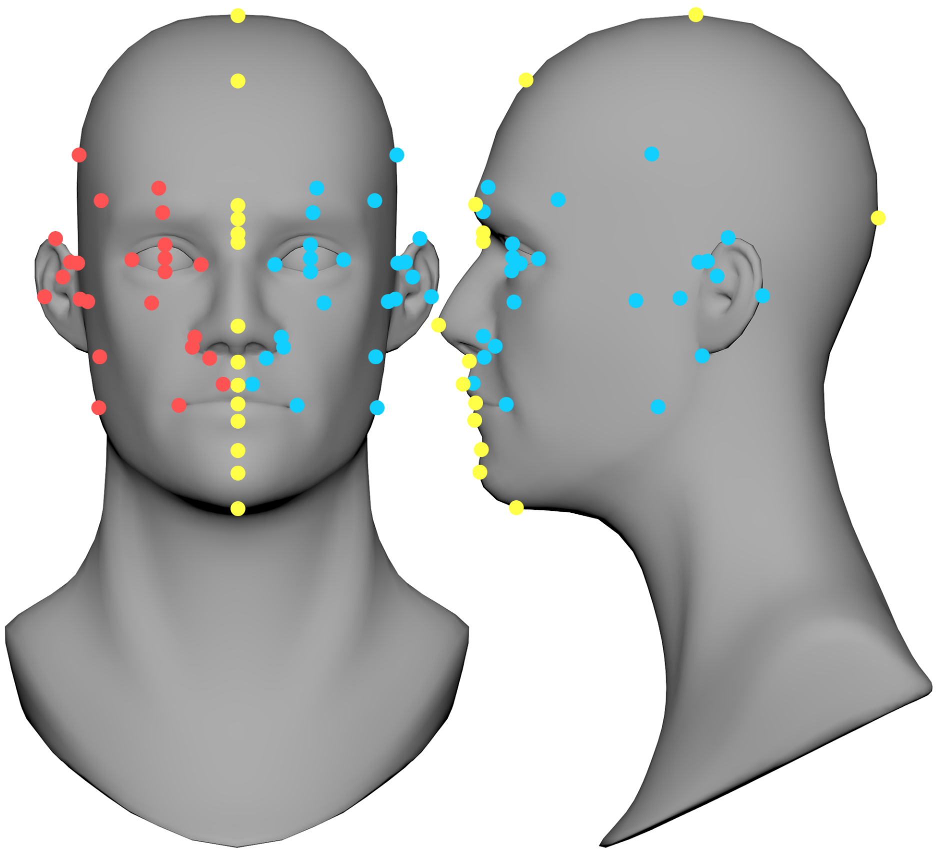


Figure 3.2: Placement of landmarks on the base facial mesh. A complete diagram showing the name, short name, and position of each landmark can be found in Appendix C.

### 3.1.4 Development of Tool Script

The procedural generation of the crowd facial models is managed and enacted by a python script that is executed within Maya. This python script was developed and formulated from the observations and concepts studied in the literature review. The design of this script is intended to employ the knowledge and data of craniofacial anthropometric research in a method that is aligned with the common practices and production techniques observed in the computer graphics industry. The script systematically produces unique facial meshes through a series of steps that can be summarized as follows: The script uses user input data to generate data-driven, randomized values for craniofacial anthropometric measurements. These measurements change the positions of craniofacial landmark linked controls that drive a uniquely designed deformation system. The deformation

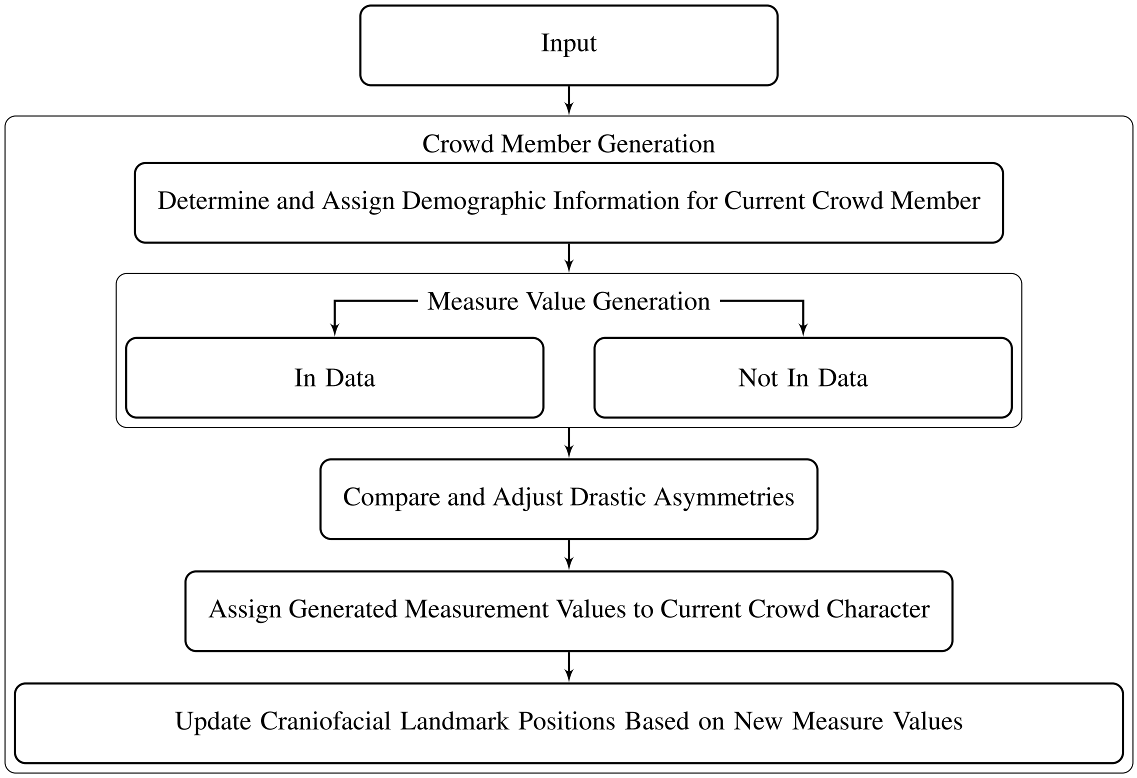



Figure 3.3: Flowchart of the core functions the script performs during its execution. A more detailed overview of the process can be viewed in Appendix B.

system changes the position of the vertices of the base facial geometry thereby resulting in a facial mesh with unique facial conformation. For the purposes of reproducibility and full-transparency, I will now outline the functions and procedures of the script in Figure B.1.

*3.1.4.1 Input Phase*

Once executed, the script displays a user interface prompting the user to input information related to the desired demographics of the crowd, such as size of the crowd, the ratio of male to female individuals in the crowd, and the percentage of each ethnicity provided by the data set within the crowd. Using this information, the script creates a number of crowd member objects equal to the crowd sized dictated by the user. These crowd member objects act as variable storage that the script will read and write to at various stages. Each of these crowd member objects are then assigned gender and ethinic information in accordance with the users input.

#### *3.1.4.2 Crowd Member Generation*

After the script has completed the process of receiving and storing the user's input, it begins the process of generating new facial meshes for each crowd member. The crowd member generation process involves the processes of measure value generation, asymmetry adjustment, value assignment, and the updating of each craniofacial landmark control's position. It is important to note that the script completes each of these steps for one crowd member before moving on to the next.

#### *3.1.4.3 Measure Value Generation*

The first of these processes, Measure Value Generation, is a series of evaluations, comparisons, and calculations that is performed repeatedly until each anthropometric measurement has been independently processed. The objective and design of this process is to generate a quasi-random value for each anthropometric measurement that aligns with the craniofacial anthropometric data of the specific demographic of the current crowd member.

The process begins with the script accessing the gender and ethnic information of the crowd member object it is currently generating. This information is used to determine and assign values for each anthropometric measurement for the current crowd member. To do so, the script refers to the data collected in the IAD.

During the Measure Value Generation process the script evaluates the current measurement it is processing in relation to the craniofacial anthropometric data as determined by the specific demographic information of the current crowd member object. As displayed in the list above, the available craniofacial anthropometric data for specific demographics is limited to 13 metrics. In contrast, the total number of craniofacial anthropometric measurements used for the facial deformation system, and therefore also the number of measurements for which the script must generate a quasi-random value, is much larger (i.e., 62). To address this limitation, the script is designed to address the two cases uniquely so as to achieve the most optimal result in spite of the limitation. The first such case is when the current measurement being processed is one of the 13 for which information is provided directly in the data. The other case being when the current measurement is not. In order to address each of these uniquely, the script performs the check to see if the current measure is or is not in the International Anthropometric Data set.

3.1.4.3.1 Case 1: In the Data

If the current measurement being processed is in the data set, the script uses the following steps to determine a random value for the measurement. The script first checks to see what the mean value and standard deviation range of that measure is. The mean value and standard deviation range provides a range from which the script can select a random value. The script is designed to more frequently select a random value closer to the mean value rather than either extreme. The script splits the total value range into three ranges as shown below.

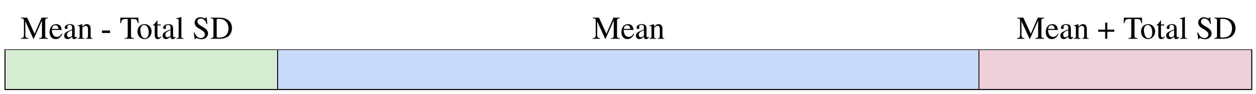



Figure 3.4: Illustration of the distribution from which random values are selected.

The script then chooses one of the three ranges and selects a random value from within that range. It is designed to select the lower and higher ranges each $22.222\,\%$ of the time and the median range $55.556\,\%$ of the time. Once a range is selected and a random value from within that range is selected, the value is then assigned to the associated measurement variable of the current crowd member object.

3.1.4.3.2 Case 2: Not in the Data

If the measurement is not listed in the data another series of procedures are needed to determine a reasonable value for the current measure. Even though the measure is not listed in the data it is important that the value it is assigned is determined in relation to the craniofacial anthropometric data available for the specific gender and ethnicity of the current crowd member. In order to do so the script evaluates the current measurement being in terms of the facial region it belongs to. By analyzing the information in the International Anthropometric Data in terms of facial region, the script is able to determine information about a general region that can then be applied in the determination of the values for other measures in the same region. For this reason each measurement that is processed by the script, both in terms of those for which values are being generated and those the script analyzes from the data set, is attached with the requisite information of the measurement's respective facial region. With this in mind, the first action of the script upon determining that the current measurement is not in the list of measurements provided in the International Anthropometric Data, is to access the current measurement's respective facial region information.

Once the facial region of the current measurement is identified the script begins the process of analyzing the information for the different measurements in the International Anthropometric Data that are of the same facial region to determine the information needed to generate an appropriate value for the current measurement being processed. Specifically the script's analysis is designed to attain statistical information related to the relationship of the mean value and standard deviation values for the measurements of the same facial region in the International Anthropometric Data. This understanding of how the mean value and standard deviation values relate for all the measurements in the same facial region of the same specific demographic, will allow the script to determine a reasonable value range, from which a random value can be selected, for the current measurement being processed. This measurement is therefore in some manner aligned with other measurements of the same facial region of the same specific demographic.

During this analysis process the script evaluates each measurement of the same region in the IAD and determines the percentage increase from the mean value (PIM) and the percentage decrease from the mean value (PDM) for each measurement.

$$PIM = (Mean + Positive\ Standard\ Deviation)/Mean$$

$$PDM = (Mean - Negative\ Standard\ Deviation)/Mean$$

Once the script determines the PIM and PDM for each of the measures listed in the data that reside in the same region of the current measurement, the script calculates the full region's average

PIM (RPIM) and average PDM (RPIM). Using the RPIM and RPDM, the script calculates the range from which the quasi-random value will be selected for the measurement currently being processed. The minimum value of the range is equal to the product of the RPDM and the initial value of the current measurement being processed, the initial value being the value collected of said measurement from the base facial mesh. Likewise, the maximum value of the range is equal to the product of the RPIM and the current measurement's initial value. Once the range is calculated, the script selects a random value from within the range similar to the process described for the case in which the current measurement being processed is in the IAD.

#### *3.1.4.4 Compare and Adjust Drastic Asymmetries*

At this stage of the script's execution, a new value has been determined and stored for each of the 64 craniofacial anthropometric measurements of the current crowd member object being processed. Before the script moves on to assigning these values to the measurement variables of the crowd member object a series of steps are performed to check and adjust the asymmetries of measurements that have values that differ from one side of the face to the other. These steps are important because it is possible for the system to assign the left and right measurements with values near the opposite extremes. While this may be potentially possible it is highly uncommon, and if left unchanged would cause many of the resulting faces produced by this implementation to exhibit drastic asymmetries.

During this Asymmetry Adjustment Phase, the script runs through all of the measurements and compares the two sides to each other. Some examples of measurements that could have different values depending on the side of the face are the Inner Canthal Width and Outer Canthal Width measurements, in which you have Inner Canthal Width Left and Outer Canthal Width Left values as well as Inner Canthal Width Right and Outercanthal Width Right values.

While the goal in this phase is to limit asymmetry, it is important to note that asymmetry plays an important role in facial uniqueness and capturing human likeness. For this reason the script is designed to retain some of the asymmetry that has resulted during the quasi-random generation of each measurement value. To do so, the script creates a randomly generated percentage range for each facial region. The maximum and minimum values for this range are designed to not exceed +$5\,\%$ and -$5\,\%$ respectively. The range is also designed to always have a minimum value that is less than zero and a maximum value that is greater than zero.

Once each facial region has been assigned a percentage of asymmetry range (PAR) the script begins the process of comparing the measurements of the left and right sides of the face. The following process is repeated until each pairing of measurements has been compared, processed, and adjusted.

This process begins with the script selecting a random value from the PAR of the region associated with the current measure being processed. The sign of the selected value will determine which side of the measurement will ultimately be larger in magnitude. For example if the sign of the selected value is negative, then the asymmetry adjustment will result in the left side of the measure being greater in magnitude. In comparison, if the sign of the selected value is positive then the asymmetry adjustment will result in the right side of the measure being greater in magnitude. Thus, either result will exhibit some level of asymmetry but the level of asymmetry will be limited by the magnitude of the value selected from the PAR.

The portion of the asymmetry adjustment process that follows the selection of a value from the PAR, is similar for both the scenario in which the sign of the value is positive and in which the sign of the value is negative. To outline this portion of the asymmetry adjustment process succinctly and clearly, I will describe a scenario in which the value that is selected from the PAR is +$2.50\,\%$. In this scenario the final result would produce the right side of the measurement being greater in magnitude than the left side of the measurement. To achieve this the script only changes the value of the left side of the current measurement being processed. First the script would set the value of the left side of the measurement to be equal in magnitude to that of the right, regardless of its current value. Next the script reduces the magnitude of the value of the left side of the measurement by $2.50\,\%$. This achieves a result in which the left side of the measurement and the right side of the measurement continue to display some level of asymmetry while also avoiding drastic asymmetry.

#### 3.1.4.5 Assignment of Values to Crowd Characters

Once the script finishes comparing and adjusting the asymmetries of the face, it has completed all the requisite procedures associated with the generation of new values for each of the 64 craniofacial anthropometric measurements used for the facial deformation system. Upon completion of this phase, the script assigns each of the new measurement values to the corresponding craniofacial anthropometric measurement variable of the current crowd character.

#### 3.1.4.6 Updating the Position(s) of Craniofacial Landmarks

The next step of the process involves the application of the newly generated measurement values to the facial deformation system. The application of the values to the deformation system will ultimately result in a new conformation of the base face mesh. The new conformation thus being a representation of the newly generated craniofacial anthropometric measurement values. At a summary level, this phase of the script's execution involves comparing the new value of each measurement with the initial value, determining how to change the landmarks of said measure to reflect the difference, and lastly updating the translational attributes of each landmark.

It is important to note that during this phase the script performs a consistent series of comparisons and calculations for each craniofacial anthropometric measurement in a predetermined order. This repetitive process is continued until the script has processed each measurement fully. The predetermined order that the script follows during this phase was designed to provide consistent and logical results and culminated through the analysis of the craniofacial anthropometric measurements outlined by Farkas to provide the most consistent and logical results. The following factors were analyzed to produce this order of operations:

- Which region or multiple regions of the face does each measurement affect.
- For each landmark, how many total measures are affected by its location.
- Does the measure affect a localized region or the face/head as a whole.

#### 3.1.4.6.1 Determine Proper Method of Change for Each Parameter's Landmarks

The repetitive process performed by the script on each measurement, begins with the comparison of the current value of the current measurement being processed to the new value that was recently assigned. Once the script computes the difference in value between the two, it performs the actions needed to update the landmarks associated with the measurement in a manner that reflects the new value of the measurement.

When determining how to change the positions of both landmarks for each measurement based on the difference in between the new value of the measurement and its original value, there are multiple scenarios the script must first factor. The first of which being - is the measurement aligned with a cardinal axis or not?

If the measurement is aligned with a cardinal axis the calculations involved in determining how to change the position of the landmarks associated are simple. In this case, the script either increases or decreases the value of the component of each landmark's translation that is associated with the cardinal axis of which the measurement is aligned. For example if the measurement is associated with facial width then it is aligned with the X-coordinate axis, for our purposes. This would mean that the script would simply change the value of the X-component of the translation attribute for one or both of the landmarks associated with the measurement. If the measurement is not aligned with a single coordinate axis then the script must evaluate the magnitude of each component of the vector so as to translate the landmarks along the vector of the measurement.

The second scenario considered by the script is whether or not one of the two landmarks that are associated with the current measurement being processed, is aligned with the facial midline. An example of this scenario would be the Nasal Ala Length measurements. One of the endpoints for both the Nasal Ala Length Left measurement and the Nasal Ala Length Right measurement, is the pronasale landmark. The pronasale landmark is one that would be considered as aligned with the facial midline. The other endpoint for both measures is the relative side's alare landmark. Both the right and left alare landmarks are not aligned with the facial midline. When one of the two landmarks is aligned with the facial midline, as is the case for the example described above, the script calculates a new position for the facial midline aligned landmark so that it continues to remain aligned. This is done to prevent landmarks along the facial midline from pushing away from the plane of symmetry that bisects the left and right sides of the face. The first step of this calculation involves the script assessing which of the current measurement's associated landmarks are designated to be changed. The script interprets these landmarks as Landmark A and Landmark B. This factor has been predetermined for both landmarks of each measurement. Similar to the predetermined order by which the script processes each measurement during this general phase, the factor of which landmarks are designated to be changed for each measurement was determined through careful analysis of the network of cause and effect connections that exists within the craniofacial anthropometric measurement structure.

In addition to the factor of which landmarks are assigned to be changed for each measurement, it is important for the script to assess if there is a center point assigned to or associated with the current measurement. For scenarios where this is true, the landmarks of the measurement will be adjusted in relation to the center point. This process is designed to retain the existing proportional relationship of the landmarks and the center point of a specific measurement. A prime example of this would be some of the orbital measurements. Some of these measurements span the upper and lower sections of the orbit, meaning they cross the pupil. For these measurements it would be beneficial to maintain the proportions of these measurement's landmarks in relation to the pupil.

In these scenarios the script calculates the difference between the each landmark and the center point. These values are then compared to the full value of the measure to understand the proportional relationship of the landmark A to center vector's value and the landmark B to center vector's value. As the script calculates new positional values for the landmarks of the current measure it maintains these proportions.

#### 3.1.4.6.2 Calculate the New Translational Values of the Affected Landmarks

After each of these factors is considered the script determines and assigns new positional values for the designated landmarks associated with the current measurement so that the distance between the two landmarks is equal to the new value of the measurement while maintaining the spatial relationship of the two landmarks as outlined by the factors.

3.1.4.6.3 Store and Assign the New Translational Values To Facial Mesh Deformation System

The positional values for each landmark directly drive the three dimensional position of the controls of the deformation system that changes the conformation of the base face mesh; meaning that once the script has processed each measurement in the manner described and per the order previously outlined, the conformation of the base face mesh has been changed to reflect the new and unique values of each measure.

### 3.1.5 Development of Rig:

There are many tools and processes available for the purposes of manipulating a mesh into a new form. After performing an evaluation of the available tools, software, and techniques, the following method for deforming the base facial mesh was developed. This method utilizes tools produced by Autodesk in the software Maya and a tool created by the company Tool Chefs as a plugin for Maya.

This deformation system, hereafter referred to as the rig, is a combination of three subsystems. The results of each subsystem are later combined together to create a more complex and effective result. Each subsystem is designed to address key aspects of the desired overall deformation of the face. The highest level of control for the rig contains a set of locators for each of the craniofacial landmarks associated with the selection of anthropometric measurements previously outlined. Each subsystem contains joints that also relate to each craniofacial landmark and are positionally constrained to the locators previously described. As new crowd faces are generated and the positional values for each landmark are updated, the position of each locator - and subsequently each joint - will be updated to match the corresponding landmark.

The rig is designed to achieve a desired value both anatomical accuracy and visual believability. In the instances in which these two interests contradict each other the rig is designed to create a result that balances the objectives. This may mean that the resulting face is anatomically informed rather than wholly accurate.

For a resulting face to be anatomically accurate, the locators must reflect the anatomical qualities of the associated craniofacial landmarks, such that the rules and anatomical definitions used to identify and locate each craniofacial landmark would need to remain true for the locators in context to the resulting facial mesh. For example, anatomical landmarks by their nature, do not exist off the surface of the face. As the position of each landmark is changed during the generation of faces, the associated locator should remain connected to the surface of the resulting facial mesh. The task of keeping the surface of a deformed object attached directly to the driving object is a difficult task when performed concurrently with the task of creating smooth, pleasing deformations.

Smooth deformations are produced through blending and weighting different deformers of the same target together. When more than one joint is given influence over the vertices of a mesh, the joints are each assigned a certain percentage of the total deformation influence for each vertex. If a joint has full influence on a vertex, the vertex inherits all of the transformations applied to the joint. Otherwise, the vertex inherits a portion of the transformations applied to the joint (i.e., a weighted deformation). This feature allows for a rigger (i.e., an individual that creates deformation systems) to create smooth and complex deformations. The byproduct of this is that as a joint changes position, its distance from the surface of the mesh grows proportionally. This result contradicts the objective of the landmarks remaining connected to the surface of the face.

This example clearly illustrates some of the difficulties that arose in the effort to achieve both anatomical accuracy and visual believability through the rig. The tools that allow for smooth and complex deformations at times worked in opposition towards the objective of anatomical accuracy. Through experimentation it was determined that the disparity between these two objectives was magnified in instances in which multiple deformation objectives were desired from a single deformer. By limiting the number of deformation objectives a single deformer was designed to achieve then blending together the results of multiple such deformers a more desirable balance between anatomical accuracy and visual believability was possible. This experimentation ultimately resulted in the rig including a total three main subsystems, each designed to achieve a narrowed deformation objective. Some subsystems include further subsystems that address increasingly narrowed deformation objectives.

The first subsystem, as depicted in Figure 3.5, focuses on the deformation of the mouth and orbital regions of the face. It includes several subsystems that each control a different direction of deformation for one of the regions. These include the horizontal and vertical deformation for the eyes and mouth.

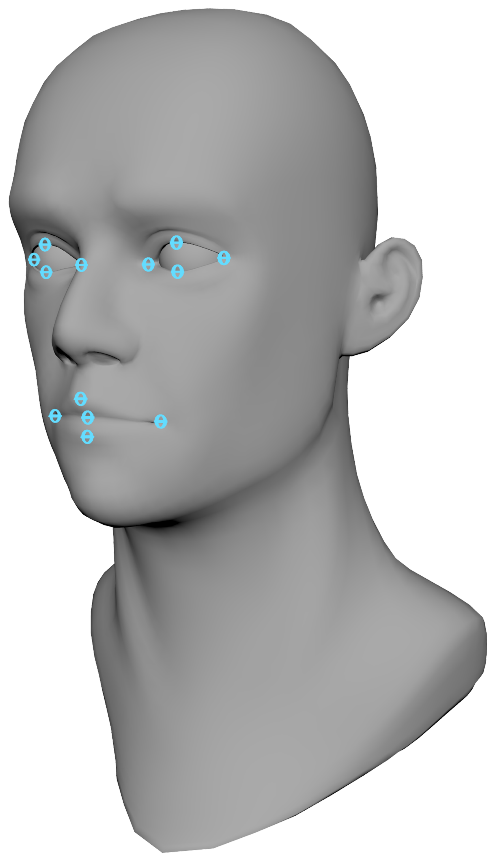


Figure 3.5: Deformation subsystem #1: Orbital and Oral regions of the face. The light blue colored circles represent the joints of the first subsystem.

The other two subsystems are designed to address holistic deformations of the face. One of the subsystems maximizes the local influence of landmarks that pertain to a small portion of the face's surface area. The other subsystem maximizes the global influence of landmarks that pertain to a larger portion of the face's surface area. Both subsystems utilize a complex lattice deformer

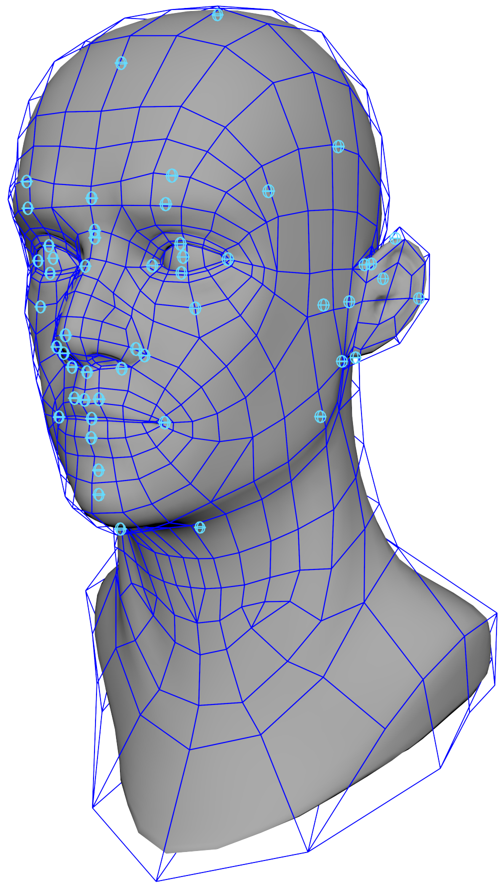


Figure 3.6: Deformation subsystem #2: Local deformation of the face. The blue lines represent the three dimensional form of the Locally Focused Deformation Cage.

created by the company ToolChefs. The deformer is called a harmonic deformer and "uses a closed mesh (cage) to deform a higher resolution mesh." The implementation of the tool was based on the siggraph paper produced by Pixar Animation Studios, titled Harmonic Coordinates for Character Articulation (Joshi, Meyer, DeRose, Green, & Sanocki, 2007).

In the development of the locally focused subsystem, as depicted in Figure 3.6, a cage mesh for harmonic deformer was created by duplicating the base face mesh and lowering its resolution by one level. The landmark based joints for this subsystem were then bound to the cage mesh so that as the position of the landmark joints change, the cage would deform, resulting ultimately in a localized deformation of a duplicate of the initial base mesh. This structure allows for the landmark joints to influence the portions of the face they pertain to without significantly influencing areas

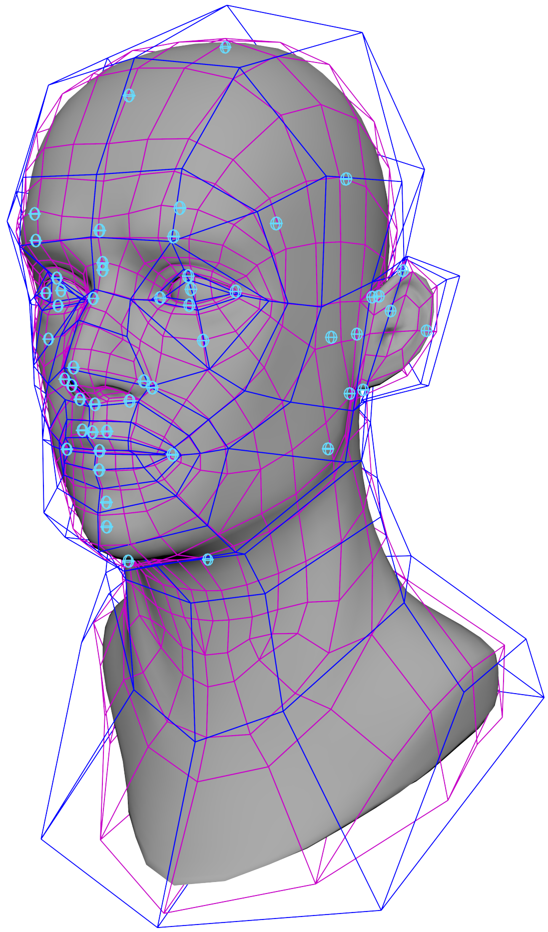


Figure 3.7: Deformation subsystem #3: Local deformation of the face. The blue lines represent the three dimensional form of the Globally Focused Deformation Cage.

under that pertain to other landmark joints.

In the development of the globally focused subsystem, as depicted in Figure 3.7, this abstraction process was continued. The cage for the system was modeled to be lower than the resolution of the original face mesh by two levels. This cage would then deform a duplicate of the cage mesh from the previous subsystem, which would intern deform an additional duplicate of the base facial mesh. This means that the landmark based joints of this subsystem influence the second level cage mesh, thus deforming the duplicate first level cage and ultimately a duplicate base facial mesh. This structure of deformations allows for the influence of each joint to have a broad effect over the deformation of the face.

Once each component of the rig was complete the results of each component were set up as blendshape deformation drivers of an additional duplicate of the base face geometry. A blendshape deformer connects together two geometries with the same topology in a driver and driven relationship. The driving geometry is typically a duplicate of the driven geometry that has either been modeled with deformations or is deformed by joints or some other deformer. The blendshape deformer provides the ability to move each vertex of the driven geometry towards the position of the analogous vertex of the driving geometry. The blendshape deformer has an attribute that controls the influence of the driving geometry over the driven geometry. For example if the blendshape is set with fifty percent influence, each vertex of the driven mesh will be changed to the midpoint along the linear vector between the position of the corresponding vertex on the driven mesh and the driving mesh.

When more than one geometry is assigned as a blend shape, the influence of each blendshape is added together. This feature allowed the different deformation results of each component of the rig to be combined to create a final deformation result that blends together the desired deformations of each system.

In summary the complete rig, including the three subsystems and final resulting mesh, takes the position of each landmark as input to change the shape of the base facial mesh into a new form. This new form of the base facial mesh is now a unique shape sharing the same topological structure of the original shape. It is a representation of the changes to the craniofacial anthropometric measurement values that were assigned during the crowd face generation performed by the script previously described.

The result of this implementation is a tool that allows a user to create multiple facial meshes that are anthropometrically unique from one another. In utilizing the tool, the user would begin by entering the following information into the fields provided in the first window of tools user interface, as depicted in Figure E.1 in Appendix E.

- Crowd size
- Name of crowd

- Gender ratio of crowd members
- Demographic distribution of crowd

Once the user selects the button labeled “Generate Crowd” the script performs the generation process previously described for the number of crowd members requested by the user. This process includes the generation of ethnically based randomized values for each measurement, the assignment of those values to the corresponding crowd character, the importation of the facial rig, and lastly the process of updating the positions of each landmark for each measurement per the measurement’s newly assigned value. This results in a chain of automatic successive updates to the facial rig which ultimately results in the visual result of a new and unique facial mesh that represents the randomly generated craniofacial anthropometric measurements assigned to that crowd member.

Once the tool has completed the generation process of each crowd member it displays a new window in the tool’s user interface, as depicted in Figure E.2 in Appendix E. The left portion of this window contains a list of each crowd member of the current crowd . The right portion of the window displays the values for each measurement of the selected crowd member. When the user selects a different crowd member in the left portion of the window, the right portion updates to show the corresponding measurement values. The right portion of the window also contains controls that allow the user to manually change each anthropometric measurement value. This ability allows the user to artistically tweak the results of the initial randomized generation. This feature is essential for the tool to be useful and effective in a production environment, in that it allows for the results of the tool to be crafted towards the specific visual result desired of the crowd which varies depending on the artistic direction of each production.

### 3.2 Results

The common methods for creating facial models for crowd characters lack the flexibility to remain effective in circumstances where demands frequently change and factors affecting the manner in which the characters’ faces are viewed by the audience are difficult to predetermine. These circumstances are taxing on modeling departments and often require the redistribution of previously allocated resources. In this research, a flexible and innovative method for generating crowd character facial models has been described and implemented. This method involves the use of a procedural, parametrically driven system built around craniofacial anthropometric metric parameters, to deform a base facial mesh into new conformations. By quasi-randomly generating values for each craniofacial anthropometric parameter - in accordance with existing craniofacial anthropometric data of the specific gender and ethnic combination of each crowd character - the base facial mesh of the system is deformed into a new conformation that matches in the previous in terms of local topology but varies in that it now represents the new values for each craniofacial anthropometric parameter.

The level of variability in the conformation of the resulting facial meshes produced by the method implemented in this research, is illustrated in the following figures - Figure 3.8, Figure 3.9, Figure 3.10, Figure 3.11, Figure 3.12, and Figure 3.13. Each of these figures depicts the visual results of specific two dimensional plots of different craniofacial anthropometric parameter combinations. Figures 3.8-3.10 show the difference in the conformation of the orbital regions of different resulting facial meshes for which the parameters of Innercanthal Width, Palpebral Fissure Length, and Eye Fissure Height have been assigned different values. Figure 3.11 shows the difference in the conformation of the cranium of different resulting facial meshes for which the parameters of Maximum Cranial Width and Morphological Facial Height have been assigned different values. Figure 3.12 shows the difference in the conformation of the oral regions of different resulting facial meshes for which the parameters of Labial Fissure Width and Upper Lip Height have been assigned different values. Figure 3.13 shows the difference in the conformation of the nasal regions of different resulting facial meshes for which the parameters of Nasal Width and Nasal Protrusion have been assigned different values.

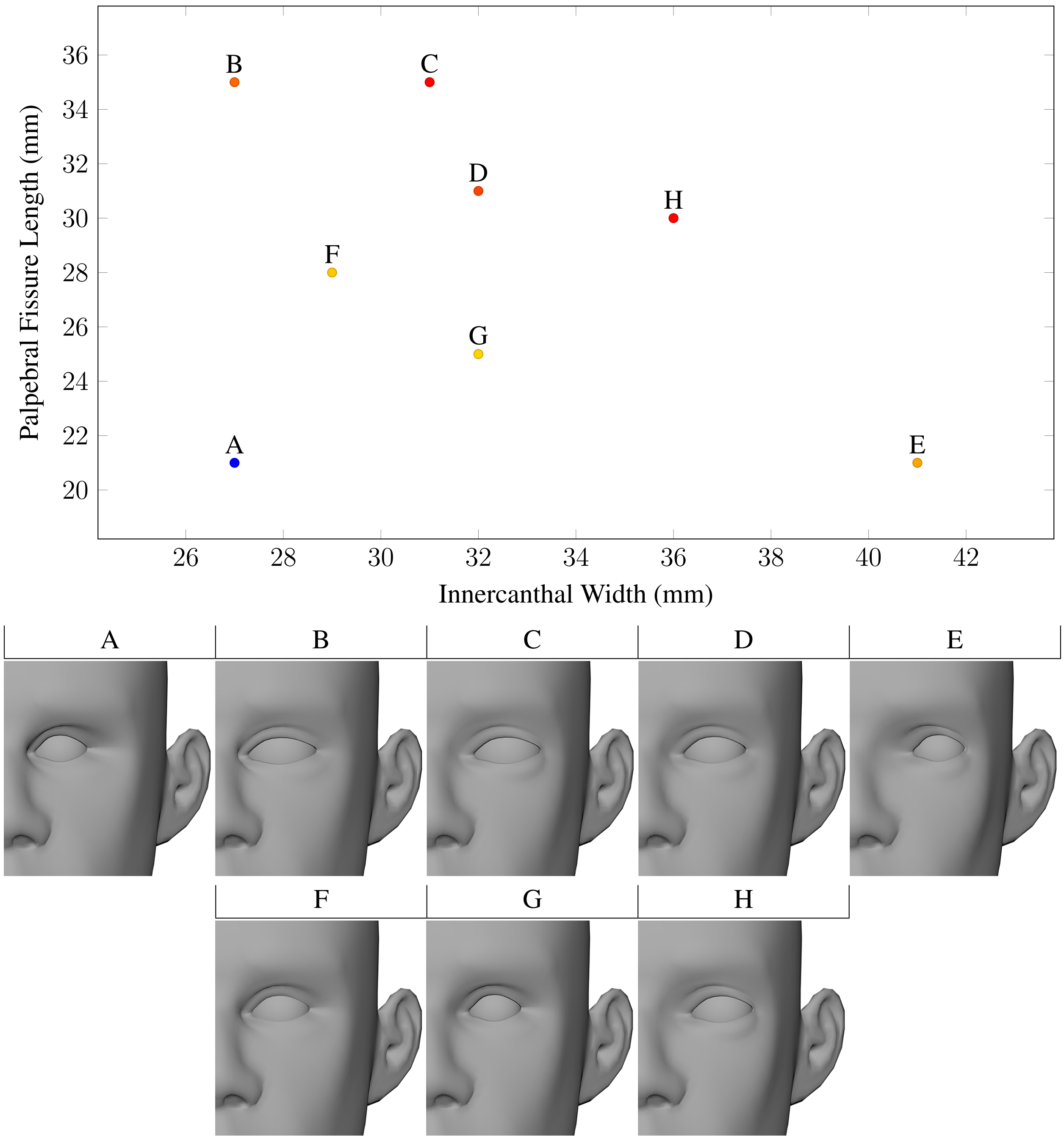



Figure 3.8: Deformation possible with parametric variation: Innercanthal Width and Palpebral Fissure Length

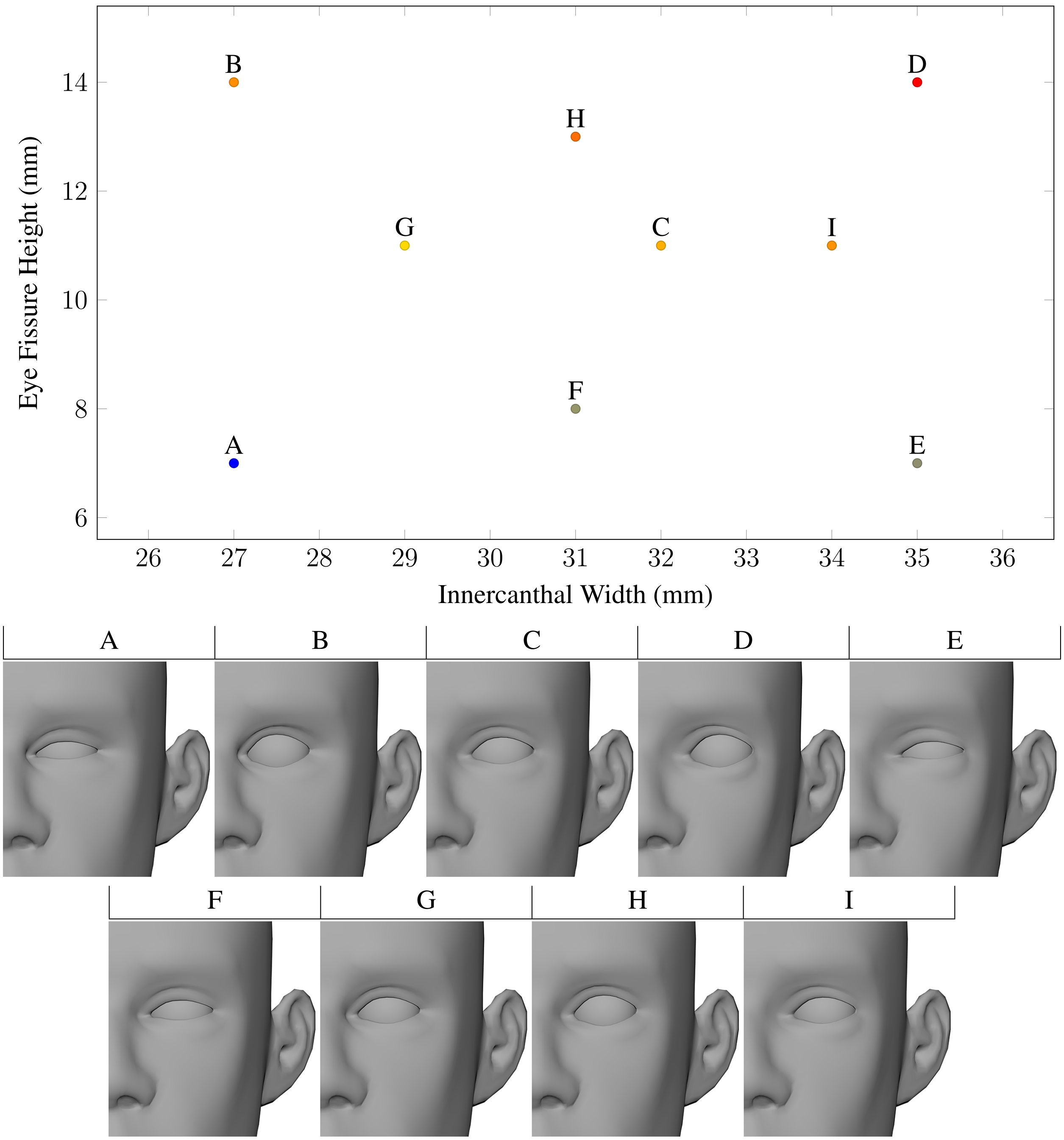



Figure 3.9: Deformation possible with parametric variation: Innercanthal Width and Eye Fissure Height

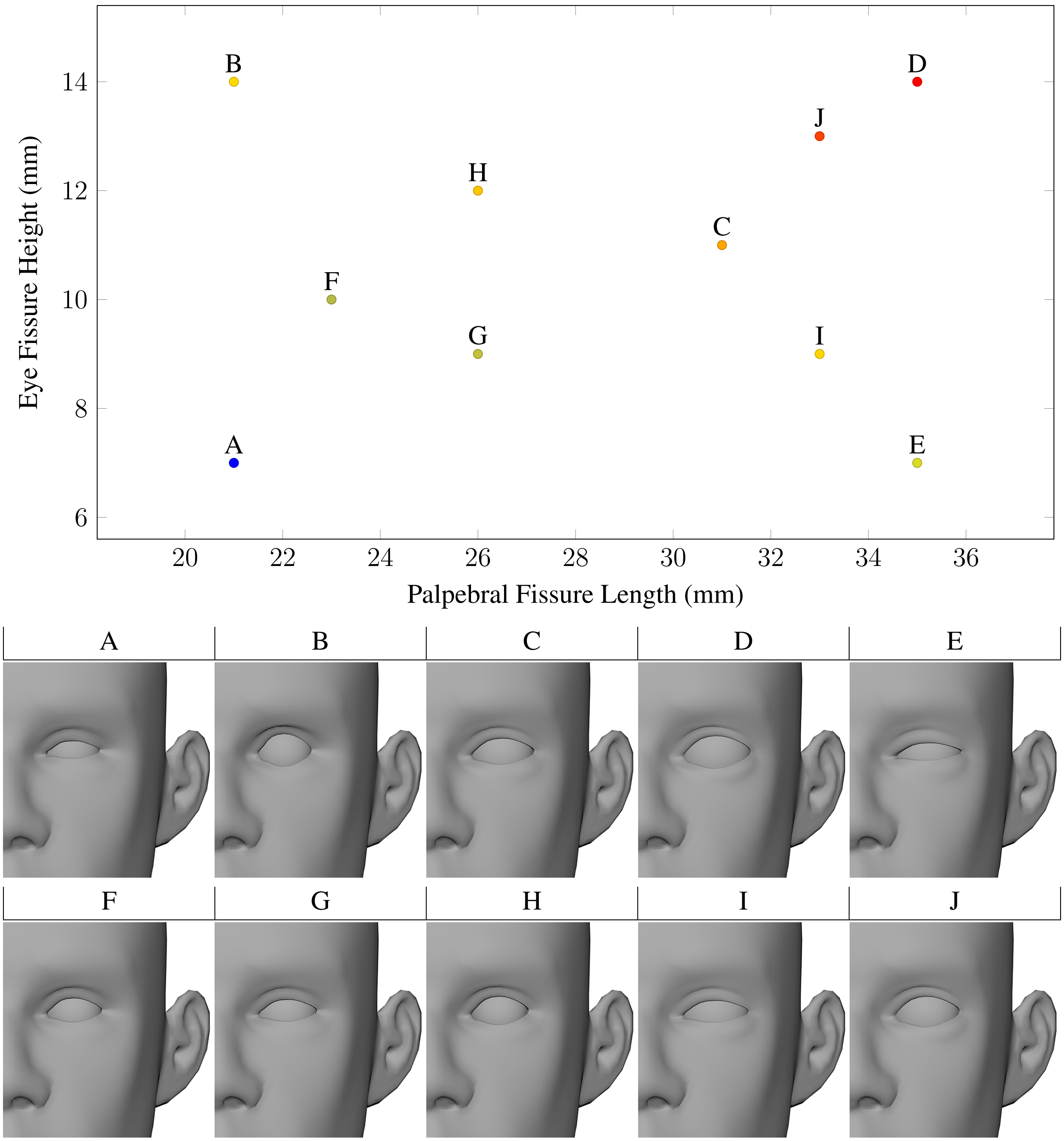



Figure 3.10: Deformation possible with parametric variation: Palpebral Fissure Length and Eye Fissure Height

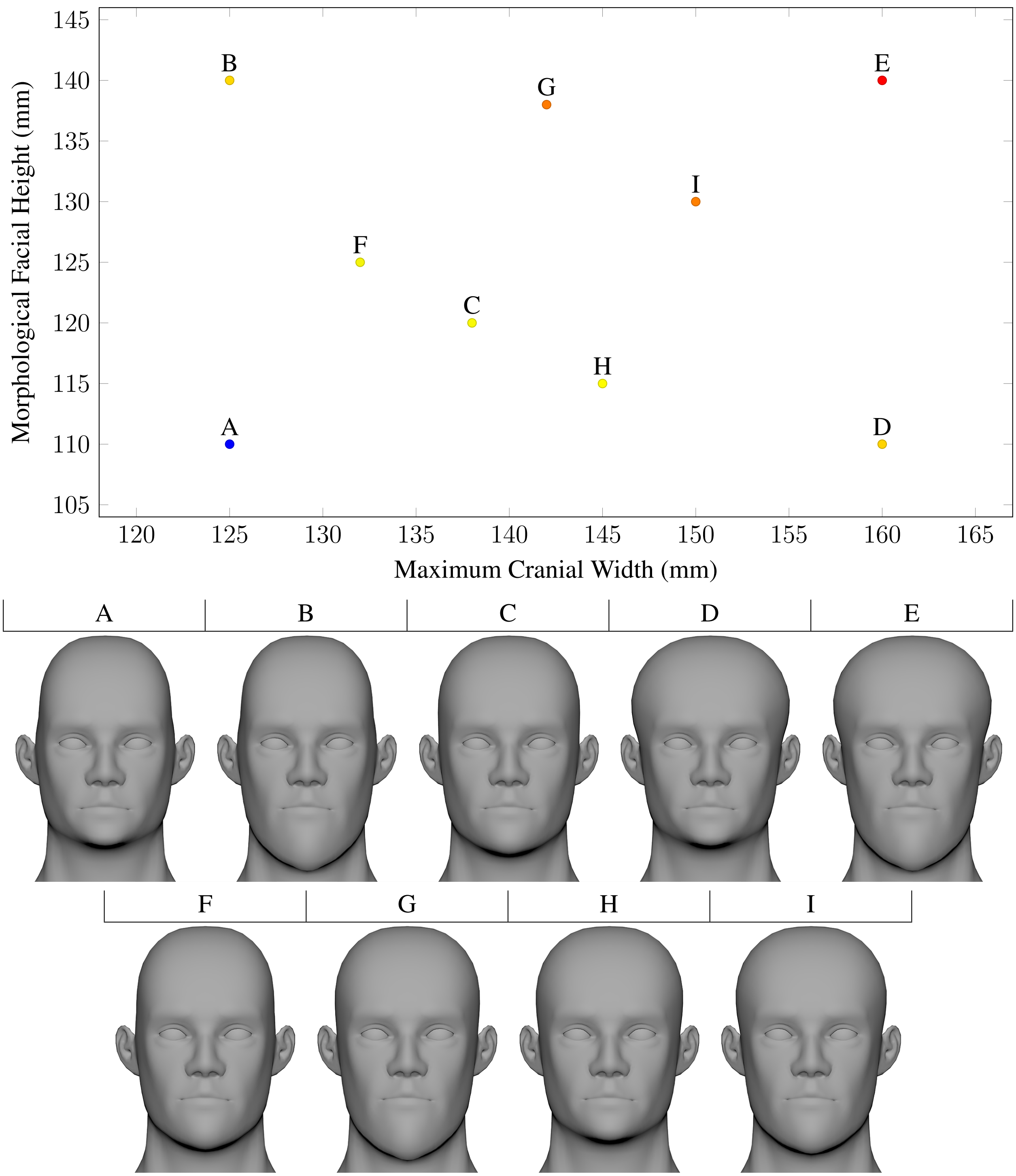



Figure 3.11: Deformation possible with parametric variation: Maximum Cranial Width Width and Morphological Facial Height

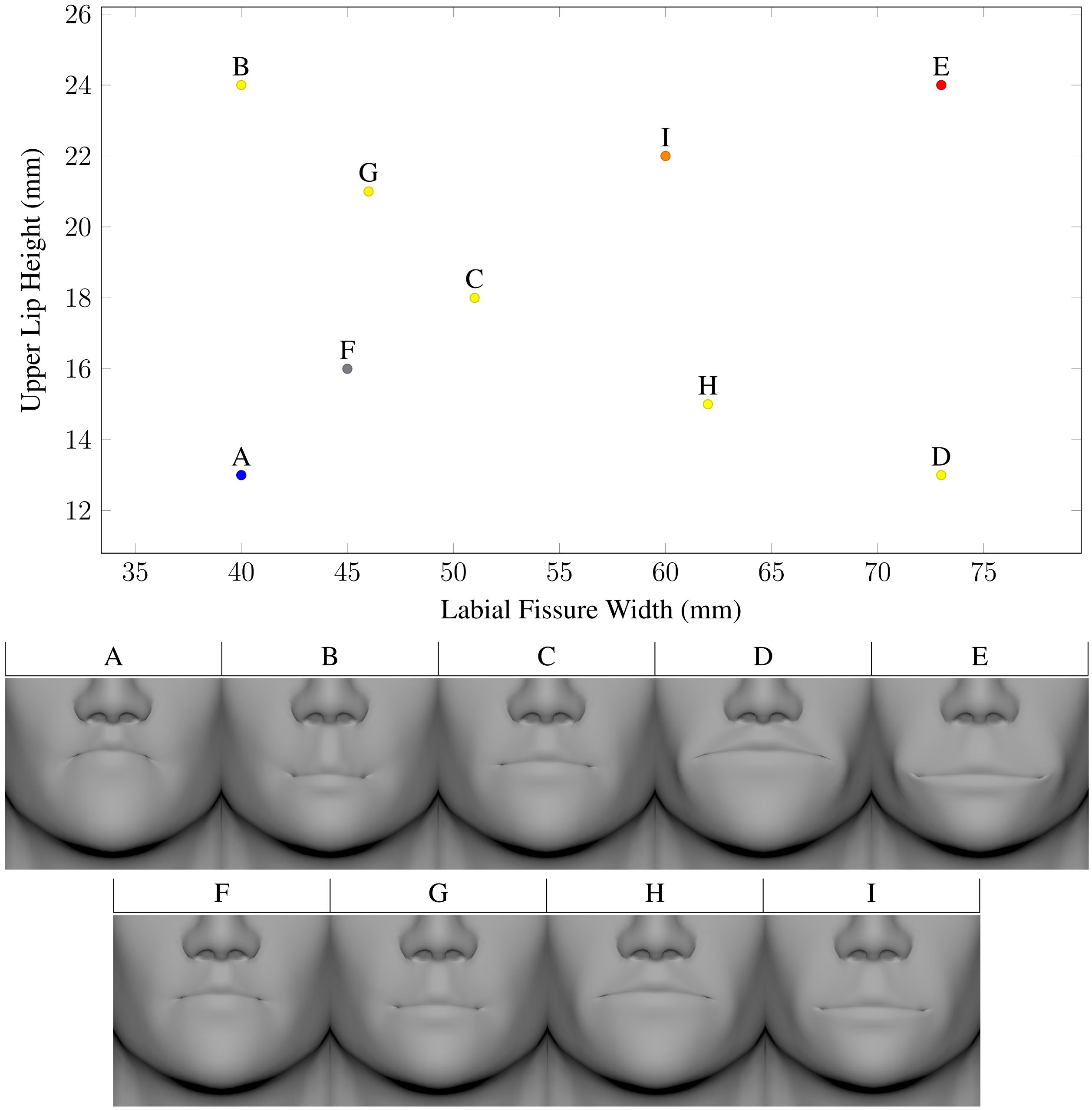



Figure 3.12: Deformation possible with parametric variation: Labial Fissure Width and Upper Lip Height

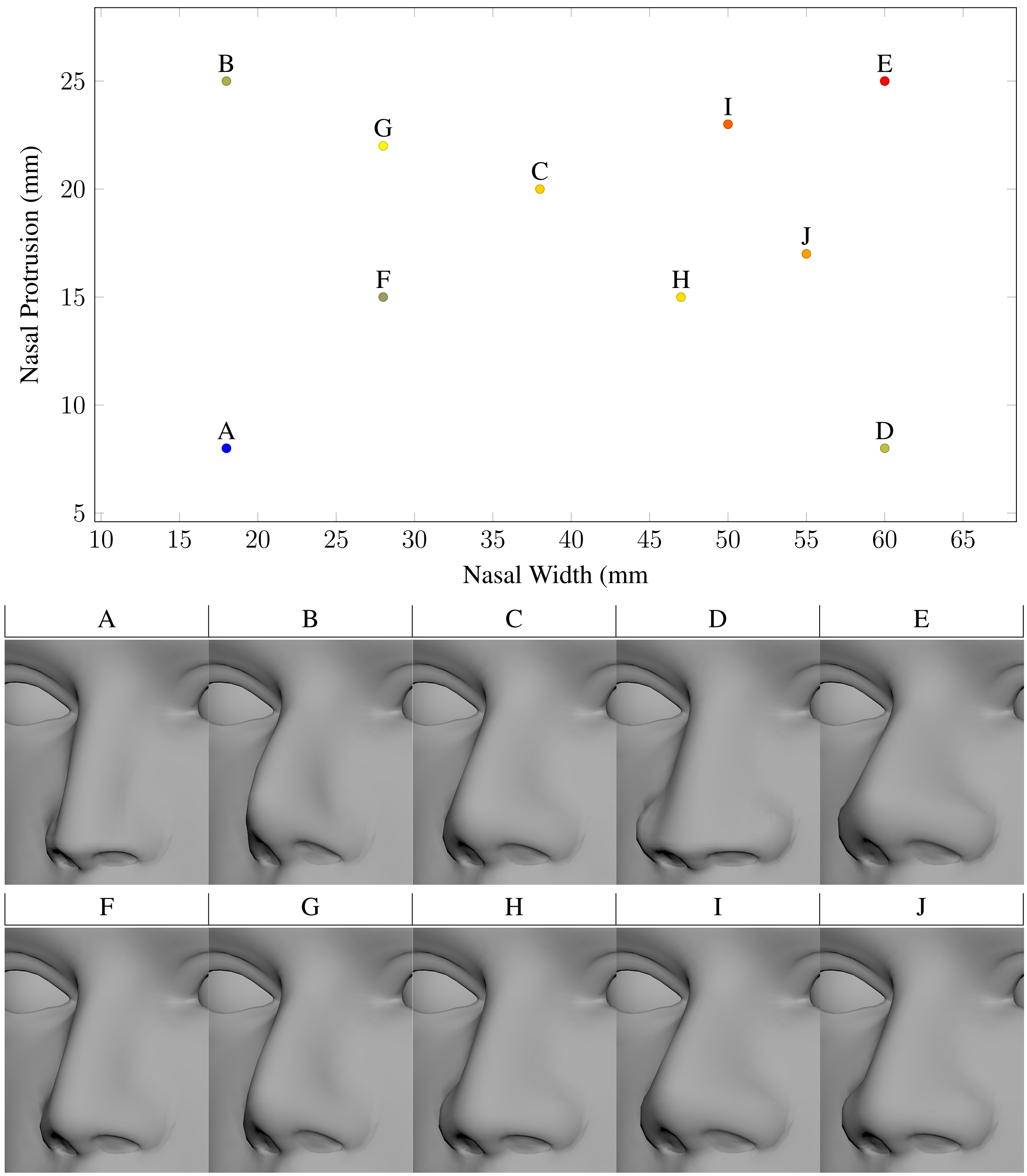



Figure 3.13: Deformation possible with parametric variation: Nasal Width and Nasal Protrusion

A significant level of visual variety and distinct uniqueness can be observed in the different facial conformations of the results displayed in Figures 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13. These figures illustrate the level of variability in the conformation of different facial meshes produced by the system when only two or three parameters are changed. The method implemented in this research uses a total of 64 parameters to deform the face in accordance with craniofacial anthropometric measurements described by Farkas (1994). The figure below - Figure 3.14 - depicts the

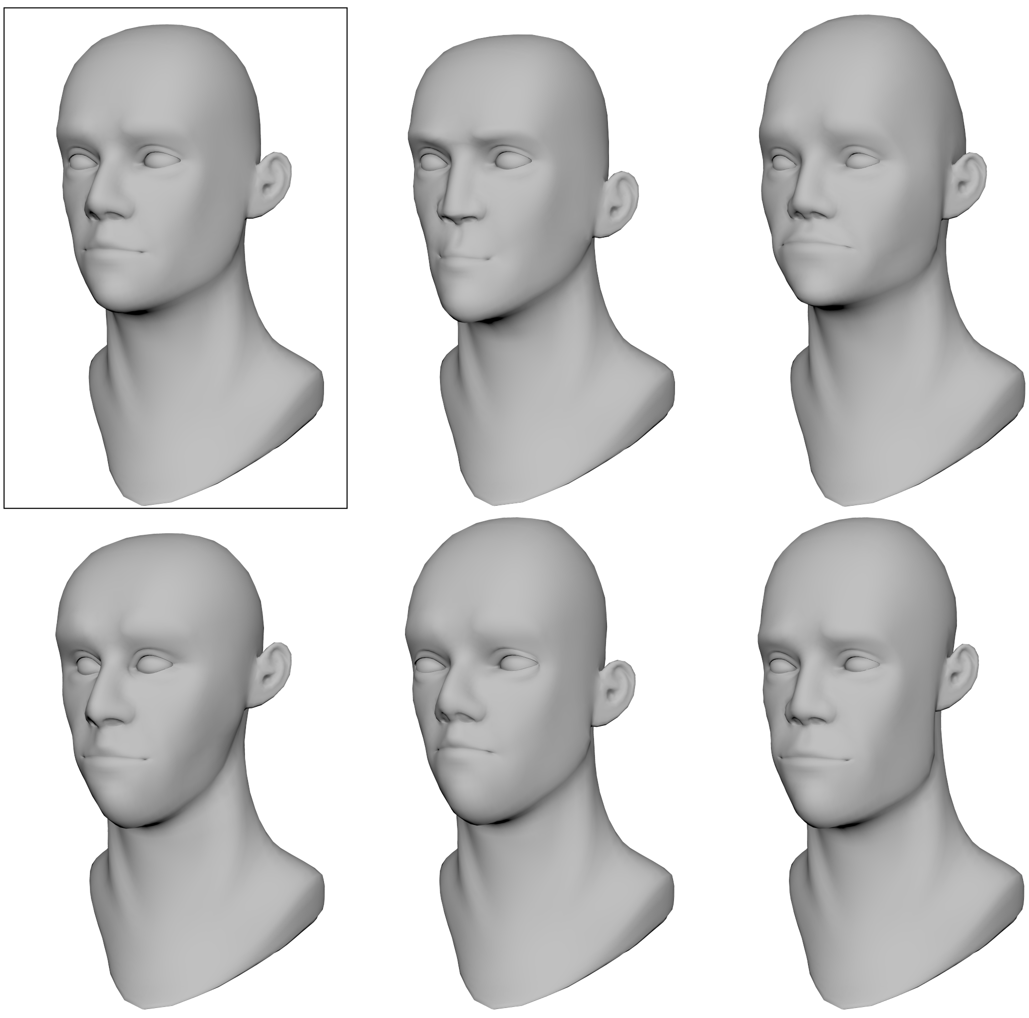


Figure 3.14: A collection of facial mesh results that are each unique plots in the full 64 dimensional parametric space of the procedural deformation system. The framed face is the base facial mesh.

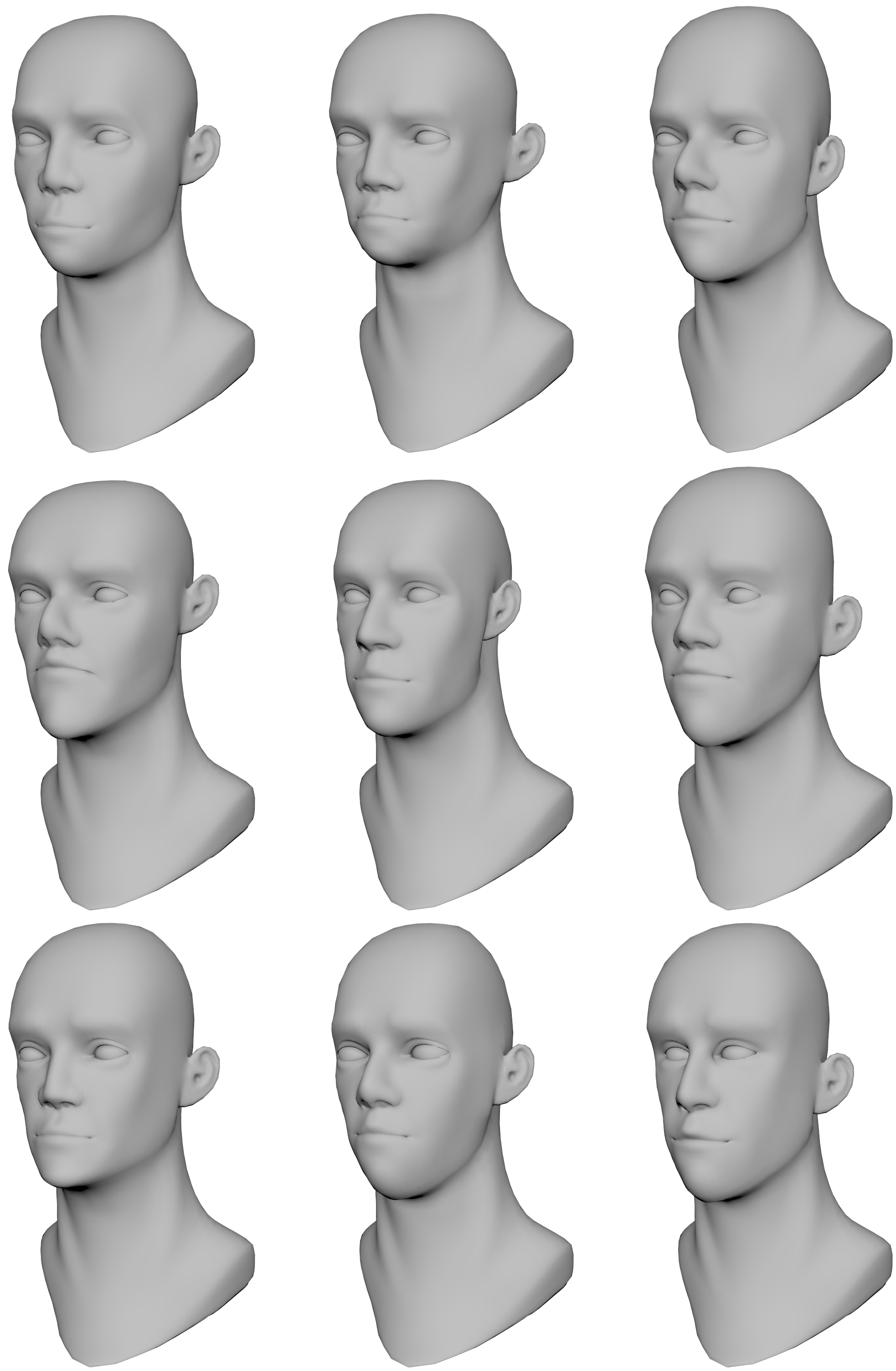


Figure 3.14 (cont.): A collection of facial mesh results that are each unique plots in the full 64 dimensional parametric space of the procedural deformation system. The framed face is the base facial mesh.

level of variability in the conformation of different facial meshes produced by the system when each of the 64 parameters are randomized. Therefore, the faces depicted in this figure are examples of random plots in the full 64 dimensional space producible by the method implemented in this research.

A noteworthy observation in the results depicted in Figure 3.14 is that some of parameter values, particularly those of more significant difference from the initial value represented in the base facial mesh, translate to final deformation results that appear to break away from an anatomically believable underlying cranial skeletal structure.

Additionally, it is important to observe in the results of Figure 3.14, how the deformation result is influenced by the explicitly added rules that determine how the positions of landmarks are changed to reflect the newly assigned values of each parameter. The measurement system provided by Farkas (1994) was originally designed for the purposes of sampling a complex shape (i.e., the human face). It was not Farkas's intent to use the system of metrics and landmarks to recreate a human face, nor generate a new and unique human face, therefore he does not provide an explicit rule set to define how the system could be repurposed for the method discussed in this research. For this reason, additional rules were developed in the implementation of the method of this research, to explicitly dictate how the complex facial shape would change as a result of the change of each craniofacial anthropometric parameter. Through the development of the specific method implementation described in this research, various scenarios were encountered for which new rules, designed to mimic the complex implicit structures and connections of the real human face, were identified. While these rules may be effective in addressing specific scenarios they are limited in their ability to adequately address every scenario in a manner that accurately reflects what is true of a real human face. These limitations can be observed in the results of Figure 3.14.

In addition to analyzing the level of visual variety and distinctive uniqueness in the resulting facial meshes produced by the method implemented in this research - as displayed in the previous figures - a collection of facial mesh results produced within the full dimensional space of the parametric system were given textures and placed into a three-dimensional scene to be evaluated in a real-time context. As described previously, the existing modeling methods used for creating crowd characters become costly in scenarios in which the factors affecting the manner in which crowd characters' faces are viewed by the audience are difficult to predetermine. For some scenarios, these factors are difficult to predetermine because of the changes within the production context. As illustrated in the introduction, the most common scenarios for which these factors are difficult to predetermine are those of interactive real-time experiences. The cause for this being the nature of the user's ability to control the camera of these experiences. The method presented in this research seeks to address this issue and for that reason it is important to observe the results of the method implemented in this research in a real-time interactive context.

Figure 3.15 depicts a selection of the crowd facial meshes - produced by the method implemented in this research - as sculptures in a gallery. Because each face mesh shares the same topology and UV layout, they are able to utilize the same textures, reducing the production costs for each individual face as they are moved beyond the modeling stage through the rest of the production pipeline.

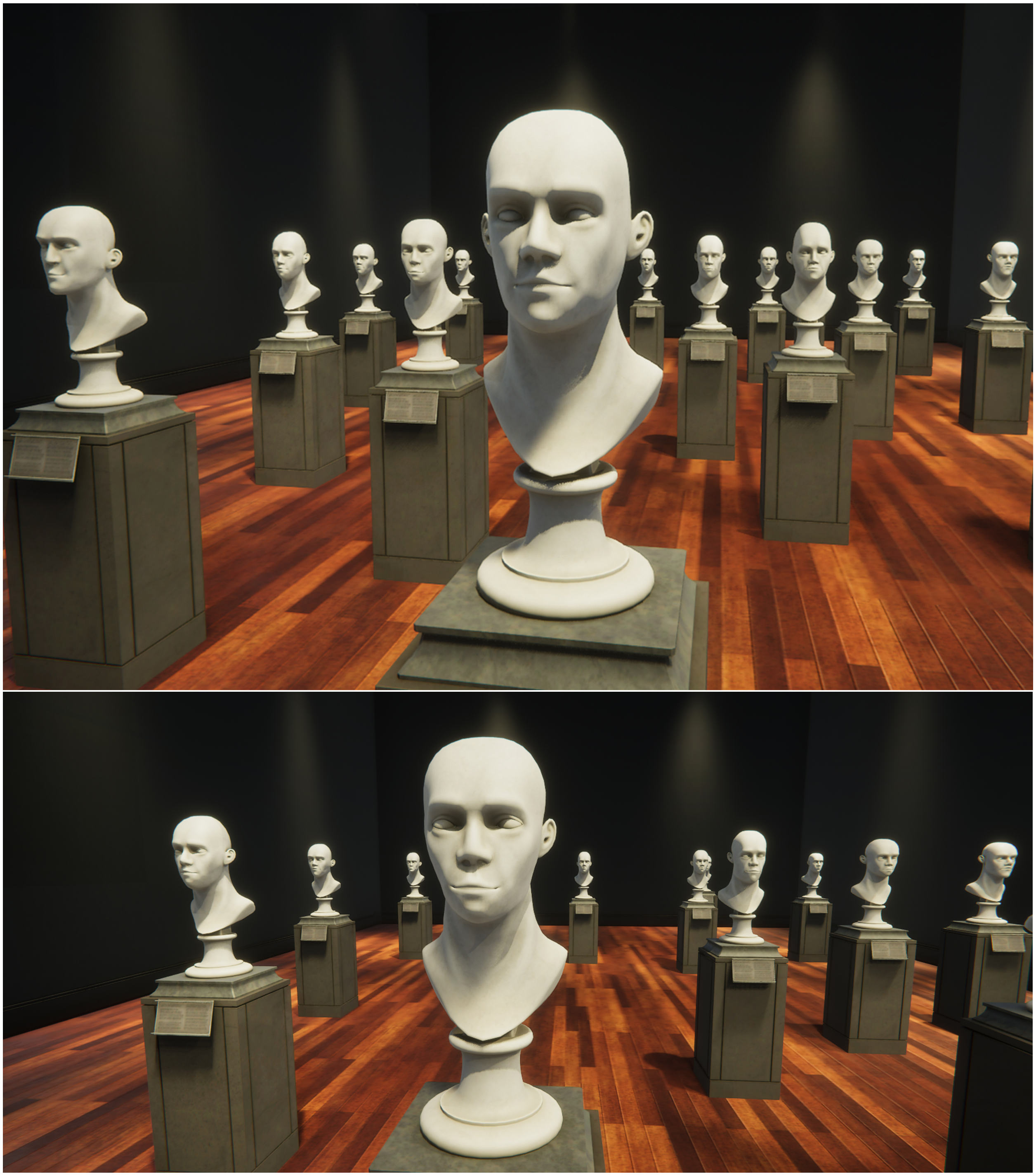


Figure 3.15: Images captured from a real-time interactive experience produced in Unity. In this scene a collection of final facial models are displayed as statues in a museum.

Figure 3.15 (cont.): Images captured from a real-time interactive experience produced in Unity. In this scene a collection of final facial models are displayed as statues in a museum.

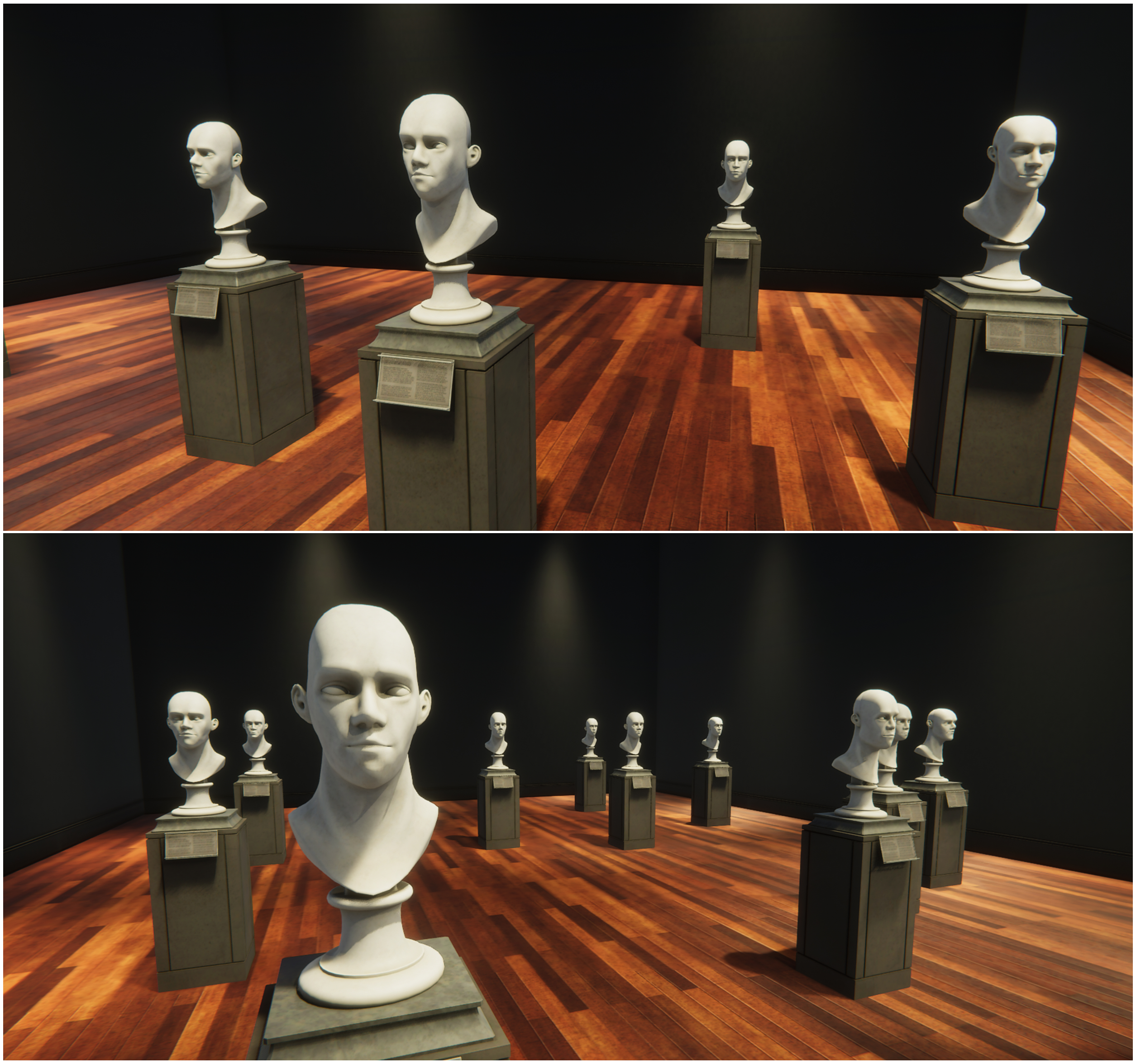


Figure 3.15 (cont.): Images captured from a real-time interactive experience produced in Unity. In this scene a collection of final facial models are displayed as statues in a museum.

In Figure 3.15, it can be observed that the resulting facial meshes produced by the method implemented in this thesis, hold up under the changing factors of time on screen and screen space contribution. This can be attributed to each result sharing the same level of fidelity as the base facial geometry. Additionally each face appears distinct from one another, even when the viewer chooses to isolate a specific face, thereby temporarily promoting it to a hero character in terms of the visual standards it is now subject to.

The resulting facial meshes produced by the system developed during the implementation of the method of this thesis also provides the artist with the ability to refine each facial mesh produced. As represented in Figure E.2, the system provides the artist with the ability to manually adjust the value of each craniofacial anthropometric parameter for each member of the generated crowd through a simple user interface. Additionally the artist could refine the shape of the face further by selecting the landmark controls directly and changing their position, thereby altering the conformation of the resulting facial mesh.

### 3.3 Discussion

The findings of this research illustrate that the method explored in this research is effective in producing crowd character facial models that exhibit substantial variability. In addition, the results demonstrate this method to be significantly more expedient at producing faces of notable uniqueness in comparison to the traditional process used for modeling faces. The ability to create considerable uniqueness quickly, improves an artist's ability to iterate within the modeling process. By proceduralizing the shaping of a human facial model, an artist is provided with the ability to quickly generate a result that is unique from the previous model.

The method of this research allows artists to quickly explore possible variations with limited risk, thereby encouraging iteration. If one result is not desirable to the modeler another result can be generated quickly without significant expense of his or her time and limited resources. This allows a modeler to engage in significantly more rounds of iteration in a limited amount of time as compared to the traditional modeling process.

### 3.3.1 Limitations

The task of creating character models for a crowd simulation is one with various complexities and challenges. One observation that resulted from this research is that the use of anthropometric measurements in the generation of facial models for crowd simulations is insufficient, in itself, to solve all of the problems specific to this task. One reason for this is that compression, aliasing, and distortion were observed in the results - as is often the case when reconstructing a complex continuous signal from samples.

The process of collecting craniofacial anthropometric measurements is akin to sampling the information and detail that describes the faces form and shape. This sampling results in a limited set of discrete information that describes the faces shape in part but not in full. It is therefore logical that any new faces constructed from the discrete samples collected would exhibit visual differences from that of actual humans. This is particularly evident when comparing the results of this research to the sample source - the sample source in this case being the faces of real people. There are many details and visual complexities in the faces of real people that are not evident in the results of this research. This is not surprising considering, there are many visual attributes of human faces that determine the uniqueness and identifiability of any given face and craniofacial anthropometric measurements are simply one such category of visual attributes.

#### *3.3.1.1 Factors Affecting Complexity Limits*

After analyzing the results of my research, I identified two key factors that limit the complexity of the samples obtained from collecting craniofacial anthropometric measurements. One key factor is related to the three dimensional data available for each landmark. After combining the craniofacial anthropometric data provided by Farkas (1994), it was observed that there is not complete three dimensional data for each landmark. As described previously in the implementation section of this document, the positional values of each landmark are extrapolated from the information associated with the anthropometric measurements provided by Farkas.

Each measurement is described and identified via the following information - the two landmarks the measure is between, the anatomical name that defines what aspect of the face the measure describes; and an indication of which three dimensional axis the measurement pertains to if applicable (i.e., there are some cases in which the measurement is not specific to a single axis but is rather a three dimensional linear vector). For example, the "anterior head height" measurement is defined as the difference in height between the vertex landmark and the nasion landmark. Through extrapolation, this measurement can be used to determine the vertical axis component of the vertex and nasion landmarks' translation. Table D.1 in Appendix D, presents the available cardinal axis information for each landmark and the names of the measurements from which that data is extrapolated. Of the information contained in this table, $73.71796\,\%$ of the cells are empty, indicating a significant loss of information as a result of the sampling process conducted through collecting anthropometric measurement data from human subjects.

The second key factor that limits the complexity of the samples resulting from the collection of craniofacial anthropometric measurements, pertains to the measurements and landmarks currently collected by anthropologists. Both in terms of their quantity and design, the measurements and landmarks collected lack the ability to effectively capture the nuances of the shape of the human face with a high level of fidelity. This is particularly true in relation to the features of the face such as the eyes, mouth, ears, and nose, that play vital roles in the visual quality of a face's overall uniqueness. A trained portraitist can attest to the fact that the key to capturing the likeness of an individual's face is identifying, capturing, and sometimes exaggerating the nuances that exist in the features of the human face. If the portraitist is unsuccessful in doing so, the result often feels generic and lacking a direct connection to the subject.

The observation of the limitations previously described presents several specific opportunities for future research. To address and supplement the limitations of the three dimensional information currently available for the landmarks presently used in the fields of craniofacial anthropometric measurement, the following future research is suggested:

- Use innovative machine learning techniques and artificial neural networks in the analysis of images/videos of human faces to locate and translate the positions of craniofacial landmarks from the input material into a CG three-dimensional space.

- Perform three-dimensional landmark data collection on human subjects using 3D scanning technology.

- Establish additional craniofacial anthropometric measurements to supplement and elaborate on the current set commonly used by researchers. Then perform collection studies of different ethnic groups using said measures.

To address the limitations in capturing the nuances of the human face, the following future research is suggested:

- Implement other forms of anatomical information into the generation of crowd faces. Examples for this include information such as age, BMI, skeletal and soft cartilage structure, musculature, dermatological features such as wrinkles, moles, etc...

- Perform facial feature specific anthropometric studies to better understand and describe through measurement the visual nuance that makes the features of each individual distinct.

- Perform anthropometric studies focused on facial asymmetry to identify possible correlations and common ranges of asymmetry for each facial feature.

# 4. CONCLUSION

This thesis has described a method for utilizing craniofacial anthropometric measurements in the development and generation of digital 3D crowd character facial models that exhibit unique and anatomically based diversity. This method, as evidenced by the results previously displayed, is objectively effective in the creation of crowd character facial models that are distinct from one another. In addition, this method provides an advancement to the traditional modeling process in terms of its expediency and randomization capabilities.

While the implementation presented in this thesis, specifically addresses the creation of facial models for crowd characters, the method itself could be applied to creation of the full bodied crowd character models with additional research and development. Additionally, it is important to qualify the results of this method by noting that the specific visual qualities of the results produced through the implementation previously described, are not anticipated to satisfy the specific art direction desired for every CG crowd simulation. While it is impractical for these specific models to be sufficient for every artistic context, the results displayed in figures 3.15 illustrate that the method itself can effectively be applied through a different implementation to produce models in other distinct artistic styles. The method of using anthropometric measurements as parameters for the procedural generation of faces for characters in a crowd simulation, is not limited to the particular context implemented in this research. With further research and development, the same method can be applied to other base facial models that are stylistically unique.

## 4.1 Future Work

The processes used for the development of crowd characters are continually evolving and as is evidence of the research recently performed, both commercially and academically, it can be expected that the amount of research in this area is likely to grow and expand. In collaboration with this ever increasing pursuit, I present the following recommendations for future work and additional possible applications of the work performed in this thesis.

One such application is in the area of real-time, interactive media, such as virtual reality experiences and video games. The method described in this research provides an innovative and unique storytelling device for these forms of CG media. It would be very interesting to integrate the anthropometrically driven facial rig described in this research into an interactive experience to dynamically change the faces of CG characters as players/viewers interact with and observe them. An application such as this, allows for the telling of many unique and compelling stories that otherwise would not be possible. Specifically, this would allow for the creation of novel, informative, and explorative experiences that challenge a user's social and societal perspective by dynamically changing the race, ethnicity or general appearance of other characters in response to the user's actions or inaction. The complex deformers used in the deformation system implemented in this thesis are not a standard compatible feature for most real-time engines, such as Unity, and thus additional research is needed to make such integration possible.

One possible avenue for this could be a procedural translational of the current deformation system into another form that is compatible with standard real-time engines, such as a blendshape deformation system. Blendshape driven facial rigs are commonly used for the animation of characters in real-time experiences. One example of an automated system that generates blendshape driven facial rigs is the online, 3D animation service platform, Polywink by Eisko™ (2021). One of the services provided by Polywink is the automation of the blendshape creation process for facial rigging. Polywink claims to be capable of generating FACS facial blendshapes for the character facial models of their customers "in mere hours" — a process "that would otherwise take weeks to achieve" as a result of "in-house procedural technology and through years of accumulated data powered by machine learning," (Eisko™, 2021). While the application of the Polywink platform is primarily concerned with procedurally rigging a customer's facial mesh for the purposes of expressive facial animation, machine learning could be used in a similar manner to procedurally translate the conformation focused deformation system of this thesis into a blendshape deformation system that is capable of similar functionality.

In addition to storytelling applications, the method described here would be effective as a tool in the creation of avatars for virtual representation in a variety of contexts such as, video games, social media, virtual tele-communication, etc... It would be particularly interesting to develop a system that takes information from a user about his or her face, such as photo or video reference, to create a three-dimensional avatar in the user's likeness. Such a system could use modern facial tracking tools, which are often driven by machine learning and data analysis algorithms, to identify and track craniofacial anthropometric landmarks on a user's face. The facial tracking information for each landmark could then be used to drive the landmark positions of the deformation system presented in this thesis.

Additionally, there are many other potential applications of machine learning to adapt and expand the capabilities of the craniofacial anthropometric deformation system presented in this thesis. Other opportunities include using machine learning and artificial neural networks to adapt the deformation system to accommodate a wide range of base facial geometries, characters of various artistic styles, and even potentially anthropomorphic characters.

REFERENCES


Bandini, S., Federici, M. L., & Vizzari, G. (2007). Situated cellular agents approach to crowd modeling and simulation. *Cybernetics and Systems: An International Journal*, *38*(7), 729-753.

Bandini, S., Manzoni, S., & Vizzari, G. (2004). Crowd modeling and simulation. In (p. 161-175). Springer.

Bandini, S., Manzoni, S., & Vizzari, G. (2006). Crowd modeling and simulation. In (p. 105-120). Springer.

Benton, A. L. (1980). The neuropsychology of facial recognition. *American Psychologist*, *35*(2), 176.

Bertillon, A., & McClaughry, R. (1896). *Signaletic instructions including the theory and practice of anthropometrical identification*. Werner Company. Retrieved from `https://books.google.com/books?id=KG1Nl2ytKcIC`

Bon, G. L. (2009). *Psychology of crowds (annotated)*. Sparkling Books.

Braun, A., Musse, S. R., de Oliveira, L. P. L., & Bodmann, B. E. (2003). Modeling individual behaviors in crowd simulation. In *Computer animation and social agents, 2003. 16th international conference on* (p. 143-148). IEEE.

Brinkley, J. F., Fisher, S., Harris, M. P., Holmes, G., Hooper, J. E., Jabs, E. W., ... Maas, R. L. (2016a). The facebase consortium: a comprehensive resource for craniofacial researchers. *Development*, *143*(14), 2677-2688. (pmid:27287806)

Brinkley, J. F., Fisher, S., Harris, M. P., Holmes, G., Hooper, J. E., Jabs, E. W., ... others (2016b). The facebase consortium: a comprehensive resource for craniofacial researchers. *Development*, *143*(14), 2677–2688.

Cassol, V. J., Musse, S. R., Jung, C. R., & Badler, N. I. (2017). *Simulating crowds in egress scenarios*. Springer.

Davis, J. P., Valentine, T., & Davis, R. E. (2010). Computer assisted photo-anthropometric analyses of full-face and profile facial images. *Forensic science international*, *200*(1-3), 165-176.

DeCarlo, D., Metaxas, D., & Stone, M. (1998). An anthropometric face model using variational techniques. In *Proceedings of the 25th annual conference on computer graphics and interactive techniques* (p. 67-74). ACM.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107.

Dirksen, N., Fischer, J., Kim, J.-H., Vassey, K., & Vogt, R. (2012). Improving crowd quality through interdepartmental collaboration on madagascar 3: Europe's most wanted. In *Proceedings of the digital production symposium* (p. 15-20).

Durupinar, F., Allbeck, J., Pelechano, N., & Badler, N. (2008). Creating crowd variation with the ocean personality model. In *Proceedings of the 7th international joint conference on autonomous agents and multiagent systems-volume 3* (p. 1217-1220). International Foundation for Autonomous Agents and Multiagent Systems.

Eisko™. (2021). *Polywink - automatic expressions blendshapes and facial rigs* (Vol. 2021) (No. March 14,). Retrieved from `https://www.polywink.com/`

El-Ali, M., Tong, L., Richards, J., Nguyen, T., Ros, A. L., & Joseph, N. M. (2016). Zootopia crowd pipeline. In *Acm siggraph 2016 talks* (p. 59). ACM.

Evison, M. P., & Bruegge, R. W. V. (2010). *Computer-aided forensic facial comparison*. CRC Press.

*Faces 4.0 - iq biometrix - faces software* (Vol. 2018) (No. October 03,). (n.d.). Retrieved from `http://www.iqbiometrix.com/products_faces_40.html`

Farkas, L. G. (1994). *Anthropometry of the head and face*. Raven Pr.

Farkas, L. G., Katic, M. J., & Forrest, C. R. (2005). International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial Surgery*, *16*(4), 615-646. (pmid:16077306)

Ferrario, V. F., Sforza, C., Schmitz, J. H., & Santoro, F. (1999). Three-dimensional facial morphometric assessment of soft tissue changes after orthognathic surgery. *Oral Surgery, Oral*

*Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *88*(5), 549-556.

Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., Mclanaghan, S., & Hancock, P. (2005). A forensically valid comparison of facial composite systems. *Psychology, Crime Law*, *11*(1), 33-52.

Ibrahim, A., Suttie, M., Bulstrode, N. W., Britto, J. A., Dunaway, D., Hammond, P., & Ferretti, P. (2016). Combined soft and skeletal tissue modelling of normal and dysmorphic midface postnatal development. *Journal of Cranio-Maxillofacial Surgery*, *44*(11), 1777-1785.

*Introduction to combinatorics* (Vol. video). (n.d.). Retrieved from `https://www.khanacademy.org/partner-content/pixar/crowds/crowds-1/v/intro-crowds`

Joshi, P., Meyer, M., DeRose, T., Green, B., & Sanocki, T. (2007). Harmonic coordinates for character articulation. *ACM Transactions on Graphics (TOG)*, *26*(3), 71-es.

Ju, E., Choi, M. G., Park, M., Lee, J., Lee, K. H., & Takahashi, S. (2010). Morphable crowds. In *Acm transactions on graphics (tog)* (Vol. 29, p. 140). ACM.

Kesner, L., Grygarová, D., Fajnerová, I., Lukavský, J., Nekovářová, T., Tintěra, J., … Horáček, J. (2018). Perception of direct vs. averted gaze in portrait paintings: An fmri and eye-tracking study. *Brain and cognition*, *125*, 88-99.

Kleinberg, K. F. (2008). No title. *Facial anthropometry as an evidential tool in forensic image comparison*.

Kleinberg, K. F., Vanezis, P., & Burton, A. M. (2007). Failure of anthropometry as a facial identification technique using high-quality photographs. *Journal of forensic sciences*, *52*(4), 779-783.

Knußmann, R. (1988). Anthropologie. *Handbuch der vergleichenden Biologei des Menschen*, *1*, 251-255.

Landau, T. (1989). *About faces*. Doubleday.

Mancusi, S. (2010). *The police composite sketch*. Springer Science Business Media.

Marazita, M. L., Weinberg, S., & Raffensperger, Z. (2017). 3d facial norms database. *FaceBase*

*Consortium*. Retrieved from `https://doi.org/10.25550/1WWP`

Massaro, D., Savazzi, F., Dio, C. D., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When art moves the eyes: a behavioral and eye-tracking study. *PloS one*, *7*(5), e37285.

Mourino, G., Evans, M., Edzenga, K., Cavaleri, S., Adams, M., & Bisceglio, J. (2017). Populating the crowds in ferdinand. In (p. 1-2).

Parke, F. I. (1972). Computer generated animation of faces. In *Proceedings of the acm annual conference-volume 1* (pp. 451–457).

Parke, F. I. (1974). *A parametric model for human faces.* (Tech. Rep.). UTAH UNIV SALT LAKE CITY DEPT OF COMPUTER SCIENCE.

Parke, F. I. (1982). Parameterized models for facial animation. *IEEE Computer Graphics and Applications*, *2*(9), 61-68. doi: 10.1109/MCG.1982.1674492

Parke, F. I., & Waters, K. (2008). *Computer facial animation*. CRC press.

Porter, G., & Doran, G. (2000). An anatomical and photographic technique for forensic facial identification. *Forensic science international*, *114*(2), 97-105.

Porter, J. P. (2004). The average african american male face: an anthropometric analysis. *Archives of facial plastic surgery*, *6*(2), 78-81.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on computer graphics and interactive techniques* (p. 25-34).

Savazzi, F., Massaro, D., Dio, C. D., Gallese, V., Gilli, G., & Marchetti, A. (2014). Exploring responses to art in adolescence: a behavioral and eye-tracking study. *PloS one*, *9*(7), e102888.

Sticco, I. M., Frank, G. A., & Dorso, C. O. (2020). Social force model parameter testing and optimization using a high stress real-life situation. *Physica A: Statistical Mechanics and its Applications*, 125299.

Taylor, K. T. (2000). *Forensic art and illustration*. CRC Press.

Thalmann, D. (2007). *Crowd simulation*. Wiley Online Library.

Thomas, F., Johnston, O., & Thomas, F. (1995). *The illusion of life: Disney animation*. Hyperion New York.

Tome, P., Vera-Rodriguez, R., Fierrez, J., & Ortega-Garcia, J. (2015). Facial soft biometric features for forensic face recognition. *Forensic science international*, *257*, 271-284.

Ulicny, B., de Heras Ciechomski, P., & Thalmann, D. (2004). Crowdbrush: interactive authoring of real-time crowd scenes. In *Proceedings of the 2004 acm siggraph/eurographics symposium on computer animation* (p. 243-252).

Vegter, F., & Hage, J. J. (2000). Clinical anthropometry and canons of the face in historical perspective. *Plastic and Reconstructive Surgery*, *106*(5), 1090-1096.

Vegter, F., & Hage, J. J. (2001). Facial anthropometry in cleft patients: a historical appraisal. *The Cleft palate-craniofacial journal*, *38*(6), 577-581.

Venkatadri, G., Farkas, L. G., & Kooiman, J. (1992). Multipurpose anthropometric facial anglemeter. *Plastic and Reconstructive Surgery*, *90*(3), 507-510. (pmid:1513900)

Villani, D., Morganti, F., Cipresso, P., Ruggi, S., Riva, G., & Gilli, G. (2015). Visual exploration patterns of human figures in action: an eye tracker study with art paintings. *Frontiers in psychology*, *6*, 1636.

Wong, J. Y., Oh, A. K., Ohta, E., Hunt, A. T., Rogers, G. F., Mulliken, J. B., & Deutsch, C. K. (2008). Validity and reliability of craniofacial anthropometric measurement of 3d digital photogrammetric images. *The Cleft Palate-Craniofacial Journal*, *45*(3), 232-239.

Woods, S. (2002). The animator's survival kit. *Film Ireland*(85), 28.

Yen, J., Gustafson, S., Lo, A., Northrup, J. D., & Sun, L. (2018). Other worldly crowds in coco. In (p. 1-2).

Zhou, S., Chen, D., Cai, W., Luo, L., Low, M. Y. H., Tian, F., ... Hamilton, B. D. (2010). Crowd modeling and simulation technologies. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *20*(4), 20.

# APPENDIX A

## CRANIOFACIAL ANTHROPOMETRIC MEASURE PARAMETER SET

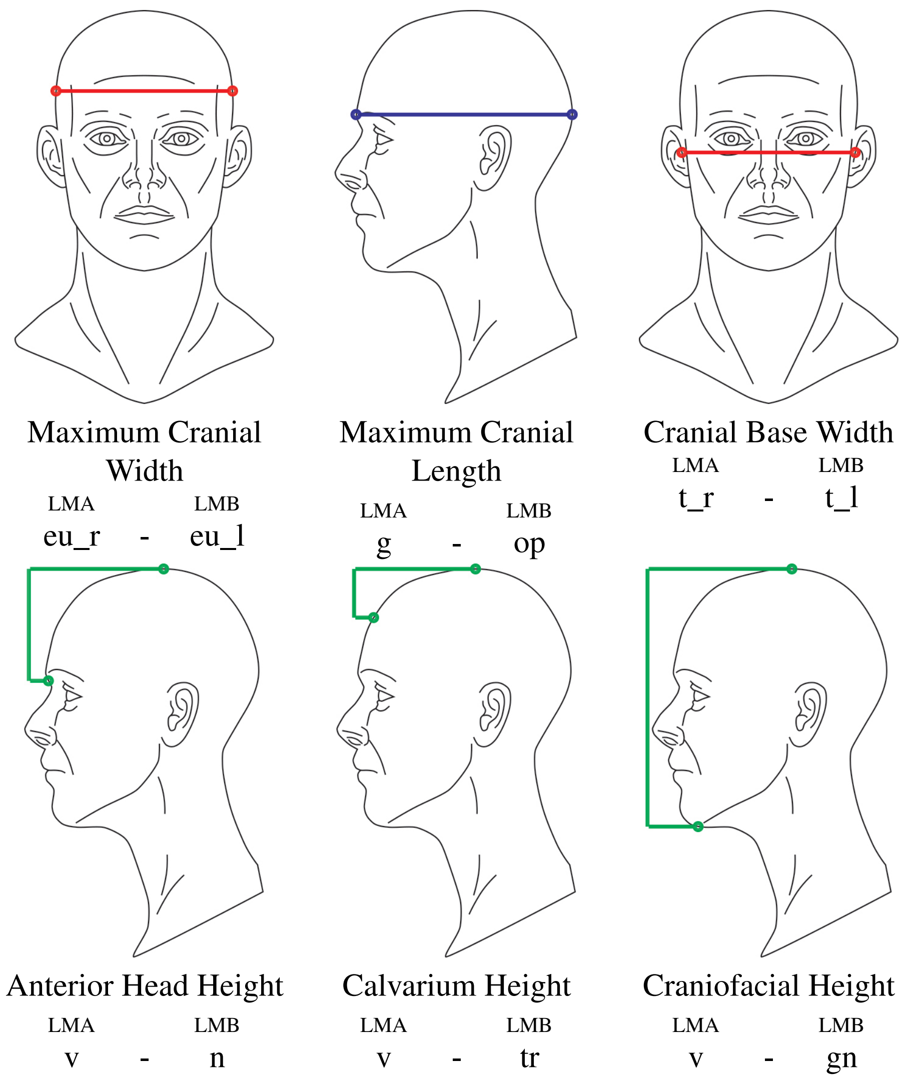



Figure A.1: The craniofacial anthropometric parameters pertaining to the head region and the landmarks associated with them.

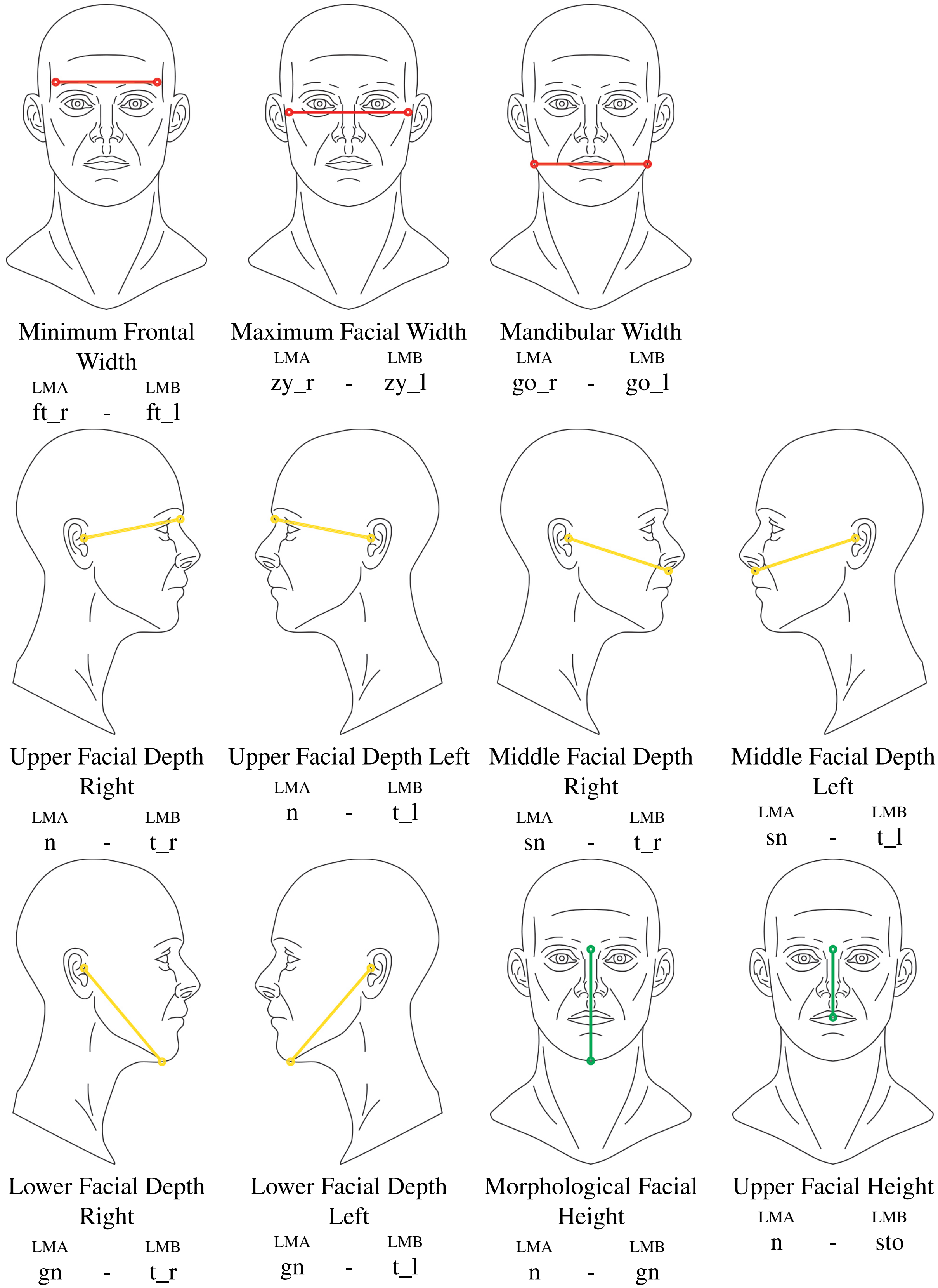



Figure A.2: The craniofacial anthropometric parameters pertaining to the facial region and the landmarks associated with them.

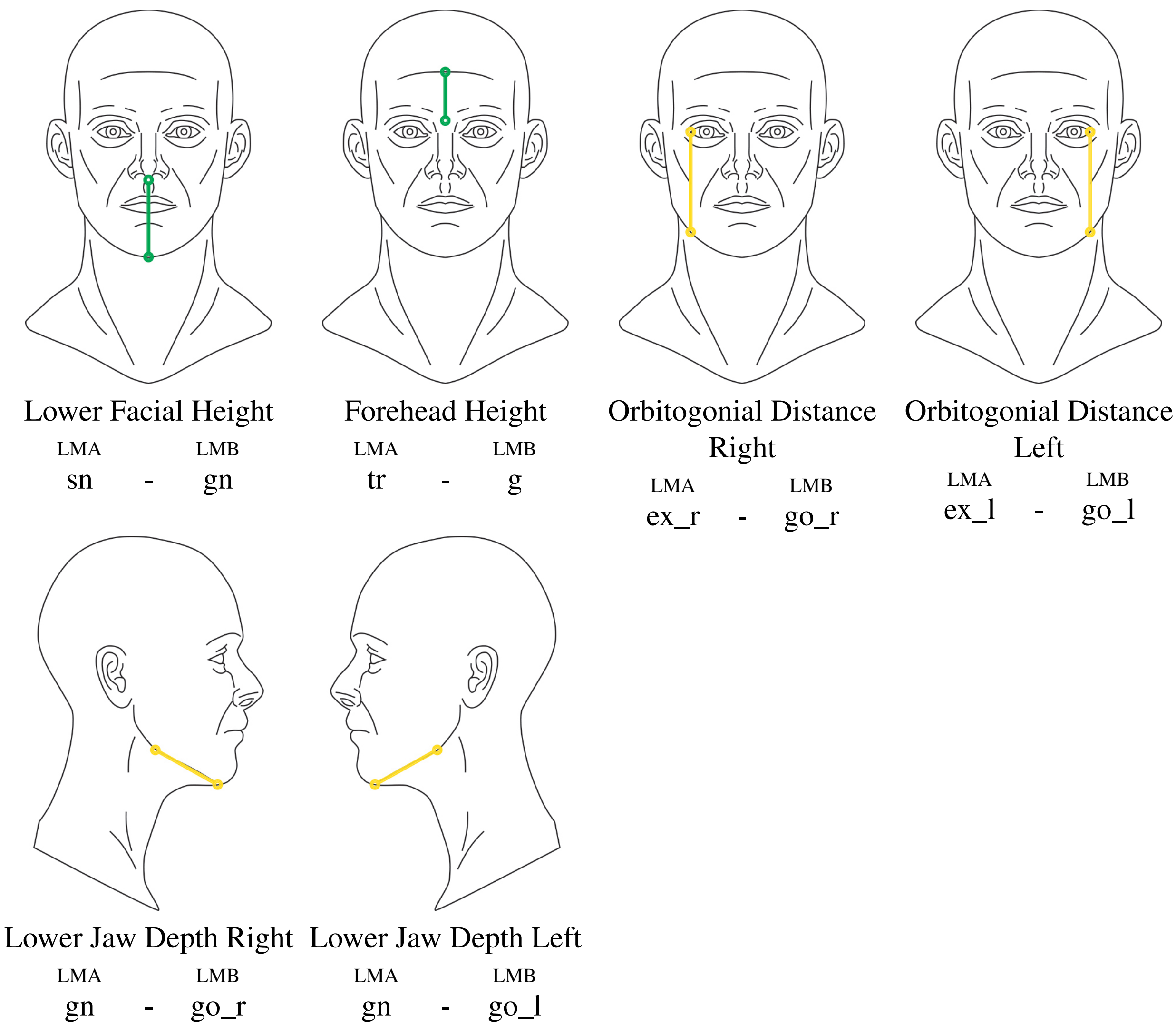



Figure A.2 (cont.): The craniofacial anthropometric parameters pertaining to the facial region and the landmarks associated with them.

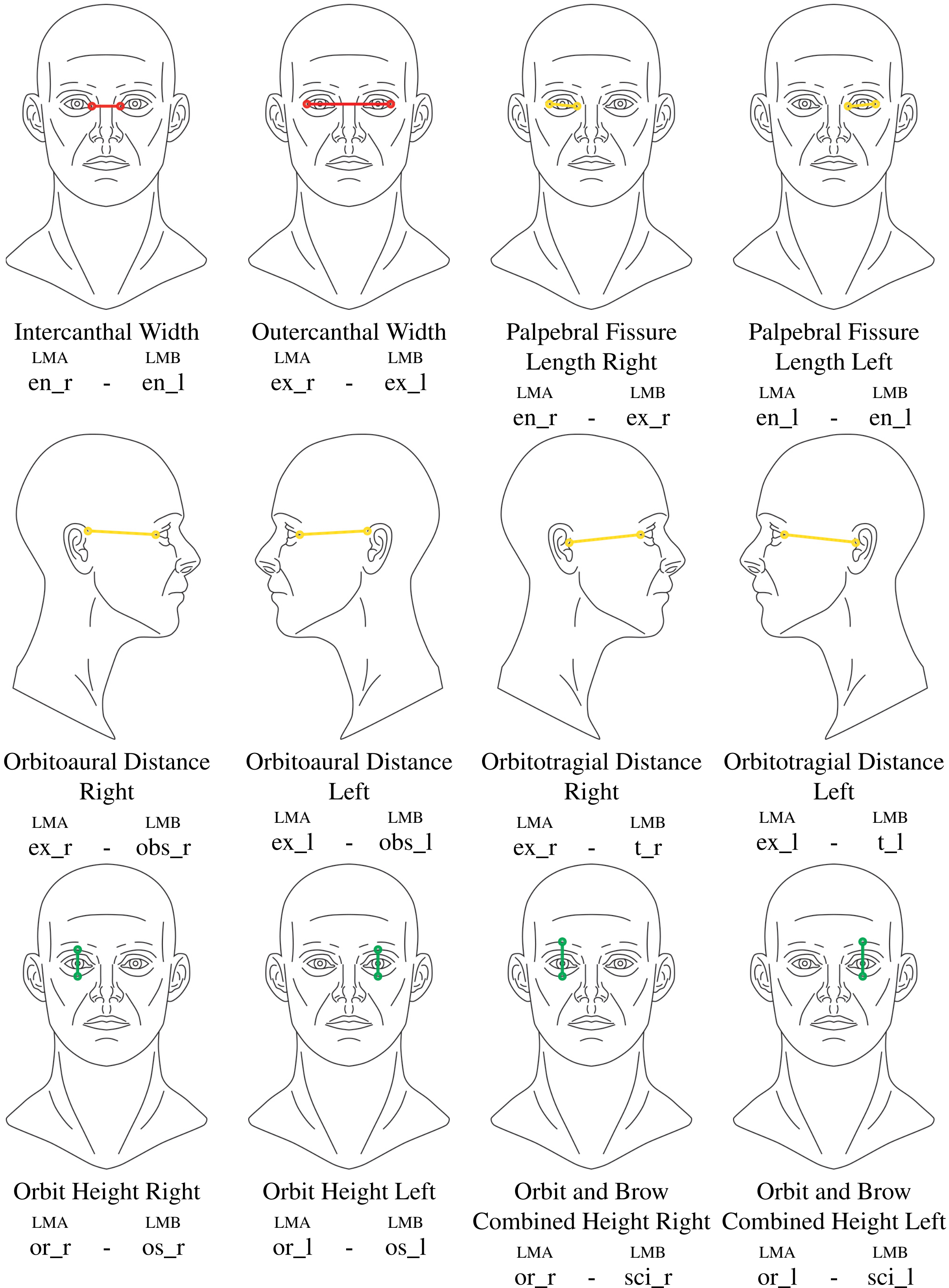
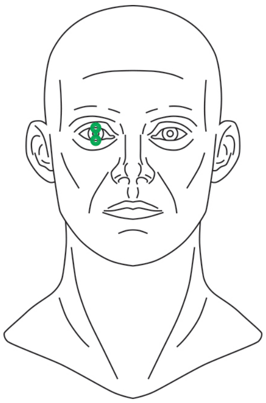
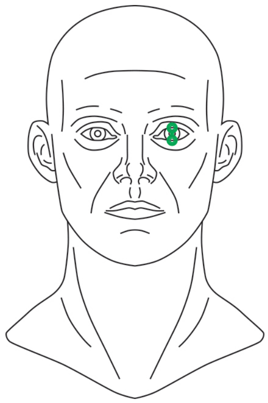


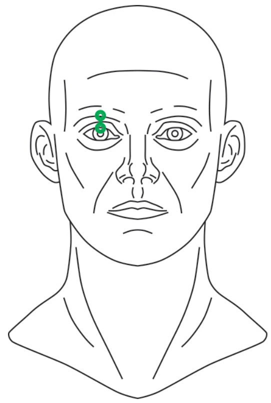

Figure A.3: The craniofacial anthropometric parameters pertaining to the orbital regions and the landmarks associated with them.



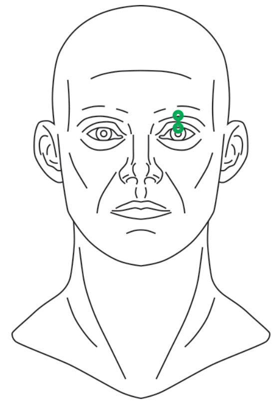
Eye Fissure Height Right
LMA ps_r - LMB pi_r



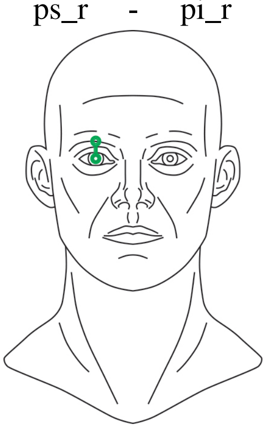
Eye Fissure Height Left
LMA ps_l - LMB pi_l



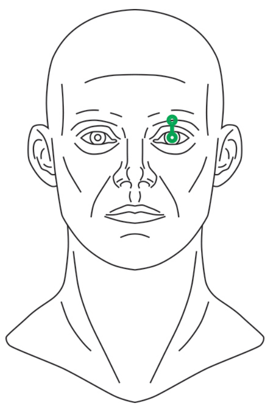
Upper Lid Height Right
LMA ps_r - LMB os_r



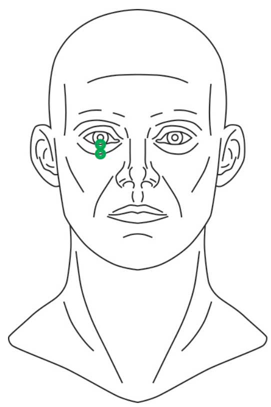
Upper Lid Height Left
LMA ps_l - LMB os_l



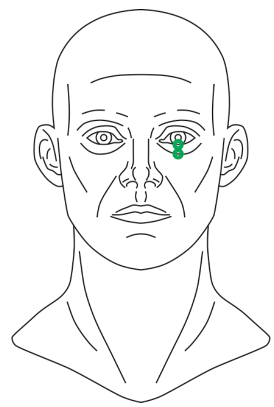
Pupil to Upper Lid Height Right
LMA pupil_r - LMB os_r



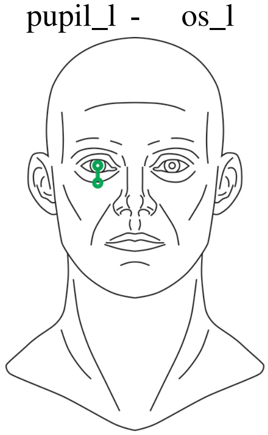
Pupil to Upper Lid Height Left
LMA pupil_l - LMB os_l



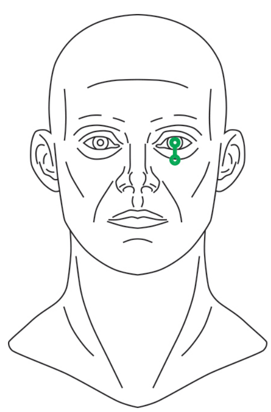
Lower Lid Height Right
LMA pi_r - LMB or_r



Lower Lid Height Left
LMA pi_l - LMB or_l



Pupil to Lower Lid Height Right
LMA pupil_r - LMB or_r



Pupil to Lower Lid Height Left
LMA pupil_l - LMB or_l

Figure A.3 (cont.): The craniofacial anthropometric parameters pertaining to the orbital regions and the landmarks associated with them.

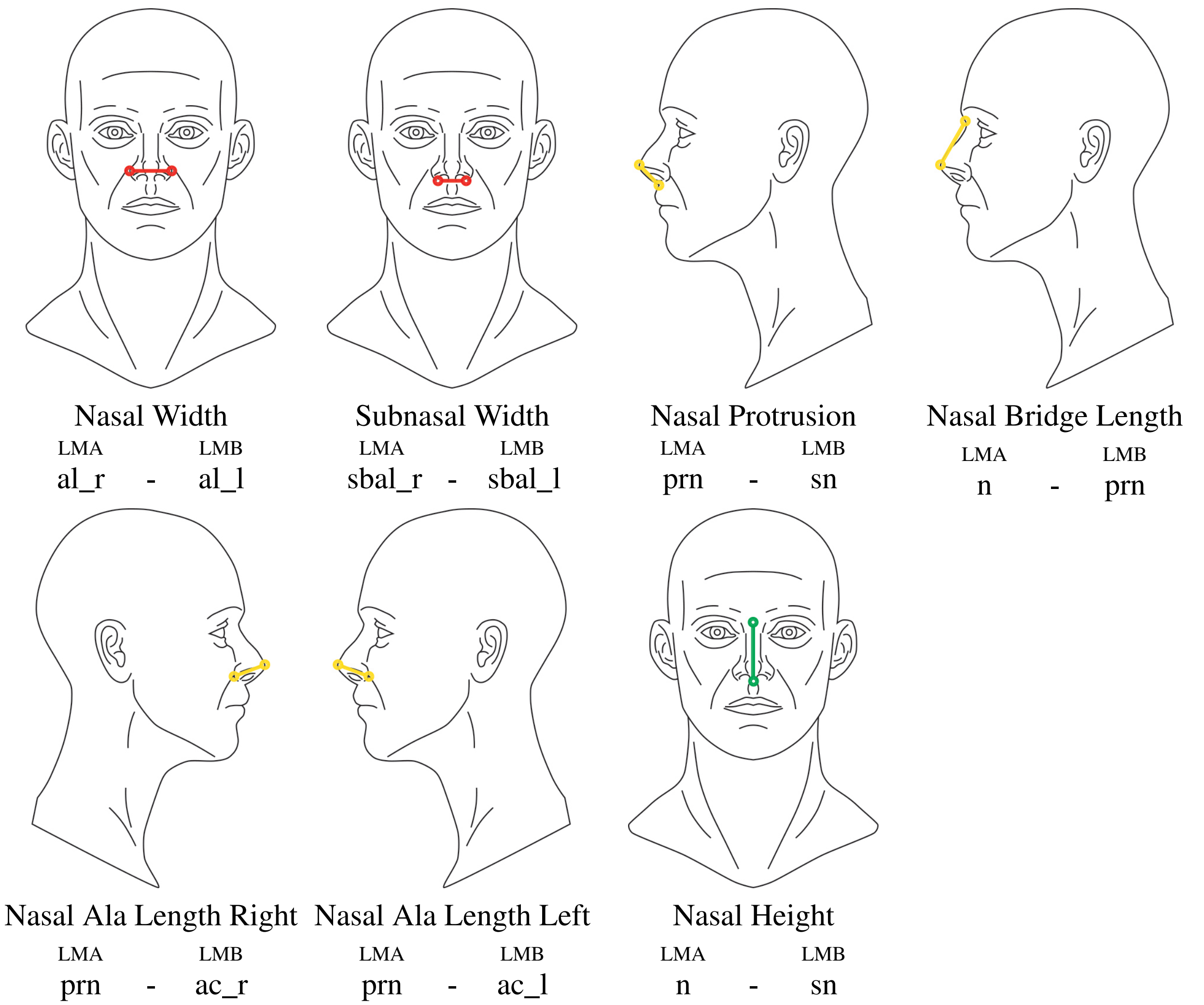



Figure A.4: The craniofacial anthropometric parameters pertaining to the nasal region and the landmarks associated with them.

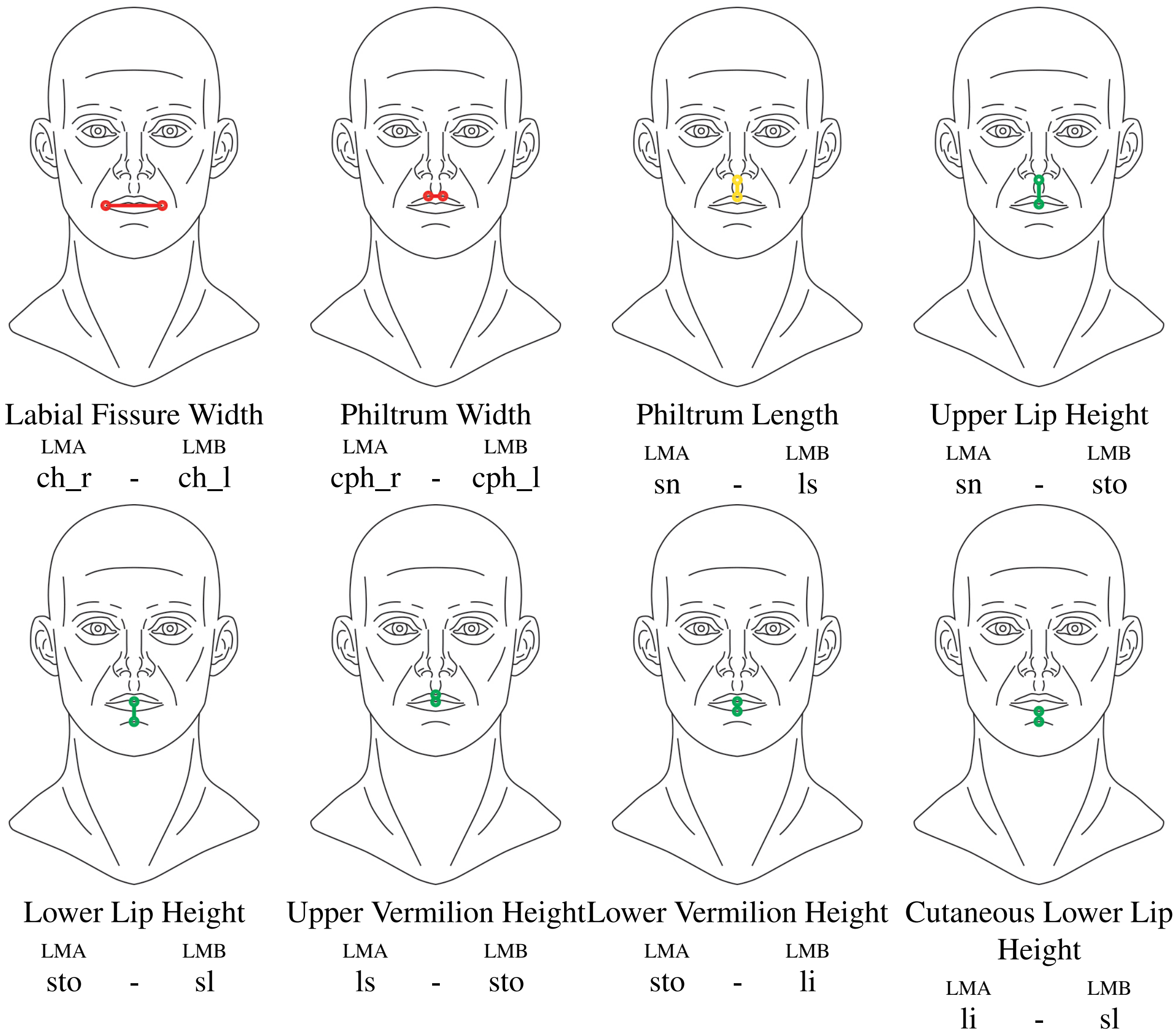
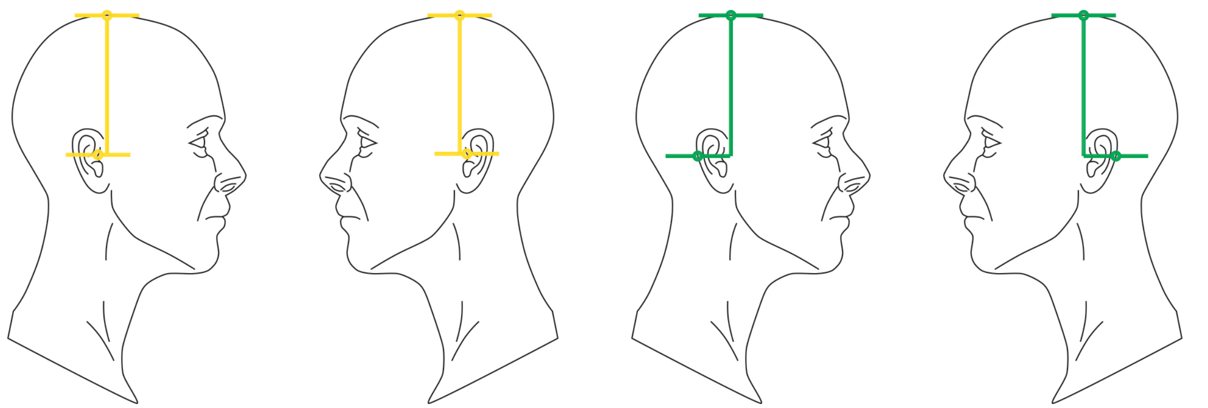



Figure A.5: The craniofacial anthropometric parameters pertaining to the labial region and the landmarks associated with them.



| Vertex Tragion Distance Right | Vertex Tragion Distance Left | Auricular Height Right | Auricular Height Left |
|---|---|---|---|
| LMA: v - LMB: t_r | LMA: v - LMB: t_l | LMA: v - LMB: po_r | LMA: v - LMB: po_l |

Figure A.6: The craniofacial anthropometric parameters pertaining to the aural regions and the landmarks associated with them.

# APPENDIX B

## PROCEDURAL CROWD MEMBER GENERATION PROCESS SUMMARY

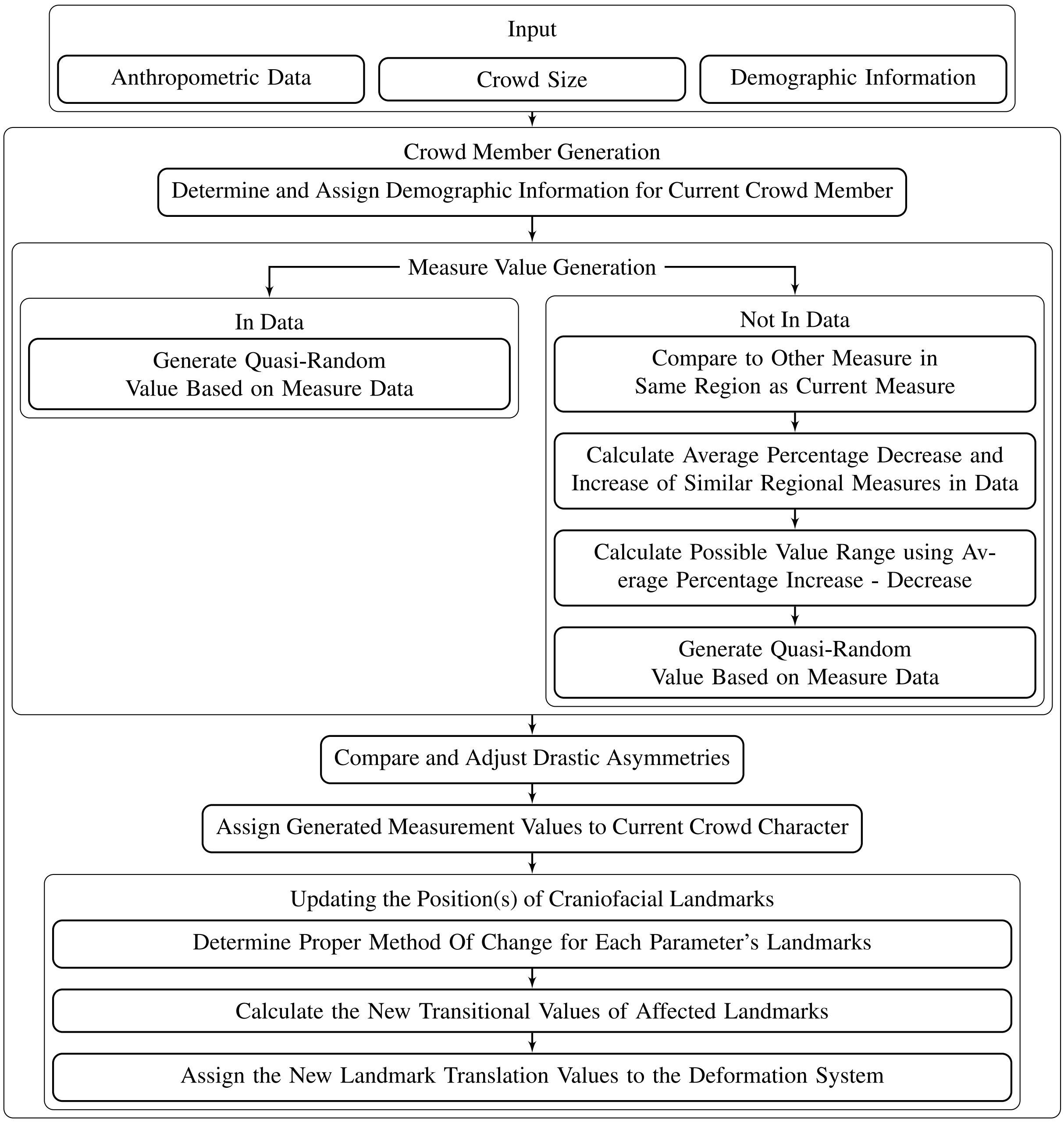



Figure B.1: Flowchart of the detailed functions the script performs during its execution.

APPENDIX C

DIAGRAM OF LANDMARKS AND THEIR PLACEMENT

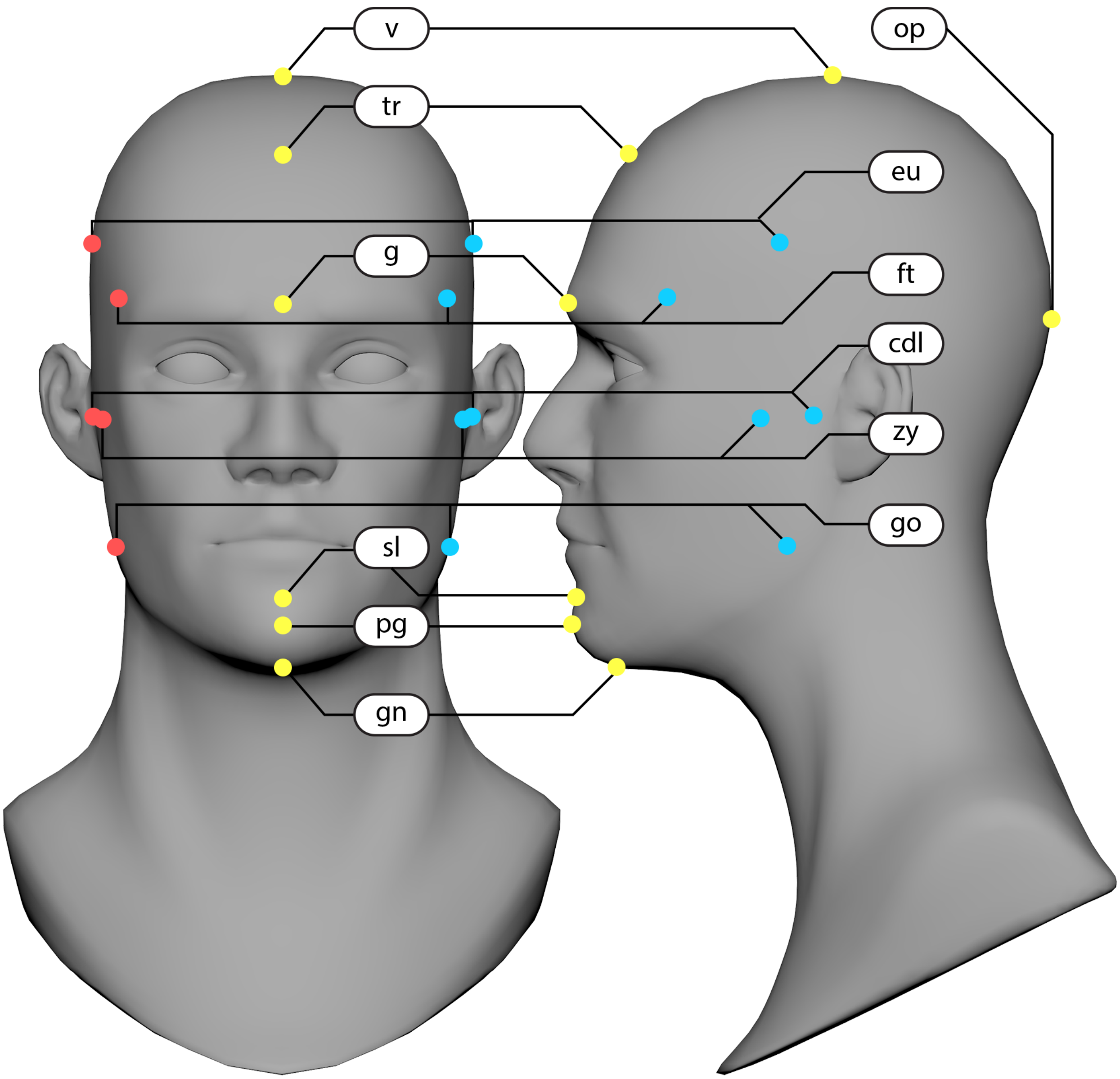



Figure C.1: The landmarks of the head and face regions.

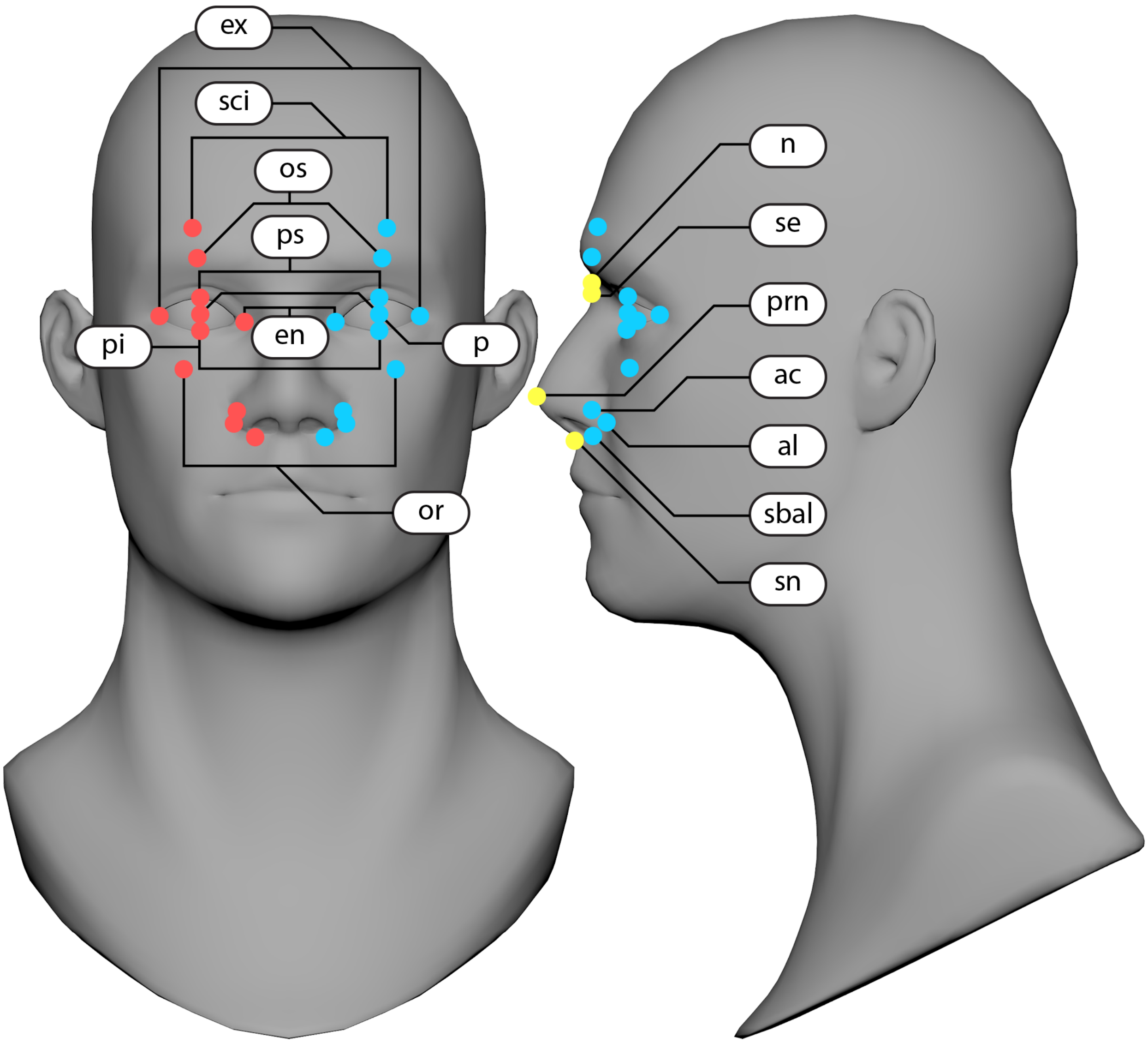



Figure C.2: The landmarks of the orbit and nasal regions.

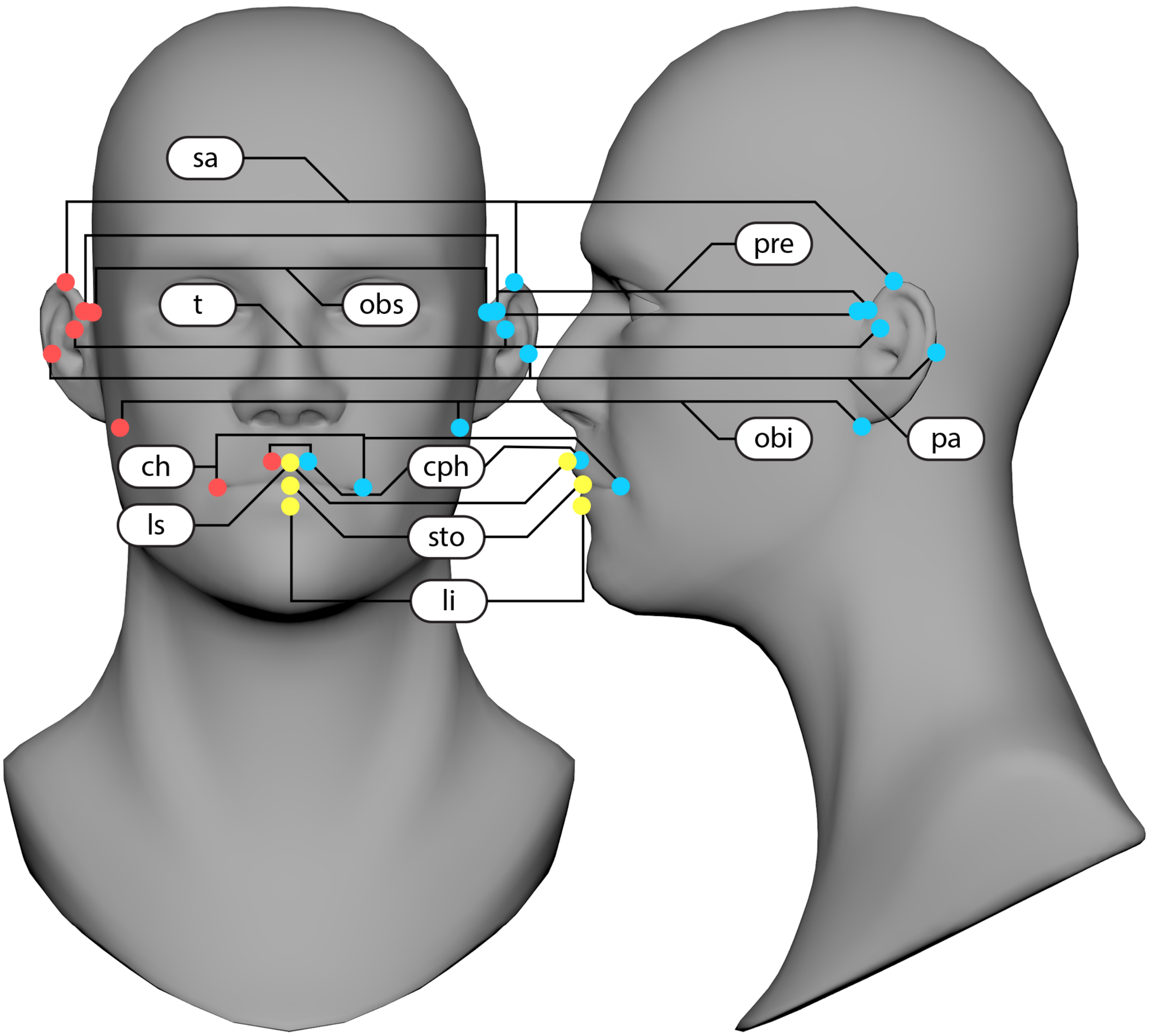
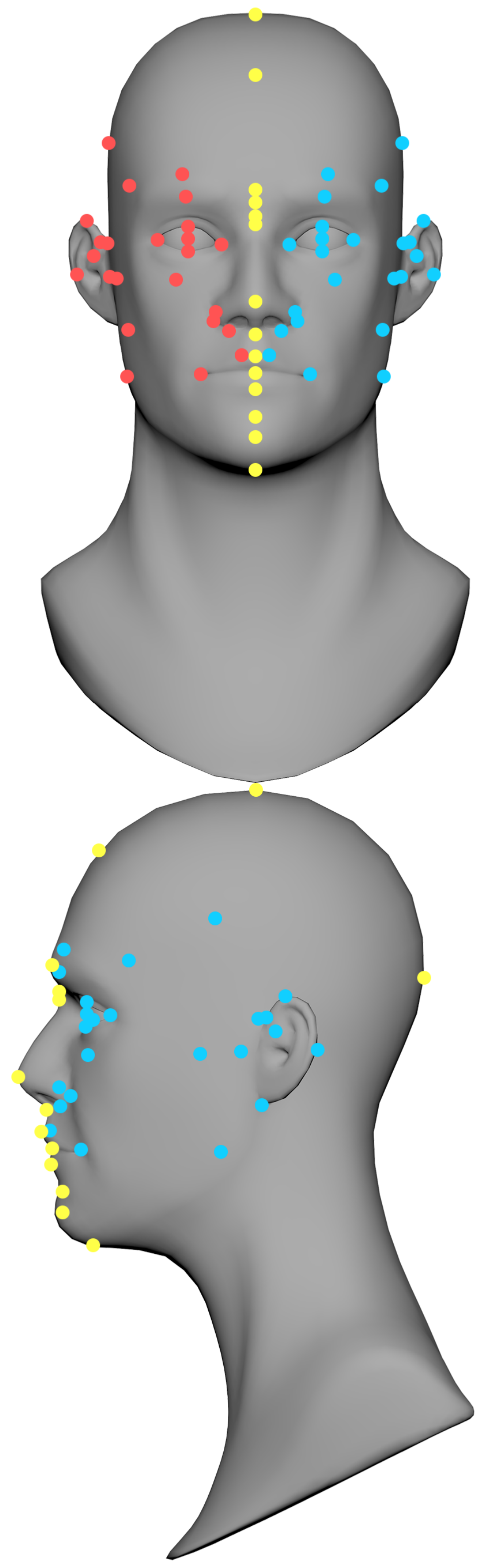



Figure C.3: The landmarks of the labiale and aurale regions.

| Landmark | | Region |
|---|---|---|
| vertex | v | head |
| trichion | tr | head |
| eurion | eu | head |
| glabella | g | head |
| opisthocranion | op | head |
| frontotemporale | ft | head |
| zygion | zy | head |
| condylion laterale | cdl | face |
| gonion | go | face |
| sublabiale | sl | face |
| pogonion | pg | face |
| gnathion | gn | face |
| superciliare | sci | orbits |
| orbitale superius | os | orbits |
| palpebrale superius | ps | orbits |
| endocanthion | en | orbits |
| pupil center | p | orbits |
| exocanthion | ex | orbits |
| palpebrale inferius | pi | orbits |
| orbitale | or | orbits |
| nasion | n | nose |
| sellion | se | nose |
| pronasale | prn | nose |
| alar curviture | ac | nose |
| alare | al | nose |
| subalare | sbal | nose |
| subnasale | sn | nose |
| crista philtri | cph | mouth |
| labiale superius | ls | mouth |
| stomion | sto | mouth |
| cheilion | ch | mouth |
| labiale inferius | li | mouth |
| superaurale | sa | ears |
| preaurale | pra | ears |
| otobasion superius | obs | ears |
| tragion | t | ears |
| postaurale | pa | ears |
| subaurale | sba | ears |
| otobasion inferius | obi | ears |

(a) Landmark Names, Shorthand, and Regions



(b) Landmark Positions

Figure C.4: Diagram of landmark information such as names, regions, and positions.

APPENDIX D

SPATIAL INFORMATION AVAILABLE FOR EACH LANDMARK

| Landmark | | X | Y | Z | Vector |
|---|---|---|---|---|---|
| vertex | v | | Anterior Head Height, Calvarium Height, Craniofacial Height, Auricular Height | | Vertex Tragion Distance |
| trichion | tr | Calvarium Height, Forehead Height | | | |
| eurion | eu | Maximum Cranial Width | | | |
| glabella | g | | Forehead Height | Maximum Cranial Length | |
| opisthocranion | op | | | Maximum Cranial Length | |
| frontotemporale | ft | Minimum Frontal Width | | | |
| zygion | zy | Maximum Facial Width | | | |
| condylion laterale | cdl | | | | |
| gonion | go | Mandibular Width | | | Orbitogonial Distance, Lower Jaw Depth |
| sublabiale | sl | | Lower Lip Height, Cutaneous Lower Lip Height | | |
| pogonion | pg | | | | |
| gnathion | gn | | Craniofacial Height, Morphological Facial Height, Lower Facial Height | | Lower Facial Depth, Lower Jaw Depth |

| Landmark cont... | | X cont... | Y cont... | Z cont... | Vector cont... |
|---|---|---|---|---|---|
| superciliare | sci | | Orbit Brow Combined Height | | |
| orbitale superius | os | | Pupil Upper Lid Height, Upper Lid Height, Orbit Height | | |
| palpebrale superius | ps | | Eye Fissure Height, Upper Lid Height | | |
| endocanthion | en | Intercanthal Width | | | Palpebral Fissure Length |
| pupil center | p | | Pupil Upper Lid Height, Pupil Lower Lid Height | | |
| exocanthion | ex | Outercanthal Width | | | Palpebral Fissure Length, Orbitogonial Distance, Orbitotragial Distance, Orbitoaural Distance |
| palpebrale inferius | pi | | Eye Fissure Height, Lower Lid Height | | |
| orbitale | or | | Pupil Lower Lid Height, Lower Lid Height, Orbit Height, Orbit Brow Combined Height | | |

| Landmark cont... | | X cont... | Y cont... | Z cont... | Vector cont... |
|---|---|---|---|---|---|
| nasion | n | | Anterior Head Height, Morphological Facial Height, Nasal Height, Upper Facial Height | | Upper Facial Depth, Nasal Bridge Length |
| sellion | se | | | | |
| pronasale | prn | | | | Nasal Bridge Length, Nasal Protrusion, Nasal Ala Length |
| alar curviture | ac | | | | Nasal Ala Length |
| alare | al | Nasal Width | | | |
| subalare | sbal | Subnasal Width | | | |
| subnasale | sn | | Nasal Height, Lower Facial Height, Upper Lip Height | | Nasal Protrusion, Middle Facial Depth, Philtrum Length |
| crista philtri | cph | Philtrum Width | | | |
| labiale superius | ls | | Upper Vermilion Height | | Philtrum Length |
| stomion | sto | | Upper Facial Height, Upper Lip Height, Lower Lip Height, Upper Vermilion Height, Lower Vermilion Height | | |
| cheilion | ch | Labial Fissure Width | | | |

| Landmark cont... | | X cont... | Y cont... | Z cont... | Vector cont... |
|---|---|---|---|---|---|
| labiale inferius | li | | Lower Vermilion Height, Cutaneous Lower Lip Height | | |
| superaurale | sa | | | | |
| preaurale | pra | | | | |
| otobasion superius | obs | | | | Orbitoaural Distance, Morphological Ear Width |
| tragion | t | | | | Upper Facial Depth, Middle Facial Depth, Lower Facial Depth, Orbitotragial Distance, Vertex Tragion Distance |
| postaurale | pa | | | | |
| subaurale | sba | | | | |
| otobasion inferius | obi | | | | Morphological Ear Width |

Table D.1: Craniofacial anthropometric landmarks and the three dimensional information effecting each as defined by the associated parameters.

# APPENDIX E

## USER INTERFACES (UI) DISPLAYED DURING SCRIPT EXECUTION

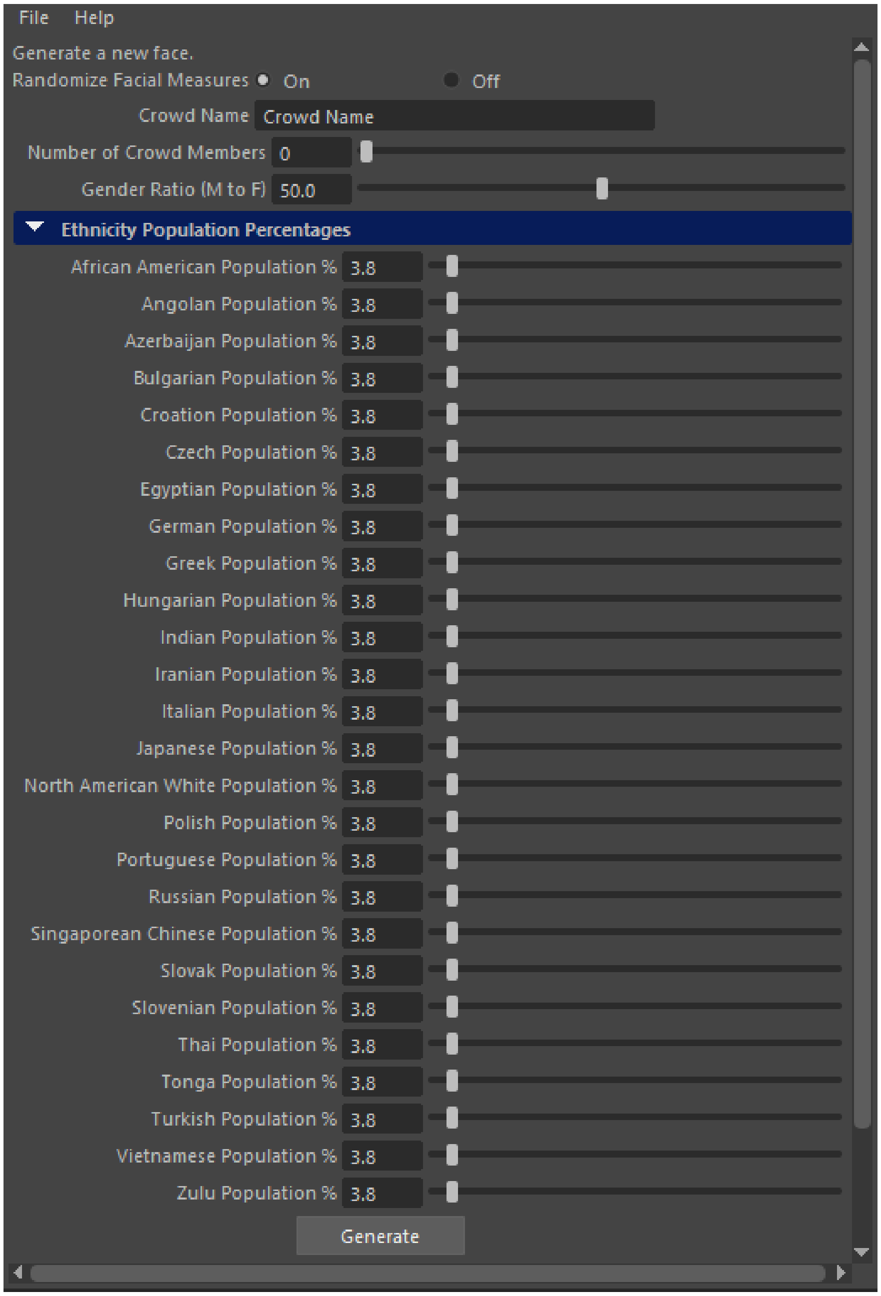



Figure E.1: The UI displayed by the python script during the "Input" phase of the script's execution.

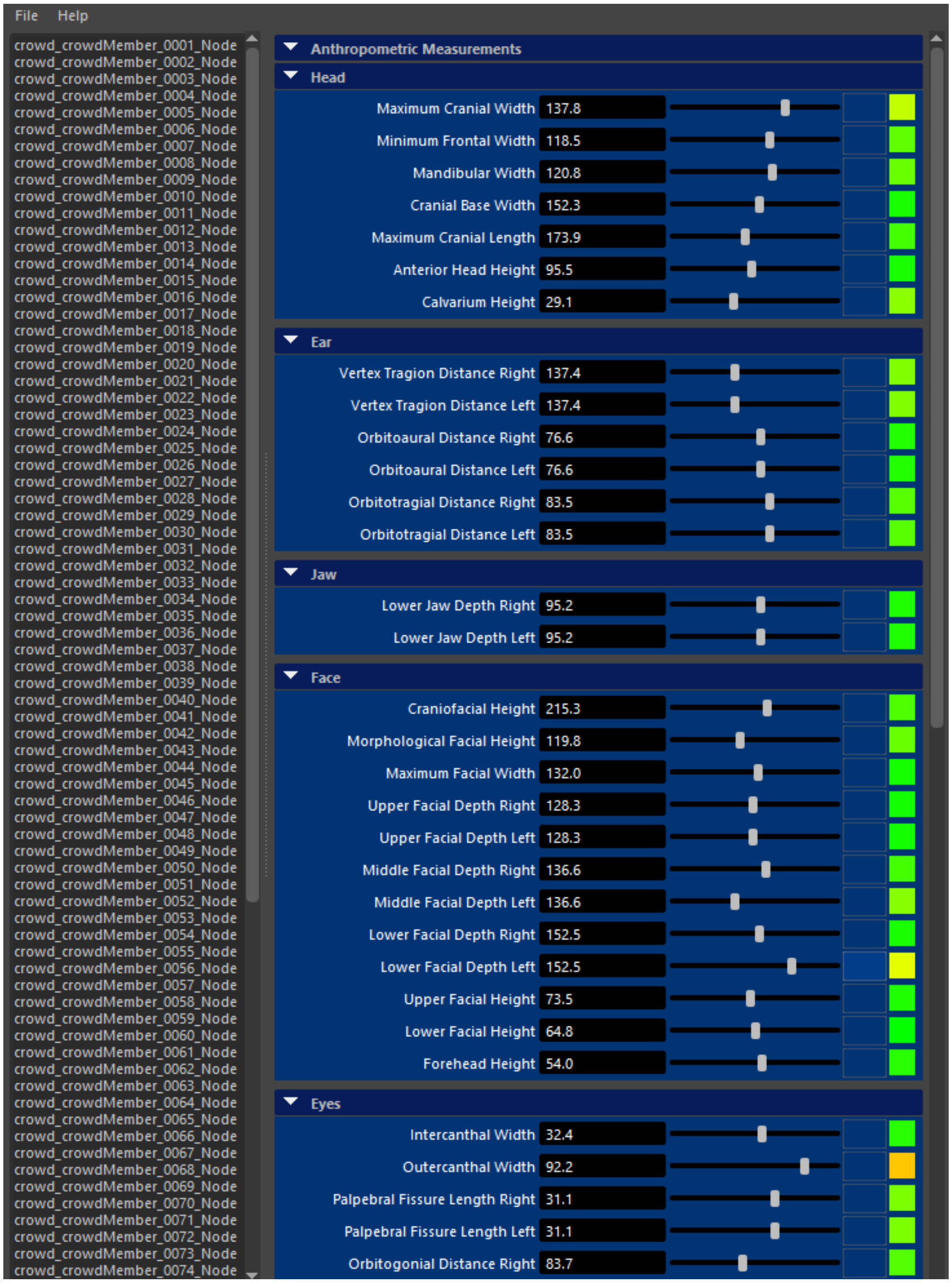



Figure E.2: The UI displayed by the python script upon successful completion of the "Crowd Member Generation" phase. The left pane of this window allows the user to select a specific crowd member. The right pane displays the current parameter values of the selected crowd member. The right pane also contains sliders that can be used to adjust the values of each parameter as desired.