Fraos
A tool to predict and communicate facial recognition after orthognathic surgeries

Master thesis

Tamara Ribbers

TU Delft

VUmc

Master thesis
Delft, August 2020

Education
MSc. Integrated Product Design
Medisign specialisation
Delft University of Technology
Faculty of Industrial Design Engineering

Supervisory team

Project chair
Prof. dr. ir. Richard Goossens

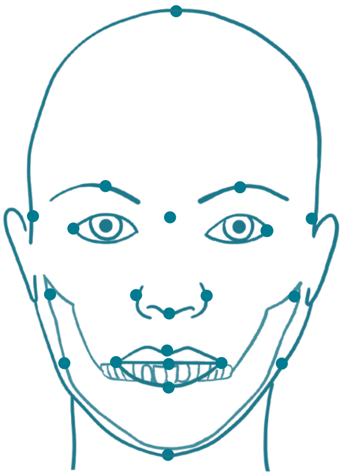
Project mentor
Dr. Yu (Wolf) Song

Company mentor

VUmc hospital
Prof. dr. Tim Forouzanfar

Author

Tamara Ribbers



Fraos
A tool to predict and communicate
facial recognition after orthognathic surgeries

Master thesis

Tamara Ribbers
September 2020

Preface

Before you lies the thesis, hopefully, provoking a conversation about facial recognition after orthognathic surgeries. The main goal of this thesis was to complete the master of Integrated Product Design and to call me an engineer (ir.) or Master Of Science (MSc.) after it. However, this project gave me much more than only a 'cool' abbreviation before my name. I never expected it to be so interesting and it brought me joy and a lot of learning experiences. During this weird, different, isolated Corona time, this project actually kept me going because I had something valuable to do. I was busy working on the project and completely forgot what was happening outside my room in the real world. When looking back it is nice to realize what I have done during my project, all behind my desk. Off course, the isolation did limit my plans sometimes, but still, it was a special experience not many graduates experienced. And a new thing to put on my CV: "Being able to finish a project behind my desk during isolated times."

I would like to thank my TU Delft supervisors, Richard and Wolf for their support. Especially at the beginning of the project, when the crisis popped up, it was great that you made extra time to make sure I was doing fine. It is a pity that we did not see each other in person, except for the kickoff meeting. But thanks to modern technologies, we still managed to lead the project in the right direction. I think we were a good team!

I also would like to thank my supervisor of the VUmc, Tim, who was always very enthusiastic about the project. His enthusiasm gave me confirmation that I was doing something valuable which motivated me.

One of the faces you can see throughout the thesis is the one of Edward, he is my roommate and also boyfriend. Luckily, we were locked in the house together during lock-down and he was willing to be my model, test participant and conversation partner and was, next to all the emotional support, of great help in being able to successfully succeed this project.

The human recognition test could not have been performed without the big amount of participants, which were the members of SRC Thor. It was great that they provided me their pictures and that they responded to my questionnaire. My experience is that online tests usually never have enough participants, but with 63 participants, I was very lucky!

Last but not least, the user test could not have been done without my family who participated in this test. During the time of the project, it was hard to meet with people in real-life. I was lucky that they were willing to help me with this test, with which I was able to validate the concept.

I hope you enjoy your reading,
Tamara

Abstract

A problem at the VUmc hospital in Amsterdam is that some patients become unrecognizable after the surgery. It is hard for surgeons to explain during the consultation how and how much the patients will change due to the surgery. Partly, because the surgeons cannot predict if the patient will become unrecognizable. Therefore, they tell them that there is a possibility they change drastically, and with this information, the patients do not know what they can expect. Even if the outcome of the surgery is optimal according to aesthetic and functional guidelines, the mismatching expectations with the outcomes can dissatisfy the patient. Some patients cannot get used to their new changes and do not recover emotionally, resulting in mental problems. Therefore, there is a need for a tool in the consultation room to predict changes in facial recognition per patient and which helps the surgeon to communicate this to the patient.

The changes in facial recognition were noticed at the VUmc, but there was no proof that the orthognathic surgeries actually influenced facial recognition, because no research in the topic was done before this project. This project proves that orthognathic surgeries can influence facial recognition which creates a new field of research. The research in this project shows that not all patients become unrecognizable, it depends on the specific face of the patient and the type of surgery the patient undergoes.

Facial recognition is complex since it depends on who is trying to recognize the patient, how familiar the patient is to this person, and a persons ability to recognize someone can differ per day and is therefore variable. Therefore, a test has been done to analyse the influence of orthognathic surgeries on facial recognition within a familiar group.

Testing a larger dataset would be too time-consuming to test among humans and, therefore, a landmark-based computer recognition method has been used to analyse a dataset of 75 pre- and post-surgical patient pictures. Both tests show that some regions of the face influence facial recognition more than others, as well as specific landmark movements which simulate orthognathic surgeries. Also, facial recognition depends on the face of the patient. Therefore, a concept of a tailored tool was created, called Fraos, showing the predictable changes in facial recognition per patient and helping the surgeon to communicate this to the patient. It also supports the patient in emotionally recovering from the facial changes.

Reading guide

Fast reading
If you want to read only the essentials it is recommended to read chapter 1, the summary. This describes the whole project in 500 words. Furthermore is recomended to read the last page of every chapter. This page contains the conclusion of the chapter and the key-points (figure 1).

Explorative reading
The research sections only show a short description with the results. More research has been done and can be found in the appendix belonging to this report. If you are interested in all research of the research section it is recommended to replace some chapters in the report with some in the appendix. Some chapters in the appendix are written in a way it consists all the information from the report, and more, written chronologically. The chapters which can be replaced by appendices are:

Chapter 4, Appendix 4
Chapter 5, Appendix 5
Chapter 6, Appendix 6

Figure 1: Example of key-points

Key-points from this chapter

- Key-point 1
- Key-point 2
- Key-point 3
- Key-point 4
- Key-point 5

Report structure
This report exists out of 14 chapters. All chapters start with an introduction. The tests all have the same structure and the method, results, discussion and conclusion are described in these chapters. The content of the chapters is described next.

Chapter 1: Summary
It summarizes the whole report in 500 words complemented with visuals.

Chapter 2: Project introduction
It describes how the project subject has been chosen, what the problem is and the approach of the project.

Research section

Chapter 3: Facial recognition
It shows the research which has been found about facial recognition after orthognathic surgery where after research has been done in human- and computer recognition to create a link between orthognathic surgeries and facial recognition.

Chapter 4: Facial landmarks
It shows different sets of landmarks and the landmark set which has been chosen to use in the landmark based approach of this project.

Chapter 5: Classification system
It shows how patients are classified currently by analysing different classification systems. It shows a new classification system which has been created to describe landmark movements happening in the patients face due to orthognathic surgery.

Chapter 6: Human recognition test
This test shows that facial recognition can be influenced when the orthognathic region is manipulated on a familiar face.

Chapter 7: Computer recognition test
This test shows an analysis of 75 real orthognathic patients and the influence on facial recognition when single and multiple landmarks move.

Chapter 8: Comparing the two tests
It shows that the results of the tests match and comparisons have been used to calculate the human recognition ratio and threshold.

Design section

Chapter 9: Problem definition
The outcomes are summarized where after the problem definition is optimized according to the research done.

Chapter 10: The concept: Fraos
This chapter presents the concept which exists out of two parts: the communication tool and the facial recognition system.

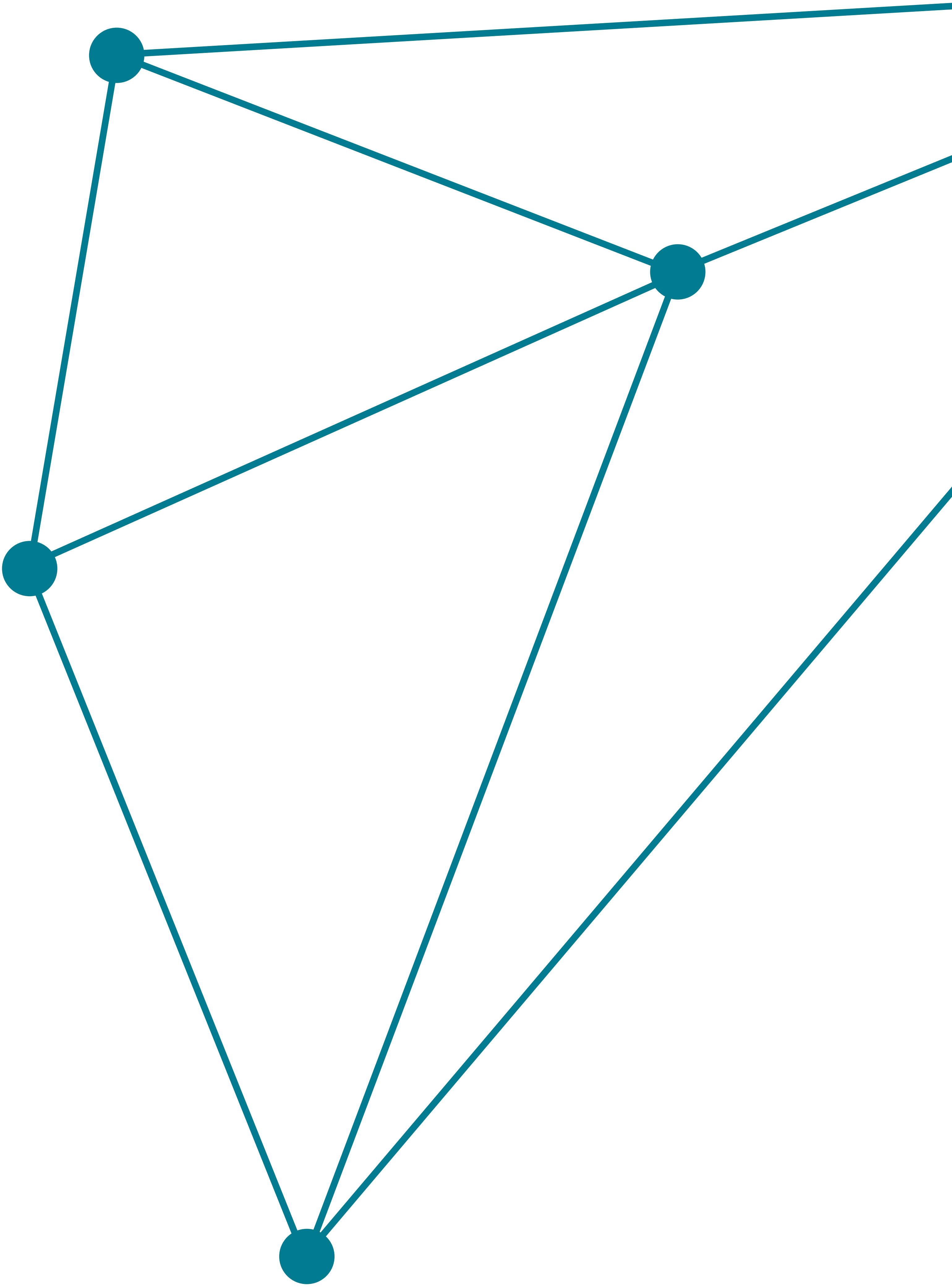
Chapter 11: Validation communication tool
The communication tool has been validated by performing a user test.

Chapter 12: Validation facial recognition system
The facial recognition system has been validated by performing a dimensions test and a proof of predictability test. Proving that the amount of landmark displacement does influence facial recognition and that facial recognition due to orthognathic surgery can be predicted.

Chapter 13: Future implementation
This chapter shows the VUmc everything they need to continue with the research and to implement the concept in the future.

Chapter 14: General Discussion
This chapter reflects on the work done and the filled research gaps. It also shows the impact of this project and the limitations.

Table of contents



Chapter 1

Summary

This chapter summarizes the project in 500 words, supported with visuals. First, it shows the framework of the project by an implementation of the model of the cognitive process and mechanisms of individual sensemaking and it briefly explains the process and approach. After, the classification system based on orthognathic surgeries and the landmark set is shown. Then, the research section and design section are summarized. All topics are placed on their own page with the belonging supporting visuals.

Process

The project was divided into two sections, the research section (45 days) and the design section (30 days). The rest of the days were used to do general tasks like writing the report and working on the showcase. The framework of the project can be seen in figure 1.1, which also summarizes the process. It shows that multiple information gaps have been filled with the research section (figure 1.3). During the design section was iterated to the research section. The approach of the project was to tackle it by looking from two perspectives, the human and computer science perspective. First, existing knowledge was found where after the gaps within the scope were identified. The gaps were filled with performing tests. From this, an ideation was done and the concept, called Fraos, which was feasible to work on according to complexity and the available time frame, was chosen and elaborated on. After this, the concept was evaluated with three more tests. The impact of the project is summarized in figure 1.2.

> ***"Now, there is proof that orthognathic surgeries do influence facial recognition"***
>
> ***"The research in the topic opened a new conversation and can provoke further research"***
>
> ***"The concept shows how research can be translated to something concrete which improves care"***

Figure 1.2: Summary of the impact of the project

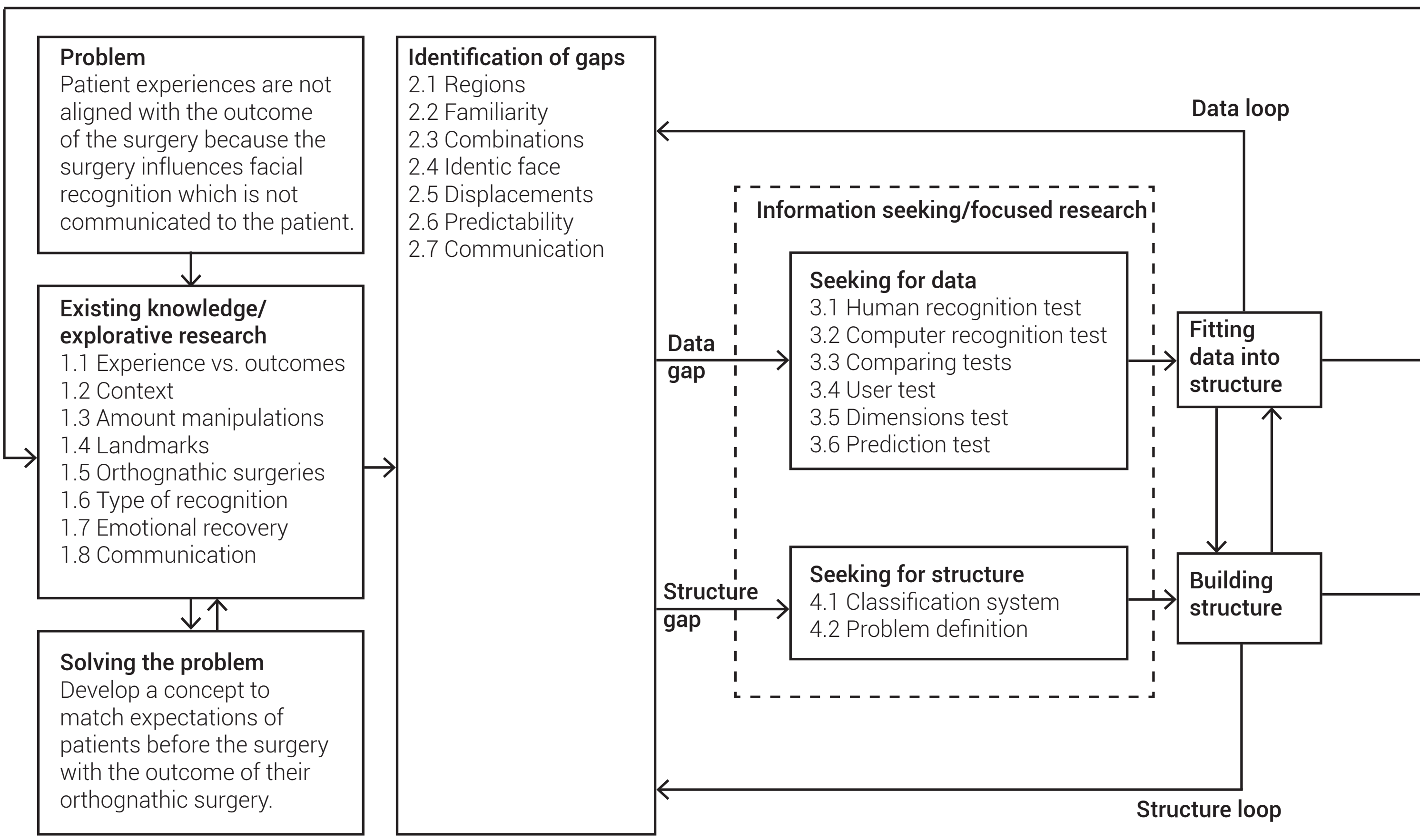


Figure 1.1: Implementation of the model of the cognitive process and mechanisms of individual sensemaking

Gaps before project
2.1 Regions
2.2 Familiarity
2.3 Combinations
2.4 Identic face
2.5 Displacements
2.6 Predictability
2.7 Communication

Gaps after project
2.8 Prediction all surgeries
2.9 Manipulate rotations
2.10 Complex surgeries familiar group
2.11 Specific area
2.12 Type of familiarity
2.13 Age
2.14 Expectations vs. similarity score
2.15 Experience Fraos
2.16 VUmc Data
2.17 3D pictures
2.18 Texture, expression, quality
2.19 Influence orthodontist
2.20 Surgery protocols
2.21 Patient evaluations

Figure 1.3: Defined gaps before and estimated gaps after the project

Outcomes updated knowledge

Accretion:
Instantiated structure
6.1 Final version of Fraos
6.2 Roadmap for future implementation
6.3 Recommendations and discussion

Tuning:
Adapted structure
5.1 The communication part of Fraos.
5.2 Facial recognition system of Fraos.

Re-structuring:
New structure
Three versions of 4.1
Two versions of 4.2

Iterative pattern
- 3.1 resulted from 1.2, 1.3, 1.4, 1.5 and 1.6
- 3.2 resulted from 1.3, 1.4 and 1.5,
- 3.3 resulted from 3.1 and 3.2
- 3.4 resulted from 1.7, 1.8 and 5.1
- 3.5 resulted from 1.4 and 5.2
- 3.6 resulted from 3.1 and 5.2
- 4.1 resulted from 1.4 and 4.5
- 4.2 resulted from all existing knowledge and 3.1, 3.2 and 3.3

Orthognathic surgeries
There was chosen for a landmark based facial recognition approach. Seen the complexity of orthognathic surgeries, they were simplified and a classification system was created to describe landmark movements happening in the patient's face (figure 1.5). The landmarks in figure 1.4 have been used throughout the project and the classification system is based on these landmarks.

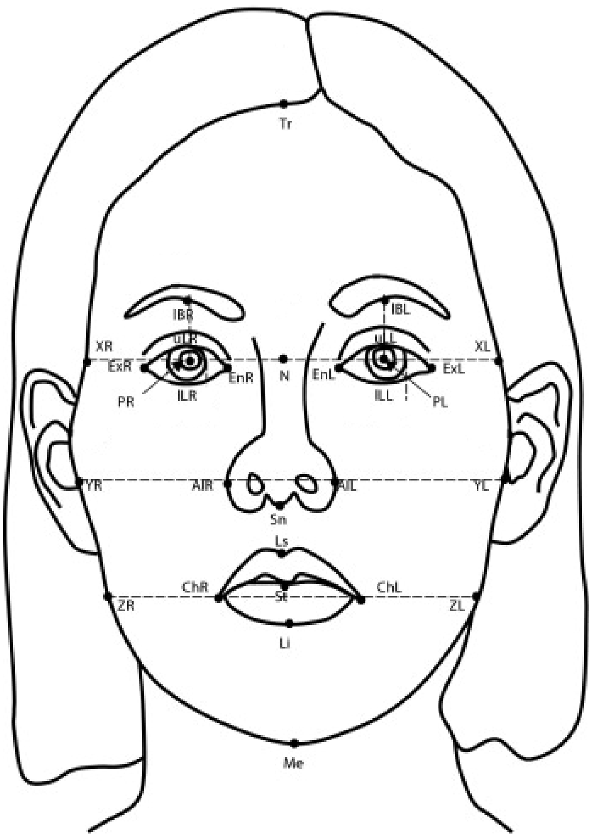


Figure 1.4: The landmarks used (Kiekens et al.,2008).

Region		Movement	Type of movement	Plane	
1	Maxilla	a	Horizontal Lengthening (L) or Shortening (S)	F	Frontal
2	Mandible	b	Widening (W) or Narrowing (N)	L	Lateral
3	Chin	c	Vertical Lengthening (L) or Shortening (S)		
4	Cheek	d	Rotating		
		e	Smoothening jaw line		

Figure 1.5: legend of classification annotations final version

Research section
During the explorative research at the start of the project, it became clear that no research in facial recognition after orthognathic surgeries was done before. Therefore, tests had to be done to be able to fill the information gaps. A human- and computer recognition test were done because the explorative research pointed out that a distinction had to be made. Although computer recognition methods are based on human perception, it does not take the familiarity of the person into account which humans do. Also, computer recognition is constant while human recognition is variable.

The human recognition test simulated changes in facial recognition after orthognathic surgeries, which were tested in a group of familiar people. The computer recognition test was done to test facial recognition among real orthognathic patients, to test more complex landmark movements and to be able to analyse a bigger data set. The outcomes of these tests are shown in figure 1.6.

Outcomes of human- and computer test

- Both tests show that movement c has the highest influence on facial recognition
- Both tests show that region 2 effects facial recognition more than region 1
- Computers are consistent in recognizing faces, humans are not
- The threshold of the computer test can be a strict recognition line, everything above will not be recognizable
- The threshold of the human test can be a recognition guideline, everything between the computer recognition line and human recognition line will be recognizable by part of the humans. Everything under the human recognition line will be recognizable by familiar humans

Figure 1.6: Outcomes of the tests in the research section

Design section

The concept, Fraos, is a tool existing out of two parts, a communication tool and a facial recognition system. The communication tool is the representation of all the information generated by the facial recognition system (figure 1.7).

The communication tool (figure 1.8 and 1.9) is a flyer which shows how recognizable a patient is predicted to be after the specific surgery. The communication tool has been validated by performing a user test.

The facial recognition system (figure 1.10) is what generates the similarity scores for all patients and it translates it to the communication tool. It has been validated by testing the effect of the amount of displacement of landmarks and the predictability with two real orthognathic patients.

The tool

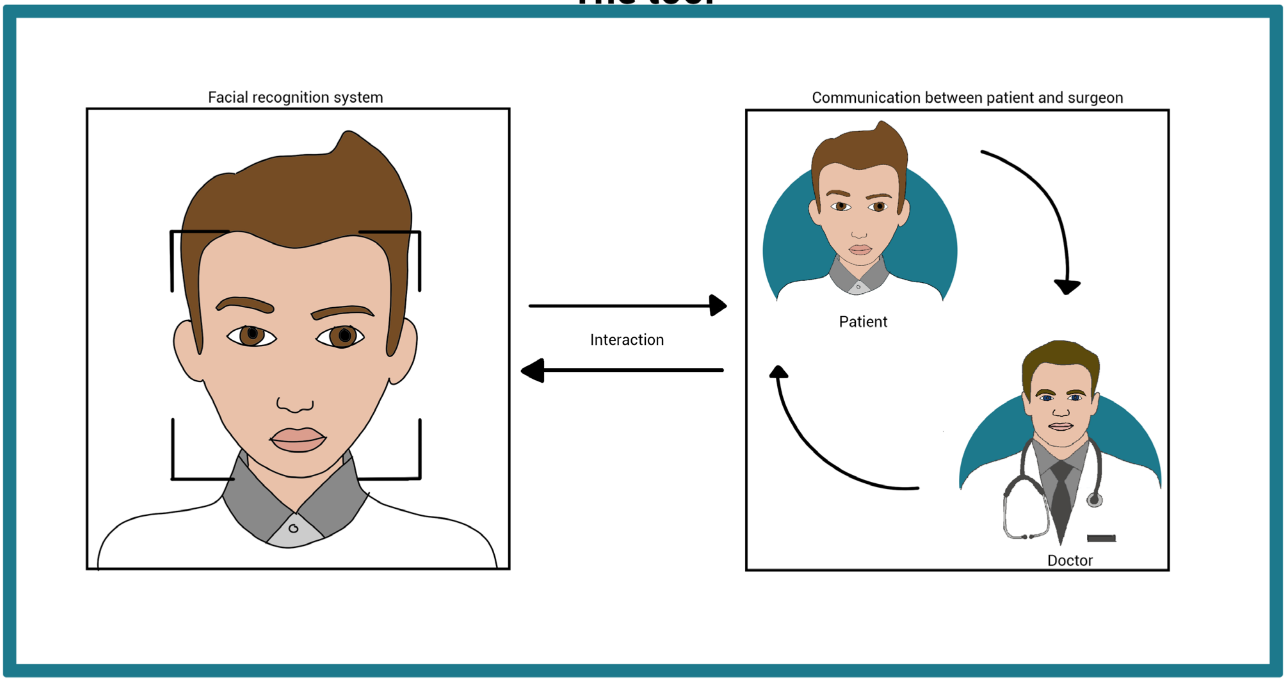


Figure 1.7: The interaction between the system and the communication

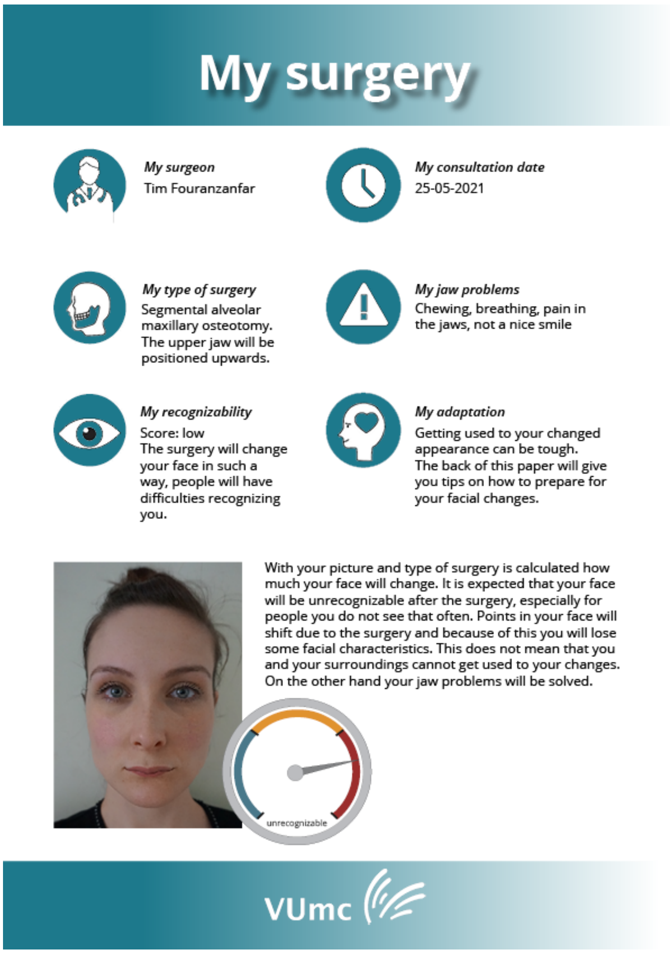


Figure 1.8 The front of the flyer

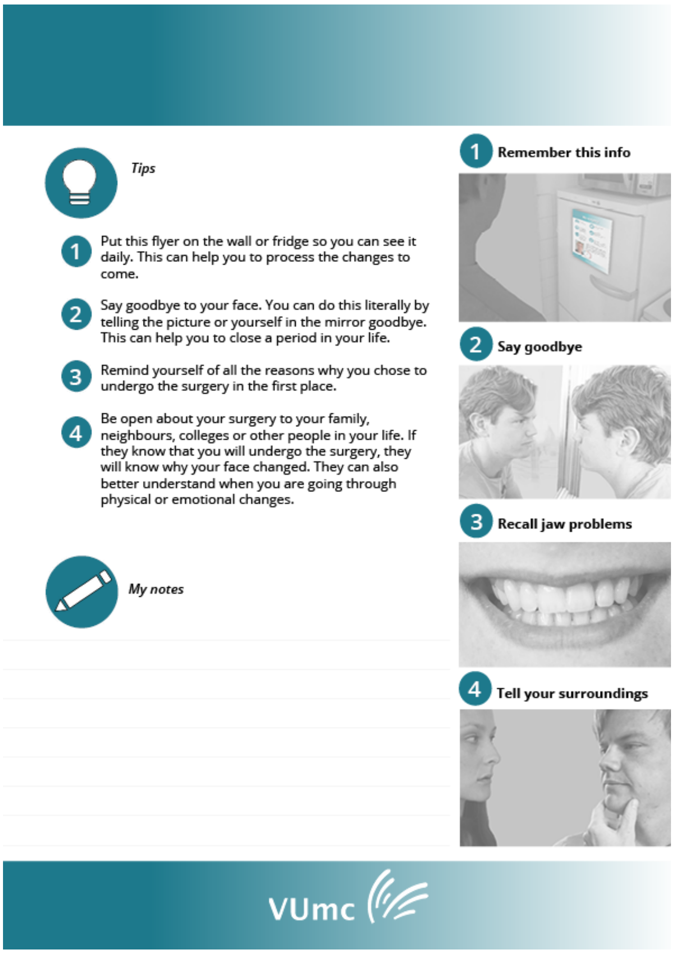


Figure 1.9: The back of the flyer

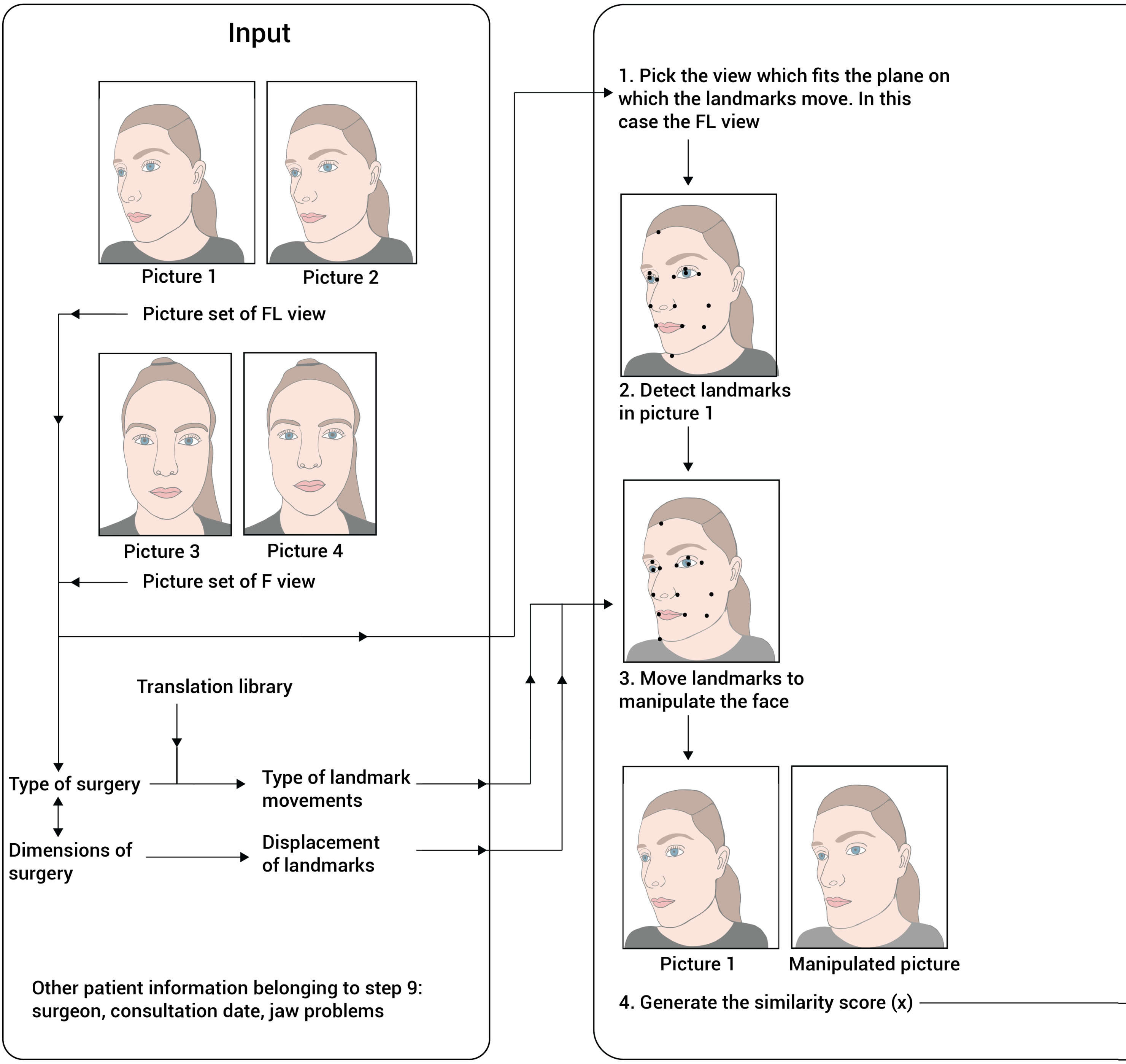


Figure 1.10: Flowchart of the facial recognition system

System

8. Put the patient information from Epic in the flyer: surgeon, consultation date, jaw problems, picture.

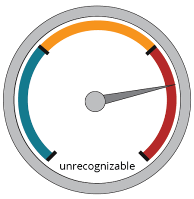
9. Save the flyer in Epic

Unrecognizable:
Appendix 25, page 112

Slightly recognizable:
Appendix 25, page 114

Recognizable:
Appendix 25, page 116

7. Activate information belonging to type of recognizability

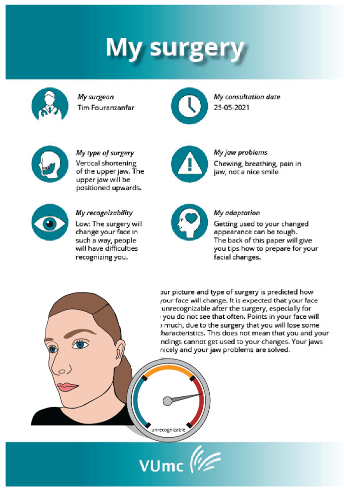


6. Find the facial recognition clock matching x

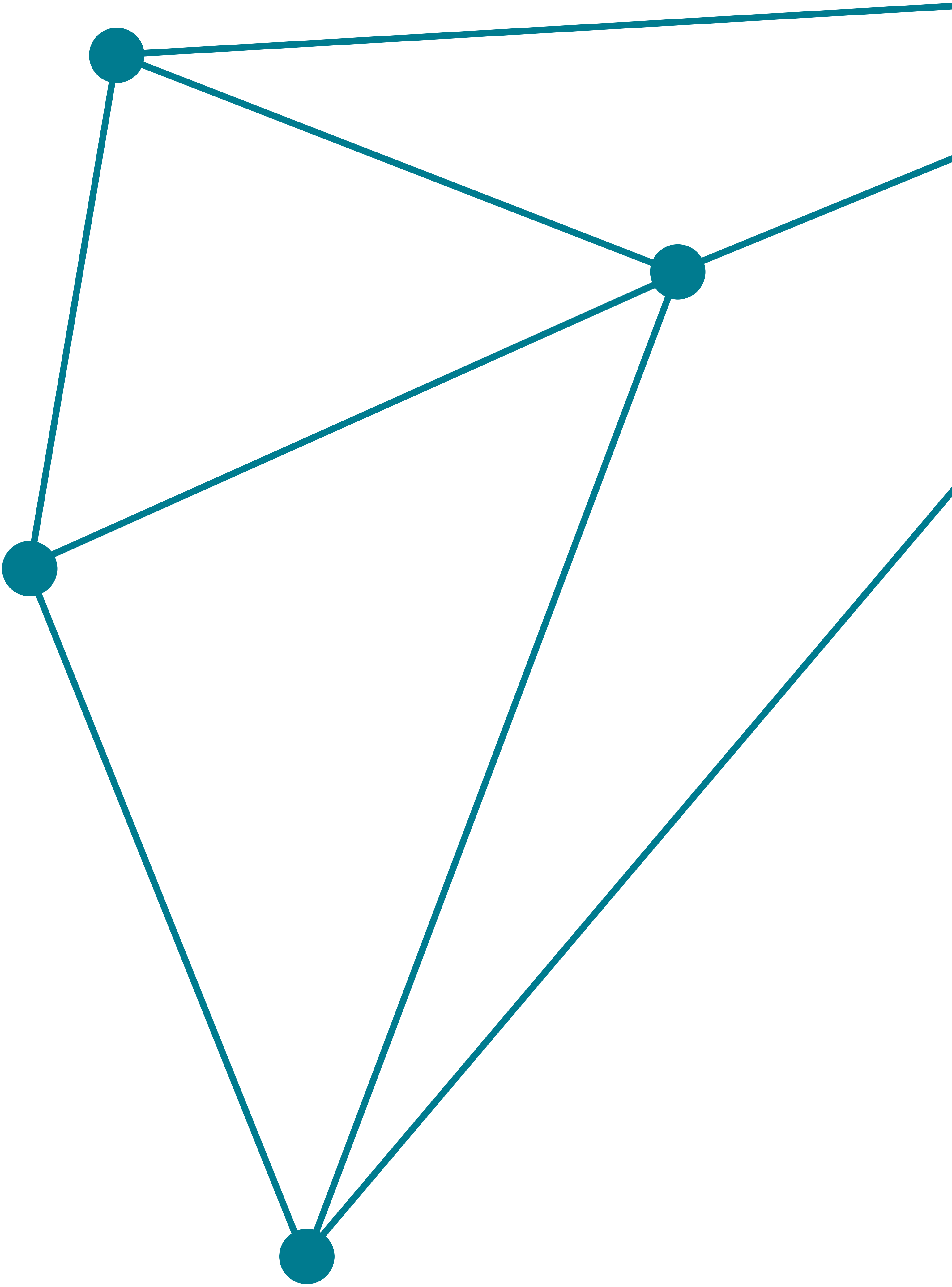
5. Define recognizability:

Recognizable x < 0.43
Slightly Unrecognizable 0.43 ≥ x < 0.45
Unrecognizable x ≥ 0.45

Output



The flyer specified to the patient



Chapter 2

Project introduction

This chapter provides an overview of the setup and approach of this project. It starts with a personal and general introduction of the project whereafter the problem statement is described. The problem statement describes the context of the project. To conclude, the chapter describes the two phases of the project, the research phase and design phase.

2.1 Introduction

2.1.1 Personal introduction
The project's subject has been chosen from personal experiences in the subject. My mom and I both underwent orthognathic surgery. Figure 2.1 shows a pre- and post-surgical picture from my double orthognathic surgery, where my mandible was horizontally lengthened and my maxilla rotated. I was still recognizable by my surroundings, but my mom's face changed drastically and she was not recognizable anymore by some people. An interview with her has been done to get some insights into her decision of undergoing an orthognathic surgery (appendix 2) and about the emotions she felt when she was not recognized. Since my mom had these experiences, I knew there would probably be more people having these experiences. When having the first meeting with Tim Fouranzanfar, an orthognathic surgeon at the VUmc hospital, I told him about my interest in facial recognition points (landmarks) and he was enthusiastic right away since some patients ended up with mental problems because they were not recognizable anymore after surgery. Therefore, this project is about facial recognition after orthognathic surgery.

2.1.2 General introduction
Around 5% of the population undergoes an orthognathic surgery to correct jaw problems (Posnick, 2013). Since the population in the Netherlands exists out of 17,4 million people (CBS, 2020), 870 thousand of these people undergo orthognathic surgery in their life. This project will be done for the orthognathic surgeons of the VUmc hospital. They are specialised in reconstructive surgeries.

A big problem within orthognathic care is the conflict of interest between patients and surgeons. The patient cares about the aesthetics of the face while the surgeon cares about the functionality of the jaw. Consequences of facial changes are therefore not always communicated during the consultation. (Bonanthaya & Anantanarayanan, 2013). Since preoperative variables are largely associated with the dissatisfaction of surgical results, it is important that the expectations of the patients are addressed during the consultation (Bellucci & Kapp-Simon, 2007; Bonanthaya & Anantanarayanan, 2013; Cadogan & Bennun, 2011).

Figure 2.1: Me before the orthognathic surgery (2011) and after the surgery (2014) with a similarity score of 0.52, which means I am unrecognizable when comparing those two pictures..

The reason why patients perform orthognathic surgery is mostly the functional indication, however, aesthetics is one of the primary motivations for surgery for patients (Andrup, 2015). The surgeons at the VUmc apply standard 'beauty protocols' when performing surgery, so they know what to do to make a face aesthetically attractive (T. Fouranzanfar, personal communication, March 2, 2020). However, some patients are still dissatisfied after the surgery, since they lose their facial characteristics which makes them unrecognizable for their acquaintances (T. Fouranzanfar, personal communication, March 2, 2020). Even if the surgical outcomes are ideal, it can lead to patient dissatisfaction when the psychological perspective has not been taken into account (Soh & Narayanan, 2013).

Surgeons in the VUmc warn the patients before the surgery since their face may drastically change (figure 2.2). However, how much the face changes and if the patient will become unrecognizable cannot be predicted by the surgeon and is therefore not communicated. Screening is done at the VUmc to screen out people with mental problems like the body dysmorphic disorder, which is 7-18% of the patients, seeking orthognathic care because of aesthetical reasons (Cadogan & Bennun, 2011; Vulink et al., 2008; T. Fouranzanfar, personal communication, March 2, 2020).

Although some patients and their surroundings can get used to the changes of the patient's face, some patients cannot get used to their different face which leads to identity problems (Bellucci & Kapp-Simon, 2007). After the surgery, the patients have to learn to recognize their adjusted face and so do their surroundings (Cadogan & Bennun, 2011). It takes time for most patients to accept the facial changes after the surgery and additional support after surgery can support emotional recovery (Bellucci & Kapp-Simon, 2007). Also, the social influence on the patient of being unrecognizable can lead to severe mental problems (T. Fouranzanfar, personal communication, March 2, 2020).

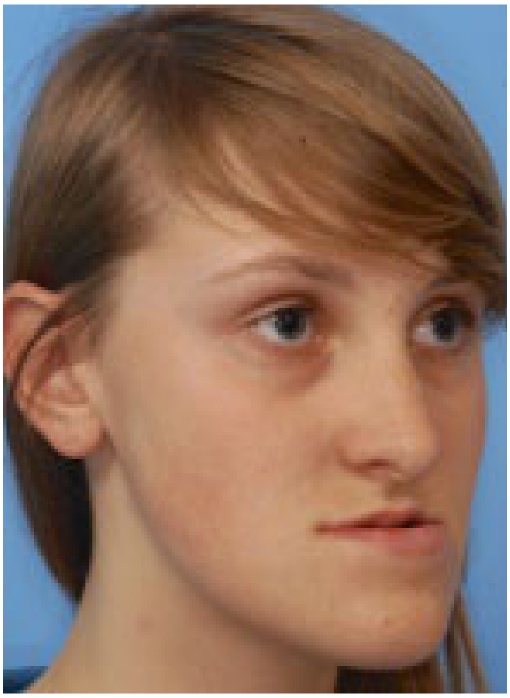


Figure 2.2: Example of a drastic change after orthognathic surgery, before and after, with a similarity score of 0.57 (Facial Sculpture Clinic, 2020)

2.2 Problem statement

2.2.1 Scope

This project is focused on facial recognition of the functional orthognathic surgeries which solves jaw deformities. A simple patient journey of the treatment is shown in figure 2.3. Most of the patients in the VUmc wear braces before, during and after the surgery which influences the positions of the teeth and jaws (T. Fouranzanfar, personal communication, March 2, 2020). The treatment at the orthodontist is out of the scope. The type of functional surgeries and the change in aesthetics is discussed in this project but is not the focus. The focus lies on facial recognition.

2.2.2 Problem definition

The expectations of the patients cannot be aligned by prediction of the appearance because this is inaccurate even when using prediction software. After all, the interaction between the soft- and hard tissues are different per patient and can only be determined during the surgery (T. Fouranzanfar, personal communication, March 2, 2020). The patient's perception and expectations about the surgery should align with the treatment plan, if not, this should be resolved before proceeding the treatment (Bellucci & Kapp-Simon, 2007).

If the surgeon could promise the patients that they would be recognizable after the surgery, it is meeting the 'before' image or perception of the patients of their face after the surgery, as well as meeting (part) of the expectations. Not being recognizable after a functional orthognathic surgery happens often at the VUmc. However, they do not know how to change the orthognathic surgery to keep patients recognizable or to communicate changes in facial recognition to the patient. There is a need for optimizing the orthognathic treatment to align patient experiences with the outcomes. Since there was a knowledge gap in facial recognition after functional orthognathic surgeries, this knowledge needed to be obtained.

2.2.3 Design assignment

The design assignment in figure 2.4 is the goal of the project, which is created by taking the scope and problem definition into account. The design assignment should solve the problem stated in the problem definition.

"Develop a concept to match expectations of patients before the surgery with the outcome of their orthognathic surgery."

Figure 2.4: The design assignment

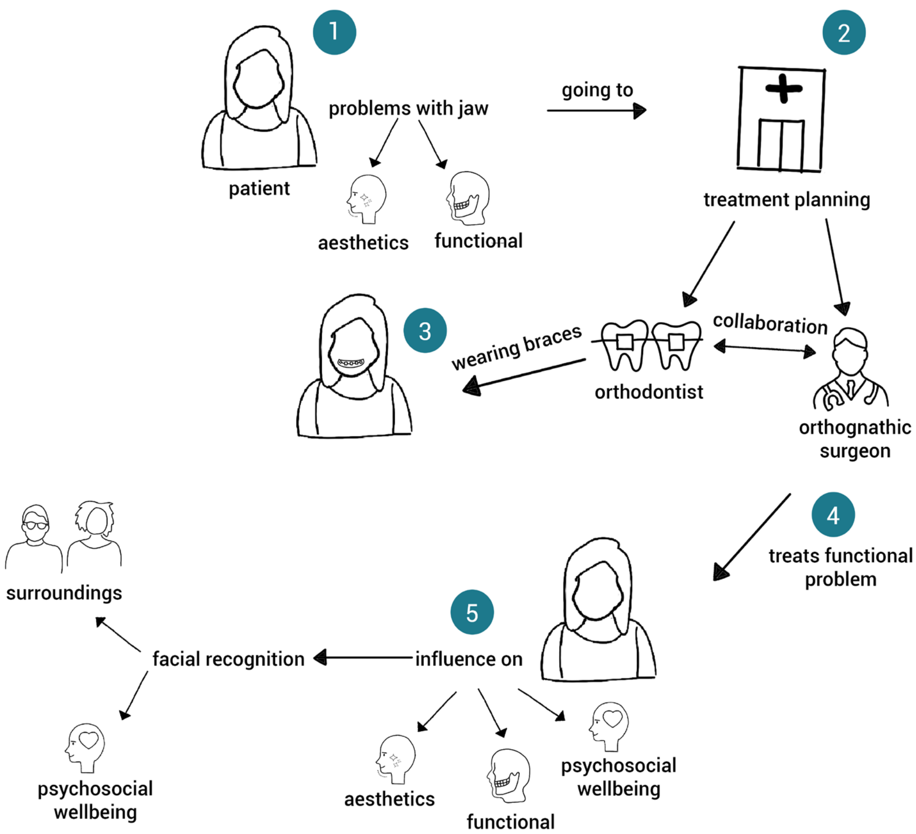


Figure 2.3 The context

2.3 Approach

2.3.1 The framework

The process of this project can be explained with the sensemaking model in figure 2.5. During the explorative research phase was decided to research two perspectives: humans and computer science. The ways to gather data via the human recognition test and computer recognition test were also determined during the explorative research phase. This has been described in chapter 2.3.2. First, a problem was formulated (chapter 2.2.2) with a possible solution (chapter 2.2.3). Existing knowledge was researched in literature where after a scope was set-up, on which the information gap to fill during the tests, was decided on. The data gap was filled by the two recognition tests and was later supplemented with further tests to validate the concept. The structure of the research was created beforehand (figure 2.6) and the details were filled in along the way. All the research was revised, where after a design vision was created which was feasible within the time frame of the project. During the design phase was iterated to the research and new research was done to fill the gaps which were needed to create a feasible design.
The sensemaking model (chapter 1 and 14) has been filled in at the end of the project to revise on all the iterations which were made, to verify if all the chapters of the report were referring to one another and to have an overview of the filled gaps and impact of this project.

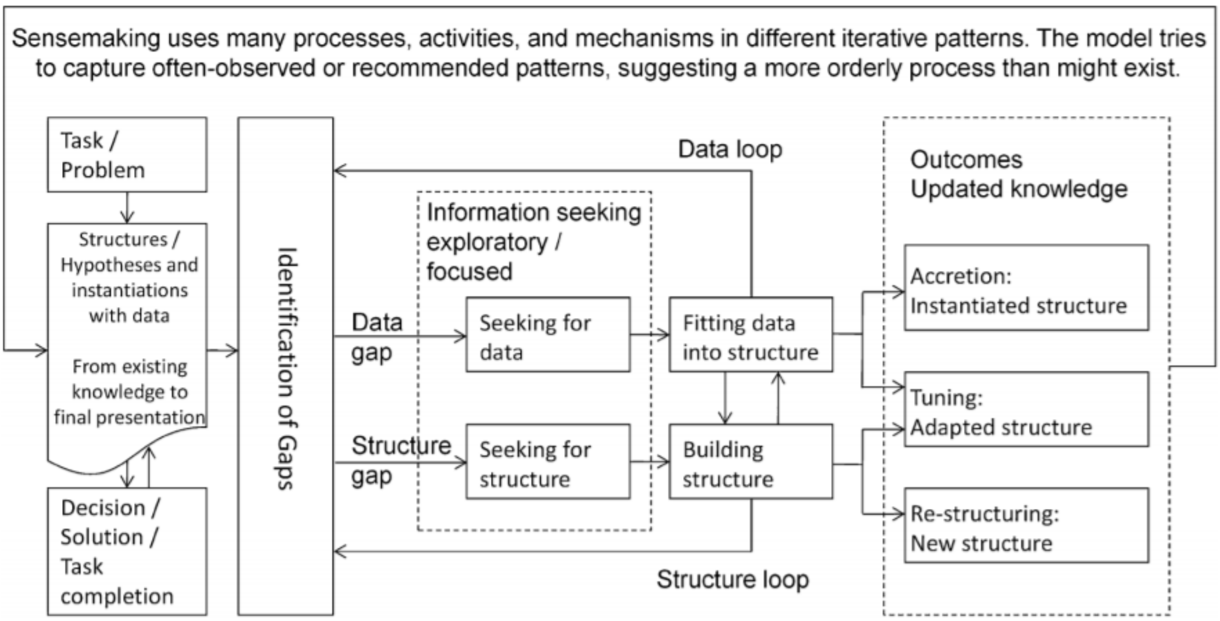


Figure 2.5: The model of the cognitive process and mechanisms of individual sensemaking (Aselmaa, 2017)

2.3.2 The two phases

The project is divided into a research phase and a design phase. The research phase is done first to create a more focused problem statement which can be used as a foundation for the design phase. Before starting this project, tests had to been done in facial recognition after orthognathic surgery, since this did not exist yet. The research phase was, therefore, divided into three types of research: literature research, a human recognition test and a computer recognition test (figure 2.6). There has been chosen to combine research of the human perspective and computer perspective to create a complete image of facial recognition.

The human perspective is needed since human recognition is the context of the patient, it is variable and it depends on familiarity. However, the computer perspective is easier and faster to examine, since it can always be used to examine a different type of data and the VUmc can do this on their own patient data as well. Combining the human recognition test and computer recognition test has been done to compare the human recognition threshold and computer recognition threshold. With this, the VUmc can always link their computer recognition results of their data to human recognition, which is applicable to the patient's context.

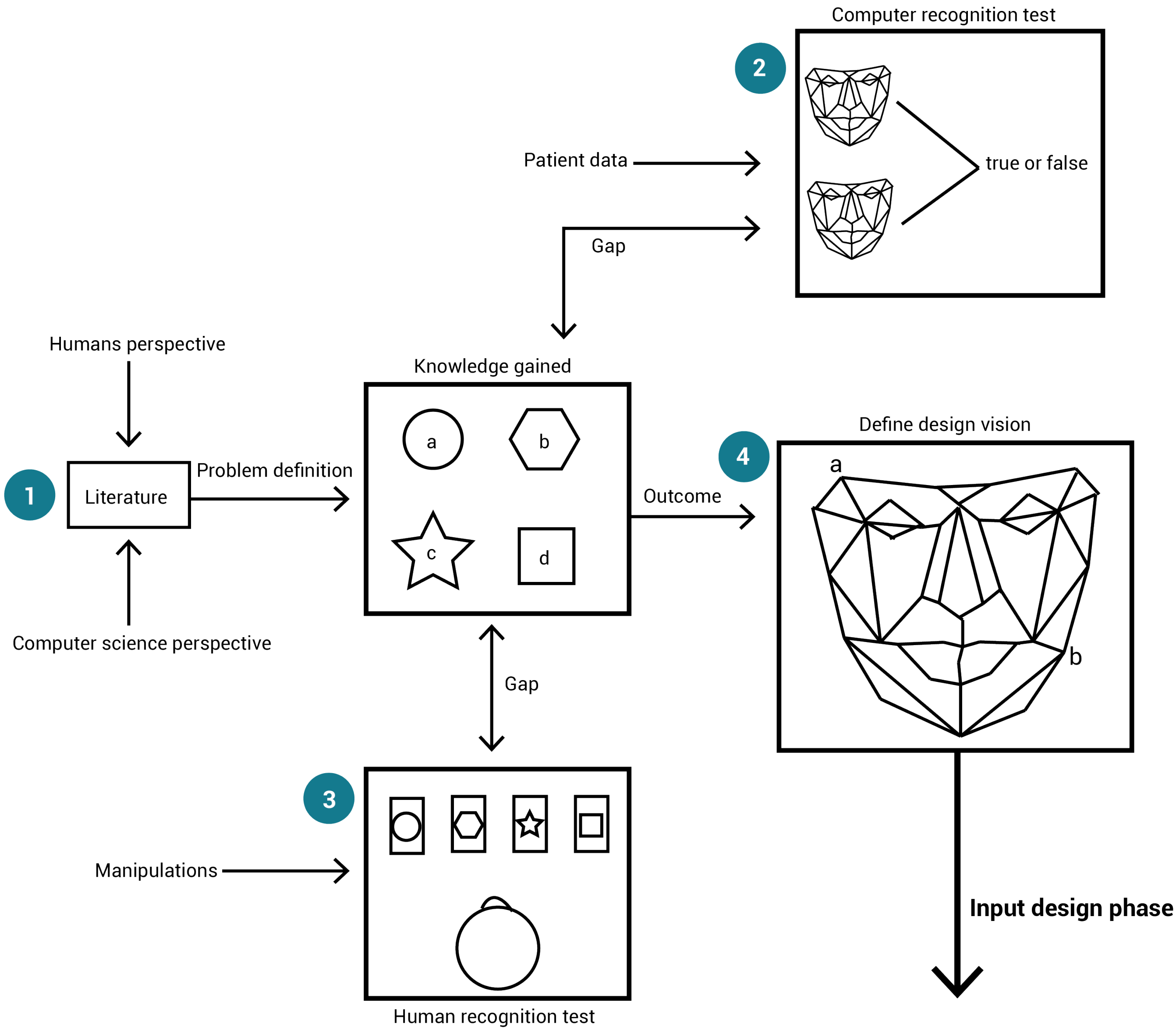


Figure 2.6: A flowchart of the research phase

2.3.3 Design process

Figure 2.7 shows the overall design process used in this project. The double diamond shape has been used, which is a visual representation of the design process. The shape exists out of four stages: discover, define, develop and deliver. The discover stage is part of the research phase, which focuses on divergent thinking and explores the research field. The define stage is also part of the research phase, which focuses on convergent thinking and is a more focused research continuing on the insights gathered in the discover stage. The develop stage is the divergent thinking and the deliver stage is the convergent thinking in the design phase (Design Council, 2019).

During the design phase, the research through design method has been used, since new knowledge has been gained during designing (Stappers & Giaccardi, 2017). Some more research has been done in order to make design decisions. Also, there has been iterated to the research phase.

2.3.4 Research questions

The research phase answers research questions to create a focused research process. The main questions per subject are shown and its sub-questions can be seen in appendix 3. The human recognition test and computer recognition test have their own more focused research questions in its chapters (6 and 7).

Main question
How and how much do orthognathic surgeries influence facial recognition?

Main subject questions
1. Humans
How and how well do humans recognize faces?

2. Computer
How and how well do computers recognize faces?

3. Orthognathic surgery
What is the influence of orthognathic surgeries on facial recognition?

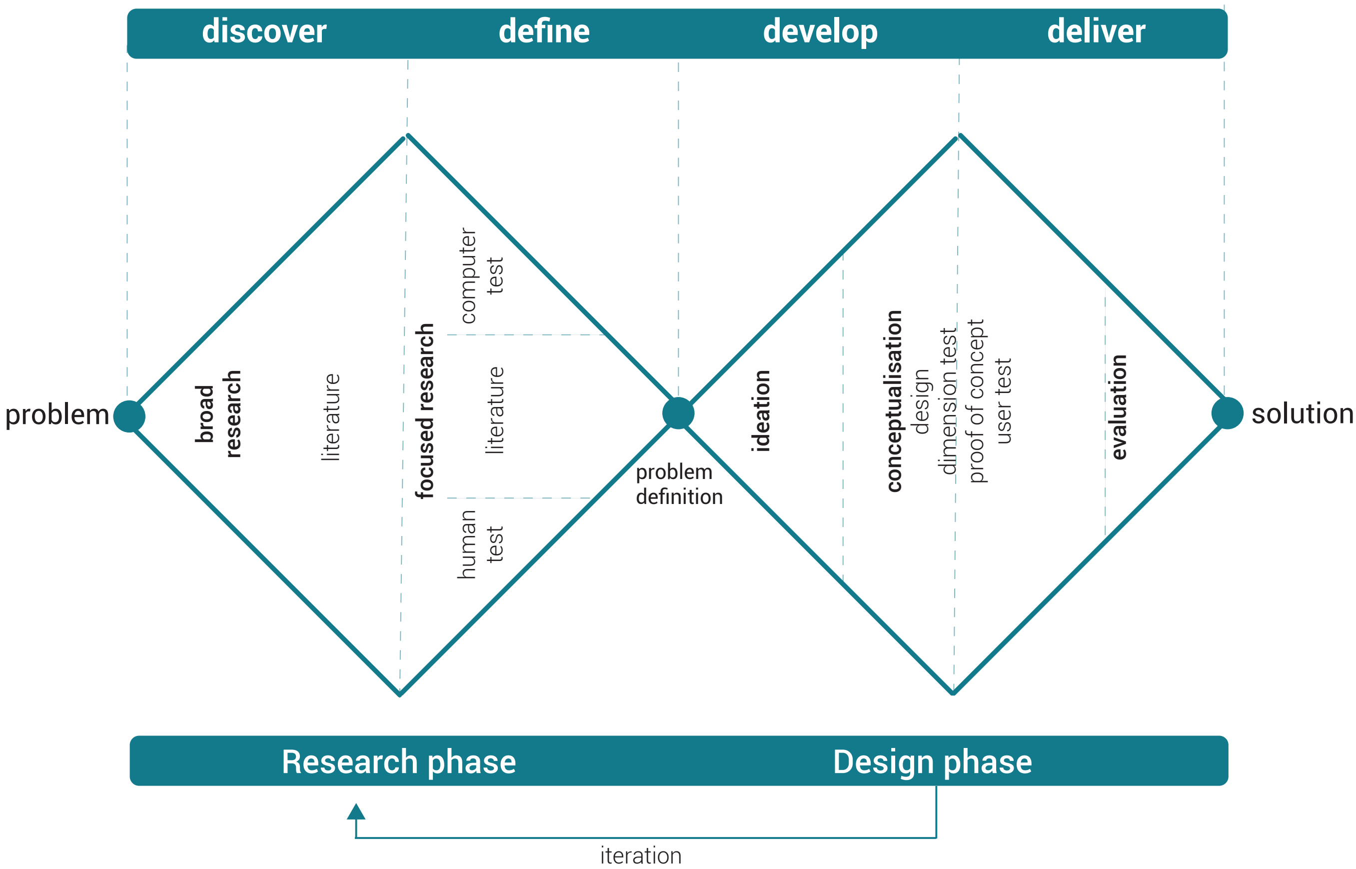
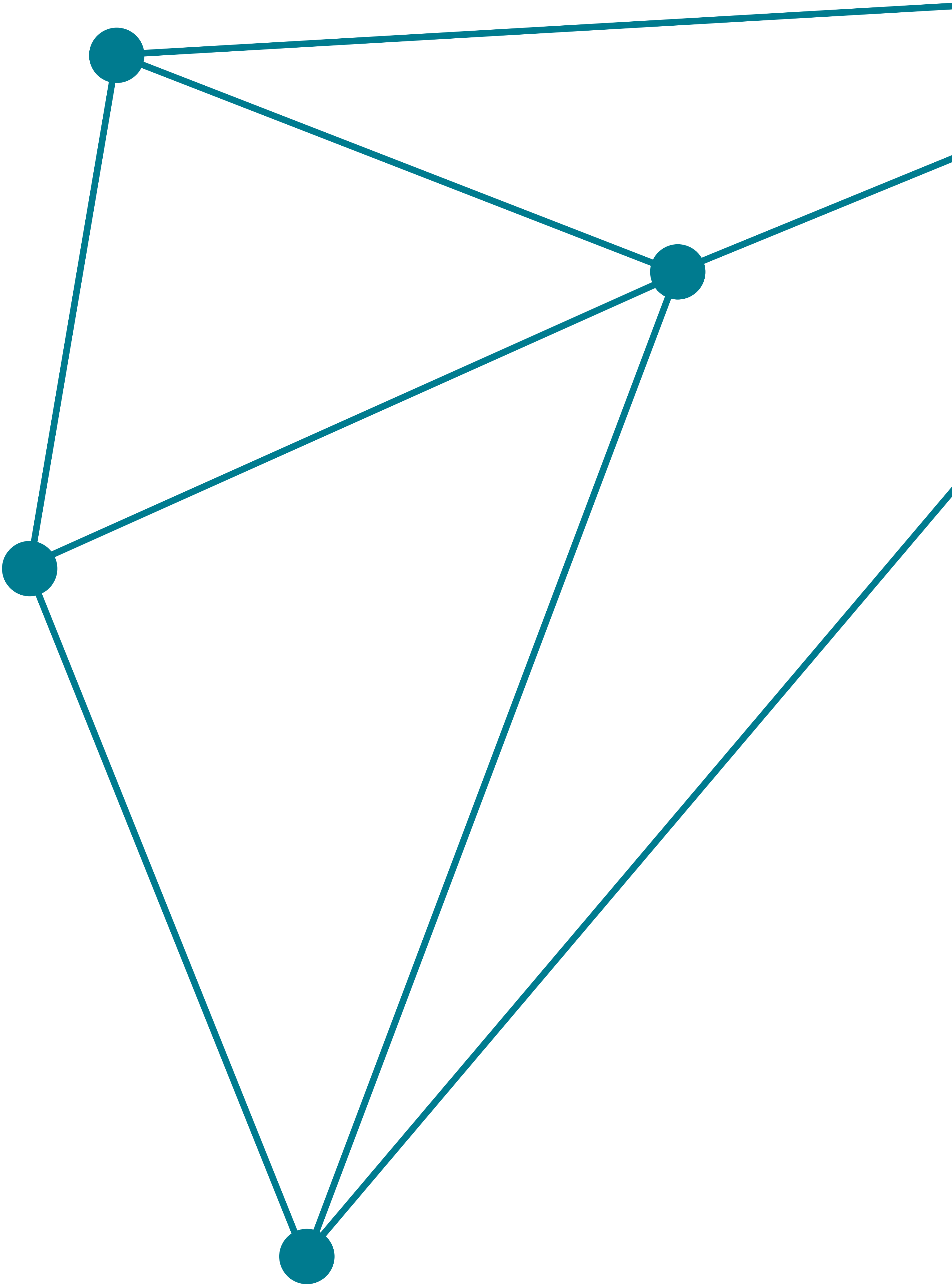


Figure 2.7: The design process during the whole project

Research section



Chapter 3

Facial recognition

This chapter shows the outcome of the literature research which has been done about facial recognition after orthognathic surgery, and facial recognition of humans and computers. The full literature research can be found in appendix 4. At first, the research about facial recognition after orthognathic surgery is described. After, the research of facial recognition of humans and computers have been described which is the foundation of this project. The research of facial recognition of humans has been used to create a research plan to create a link with facial recognition after orthognathic surgeries. Since 75 pre- and post-surgical pictures of orthognathic patients have been found, it is valuable to analyse this data to know if and how it influences facial recognition. However, analysing it by hand is time-consuming considering the number of pictures and the complexity of landmark movements due to the surgery. Therefore, research in computer facial recognition has been done which has been the foundation for the computer recognition test of chapter 7. In the computer research, iterations back to human facial recognition have been made, to find the differences.

3.1 Facial recognition after orthognathic surgery

Literature research has been done on how orthognathic surgery influences facial recognition. Unfortunately, only a little research has been done on facial recognition after orthognathic surgery. In total, one test has been done and this is done by Keshtgar et al., (2019). The result of this test is that it shows orthognathic surgery can influence facial recognition, other conclusions cannot be made from this research. Appendix 4.1 shows an elaborated analysis of this research and on another comparable research done in facial recognition after plastic surgery.

3.2 Human recognition

To understand how humans recognize other faces, literature research has been done. No research about the change of human facial recognition after orthognathic surgery can be found in the literature. However, understanding how humans process faces can help to estimate the influence of orthognathic surgery on facial recognition. After reflecting on this analysis and the facial recognition after orthognathic surgery analysis from chapter 3.1, a human recognition test is set up (chapter 6) to fill up the gaps in this research. In this chapter, the outcomes of the literature research are shown and links with computer recognition are made. The full literature research about human facial recognition can be found in appendix 4.2.

3.2.1 Identifying a person

In this project, the influence of orthognathic surgeries on face recognition is researched which is done by researching only recognition of the face. However, identifying a person is not only done by face recognition. Research in multiple cognitive models has been done to find out if separating research by only taking the face into account, can be done. Apparently, recognition of a person can be done through face but also through other multiple ways such as voice, name, personal belongings, body physique, body motion and handwriting (Barton & Corrow, 2016). Also, the contextual information affects facial recognition of familiar faces (Bruce & Young, 1986). Nevertheless, research of Barton & Corrow (2016) shows evidence that familiarity for faces, names and voices can be affected independently. Therefore, in this project, researching facial recognition related to orthognathic surgery by only taking the face into account, could be done.

3.2.2 Facial features

Facial recognition of humans is based on the facial features and their relation to each other (eyes, nose, mouth, ears) (Robson, 2014) (SciShow Psych, 2018) (Atkinson & Adolphs, 2011). When you see a person's face, the facial features are combined by the brain in such a way you can recognize the person. Research shows that recognition is not only about analyzing the separate facial features but also by creating relationships among features (Gold et al., 2012) (Atkinson & Adolphs, 2011). Which means, the perception of a face is the sum of its parts. This is called holistic face processing (Fuentes-Hurtado et al., 2019) (Gold et al., 2012). Humans first recognize the face as a face, before identifying it (Bruce & Young, 1986). This is also what is done in most face recognition computer techniques and in the computer method used in this project (chapter 3.3, figure 3.4). The position and shape of facial features play a big role in facial recognition and since the relationships among features are how human recognize faces, the distances between features play a role as well (Atkinson & Adolphs, 2011). In orthognathic surgery, the shape of the jaw can change and the positions of landmarks (chapter 4) of facial features change. Therefore, facial recognition can change due to orthognathic surgery. The way the face is processed depends on the familiarity of the face (Gold et al., 2012). When researching facial recognition, there should be a clear indication of using familiar or unfamiliar faces as stimulus material (Ellis et al., 1979). The context of this project is about the people who know the patient and therefore the patient's face is familiar. Therefore, only facial recognition research about familiar faces is done, which is described in chapter 6.

Also, research in the effects of the inner and outer features on facial recognition of familiar faces has been done (appendix 4.2.5). Apparently, the relationship between the inner and outer features play a big role in the identification of familiar faces (Ellis et al., 1979). Since orthognathic surgeries are changing one or multiple relationships between the outer features (e.g. chin, jawline) and the inner features (e.g. lips, maxilla, mandible) but also within the outer features and inner features, it could influence facial recognition. Especially the surgeries influencing multiple features in the face, like on both the mandible and maxilla, it is expected that these surgeries will influence facial recognition the most. This matches the outcomes of the human- and computer recognition tests in chapters 6 and 7, which shows orthognathic surgeries do indeed influence facial recognition. Manipulating multiple features is done in the computer test and manipulating single features in the human test. When a facial feature is reshaped, the local skin texture around this feature may also change and when changing a facial feature, the whole face geometric structure and appearance will be disturbed (Liu et al., 2013).

3.2.3 Face-space framework

To explain face identification in a more concrete way, the 'face-space' framework is used. The face-space (figure 3.2) is a multidimensional vector spaced framework to understand face identification. In this space, a database of individual faces exists and they are all compared to the average of all faces present in the database. This is also how it works in the human brain. The database of a person exists out of all the faces the person has ever seen. The distances among faces in this database represent their perceived similarity (Nishimura et al., 2010) (Valentine, 2001). Different computer recognition methods use the face-space framework (Valentine, 2001) but the computer recognition method used in this project does not use it (chapter 7).

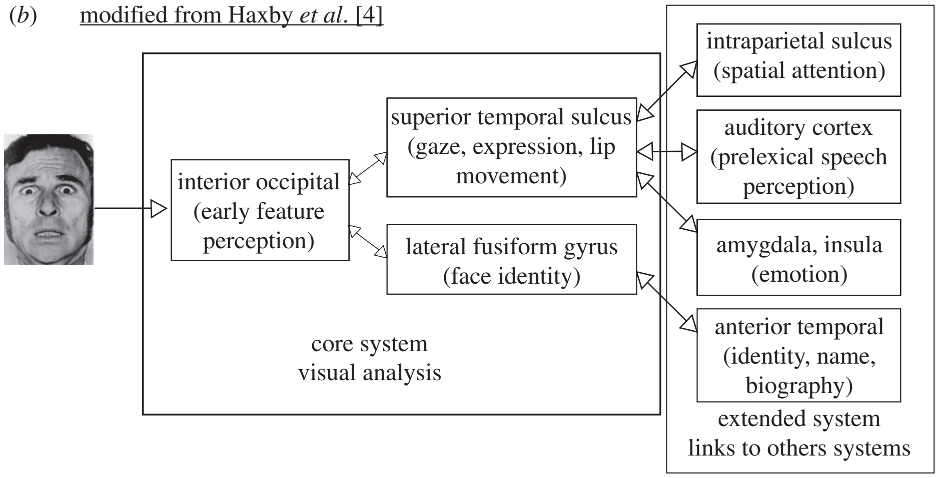


Figure 3.1: Neuropsychology of processing a face (Atkinson & Adolphs, 2011)

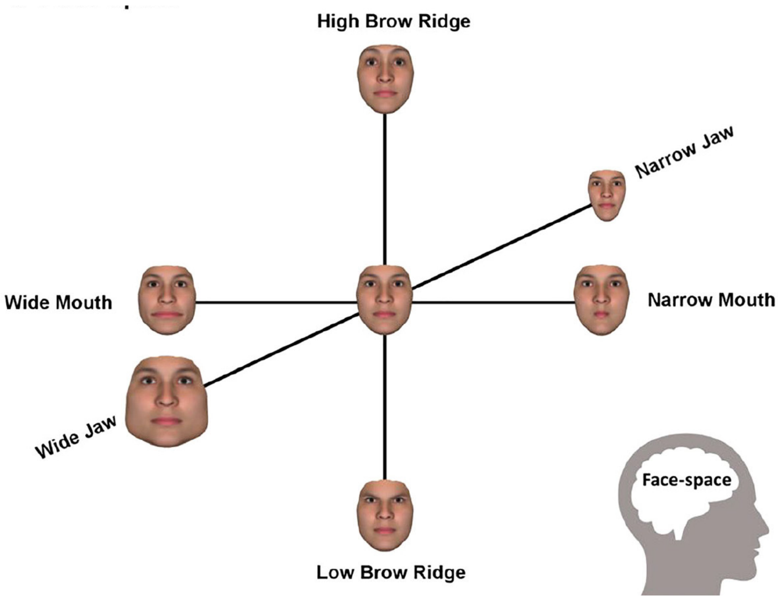


Figure 3.2: A face-space representation with facial feature dimensions. The average face in the center (Conway et al., 2019).

3.3 Computer recognition

Research in computer recognition is done as preparation for the data analysis in chapter 7. This data analysis uses pre and post-surgical pictures to analyze complex pictures and a bigger amount of data with the use of a computer recognition method (appendix 4.3.1). Since many methods for computer recognition are available, it was not possible to analyze and test them all within this project. Figure 3.3 shows how a face-recognition system works. The method which is chosen to be used for the computer test uses 'feature geometry' as an approach. Therefore, the skin colours and textures have not been taken into account. How the research in computer recognition systems is done where after the computer recognition method has been chosen can be found in appendix 4.

3.3.1 Similarity score

The goal of the computer recognition test was to find out how much orthognathic surgeries influence facial recognition. A similarity score can be used to give information about how similar two faces are, instead of only giving a true or false score. The similarity score in the chosen method is the difference in face distance between the pre and post-surgical face which is determined by comparing 128 'random' face values. The face_recognition and dlib libraries are used to achieve this. In short, the dlib library is trained to learn how to map the characteristics of a human face to a face embedding, which is a feature vector with 128 values called the 128-d embedding (figure 3.4) (King, 2009). The face_recognition library is built upon the dlib library (figure 3.5). In this library, the similarity score is the Euclidean distance between the two encoded faces (Geitgey, 2018; Ahdid, 2017) (figure 3.6). The landmarks do not play a role in creating the similarity score, they are only created to scale and position the face where after the face is encoded.

3.3.2 Comparison to human recognition

According to chapter 3.2, humans recognize faces by comparing them to the average face in their database. Faces from other races are more difficult to recognize because not a lot of these people are present in this database. In other words: the face is recognized based on where it is trained on. This is exactly what happens in computer facial recognition as well. In the method, the landmarks are not used to identify the face, but only to detect the face by detecting the facial features. This differs from human recognition since humans process faces by measuring the distance between facial features while the dlib and face_recognition databases compare faces through the 128 values. According to the founder of the face_recognition library, Facebook can recognize faces as good as humans can, with an accuracy of 98% (Geitgey, 2018). However, by using the dlib library, the accuracy is 99.38% (Geitgey, 2017). This means the method used, which includes human perception, has higher accuracy than only human recognition. What should be taken into account is that computer recognition is constant while human recognition differs per individual and also within individuals (appendix 4.3.4). Therefore a fixed accuracy cannot be set for human recognition.

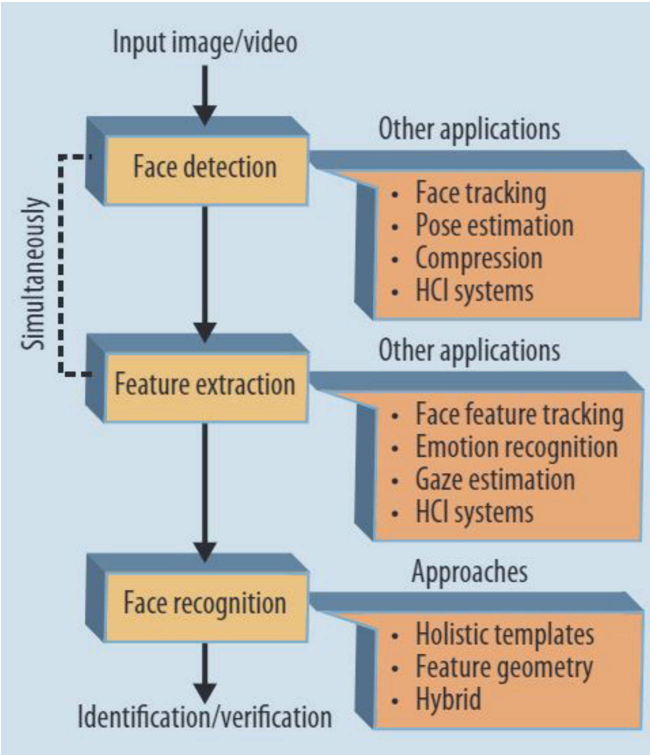


Figure 3.3: Generic face-recognition system (Chellappa et al., 2010)

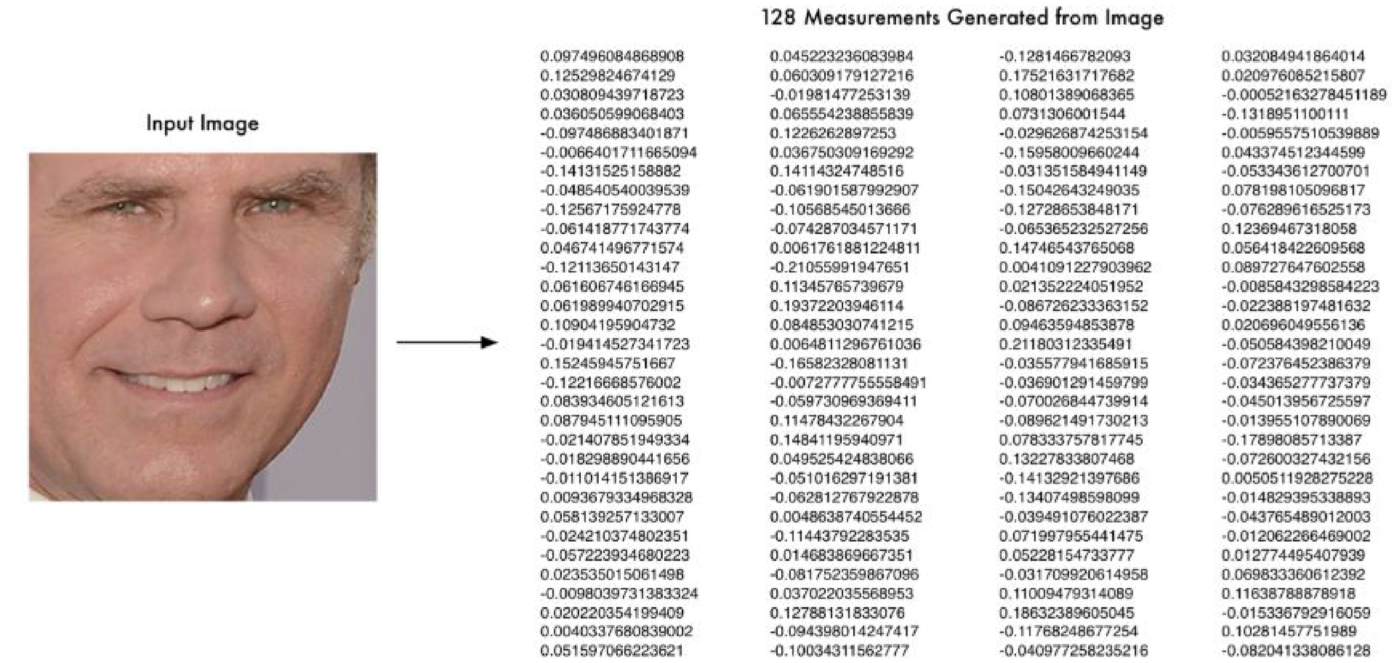


Figure 3.4: 128d embedding (Geitgey, 2018).

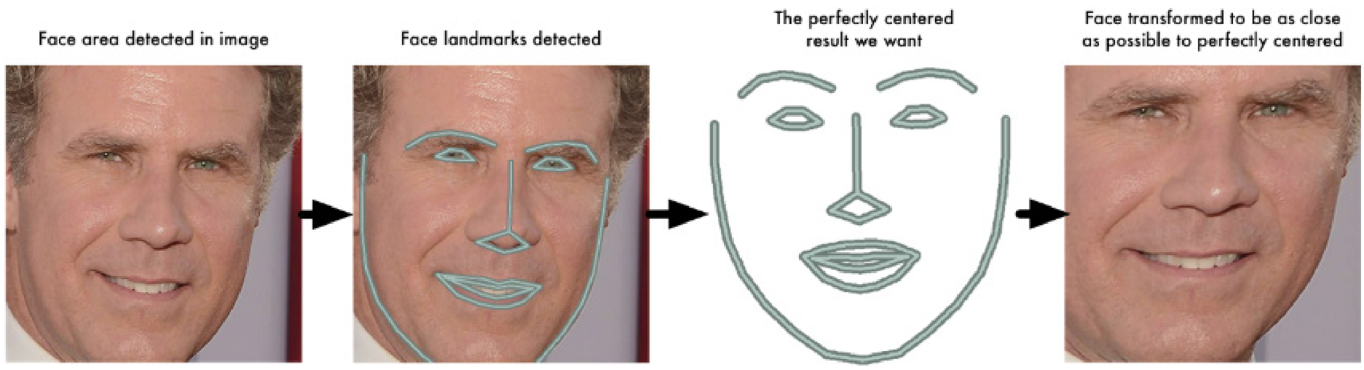


Figure 3.5: What the dlib does (Geitgey, 2018)

$$n=128 \qquad \begin{bmatrix} p1 \\ p2 \\ \vdots \\ pn \end{bmatrix} \qquad \begin{bmatrix} q1 \\ q2 \\ \vdots \\ qn \end{bmatrix}$$

Encoding face 1 Encoding face 2

$$\text{Euclidean distance} = [(p1-q1)^2 + (p2-q2)^2 + \ldots + (pn-qn)^2]^{\frac{1}{2}}$$

Figure 3.6: How calculating the Euclidean distance from the encodings of the pre- and post-surgical face is done

3.4 Conclusion

The research shows that the link between human and computer facial recognition and orthognathic surgeries is missing. The research done in facial recognition after orthognathic surgery is scarce and from this research, it seems that orthognathic surgery can influence facial recognition but is not clear how facial recognition is influenced.

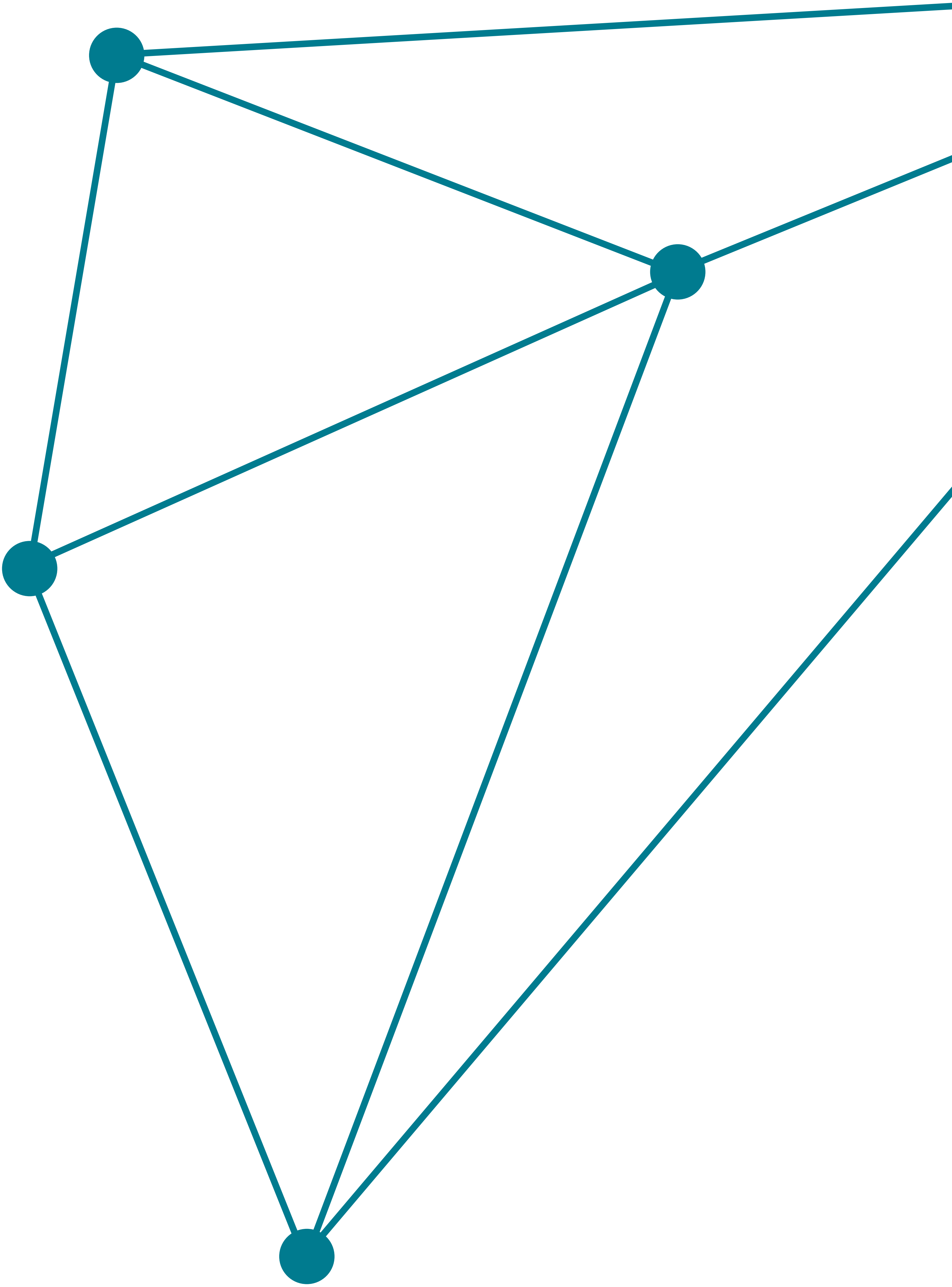
Human and computer recognition methods are performed differently besides that the distances within the face are used by both methods. Also, computer methods do have higher accuracy in recognizing a face than humans do. However, all computer recognition methods can be seen as psychologically plausible, since they are trained on human perception. Before starting this project, a research plan was made to fill the research gaps, by performing a human recognition test and computer recognition test. The plan of performing a human recognition test was initially initiated to implement the 'human' factor more since human recognition is what is designed for. However, after literature research, the testing plan changed. The human recognition test is performed to test how single configurations happen in different types of orthognathic surgeries and how this influences facial recognition of familiar faces. This had to be done in a group where most people knew each other. From this test is known what landmark movements representing specific types of surgeries influence facial recognition between familiar people (chapter 6).

The plan of performing a computer recognition test was initially initiated to analyse a large and complex amount of data consisting out of multiple configurations happening in multiple types of surgeries. This analysis would be used to compare the computer and human threshold of facial recognition. This is done, however, the focus of this computer test was to compare the similarity of a complex amount landmark movement due to certain orthognathic surgeries (chapter 7).

When changing a facial feature, the whole face geometric structure and appearance will be disturbed. The computer method which is chosen for the computer recognition test in chapter 7 is based on feature geometry. However, when a facial feature is reshaped, the local skin texture around this feature may also change which can influence facial recognition. Mark that the method does not take skin texture into account. The influences from this on facial recognition is out of the scope in this project but could be explored in further research.

Key-points from this chapter

- There is indeed a research gap in facial recognition after orthognathic surgery
- The human recognition test is performed, based on this chapter, to test how single configurations happening in different types of orthognathic surgeries influence facial recognition of familiar faces
- It is expected that surgeries on both the mandible and maxilla will influence facial recognition the most
- The dlib and face_recognition libraries are used for the computer recognition test
- Computer recognition methods are psychologically plausable since they are trained with human perception



Chapter 4

Facial landmarks

A facial landmark is a point lying somewhere in the face. The landmark is defined by a coordinate pair of x and y. In literature, different configurations of landmarks within the face can be found. Landmarks in the face are easy to find by looking at the facial features. The landmarks are also called: facial nodes, facial recognition units, facial key points, facial feature points or nodal points. However, the most common name is landmarks and therefore in this project, they are called landmarks. Research is done to define what configuration is most applicable for this project. In this project two types of landmarks are used, the landmarks used by the computer recognition method and the landmarks used to describe orthognathic changes in the face, which is used in the human recognition test. An analysis has been done in different sets of landmarks, in the areas of computer science and orthognathic care. This chapter only describes the two sets of landmarks which are used throughout the rest of the project. The full analysis can be seen in appendix 5.

4.1 In computer science

The method which is chosen for the computer recognition test (chapter 7) uses 68 landmarks (figure 4.1). This is the amount of landmark used the most in literature and software. The 68 landmarks are based on the default amount of 27 landmarks used in software (figure 4.2). In this project, the 68 landmarks used by the computer method are only used to detect the face by detecting the facial features. More about these landmarks is explained in appendix 5.

4.2 In orthognathic care

In orthognathic treatment planning, anthropometric facial analysis is used by some different type of physicians for the soft-tissue facial analyses (Kiekens et al., 2008) (Sforza & Ferrario, 2006). However, in the VUmc, the orthognathic surgeons use cephalometric landmarks of the hard-tissue for treatment planning. The guidelines of the VUmc are mostly functional, and their guidelines only consist of advise to optimize aesthetics. During the orthognathic surgeries in the VUmc, the functional part which belongs to the hard-tissue is the focus and the aesthetics which belongs to the soft-tissue are done instinctively after the functional part is done since a prediction of the movement of the soft-tissue cannot be done and differs per patient. The surgeons know how to influence the soft-tissue to make the patient look aesthetically good, but it is not explicitly mentioned in their guidelines (T. Fouranzanfar, personal communication, March 31, 2020). In future research, it could be valuable to find a balance between facial aesthetics and recognition. An optimal position of landmarks could be found, taking the recognition margin into account. This type of research could change the guidelines of surgeries in the future. The anthropometric and cephalometric landmarks are used in treatment planning and are therefore interesting to further examine when a tool is developed for the surgery itself. However, in this project, nothing will change in the treatment planning and the orthognathic anthropometric landmarks are only used to describe changes happening in the face.

When further research is done in facial recognition due to orthognathic surgeries, more information about anthropometric and cephalometric landmarks can be found in appendices 5.2.2 and 5.2.3.

The landmarks which are used in this project for describing movements of landmarks due to orthognathic surgeries are shown in figure 4.3, 4.4 and 4.5. These anthropometric landmarks exist out of 29 frontal and 16 lateral points. The landmarks are based on research on all types of orthognathic patients from Kiekens et al., (2008). Since the landmarks are based on a wide range of orthognathic patients, it is assumed that these landmarks cover the landmarks which will be influenced by the orthognathic surgeries on the different patient classifications (chapter 5) which are used as data in chapter 7. Kiekens et al., (2008) used the landmarks of figure 4.3 to create ideal ratios and angles which create an average ideal face which can be used as a guideline in orthognathic care. However, this aesthetic focus is outside the scope of this project but could be interesting for further research. Appendix 8 describes the details of the research of Kiekens et al., (2008) and shows an overview of the ideal ratios and angles.

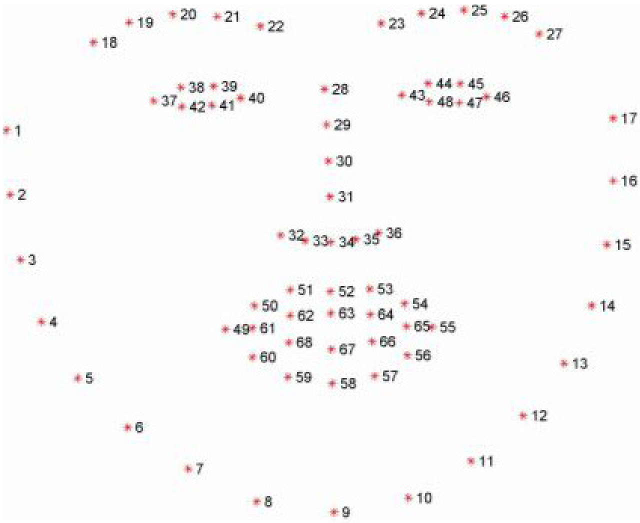


Figure 4.1: The amount of landmarks used the most in literature and software is 68 (Amato et al., 2018)

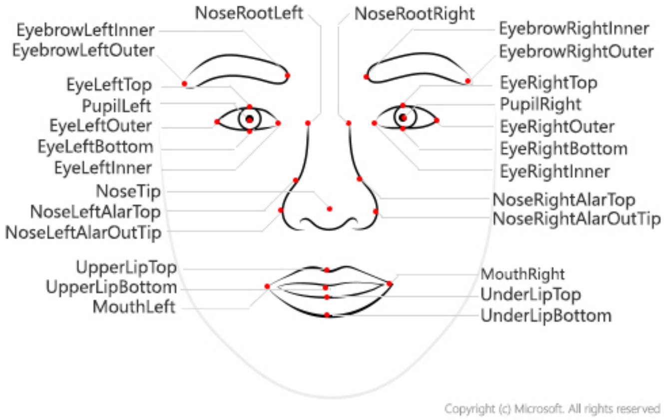


Figure 4.2: Default 27 landmarks in software (Farley, 2019)

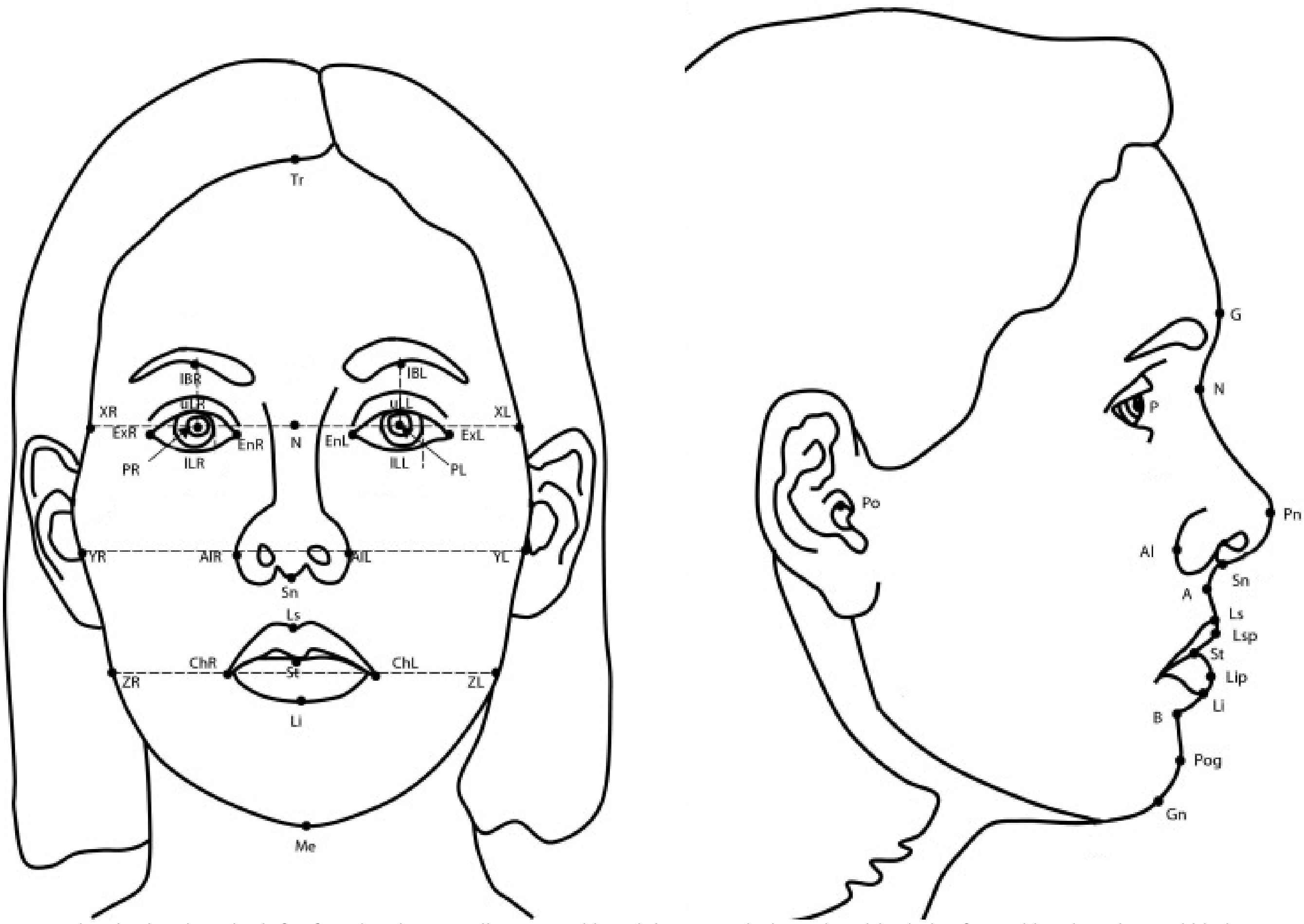


Figure 4.3 a,b: The landmarks left after the data quality control by Kiekens et al., (2008). With a) the frontal landmarks and b) the lateral landmarks.

List of names of figure 4.3

a), TR, trichion; IBR, lower border of the eyebrow on the right side; IBL, lower border of the eyebrow on the left side; N, skin nasion at bipupil line (constructed point); ExR, exocanthion on the right side; ExL, exocanthion on the left side; EnR, endocanthion on the right side; EnL, endocanthion on the left side; uLR, upper limbus the right side; uLL, upper limbus on the left side; ILR, lower limbus on the right side; ILL, lower limbus on the left side; PR, middle of the pupil on the right side; PL, middle of the pupil on the left side; AlR, alare on the right side; AlL, alare on the left side; Sn, subnasale; St, stomion; ChR, cheilion on the right side; ChL, cheilion on the left side; Ls, labrale superior; Li, labrale inferior; Me, menton; XR-XL, face width at bipupil line (XR and XL = constructed points); YR-YL, face width at alare (YR and YL = constructed points); ZR-ZL, face width at stomion (ZR and ZL = constructed points).

b) Landmarks on the lateral photograph: G, glabella; N, nasion; P, pupil; Pn, pronasale; Al, alare; Sn, subnasale; A, soft-tissue Point A; Ls, labrale superior; Li, labrale inferior; St, stomion; Lsp, most protruded point of upper lip; Lip, most protruded point of lower lip; B, soft-tissue Point B; Pog, pogonion; Gn, gnathion;

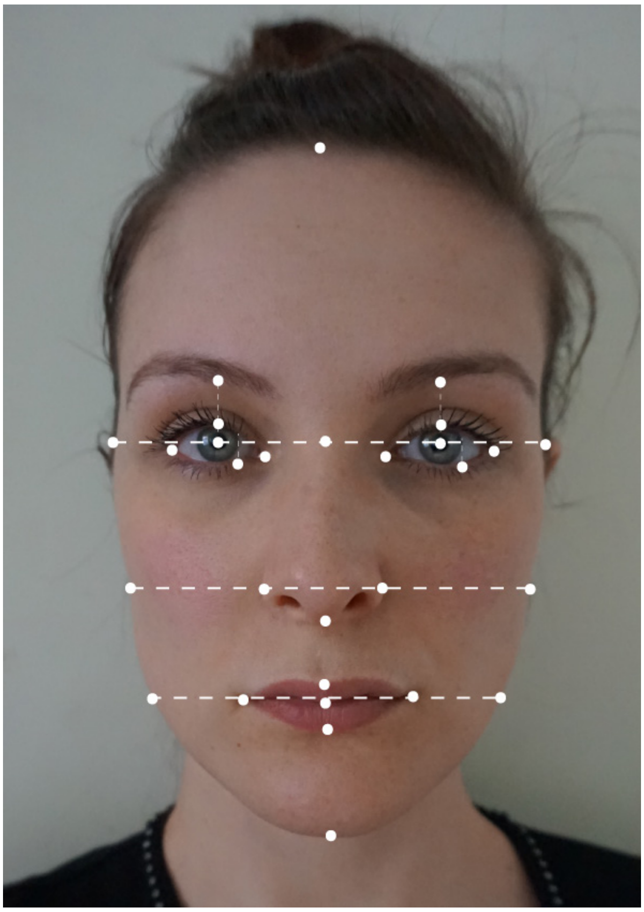


Figure 4.4: The frontal landmarks

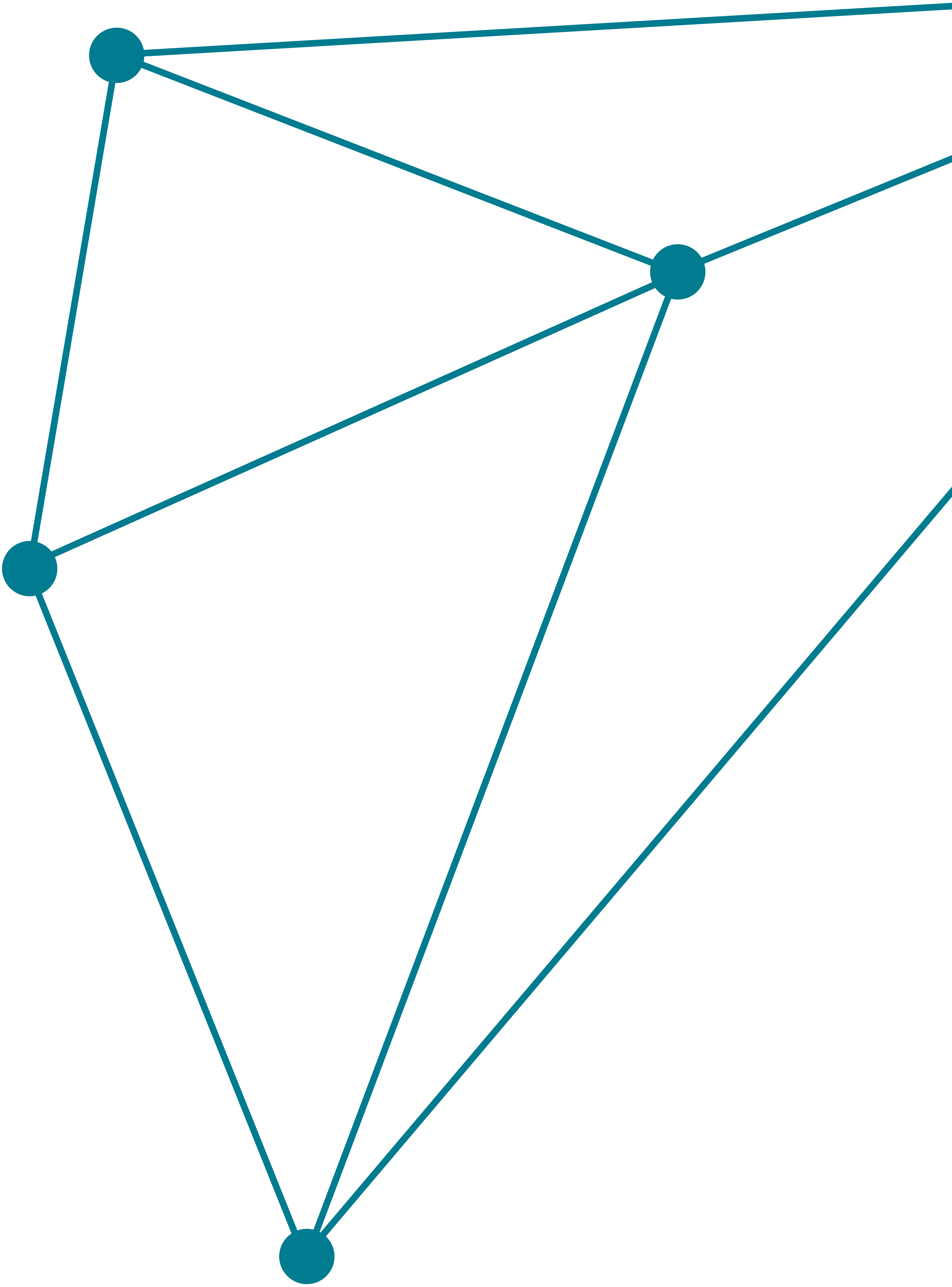
Figure 4.5: The lateral landmarks

4.3 Conclusion

In computer science, the default amount of landmarks is 27, which excludes the jaw region. The jaw is localized with use of the default landmarks, which is part of the 68 landmark model which is mostly used in literature and software and which is the model which is used in the computer recognition test (chapter 7). In describing landmark movements in this project, the landmarks of Kiekens et al., (2008) are used. These anthropometric landmarks are based on the research of orthognathic patients in the Netherlands and are, therefore, optimal for describing what happens after orthognathic surgeries. The movement of the landmarks in the soft-tissue is the focus of this project, therefore the cephalometric landmarks are only shown in appendix 5.2.3 as a foundation for further research.

Key-points from this chapter

- The 68 landmark set is used in the computer recognition test
- Anthropometric landmarks are landmarks of the soft-tissue, which are not used by the VUmc since the prediction of soft-tissue is inaccurate
- Cephalometric landmarks are landmarks on the hard-tissue, which are used by the VUmc
- The landmarks of Kiekens et al., (2008) have been chosen as landmarks for describing movements in the orthognathic region. Existing out of 29 frontal and 16 lateral points



Chapter 5

Classification system

Different types of orthognathic surgeries exist. In this project, only the functional surgeries will be taken into account with the focus on jaw deformities. Classifying the patients for the computer and human test is needed. Figure 5.1 shows a classification of the jaw deformities where the osseous deformities affect the jawbones and the dental deformities affect the teeth. This scheme represents the complexity of orthognathic care. This type of classification is too complex to examine in this project. Therefore, a simplified classification system is introduced in this chapter. The patients in this project should be classified according to the movements of the landmarks. However, the classification system of the VUmc does not apply this and neither do classification systems found in research. In this chapter, the new classification system is introduced and an example of how to apply this system is shown.

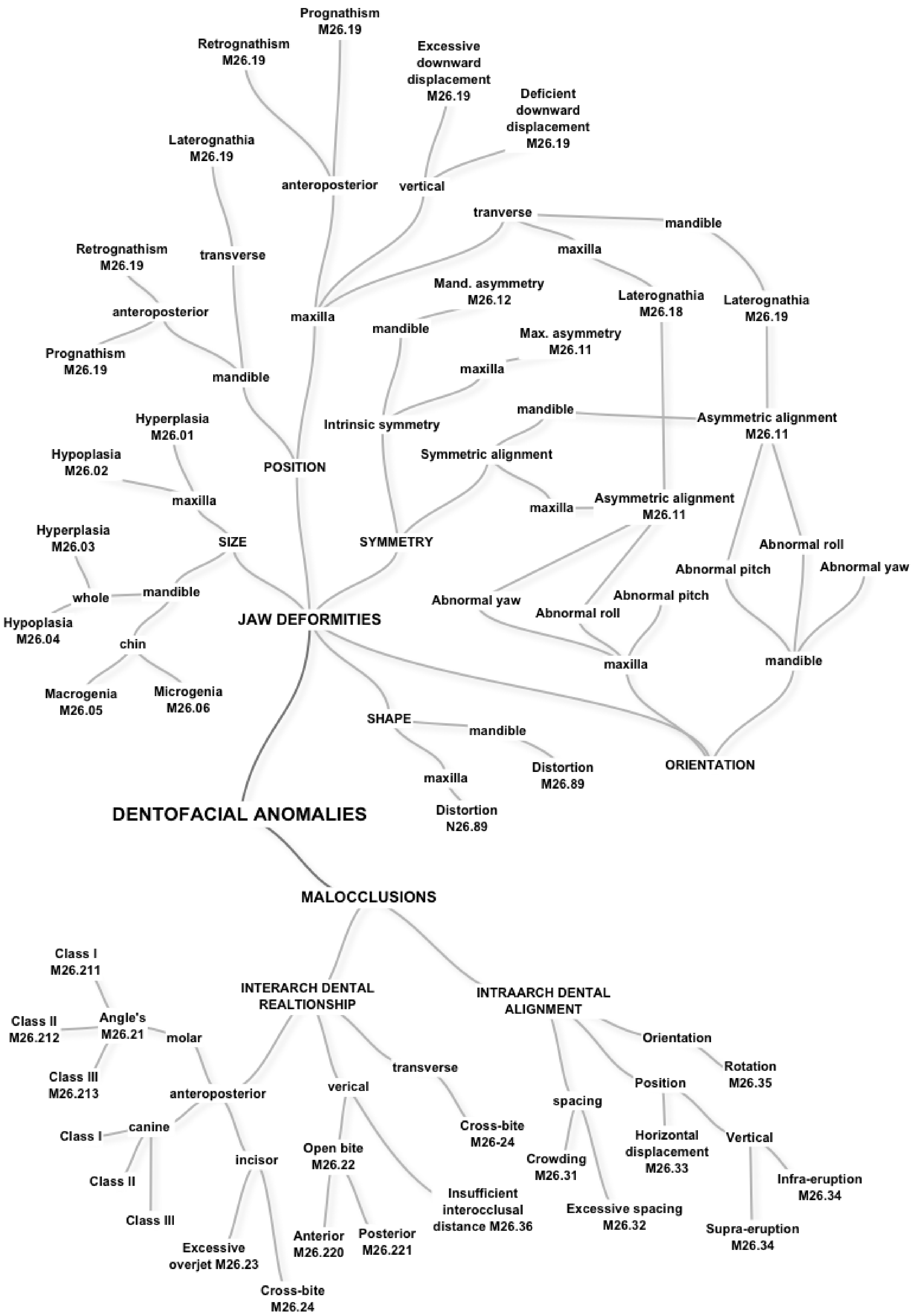


Figure 5.1: Mindmap of the dentifactial anomalies (Gateno & Xia, 2015)

5.1 The new classification system

The landmarks which were manipulated in chapter 6 were selected based on what landmarks change due to certain othognathic surgeries. This type of classification system did not exist yet and is created during this project. The reason for this new system is to classify patients on type of landmark movements. The new classification system is created after analyzing the classification system in the VUmc, other classification systems in literature, after observing the pre- and post-surgical pictures of patients and analyzing the most common performed surgeries (appendix 7).

The new classification system which is used for all the research is shown in figure 5.2. How this system can be used is shown in an example in chapter 5.2. After the research and design phase, an improved classification system was created (figure 5.3), taking the specific type of movement into account. The 'type of movement' in the classification system shown in figure 5.3, makes a distinction between lengthening or shortening and widening or narrowing. This has been done because of the outcomes in the dimension test of chapter 12, pointing out that facial recognition is depending on the type of movement. However, this test has been done after creating and applying the first version of the classification system and the rest of this report will, besides chapter 10, use version 1 (figure 5.2) of the classification system.

Region		Movement		Plane	
1	Maxilla	a	Horizontal Lengthening/Shortening	F	Frontal
2	Mandible	b	Widen/Narrowing	L	Lateral
3	Chin	c	Vertical Lengthening/Shortening		
4	Cheek	d	Rotating		
		e	Smoothening jaw line		

Figure 5.2: legend of classification annotations version 1

Region		Movement	Type of movement	Plane	
1	Maxilla	a	Horizontal Lengthening (L) or Shortening (S)	F	Frontal
2	Mandible	b	Widening (W) or Narrowing (N)	L	Lateral
3	Chin	c	Vertical Lengthening (L) or Shortening (S)		
4	Cheek	d	Rotating		
		e	Smoothening jaw line		

Figure 5.3: legend of classification annotations version 3

The new classification system can be used for 2D images from the frontal and lateral view (figure 5.4), or frontal-lateral view (figure 5.6). Mark that the depth of the images is not taken into account when only the frontal or lateral view is used. However, all the movements in the frontal plane could affect the face in the lateral plane and vice versa but the classification system only describes in what plane the movement itself happens (figure 5.7). Since this project is looking into what landmarks change due to the surgery, different orthognathic regions are created to be able to classify the patients according to landmarks. The regions are enclosed by landmarks and only address movements of the soft-tissue. Since the movement of the soft-tissue cannot be predicted, the soft-tissue regions are estimated and are only used as a communication tool to see which type of surgery addresses which part of the face and to see which regions influence facial recognition. The regions are bounded by the landmarks of Kiekens et al., (2008)(figure 5.8).

What could be noticed is that movements over the x-axis only influence the landmarks of the corresponding frontal or lateral plane and the movements over the y-axis influence the landmarks of both planes (figure 5.4).

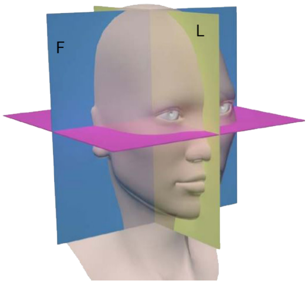


Figure 5.5: The blue frontal and yellow lateral plane (Gateno & Xia, 2020, edited image)

Figure 5.6: The frontal-lateral view.

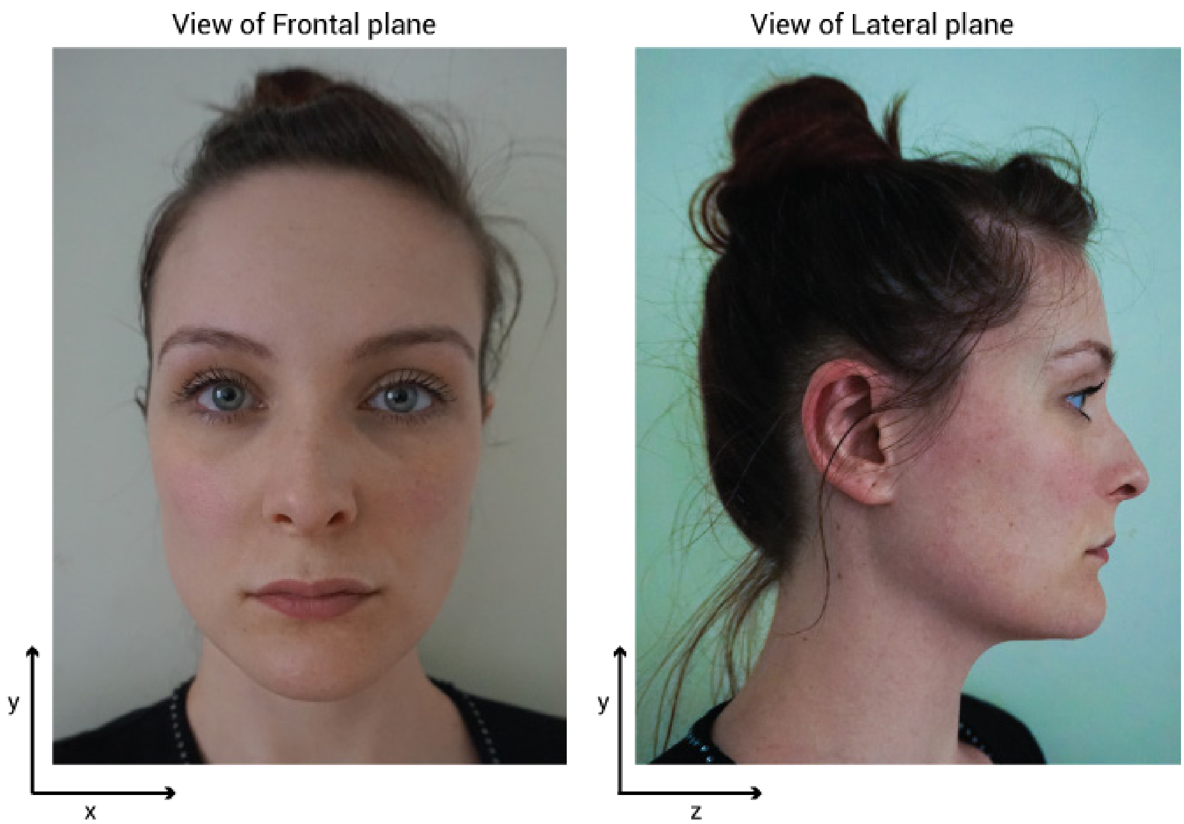


Figure 5.4: The axes of both planes.

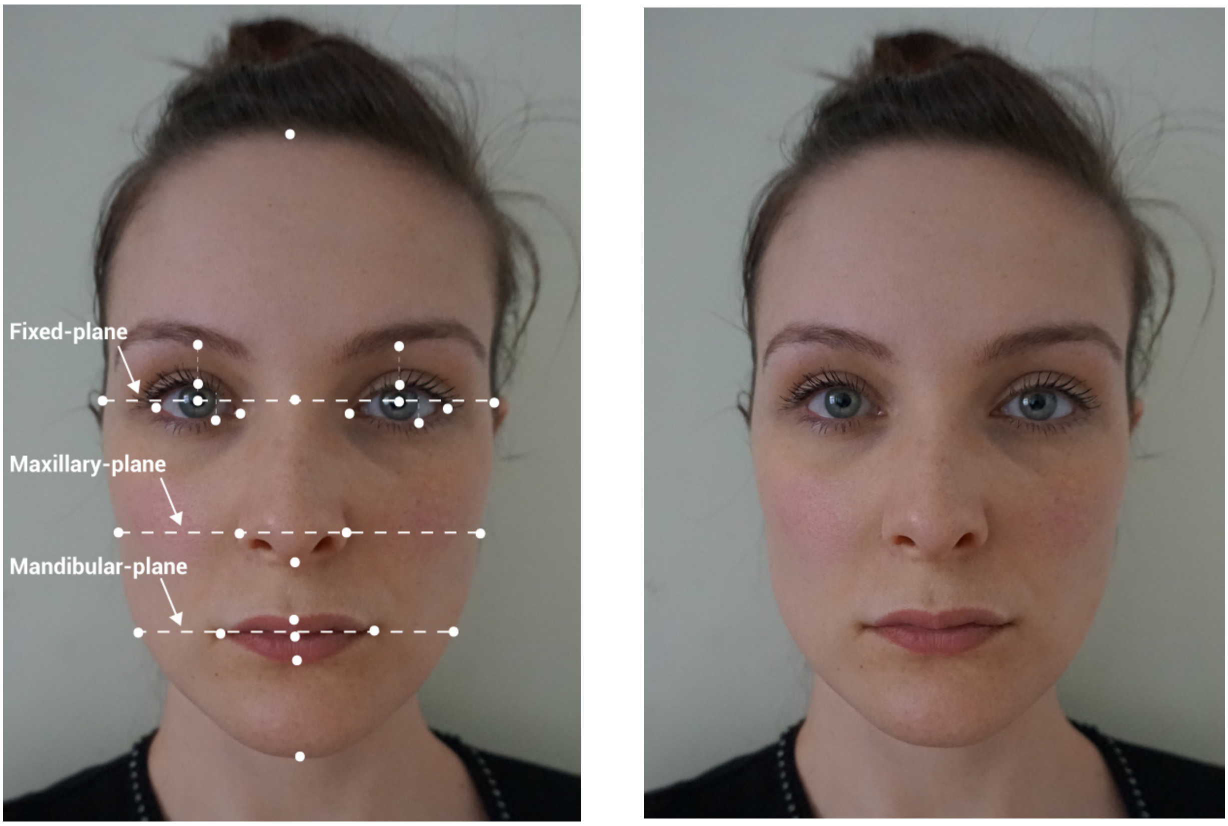


Figure 5.7: The planes defined by the landmarks and the original face

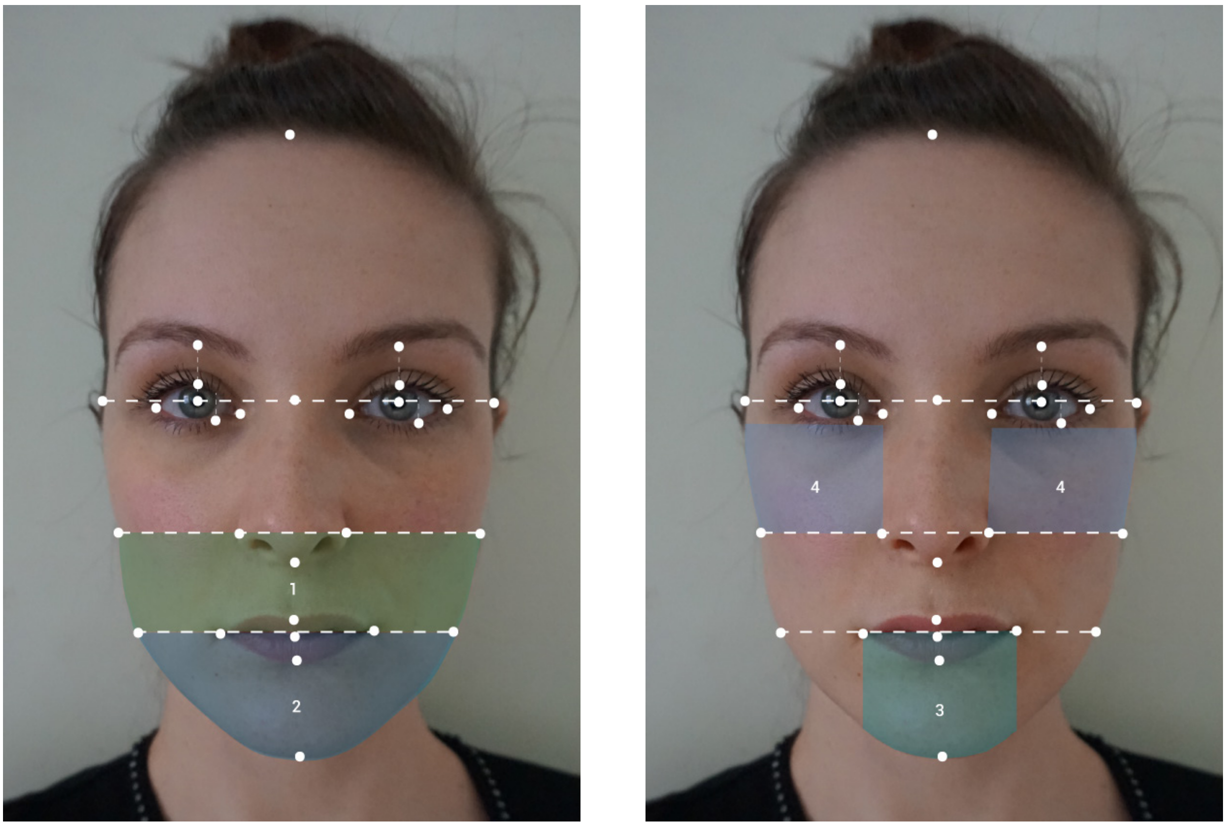


Figure 5.8: The four regions bounded by the landmarks

5.2 Example of using the system

The classification system is used throughout this report and therefore an example of how to use the system is shown to understand how to work with this system.

A patient who underwent multiple surgeries is shown in figure 5.9. This patient had a receding mandible, protruding maxilla and her maxilla was too broad. Therefore she received these type of surgeries according to the Facial Sculpture Clinic (2020) where she was treated:

- Advancement lower jaw (BSSO)
- Setback upper jaw (Le Fort I)
- Chin surgery (Sliding genioplasty)
- Transversal narrowing upper jaw

Instead of calling the type of surgeries like this, the new classification system simplifies the description. The VUmc will classify this patient as a type C patient. After applying the new system, this patient will get the classification: 1abFL, 2aF, 3aL. Where the numbers describe the amount and type of regions affected, the letter describes the movement of the landmarks in this region and the F and L describes in what plane the landmark is moving.

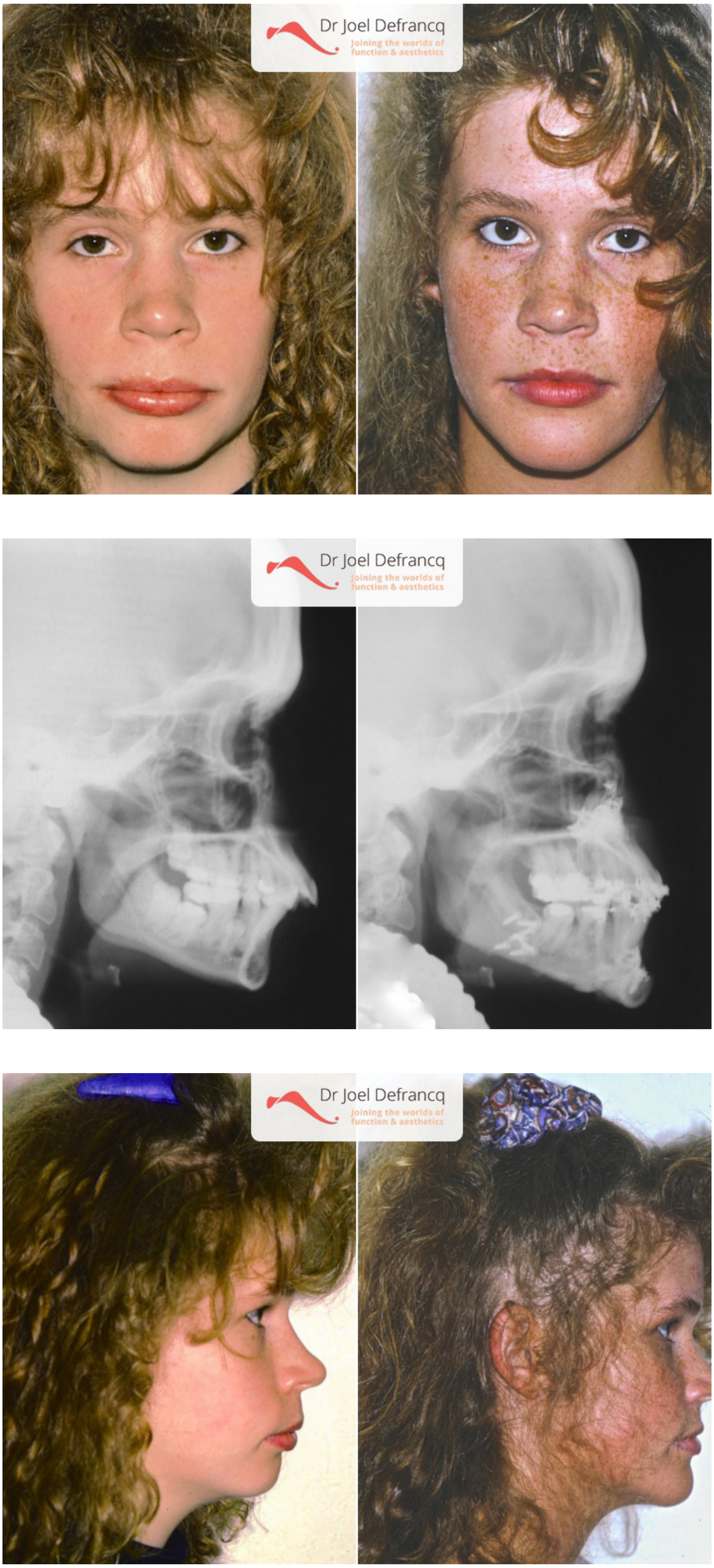


Figure 5.9: A patient who underwent multiple surgeries (Facial Sculpture Clinic , 2020)

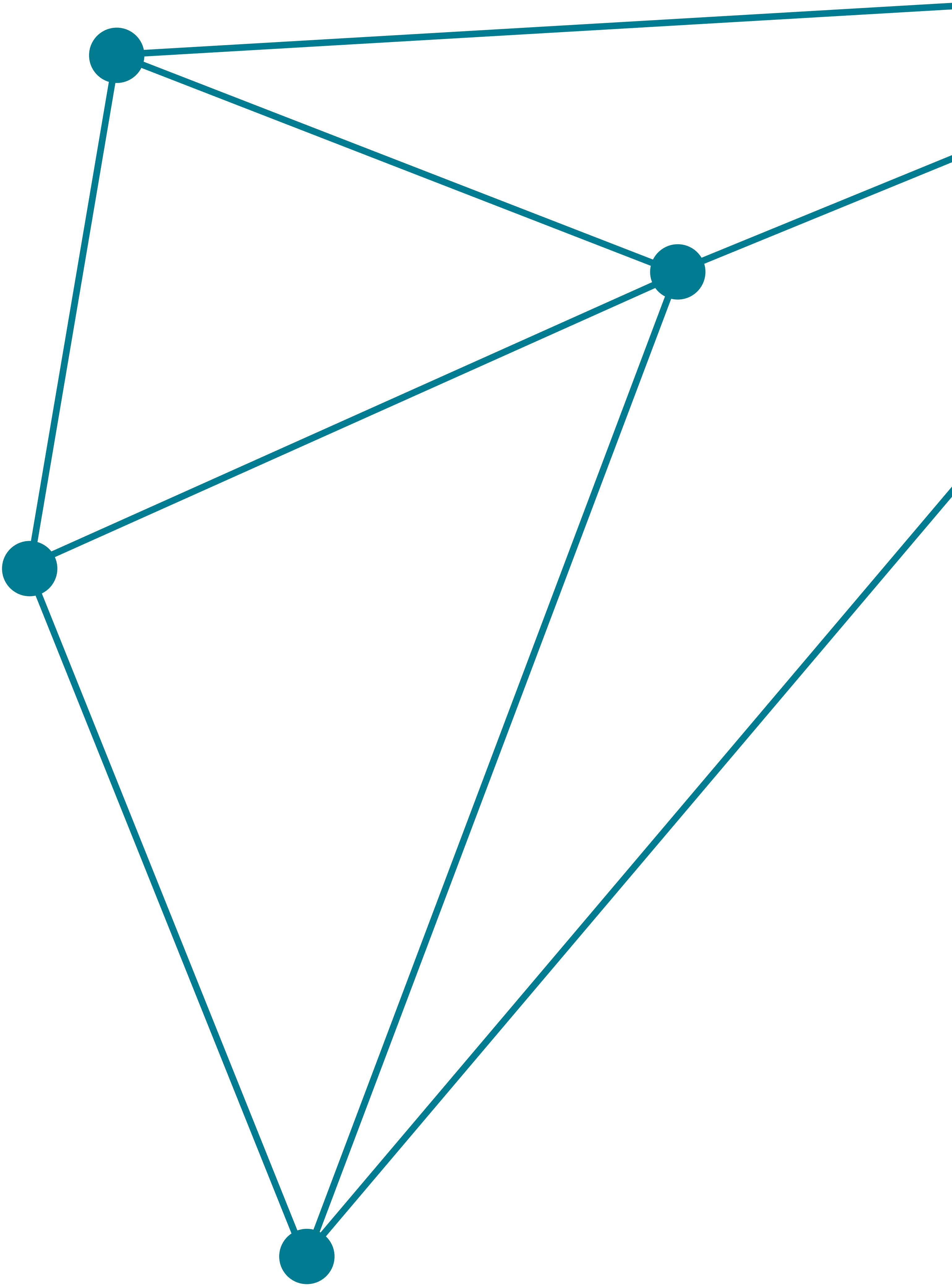
5.3 Conclusion

The VUmc classification is analysed to see what landmarks they use and to get more insight into their treatment planning. Within their patient classifications, different patients exist, needing different surgeries. To understand how landmarks move per surgery, the patients needed to be divided into multiple landmark movement groups. The VUmc classification is therefore not usable for describing landmark movements.

A new classification system has been created (figure 5.2) describing the region wherein the movement of the landmarks is happening, the type of movement and the type of plane where the movement is happening. The new classification system is created after analyzing the classification system in the VUmc, other classification systems in literature, after observing the pre- and post-surgical pictures of patients and analyzing the most common performed surgeries

Key-points from this chapter

- The classification system of the VUmc will not be used for describing the movement of landmarks
- A new classification system has been created to divide patients into groups which describe landmark movements
- The landmark movements happening in the patient's face are described by region, type of movement and on what plane the movement happens



Chapter 6

Human recognition test

The human recognition test is performed to test how single configurations happening in different types of orthognathic surgeries influence facial recognition of familiar faces. Research shows that the way the face is processed depends on the familiarity of the face (chapter 3.2). No research has been done with the acquaintances of a patient after orthognathic surgery. However, consultations in the VUmc do point out problems with facial recognition of the patients faces by the patients' acquaintances. It is unknown to what type of patients this is happening. Therefore, the effect of orthognathic surgery on facial recognition had to be examined with a human recognition test between familiar people. Landmarks in the face can be used to see what type of movement of the soft-tissue influences facial recognition. This test focuses on six different manipulations, three of the mandible and three of the maxilla. Faces of familiar people of a group have been collected, manipulated and tested on facial recognition. This chapter shows the test setup. The details of how the test is performed exactly and the decisions behind the research setup can be found in appendices 9 till 14.

Research objective
How does the movement of the soft-tissue in the maxilla and mandible region on the front plane influence facial recognition?

Research questions
• Does the direction of the landmark movement over the x or y-axis influence recognition?
• Is there a difference between moving landmarks over the x or y-axis?
• Is there a difference between influencing the maxilla or mandible region?

6.1 Method

1 Participants
Pictures from active members within a rugby student society in Delft, the Netherlands, have been collected. The consent form can be found in appendix 13. Pictures from the frontal plane of the face from 15 women and 21 men between the age of 18 and 32 have been manipulated. The participants who participated to the human recognition test were also members from the same rugby student society and existed out of 41 men and 22 women between the age of 18 and 32 from Europe or European culture. 31 participants whos pictures were used also participated in the facial recognition test.

2 Stimuli
The pictures shown to the participants were manipulated to the classification codes: 1.b (widening), 1.c (up), 1.c (down), 2.b. (widening), 2.c (up), 2.c (down) (figure 6.1). The landmarks of Kiekens et al., (2008) have been used (figure 6.3). From the 36 pictures collected, six pictures were manipulated per classification group. The hair in the picture was removed to let the participants focus on the jaw (appendix 12).

3 Apparatus
The test happened online with the use of a Google Form. The type of questions can be seen in appendix 14.

4 Procedure
The link of the online test was sent out to all the members of the rugby student society. A description was shown first to tell they should do the test alone, should not look into pictures on the internet and to not share any pictures with other people until the test was closed. First, a test question was shown so they could prepare for the type of questions. Each participant received the exact same test.

5 Measures
The results were saved automatically in an Excel file by Google Form. How the measures are processed to get to the results is explained step by step in appendix 9.

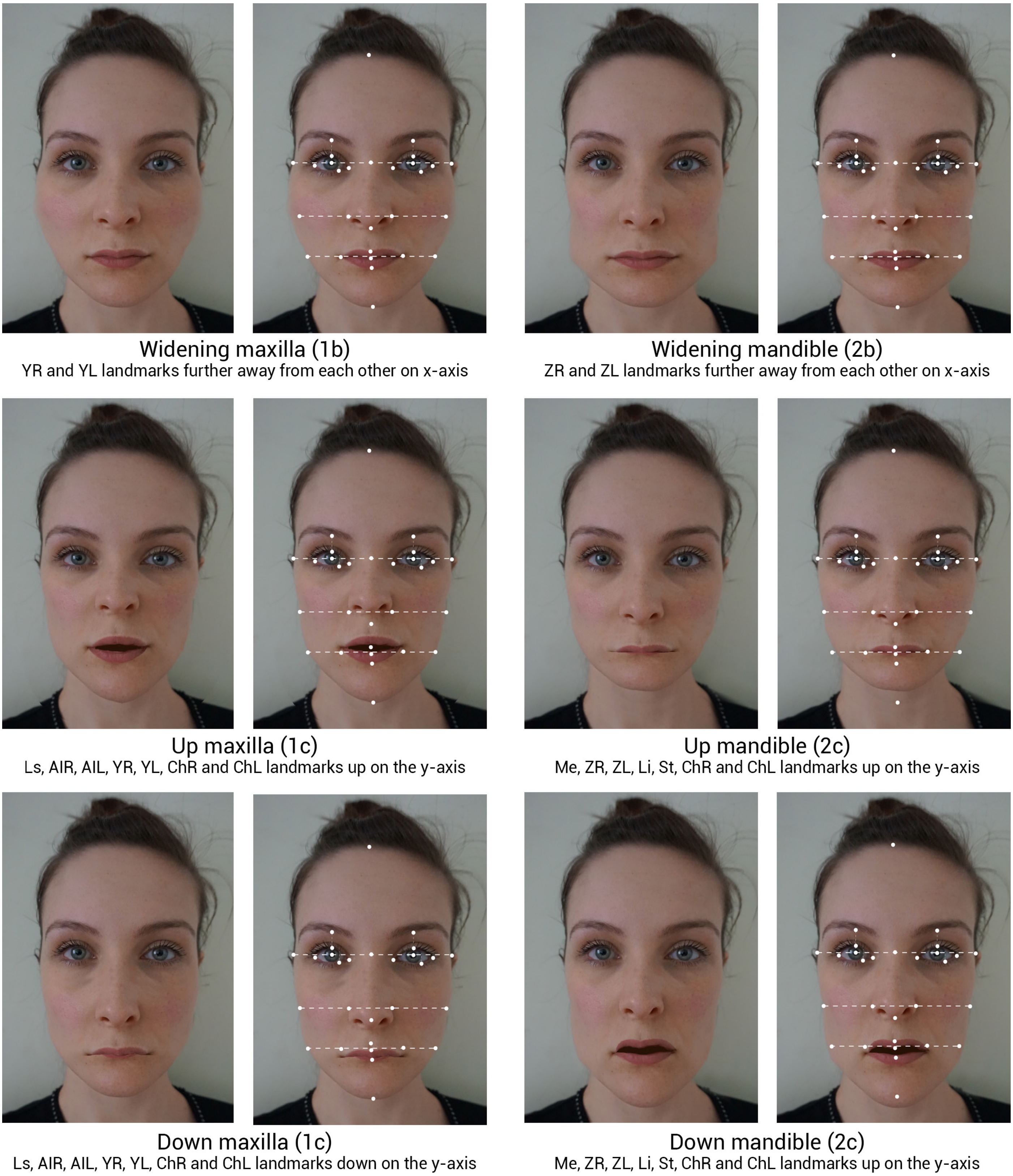


Figure 6.1: Type of stimuli. The names of the landmarks are shown in figure 6.3.

6.2 Results

Figure 6.2 shows the percentages of the respondents who did not recognize the face of the familiar person after manipulation of one of the groups. Apparently, all 6 groups can influence facial recognition. Before this research was expected that the widening manipulations would not have any effect on facial recognition, however, the results of the test point out they do. The up manipulation of the mandible has the biggest effect because 5.25% of the respondents did not recognize the familiar person.

The results from the manipulations on the mandible are all statistically significant. Only the up movement on the maxilla was statistically significant (appendix 9).

6.3 Discussion

The manipulations which are done on the pictures are each based on one type of surgery and therefore these results can say something about whether a patient undergoing one of these surgeries will be recognizable or not after the surgery. However, it happens regularly that patients undergo multiple types of surgeries. The combinations of these types have not been tested in this research. However, these patients are, among other patients, tested in the computer recognition test (chapter 7).

Before the test, there was expected that the widening manipulations would not have an effect on the facial recognition at all and that the up/down manipulations would have a little effect of ca. 1%. However, all manipulations have an influence and at least 1% of the respondents did not recognize each of the groups (figure 6.2). Positioning the mandible and maxilla up have the highest influence on facial recognition, which means moving the landmarks up on the y-axis. Positioning the mandible up has the highest influence on facial recognition, 5.25% of the respondents did not recognize the faces after this manipulation. The manipulations on the mandible do have a higher effect on facial recognition than on the maxilla. Since all these manipulations have an influence on facial recognition, it is expected that a combination of these manipulations will lead to even less facial recognition. Not all types of landmark movements caused by orthognathic surgeries have been tested. And therefore all of the movements should be tested like this one by one in future research to extract the type of movement from the complexity of combined surgeries.
With the up and down manipulations was decided on a correct functional position of the maxilla and mandible on the lateral plane, so the maxilla is positioned more forwards than the mandible. This resulted in other manipulations. When moving the mandible up, (part of) the bottom lip disappeared under the lip of the maxilla and when moving the maxilla down, the upper lip would fall over the bottom lip as well. However, patients where the maxilla or mandible are not positioned correctly in the lateral plane exist and therefore these influence the frontal view. Future research should be done to test how this influences facial recognition.

In this research, the pictures which were collected online differed in quality, scale, illumination and head position. This could have influenced the results. However, by using these pictures they lied closer to reality than studio pictures would be (like in figure 6.1) since the scale, illumination and head position is never the same when seeing people face-to-face.

The difference between showing teeth or not in the pictures of the faces has not been taken into account, although no differences between groups with more or fewer teeth can be seen, this still could have influenced the results. Only manipulations on the jaws and the surrounded soft-tissue has been done and nothing has been done on the teeth itself. However, when a jaw is widened, it is expected that the teeth will change positions and that a gap will appear between the two teeth on the front, between teeth 8 and 9 on the maxilla or 24 and 25 on the mandible (appendix 7). This could influence facial recognition as well, although the gap between the teeth would be corrected by braces.

There has been chosen to calculate the average of the number of people who did not recognize the original picture and for them, the face was therefore not familiar. This was 4,8% of the total amount of respondents.

This amount of respondents was removed from the total amount (65 respondents), and therefore the total amount of respondents for whom every face was familiar was used for the calculation (63 respondents). However, the number of unfamiliar respondents differed per face. Therefore, not every face had the same amount of respondents. This has not been taken into account in the final results.

It was decided to not test the narrowing of the maxilla and mandible because it was expected that these would not influence facial recognition. However, to be sure about this, this human test could be performed again by taking the narrowing manipulation into account as well.

In this test was chosen to remove the hair to let the respondents focus on the jawline instead of the hair. However, the hair does play a role in facial recognition, since facial recognition is a holistic process. In the computer recognition test in chapter 7, the hair will be taken into account. However, to test how the hair influences facial recognition among human exactly, this type of human test could be redone in further research while including the hair.

What also could have influenced the results is the strength of the familiarity, how well the respondents knew the person in the picture. In this test was only tested if the face was familiar but not how familiar. It is expected that when you see your partner, whom you see every day and when you see a neighbour, whom you see once a month, you can recognize your partner faster, also when some landmarks are different. However, this is not known and there is no research done on this topic yet. Future research should point out the difference in familiarity and the effect on landmarks.

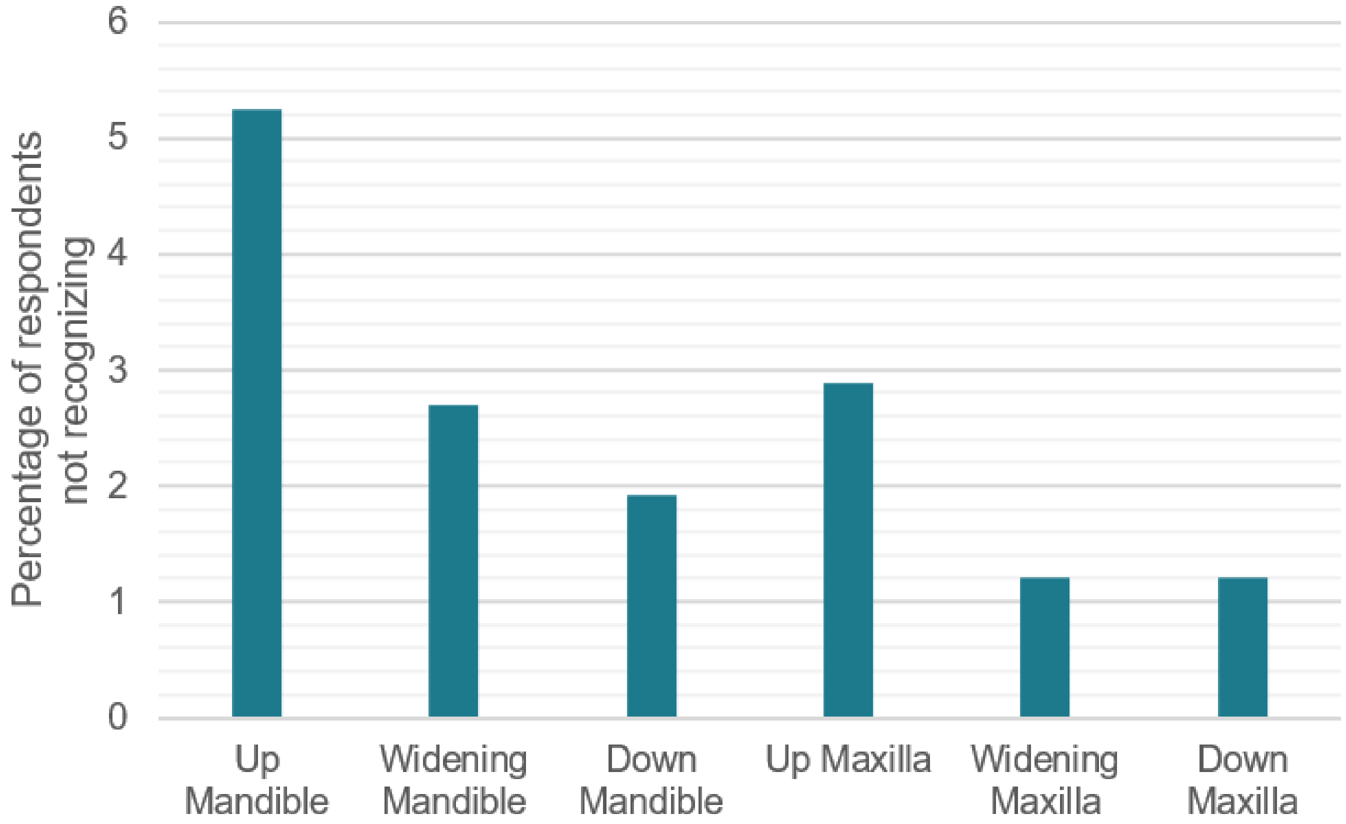


Figure 6.2: The percentage of respondents who did not recognize the group.

6.4 Conclusion

Moving the mandible or maxilla up or down on the y-axis in the frontal plane influences facial recognition. Also widening the maxilla by positioning the YR-YL landmarks (figure 6.3) further away from each other on the x-axis and by widening the mandible by positioning the ZR-ZL landmarks further away from each other influences facial recognition. The manipulations on the mandible were less recognized than the ones on the maxilla. Moving the mandible upwards, which corresponds to vertical shortening in the classification system, was the least recognized, 5,25% of the respondents did not recognize this group.

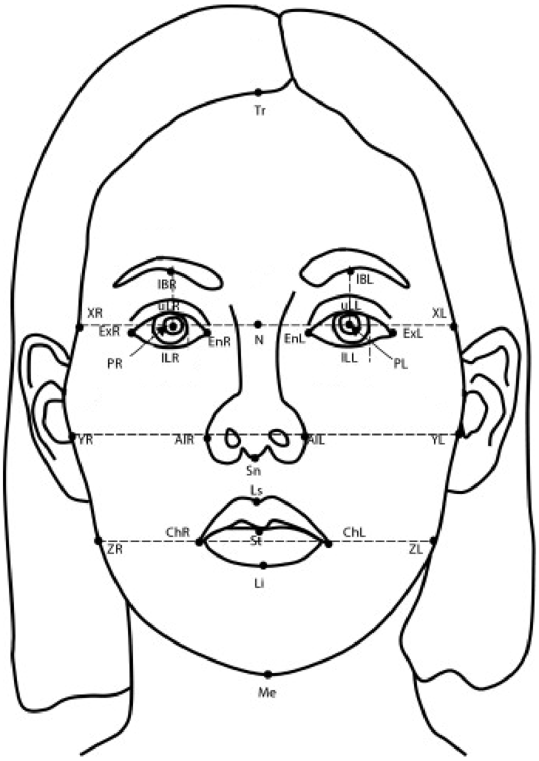
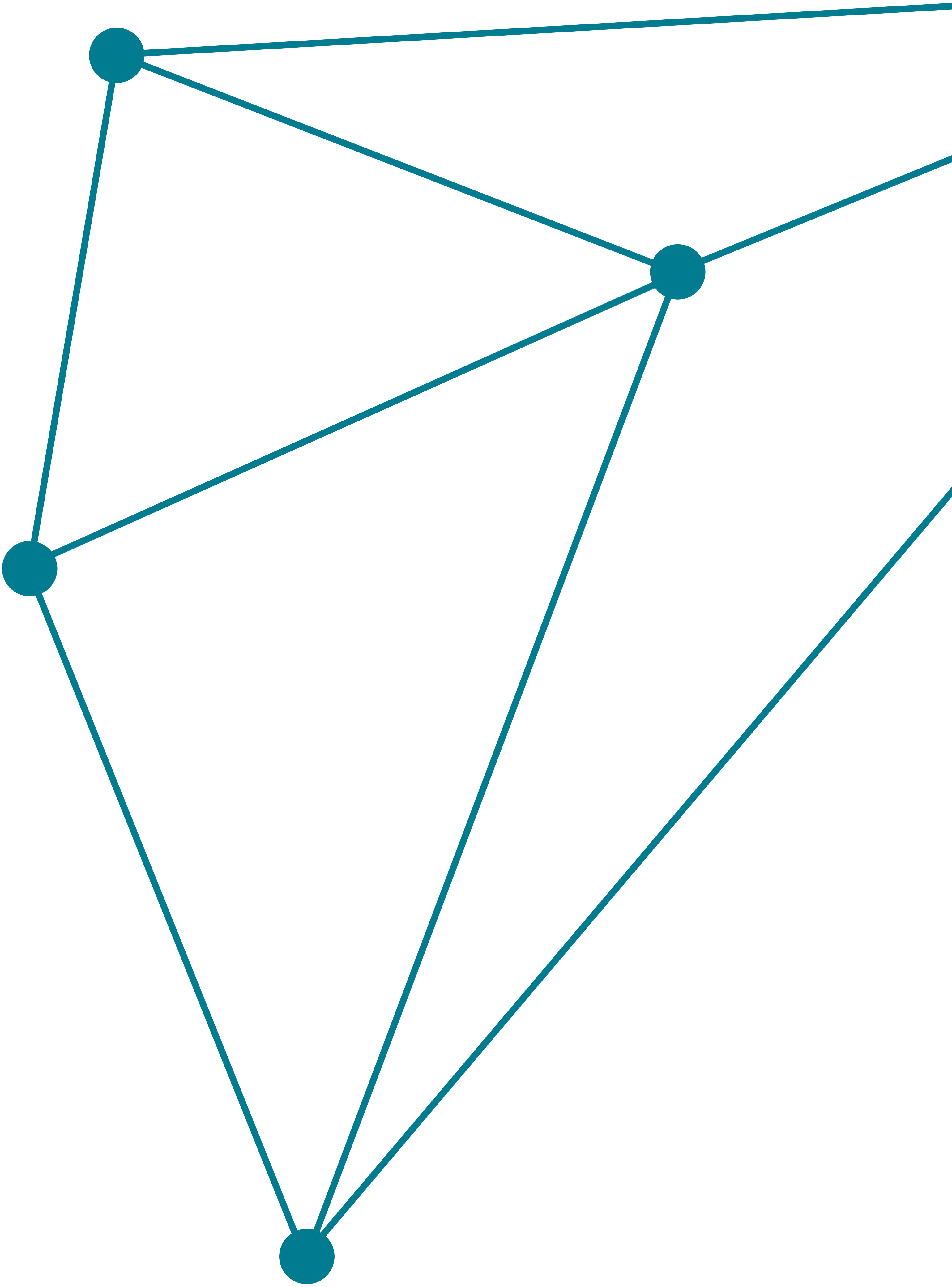


Figure 6.3: The frontal landmarks (Kiekens et al., 2008).

Key-points from this chapter

- All the tested manipulations have an influence on facial recognition among familiar people
- Moving the mandible up (2c, vertical shortening) has the highest influence on facial recognition. After this comes moving the maxilla up (1c)
- The manipulations on the mandible have the highest influence on facial recognition



Chapter 7

Computer recognition test

Multiple landmark movements are difficult to test with humans, since extracting single landmark movements out of patient data is barely possible because the soft-tissue movement is difficult to predict. Also, to be able to compare patients from different classifications with each other, more data is needed. The more data, the more time consuming the human recognition test would be. The computer recognition test is done to analyze complex pictures and a bigger amount of data than possible with a human recognition test. The similarity of pre- and post-surgical orthognathic patient pictures has been compared between patient classes based on regions and landmark movements. The similarity is represented as a score. The meaning of this similarity score and how the method works is described in chapter 3.3. The results of this test show the relations between the amount of movements and the amount of regions and the similarity score. It also shows what regions and what type of landmark movements do have the most influence on facial recognition and it shows an analysis of the region combinations and movement combinations. The next chapter, chapter 8, will continue on this test by comparing it to the human recognition test and by calculating the human recognition threshold. This chapter shows the test setup and the results of the test have been summarized, addition figures of the results can be found in appendix 16. The details of how the test is performed can be found in appendix 17 and 18.

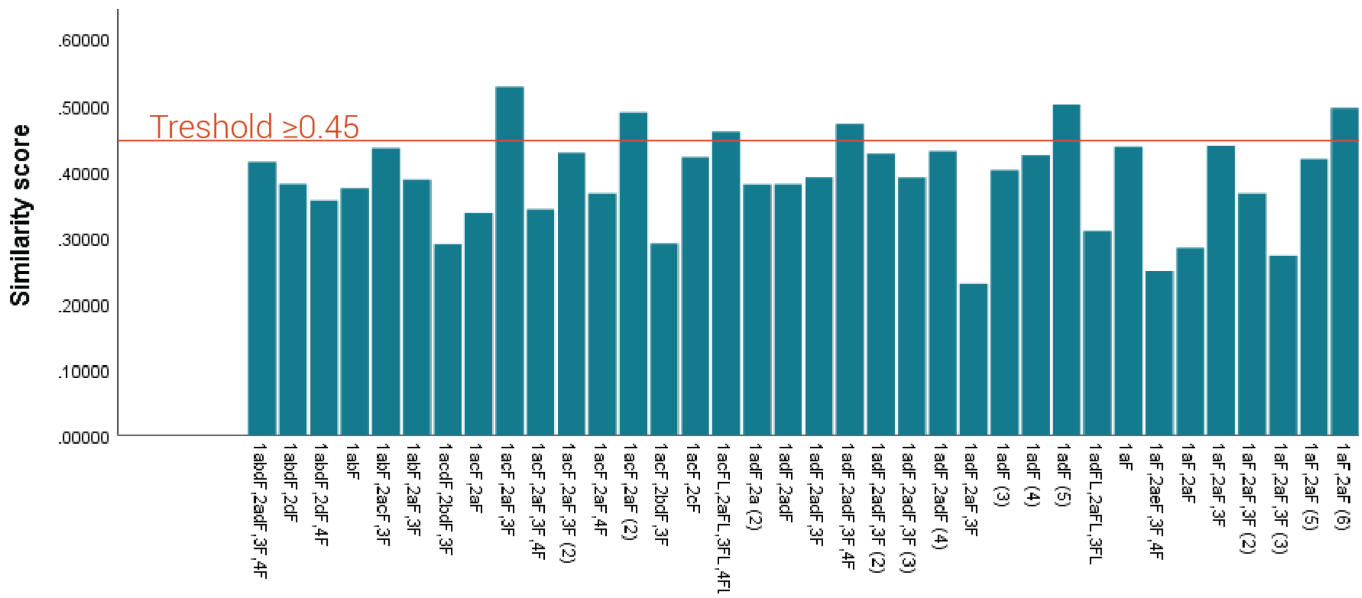
Research objective
What is the link between the regions and the type(s) and amount of landmark movements caused by orthognathic surgeries and with the percentage of unrecognizable patients?

Research questions
- Region 1 or 2 is always affected in data. Does region 1 or region 2 have the most influence on facial recognition?
- What effects do region 3 and 4 have on facial recognition?
- How does the amount of regions affect facial recognition?
- How does the amount of movements affect facial recognition?
- What landmark movements on which regions do affect facial recognition?
- What landmark combinations on which regions do affect facial recognition?

Hypothesis
The more regions and landmark movements are involved the higher the percentage of unrecognizable patients is.

This hypothesis is based on the human recognition test. This test covered only single regions and movements which already influenced facial recognition. It is expected when multiple surgeries in multiple regions are performed, multiple landmarks move and therefore facial recognition will be lower. The more landmarks move, the more the Euclidian distance of the post-surgical face will differ from the pre-surgical face, which results in lower facial recognition.

Figure 7.1: All patient pictures rated on similarity score by the computer



7.1 Method

1 Stimuli
75 patient pictures of both the pre-surgical and post-surgical state in frontal or frontal-lateral view have been collected (this also explains the F or FL classification. In this case, the F and FL only show on what view the picture is taken. It is not of importance in this analyses and has only been used to structure the data). From these pictures, the type of surgery is known and therefore an estimation of the landmark movements are known. With this, the patients are classified with using the classification system created in chapter 5, shown in figure 7.5. All patient pictures with classification can be found in appendix 19.

2 Apparatus
The test has been done in the Spyder software, with the code (appendix 15) written in the programming language of Python. The dlib and face_recognition libraries have been used (chapter 3.3). The results of Spyder are copied and processed in Excel and SPSS

3 Procedure
First, all the data was converted the same file type (jpg), this way the code is not skipping any files. Then, sets of the matching pre and post-surgical pictures were made which were put in one map together. This was done to verify if all the sets were complete and to be sure the code would compare the two pictures of the set. All the sets were manually classified with use of the classification system created in chapter 5.2.

4 Measures
The similarity scores with the corresponding file names were printed in Spyder whereafter it was copied to Excel and SPSS. How the measures are processed is explained step by step in appendix 17 and 18.

7.2 Summary of the results

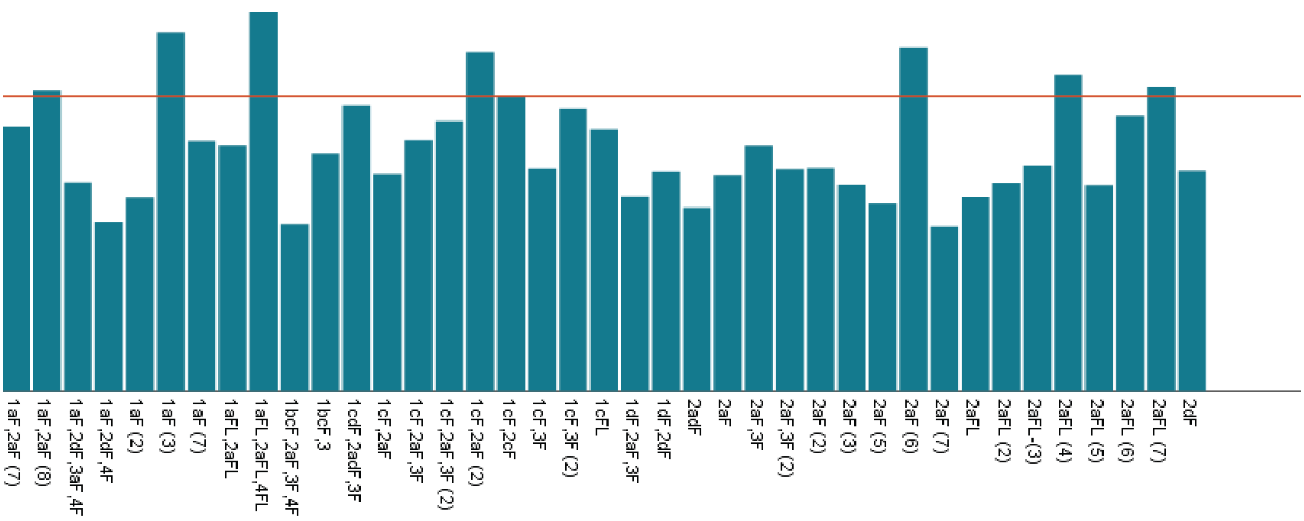
7.2.1 General results
The patients are unrecognizable by the computer when the similarity score is ≥0.45. 19% Of the total of 75 patients was unrecognizable after the orthognathic surgery by the computer. Figure 7.1 shows what patients are exceeding the 0.45 threshold. The patients who are unrecognizable are listed in figure 7.5.

7.2.2 Results regions
The number of regions
The scatter plot of figure 7.3 shows how the similarity scores are distributed per amount of region. Combining region 1 and 2 has the highest percentage of unrecognizable patients, together with combining all regions. This goes against expectations because the hypotheses expected that more regions would lead to more patients being unrecognizable. The four amount of regions has the highest percentage of patients who are unrecognizable, which meets the hypothesis. Remarkable is that the amount of three regions is the lowest of all, which means the combination of regions must have played a role here and the hypothesis is not meeting up.

Type of regions
A bar graph has been made (figure 7.4) which shows all the region combinations present in the database. Region 2 has a higher rate of unrecognizable patients over region 1, which matches the results from the human recognition test. The bar graphs show that the combination of regions plays a role since 1+2+3 score 6% and 1+2+4 score 25% of patients who were unrecognizable. Apparently, adding region 3 can bring the similarity score down and the landmark movement combinations do play a role as well. Region 4 does bring the similarity score down as well, but not as much as region 3, looking at the percentages.

The reason for this phenomenon is unknown. However, an explanation of this could be that these regions are 'correcting' the landmarks which are moved by regions 1 and/or 2, by moving them back to the original position which leads to less distance between the pre- and post-surgical landmarks, resulting in a lower similarity score. The combination of 3+4 is not affecting the percentage of unrecognizable patients.

7.2.3 Results movements
The number of movements
The hypothesis state the more movements, the higher the unrecognizable patient percentage would be. The percentage of unrecognizable patients is not increasing linearly when the amount of movements goes up, which does not meet the hypothesis (figure 7.6). There is not much to be concluded about the amount of movements, since no relation to the amount of movements and the percentage of unrecognizable patients can be found. Nevertheless, the percentage of unrecognizable patients can still rely on the type of movements. This is described next.



classifications of patients

Region		Movement		Plane	
1	Maxilla	a	Horizontal Lengthening/Shortening	F	Frontal
2	Mandible	b	Widening/Narrowing	L	Lateral
3	Chin	c	Vertical Lengthening/Shortening		
4	Cheek	d	Rotating		
		e	Smoothening jaw line		

Figure 7.2: A recap of the legend of classification annotations

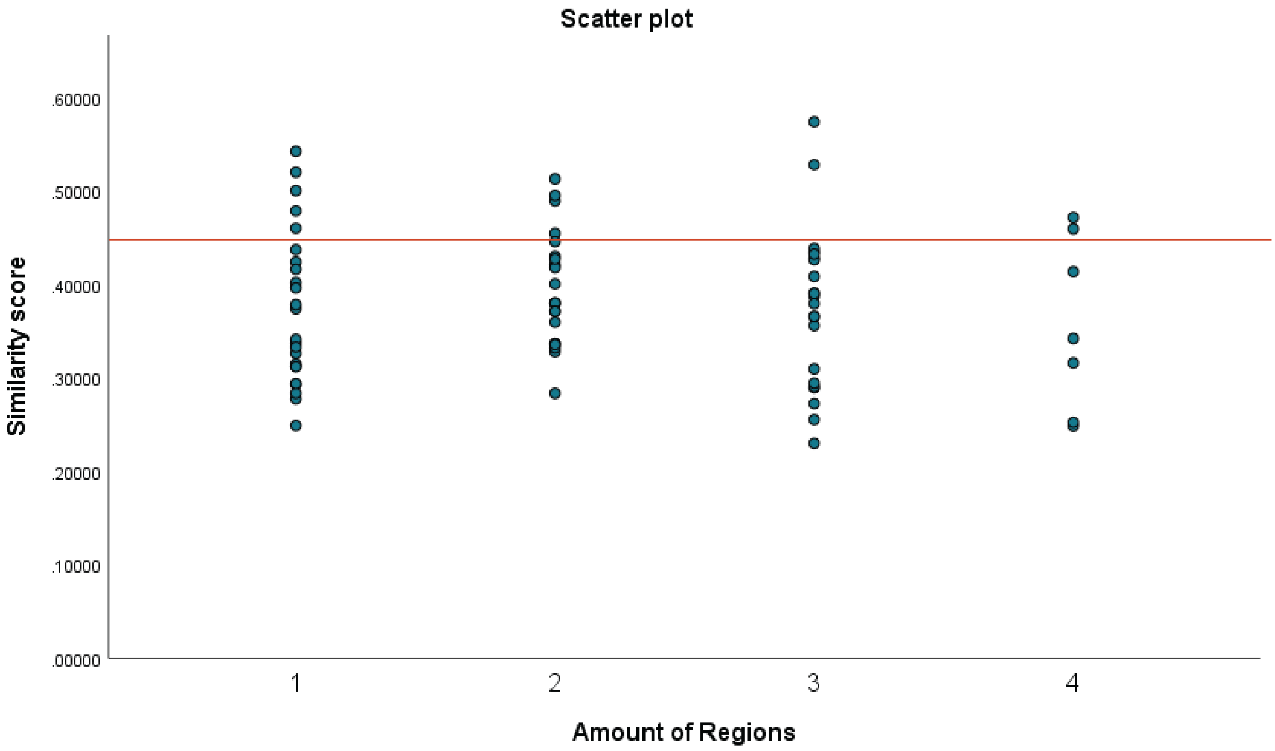


Figure 7.3: A scatter plot of the similarity scores per amount of regions

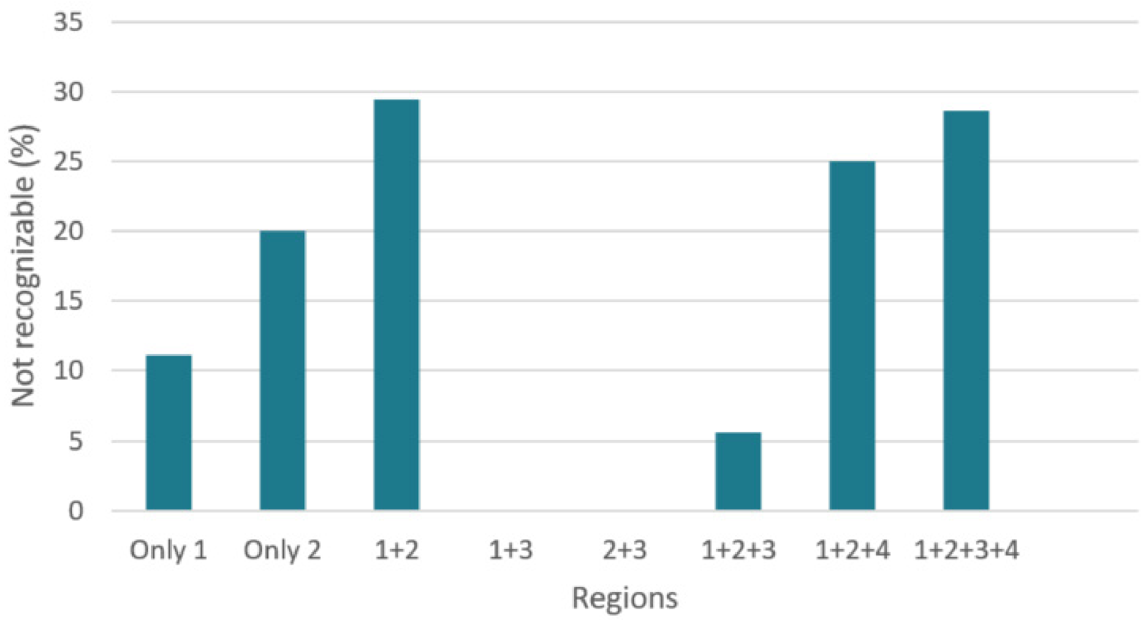


Figure 7.4: A bar graph of all region combinations and the percentage not recognizable patients in these combinations

The type of movements

Figure 7.5 shows the combinations of region and landmark movements which were unrecognizable by the computer. What can be seen, is that the 'a' movement, which is horizontal shortening/ lengthening (figure 7.2), is present in almost every classification. Therefore, the horizontal shortening/ lengthening landmark movements in both the mandible and maxilla affect facial recognition. Also, 1a and 2a can, as the only single movement, cause an unrecognizable face after surgery. However, not every patient undergoing a 1a or 2a involved surgery was unrecognizable, therefore, the results are not linear and the results still depend on the patients face.

Since most of the patients have different combinations of regions and landmark movements, a scatter plot and bar graph of the data has been made as well, to be able to see the differences in movements. Figure 7.8 shows the scatter plot of the similarity of all movements of region 1 and 2 which are involved. Right away can be seen that 1a and 2a are represented in the database the most and that they both have a patient with the highest similarity score. Figure 7.7 shows the percentage of unrecognizable patients against the involvement of movements. Here can be seen that all the movements seem to be independent of its regions since 1a and 2a have a close result, so do 1c and 2c and 1d and 2d. The c movements do have the highest percentage of unrecognizable patients, which is the vertical lengthening/shortening. This matches the outcomes of the human recognition test. Hereafter, comes 1a and 2a, the horizontal lengthening/shortening. Apparently, the lengthening and shortening surgeries have the most influence on facial recognition and they bring the similarity score up. The 1d and 2d movements, were unrecognizable for respectively 10% and 5%, which means rotating the maxilla or mandible could increase the similarity score as well.

Patient	Similarity score
1acF,2aF (2)	0.48924
1acF,2aF,3F	0.52795
1acFL,2aFL,3FL,4FL	0.45933
1adF (5)	0.50019
1adF,2adF,3F,4F	0.47151
1aF (3)	0.54234
1aF,2aF (6)	0.49514
1aF,2aF (8)	0.45459
1aFL,2aFL,4FL	0.57398
1cF,2aF (2)	0.51280
1cF,2cF	0.44573
2aF (6)	0.52001
2aFL (4)	0.47861
2aFL (7)	0.46005

Figure 7.5: All patients which are not recognized by the computer

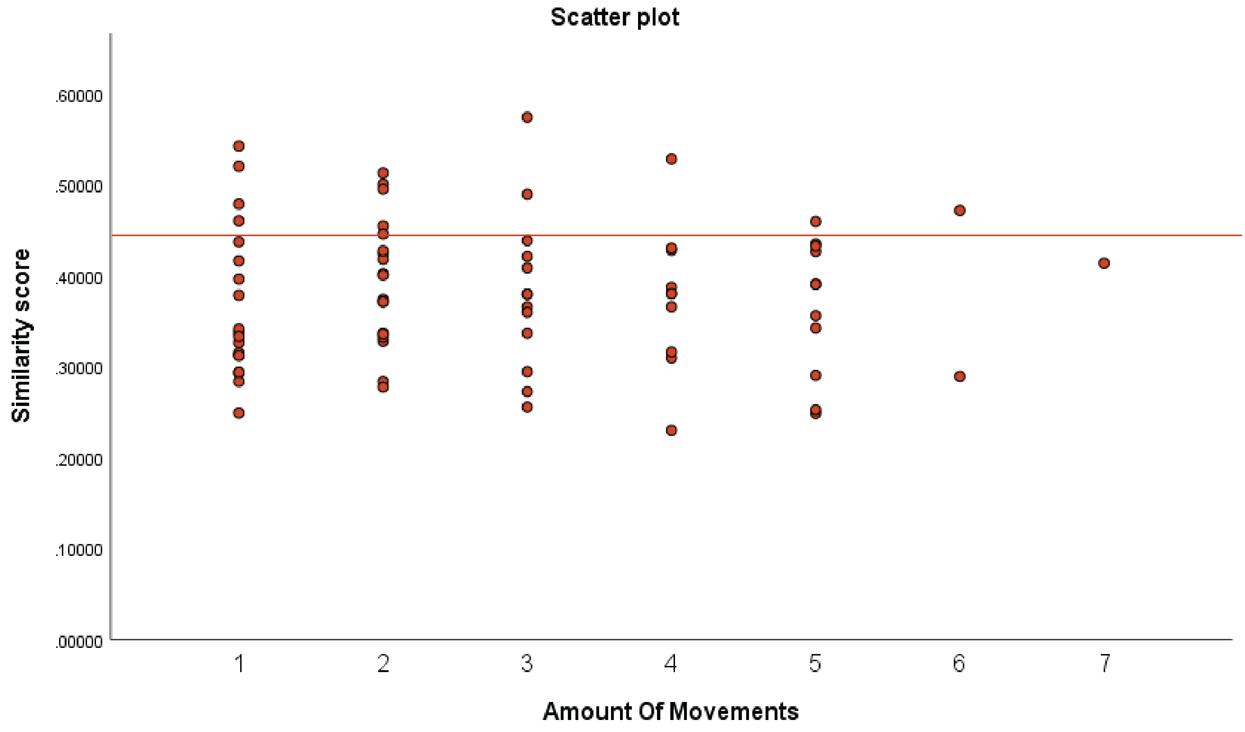


Figure 7.6: A scatter plot of the similarity scores per amount of movements

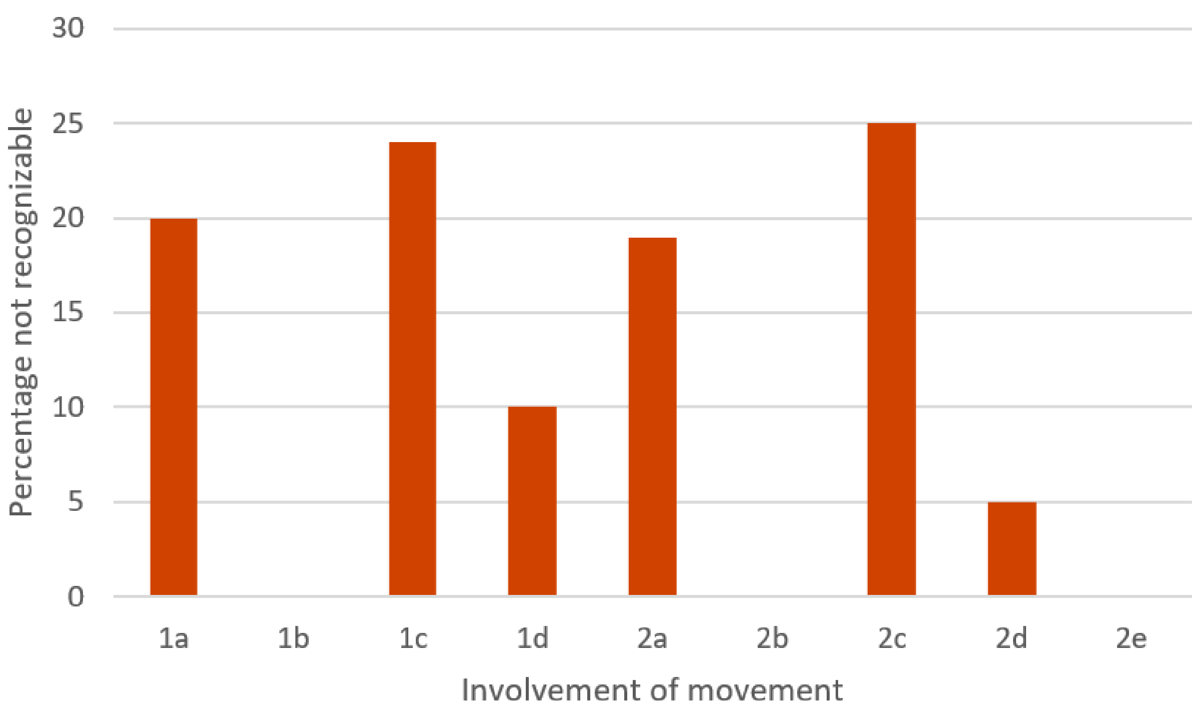


Figure 7.7: A bar graph of the movements which are involved in the patient's surgery and the percentage not recognizable patients belonging to this involvement

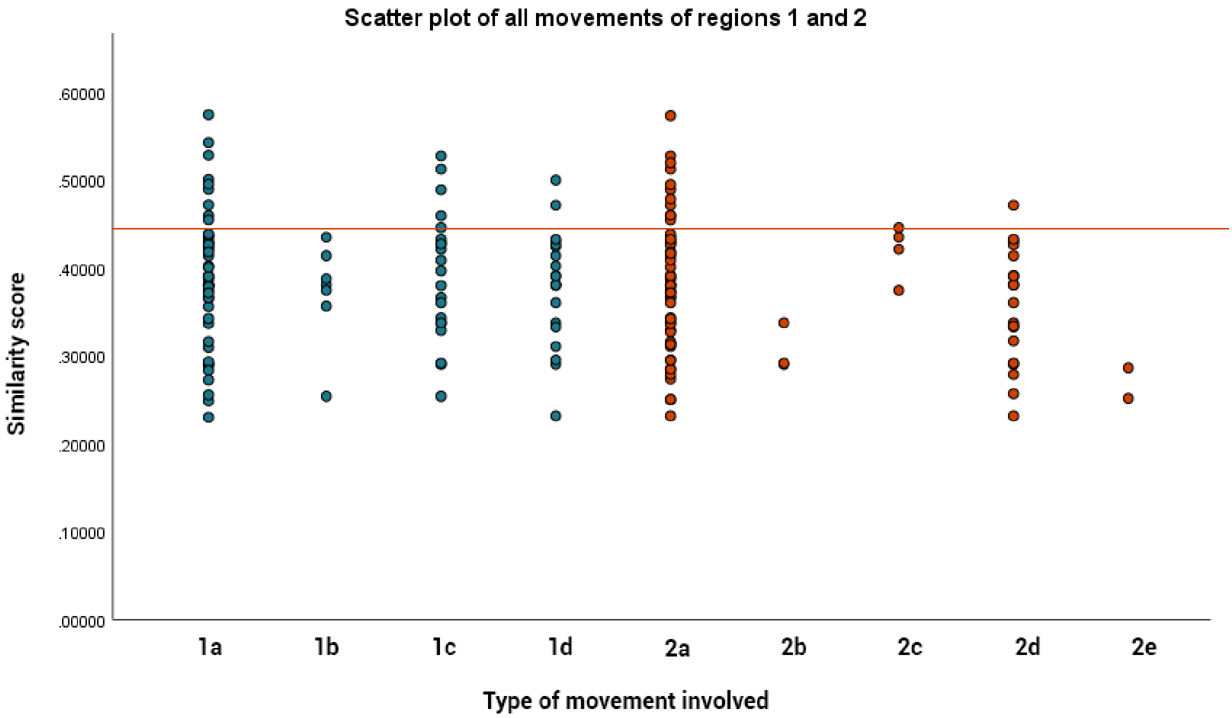


Figure 7.8: Scatter plot of type of movements

7.3 Discussion

All the movements within a, b and c have been seen as one movement since the landmarks move over the same axis within these movements. However, the landmarks within these movements do not move the same because they can still be distinguished between shortening/lengthening and widening/narrowing. However, these have not been taken into account in this research and therefore there could be differences between facial recognition of shortening and lengthening and widening and narrowing (covered in chapter 12.1). Movement d can rotate around different axis'. Because of the complexity, the type of rotation has not been taken into account.

Also, how big the dimensions of the landmark movements of these surgeries are, have not been taken into account (covered in chapter 12.1).

It is not known if the way the surgeries are performed on the patients in the data is the same as the way the VUmc performs. Therefore it is recommended to the VUmc to perform this computer test on their own patients.

The age difference between pre- and post-surgical patient pictures was unknown but it could have influenced facial recognition. However, the difference would be estimated between 2 and 4 years and since the test was only about the geometry of the face and not skin textures, it is expected that this would only have a little influence on the results.

The classes 1a and 2a can, as only single movements, cause an unrecognizable face after surgery. However, the results are not linear and the results still depend on the patients face. However, some regions and landmark movements can be seen as a risk for becoming unrecognizable.

The analysis has been done on only 75 patients with a variety of orthognathic surgeries. Therefore, not all regions and movements are equally present. Also, multiple combinations of regions and movements are present in the data, which makes it hard to find correlations. However, this wide variety in the classifications of the data also shows that most orthognathic surgeries are different and patients cannot be treated the same.

7.4 Conclusion

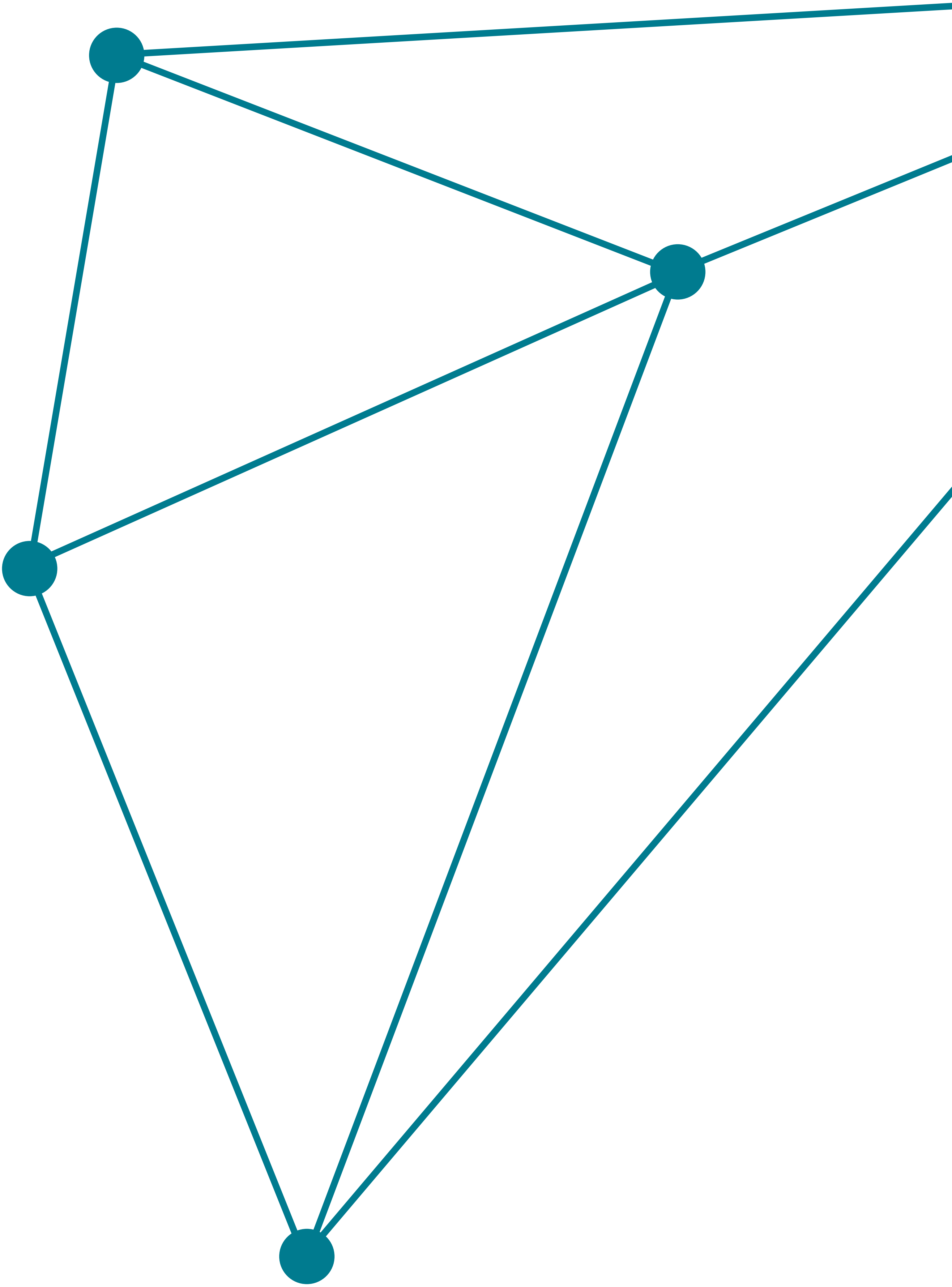
More regions or movements does not automatically mean it increases the similarity score which decreases facial recognition. Combining region 1 and 2 has the highest percentage, not recognizable patients, together with combining all four regions whereas only adding region 3 or 4 upon region 1 and/or 2 can bring the similarity score down. However, region 4 (cheek) does not bring the similarity score down as much as region 3 (chin). 1a and 2a can, as the only single movement, cause an unrecognizable face after surgery. However, this does not happen on each patient and the recognizability still depend on the patient's identical face. The c movements do have the highest percentage of not recognizable patients, which is the vertical lengthening/shortening. After comes the a movements, which is the horizontal lengthening/shortening. Therefore, the lengthening and shortening surgeries have the most influence on facial recognition and they bring the similarity score up. The 1d and 2d movements, also influence facial recognition, which means rotating the maxilla or mandible could increase the similarity score as well.

Therefore, for the movements of landmarks, these are the movements which can influence facial recognition in an orthognathic surgery on the mandible and/or maxilla with the consecution of unrecognizablility:

1. Vertical Lengthening/Shortening
2. Horizontal Lengthening/Shortening
3. Rotating

Key-points from this chapter

- Region 2 has a higher effect on facial recognition than region 1
- Adding region 3 can bring the similarity score down
- Region 4 does bring the similarity score down as well, but not as much as region 3
- The combination of 3 +4 is not affecting the percentage of not recognizable patients
- Combining region 1 and 2 has the highest percentage unrecognizable patients, together with combining all regions
- 1a and 2a can, as the only single movement, cause an unrecognizable face after surgery
- The c movements do have the highest percentage of not recognizable patients, which is the vertical lengthening/shortening. After come the a movements, which is the horizontal lengthening/shortening
- The d movements could increase the similarity score, which is the rotating movement



Chapter 8

Comparing the two tests

The database which is used in the computer test is already trained on human perception. However, to determine the similarity of a face with the computer, the face is processed in a different way than humans do (chapter 3) which leads to a higher accuracy than human recognition. The accuracy of the computer test which is done is 99.38% while human recognition has an accuracy of 95.2% (human recognition test) in a group of familiar people. Therefore, human perception is included in the computer test, but since face processing is done differently from humans, the accuracy is increased in comparison to humans. Therefore, the similarity scores from computer recognition and human recognition are not the same. Also, the pictures in the human test have been manually manipulated according to the classification system created in this project (chapter 5). Therefore, it is important to see if the outcomes of the test matches the computer test, which is described in this chapter. Also, a human-computer ratio has been found to convert the similarity scores of the computer test to human perception scores. Hereafter the similarity threshold of human recognition is found.

8.1 Comparisions and ratios

8.1.1 Results of human and computer test

The human test only tested landmark movements b and c and regions 1 and 2, therefore, only these could be compared with the computer test. Both the computer-as human test show that movement c has the highest influence on facial recognition. This means that the faces in the human test seem to be manipulated correctly, meeting the outcomes of real orthognathic surgeries.In the human test, movement b showed an influence on facial recognition whereas the computer test did not at all. Both of the tests show that the mandible has a higher effect on facial recognition than the maxilla.

8.1.2 Human-computer ratio

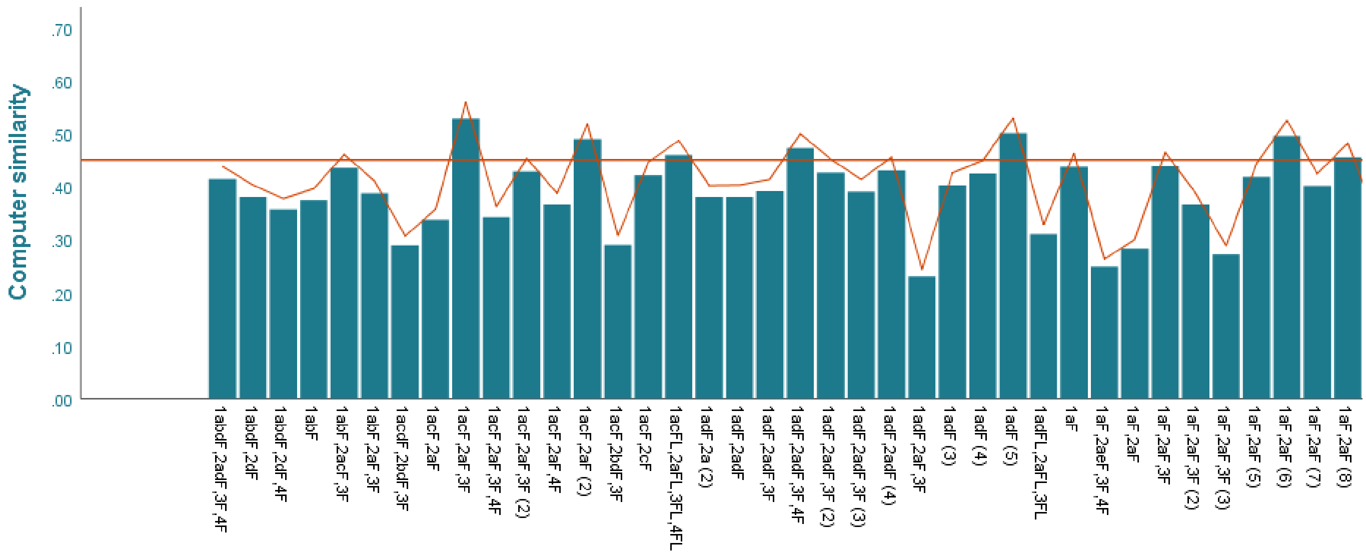
A human-computer ratio is determined to convert the similarity scores of computer recognition to human recognition. Two ratios can be found, one based on the human recognition accuracy of 98% by Geitney (2018) or based on the human recognition test. The human recognition test ratio is more trustworthy since it is based on familiar face recognition and it is done with a variety of people, which therefore includes the inconsistency of the human brains. Mark that an exact human-computer ratio does not exists since humans are not consistent in recognizing faces while computers are (Bindemann et al., 2012), the ratio differs per human and depends on the amount of familarity. However, this ratio can be used as a guideline, to get a feeling about the similarity differences between computers and humans. When further research in the influence of the amount of familiarity is done, the threshold can always be optimized.

In the human recognition test, 4.8% of the respondents did not recognize the familiar face. Therefore, the average accuracy of the people in this test was 95.2%, which is lower than the accuracy of Geitney (2018). The ratio is calculated by dividing the computer accuracy by the human accuracy. This ratio is, therefore, 99.38/95.2= 1.044. A human-computer ratio is determined to convert the similarity scores of computer recognition to human recognition. When using the ratio by multiplying the similarity scores generated by the computer test, the percentage of not recognizable faces increased from 19% to 29%. Figure 8.1 shows the difference in the similarity score of the computer and of humans when the human-computer ratio of 1.044 is applied.

8.1.3 Human recognition threshold

The similarity scores of the computer which lie close to the threshold line, pass the line when the human similarity scores are applied. Table 8.2 shows all the patients who are not recognizable by humans, the green marked scores are added upon the patients which were not recognizable by the computer. What could be seen is that most added patient classes already existed in the list or differ in the number of regions, but five new region/ movement combinations are added as well. The

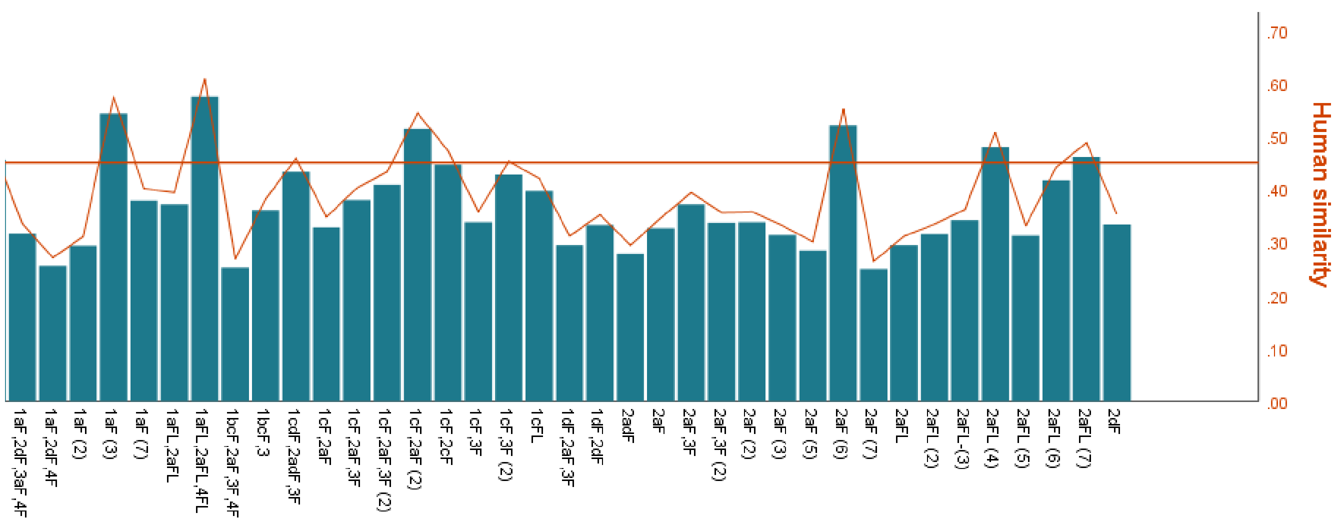
Figure 8.1: the similarity scores of the computer (green block graph) and humans (red thin line) and the thresholdline (red thick line) in one graph



threshold line for human recognition including the human-computer ratio has been calculated (figure 8.3) which is at the similarity score of 0.43. This is the threshold line which corresponds to the specific group tested within the human recognition test. When the same test is performed in another group with familiar people, this threshold line may be different since it depends on the individuals in the group and the amount of familiarity.

Figure 8.4 shows three zones, based on the computer and human threshold. Everything above the computer threshold will most likely not be recognized by humans. Everything between the computer and human threshold will be recognized by the computer but not by part of the humans, when the familiarity is comparable to the familiarity of the tested group. Everything beneath the human treshold is recognizable by most humans.
Figure 8.5 shows the specific human thresholds per tested region among familiar people in chapter 6. These thresholds could be expanded in the future when more regions are tested among familiar people. This could give more insights about certain surgerys to surgeons. With this information, they may change their surgery techniques on certain surgeries which are influencing facial recognition the most.

In practice, when the VUmc analyses her own pictures. The computer analysis can calculate the similarity score but the threshold line can be lowered to 0.43, which is the threshold for human recognition based on the human recognition test and which can, therefore, function as a guideline.



classifications of patients

8.2 Discussion human threshold

The human threshold should be further examined in further research. Only one test is not enough to create a threshold which can be used as an indication of facial recognition among patient's acquaintances during a consultation. A distinction should be made in the amount of familiarity. In reality, familiarity depends on how often you see someone and the type of relationship. When people do not see each other that often, like far neighbours, they may have difficulties recognizing each other without even undergoing orthognathic surgery.
It is therefore expected that patients will be less recognizable after the surgery for far acquaintances than close ones. This will bring the determined human threshold lower than the one determined in chapter 8.1. This is also in line with research later done in chapter 12.2 where the patient experiences were not in line with the recognizability determined by the human threshold.

Furthermore, the regions and movements tested in the human test do not take the amount of displacement or direction of movement (lengthening or shortening, widening or narrowing) into account. Later research done in chapter 12.1 shows both dimensions as the direction of the movement does influence facial recognition.

Not similar by Humans

Patient	Similarity score
1abF,2acF,3F	0.46038
1acF,2aF (2)	0.51791
1acF,2aF,3F	0.55889
1acF,2aF,3F (2)	0.453
1acFL,2aFL,3FL,4FL	0.48625
1adF (5)	0.5295
1adF,2adF (4)	0.45536
1adF,2adF,3F (2)	0.45134
1adF,2adF,3F,4F	0.49914
1aF	0.46267
1aF (3)	0.57413
1aF,2aF (6)	0.52416
1aF,2aF (8)	0.48123
1aF,2aF,3F	0.46411
1aFL,2aFL,4FL	0.60762
1cdF,2adF,3F	0.45773
1cF,2aF (2)	0.54285
1cF,2cF	0.47185
1cF,3F (2)	0.45221
2aF (6)	0.55049
2aFL (4)	0.50666
2aFL (7)	0.48701

Figure 8.2: All patients which are not recognized by humans when the similarity scores of the computer test are multiplied with the human-computer ratio of 1.044.

HST= CST – ((CST*HCS) - CST)
HST = 0.45-((0.45*1.044)-0.45)=0.43

HST=Human_SimilarityTreshold
CST=Computer_ SimilarityTreshold
HCR=HumanComputer_Ratio

It first calculates the difference of the increase of the similarity scores where after this difference is subtracted from the threshold similarity score. The Human_SimilarityTreshold is 0.43.

Figure 8.3: How the threshold is calculated

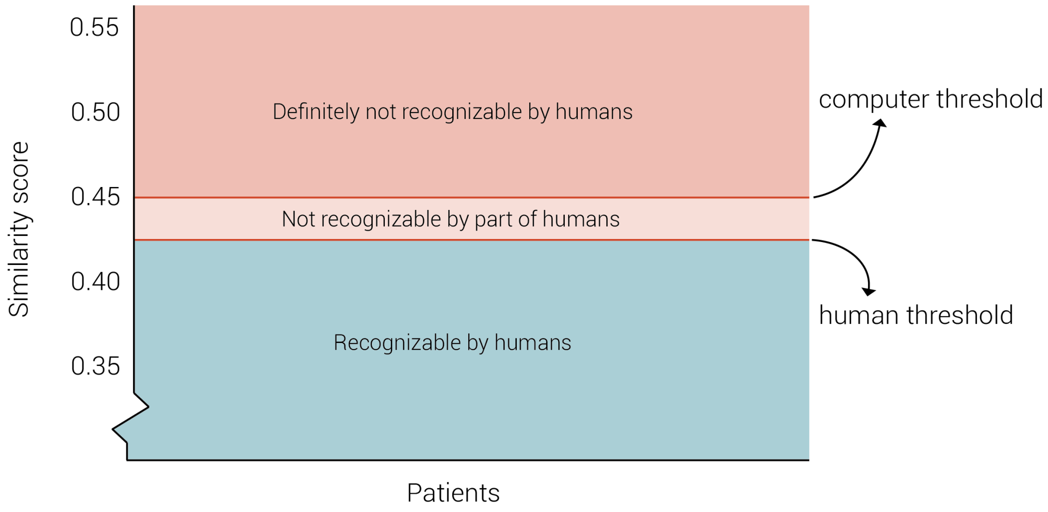


Figure 8.4: The computer and human threshold, creating three recognition zones

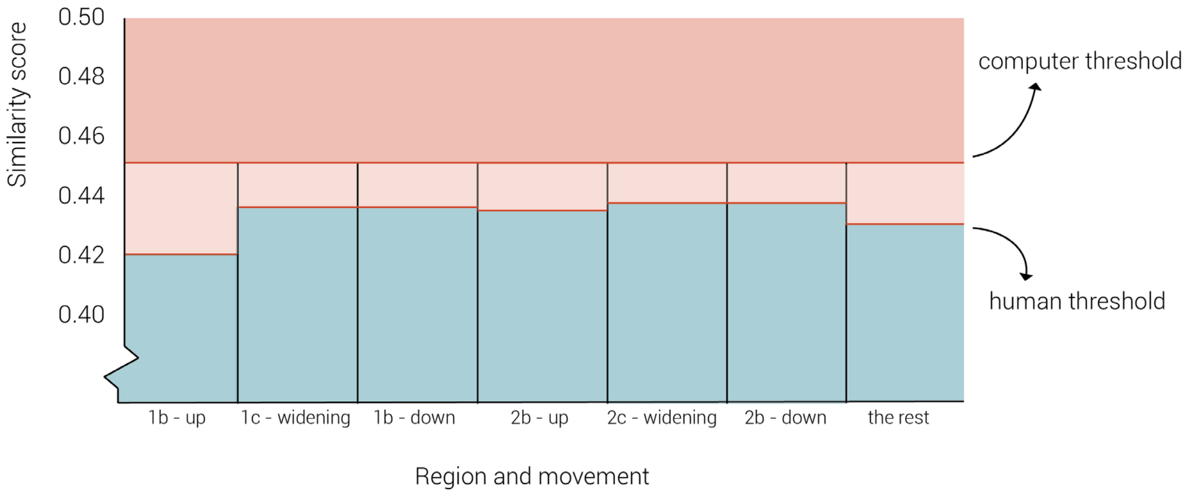


Figure 8.5:The computer threshold and specified human thresholds based on the human recognition test

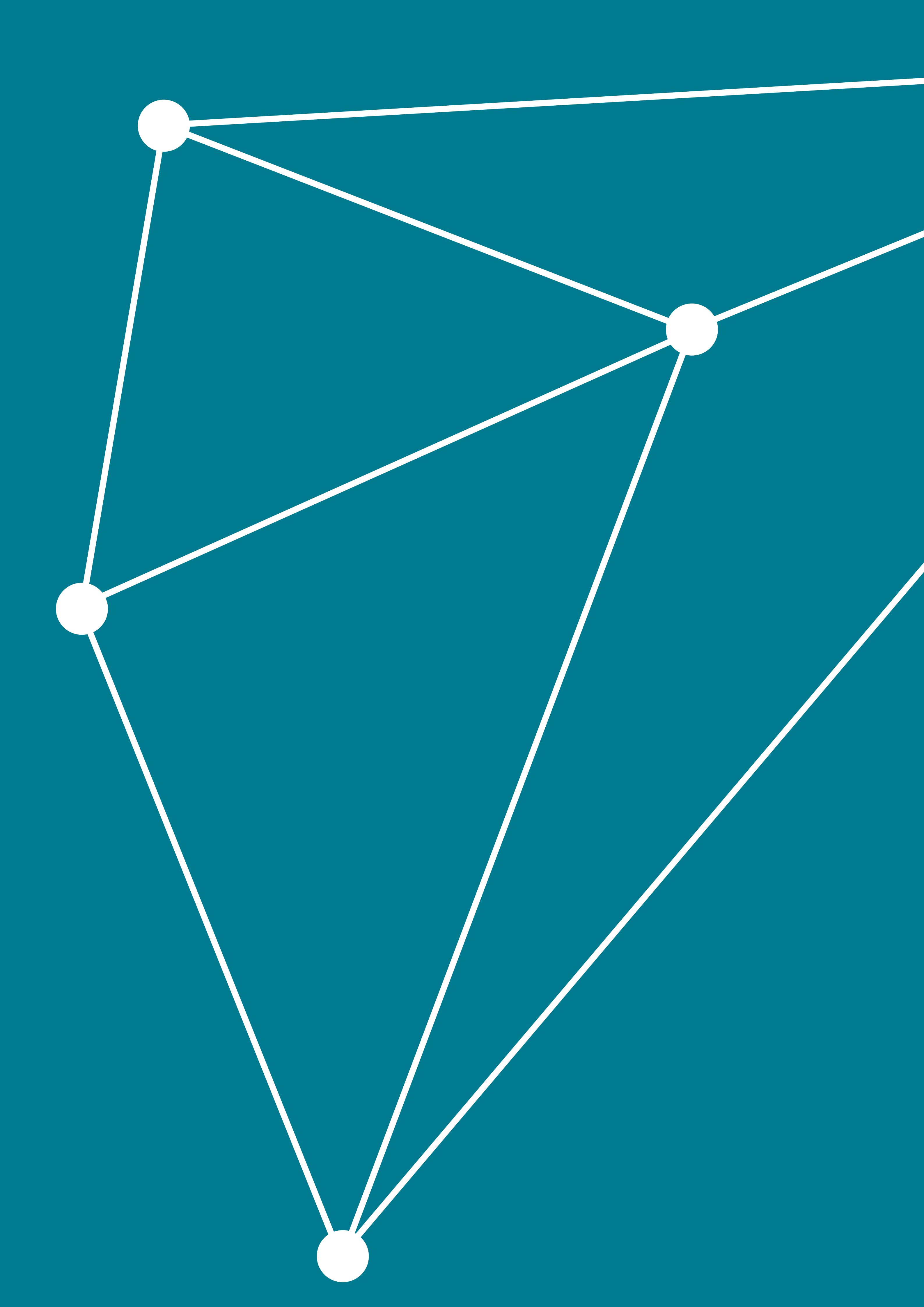
8.3 Conclusion

The results of the human recognition test and computer recognition test are aligned. Therefore, it seems that the manipulation method of the human test is giving a realistic shift in landmark movements.

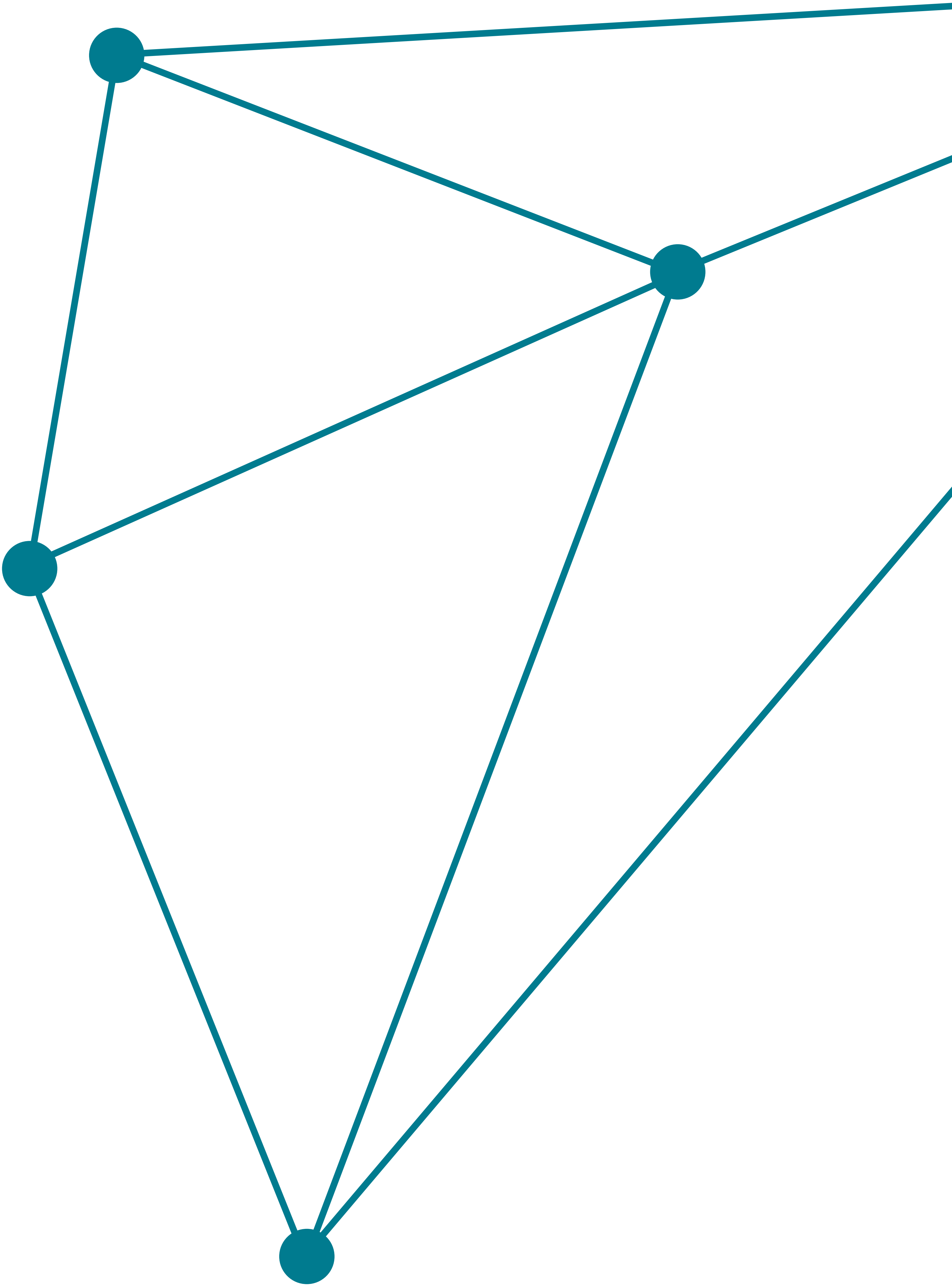
The computer test can show what faces are not recognized by the computer. Since humans recognize faces less accurate, it can be assumed that when computers do not recognize certain faces, humans do also not recognize these faces. Also, because human perception is part of the computer recognition method. An exact human-computer ratio cannot be found, since humans are not consistent in recognizing faces while computers are. However, it can function as a guideline. Figure 8.4 shows three zones, based on the computer and human threshold. These zones can give an indication of the patient is unrecognizable after the surgery. Figure 8.5 shows the specific human thresholds per tested region among familiar people in chapter 6. This could give more insights about certain surgerys to surgeons. More research about the influence of certain surgeries on facial recognition could optimize this figure. With this information, they may change their surgery techniques on certain surgeries which are influencing facial recognition the most.

Key-points from this chapter

- Both tests show that movement c has the highest influence on facial recognition
- Both tests show that region 2 affects facial recognition more than region 1
- Computers are consistent in recognizing faces, humans are not
- The threshold of the computer test can be a strict recognition line, everything above will not be recognizable
- The threshold of the human test can be a recognition guideline, everything between the computer recognition line and human recognition line will be recognizable by part of the humans. Everything under the human recognition line will be recognizable by familiar humans



Design section



Chapter 9

Problem definition

The research phase gave new insights into the influence of orthognathic surgery on facial recognition. This chapter summarizes the important insights from the research phase. After, it provides a problem definition which results in a design vision which supported the designing of a tool to improve orthognathic surgeries. The design criteria are shown after, building upon the problem definition.

9.1 Summary of the outcomes

The literature research made clear that there was a research gap in facial recognition after orthognathic surgery. Before being able to start designing a product to improve orthognathic surgeries, tests had to been done to fill a part of the research gap. The tests provided new insights into the influence of orthognathic surgeries on facial recognition. However, since this has been the first and only research in the field, there can be done a lot more research to get a better understanding of the relationship between orthognathic surgeries and facial recognition. Therefore, the tool in the design section can be improved when more research has been done on the topic. The VUmc can use the design of the tool as an inspiration for the future. It shows how the research which is done could be translated to a concrete application.

9.1.1 Design criteria

The most important outcomes of the research section which are used in the design section are shown. The text in bold are the design criteria:

1. There is a need for help in **communicating** changes in facial recognition from surgeon to patient and the surgeon cannot **predict** changes in facial recognition due to the surgery which makes it hard to communicate (chapter 2)

2. Additional support after surgery can support **emotional recovery** (chapter 2)

3. The **familiarity** of the person recognizing the patient influences facial recognition. Different thresholds can be created regarding the amount of familiarity (chapter 3 and 6)

4. How much facial recognition influences the patient due to the surgery depends on the **patient's face** and the **type of surgery** (chapter 6 and 7)

5. The results of the human recognition test and computer recognition match. This means **the approach** of the human test by manipulating the faces, seems to work (chapter 8).

The scientific value of all the outcomes of the research (chapter 9.1.2) gives a more elaborated description of the research outcomes. The list of requirements in appendix 20 is created after the ideation phase of the design section, and is an addition to the design criteria.

9.1.2 Scientific value of the outcomes

After this research, a start has been made on the topic with still a lot of questions left. The gaps of the research which were filled after the research section are shown here. A complete overview of all the filled gaps of the whole project and the research gaps which are still existing can be found in chapter 14.

1. If a patient is recognized by his acquaintances depends on more than just the face. Things like the context where the patient is seen, the voice, hair- or clothing style all play a role in recognition. Also the expression can play a role. When talking about facial recognition only, the familiarity of the person plays a big role and this should always be taken into account in further research. Also, facial recognition is never constant for humans and differs between- and within individuals. Therefore, thresholds for humans are difficult to define and are variable, it depends on the group the test is performed with.

2. Even though the perception of humans is integrated into computer recognition methods, human recognition and computer recognition cannot be seen as the same. Using them both in further research will take the errors of the variability in human recognition and influence of familiarity into account.

3. In the tested group of familiar people, the vertical shortening manipulation on both maxilla as mandible had the highest effect on facial recognition, which meets the results of the test with the similarity scores of the 75 patient pictures. Also, the manipulations on the mandible had a higher effect on facial recognition than on the maxilla.

5. Orthognathic surgeries do influence facial recognition. Some surgeries more than others. From the 75 patient pictures, the vertical shortening/lengthening surgeries (c) showed the highest percentage in unrecognizable patients, then horizontal shortening/lengthening (a) surgeries and then rotating (d) the mandible (2) or maxilla (1) have the most influence.

6. Single landmark movements, and therefore, single orthognathic surgeries on the mandible or maxilla, can cause an unrecognizable face after the surgery.

7. When both a surgery on the mandible (2) and maxilla (1) is performed, the similarity score is likely to be higher than when only one of these regions is manipulated.

8. Aesthetic surgeries, like a genioplasty (3) or cheek implants (4), which are performed with functional surgeries on the mandible (2) and/or maxilla (1), can decrease the similarity score and can, therefore, make patients more recognizable after the surgery than when this aesthetic surgery is not performed.

9.1.3 New classification system

A set of landmarks based on orthognathic surgeries from Kiekens et al., (2008) is chosen to describe the landmark movements (figure 9.1). However, the computer recognition test uses a set of other landmarks since these were already trained by the dlib library consisting out of 68 landmarks. The 68 landmarks do not matter in this project since the landmarks are only used to identify the face as a face. The landmarks used for the computer test are only of the frontal view, therefore, the computer recognition test is only done on the frontal and frontal-lateral view. A new classification system has been created to divide patients into groups which describe landmark movements (figure 9.2) instead of surgeries. The landmark movements happening in the patient's face are described by region, type of movement and on what plane the movement happens.

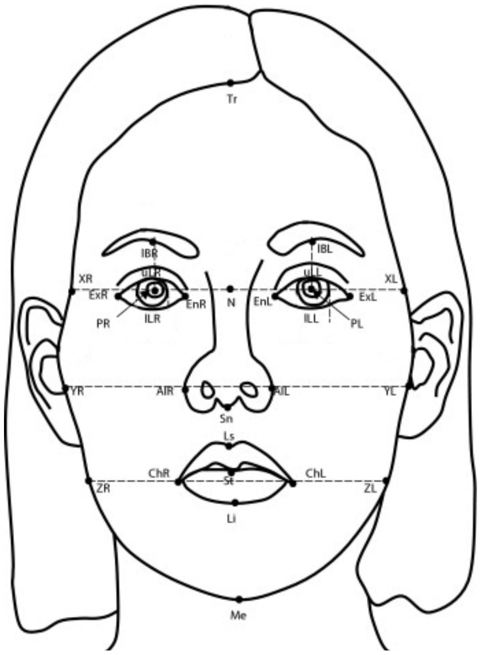


Figure 9.1: The frontal landmarks (Kiekens et al., 2008)

Region		Movement		Plane	
1	Maxilla	a	Horizontal Lengthening/Shortening	F	Frontal
2	Mandible	b	Widening/Narrowing	L	Lateral
3	Chin	c	Vertical Lengthening/Shortening		
4	Cheek	d	Rotating		
		e	Smoothening jaw line		

Figure 9.2: Legend of classification annotations

9.1.4 The human recognition test

The human recognition test has been done to test how single configurations happening in different types of orthognathic surgeries influence facial recognition of familiar faces. The literature research in chapter 3 showed that processing familiar faces are different from processing new faces. Since the patients face needs to be recognized by familiar people, a test needed to be done between familiar people. Regions 1 and 2 have been tested with movements b-widening and c-vertical lengthening and shortening (figure 9.2). Apparently, all the tested manipulations have an influence on facial recognition among familiar people. Vertical shortening of the mandible (2c) has the highest influence on facial recognition. After this comes vertical lengthening of the maxilla (1c). The manipulations on the mandible have the highest influence on facial recognition.

9.1.5 The computer recognition test

The computer recognition test has been done to analyze complex landmark movements and a bigger amount of data than possible with a human recognition test. Against expectations, more regions or movements does not automatically mean it increases the similarity score which decreases facial recognition. First, the influence of the regions on facial recognition has been analyzed. Combining region 1 and 2 has the highest percentage of unrecognizable patients, together with combining all regions. Therefore, the combination of 3+4 is not effecting the percentage of not recognizable patients. Remarkable is that adding region 3 or region 4 can bring the similarity score down.
After analyzing the regions, the movements have been analyzed. 1a and 2a can, as the only single movement, cause an unrecognizable face after surgery. The c (vertical lengthening/shortening) movements do have the highest percentage of not recognizable patients. After come the a movements (horizontal lengthening/shortening). The d (rotating) movements could increase the similarity score as well.

9.1.6 Comparing the two tests

Human perception is included in the computer test, but since face processing is done differently from humans, the accuracy is increased in comparison to humans. Both the human recognition test and computer recognition test have been compared and the threshold of human recognition has been found. The results of both tests show that the c movement and manipulations on the mandible have the most influence on facial recognition. Therefore, it seems that the manipulation method of the human test is giving a realistic shift in landmark movements.

Computers are consistent in recognizing faces, humans are not. Therefore, a threshold line for human recognition is not an exact line and is dependent on the group the human recognition test is performed with. The threshold of the computer test can be used as a strict recognition line, everything above will not be recognizable.
The threshold of the human test can be a recognition guideline, everything between the computer recognition line and human recognition line will be recognizable by part of the humans. Everything under the human recognition line will be recognizable by humans (figure 9.3 and 9.4). The thresholds can be used as a foundation for the tool.

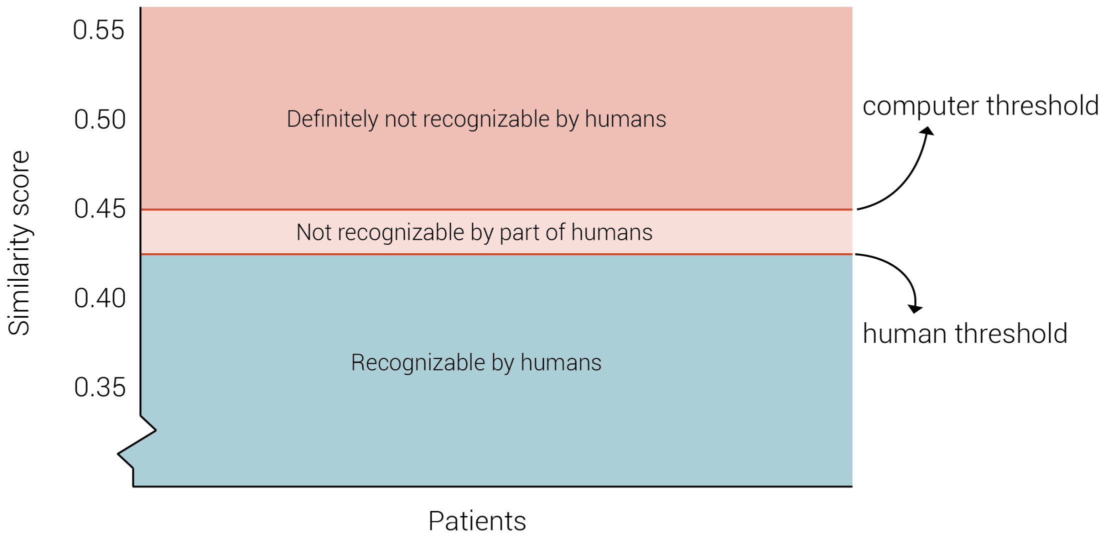


Figure 9.3: The computer and average human threshold

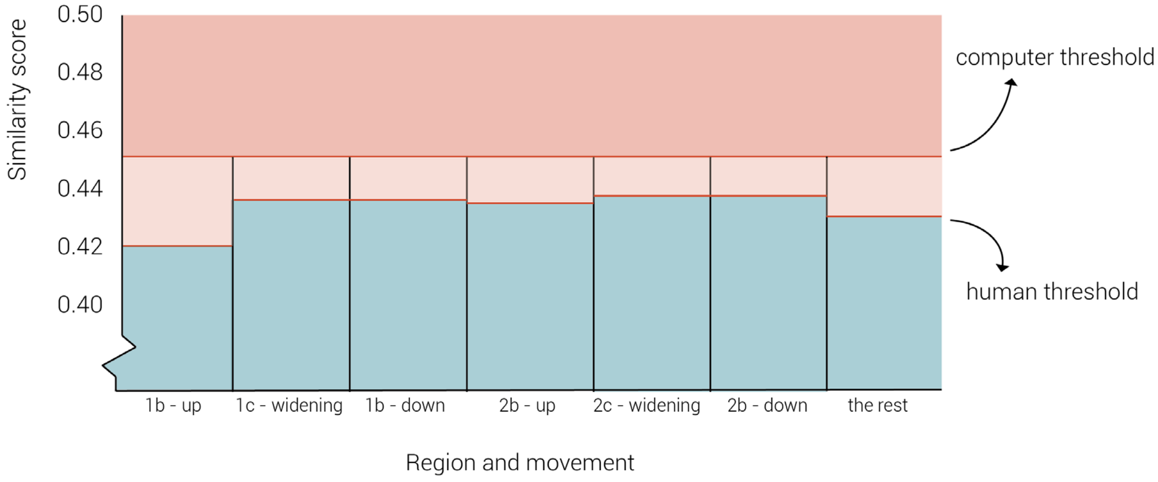


Figure 9.4: The computer and specified human thresholds

9.2 Problem definition

The research proves that orthognathic surgeries can have an influence on facial recognition. Therefore, the first plan was to optimize orthognathic surgeries. Unfortunately, orthognathic surgeries cannot be improved to preserve the patient's recognizable face, since the surgeons have to perform the surgery fast and do not have time to move the soft-tissue back to place. Also, the patient's mouth is full of equipment and sometimes already starts swelling during the surgery (figure 9.5). This makes it impossible to integrate the extra task of moving the landmarks on the soft-tissue to place (T. Fouranzanfar, personal communication, May 26, 2020). This is adding boundaries on the design space and the tool was designed for during the surgery.

Communication between surgeon and patient is important to meet the expectations of the patient, which is not happening currently at the VUmc (figure 9.10). If the surgeon could tell the patients if they would be recognizable or not after the surgery, it is meeting the 'before' image or perception of the patients of their face after the surgery, as well as meeting (part) of the expectations. The tool could help with this in the pre-surgical stage. The last part of the research (chapter 8) shows the relationship between the computer threshold and the human threshold. Therefore, computer recognition can be implemented in the tool by taking the human threshold into account. The goal of the tool is to match the expectations of the patient with the outcomes. Research has been done in how to achieve this goal which can be read in appendix 26. According to this research, effective communication has a positive impact on important outcomes including patient satisfaction. Psychological preparation for the treatment helps to prepare the patients for the post-surgery image. This can be achieved by the surgeon providing information. The tool should create a balance between the three pillars: functional, aesthetics and psychosocial (figure 9.6).

To conclude, the tool should take communication between doctor and patient into account, so the patient will understand what the doctor is saying (figure 9.8 and 9.9). After the doctor shares the facial recognition information with the patient, the patient can decide together with the doctor what to do. A scenario with three possible reactions of the patients is shown in figure 9.11. Choice 1: continue with the plan, is desired and this is the reaction to design for. This conclusion is summarized in the design challenge shown in figure 9.7.

Figure 9.5: Equipment in the patient's mouth during the surgery

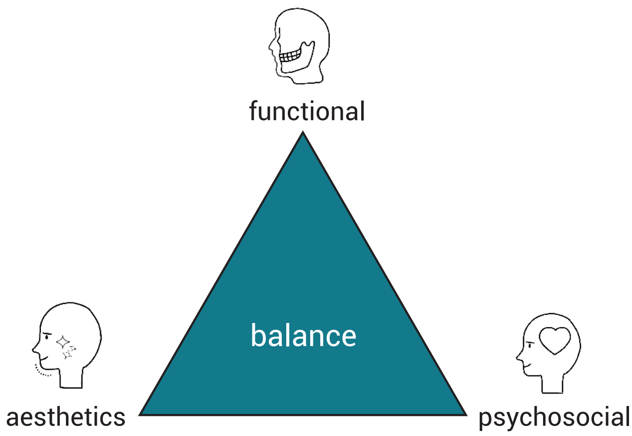


Figure 9.6: The three pillars

How to optimize communication between the surgeon and patient by making the patient aware of changes in facial recognition which will happen due to the planned orthognathic surgery?

Figure 9.7: The design challange

Figure 9.8: The problem which arises when using the facial recognition research outcomes but when not using the proper way of communication

Figure 9.9: The reaction which is desired when using the tool

Figure 9.10: The current situation at the VUmc


It depends per person. I cannot tell how or how much your face will change
I have no idea what to expect...

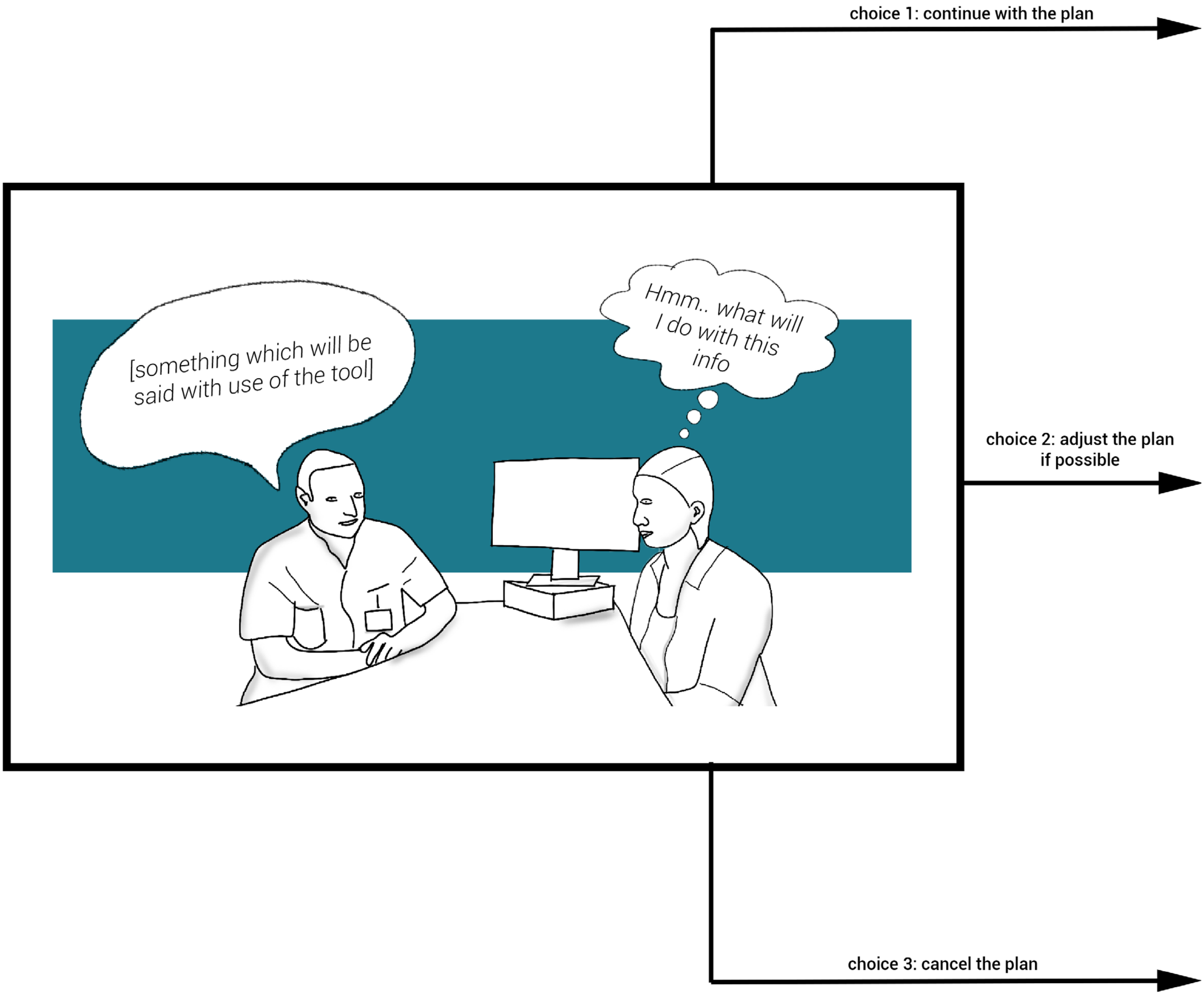
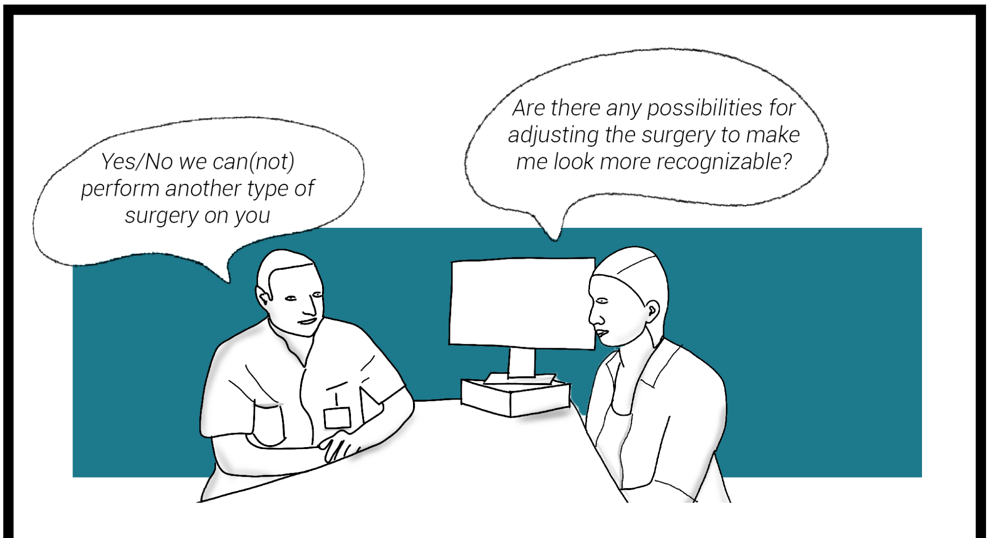


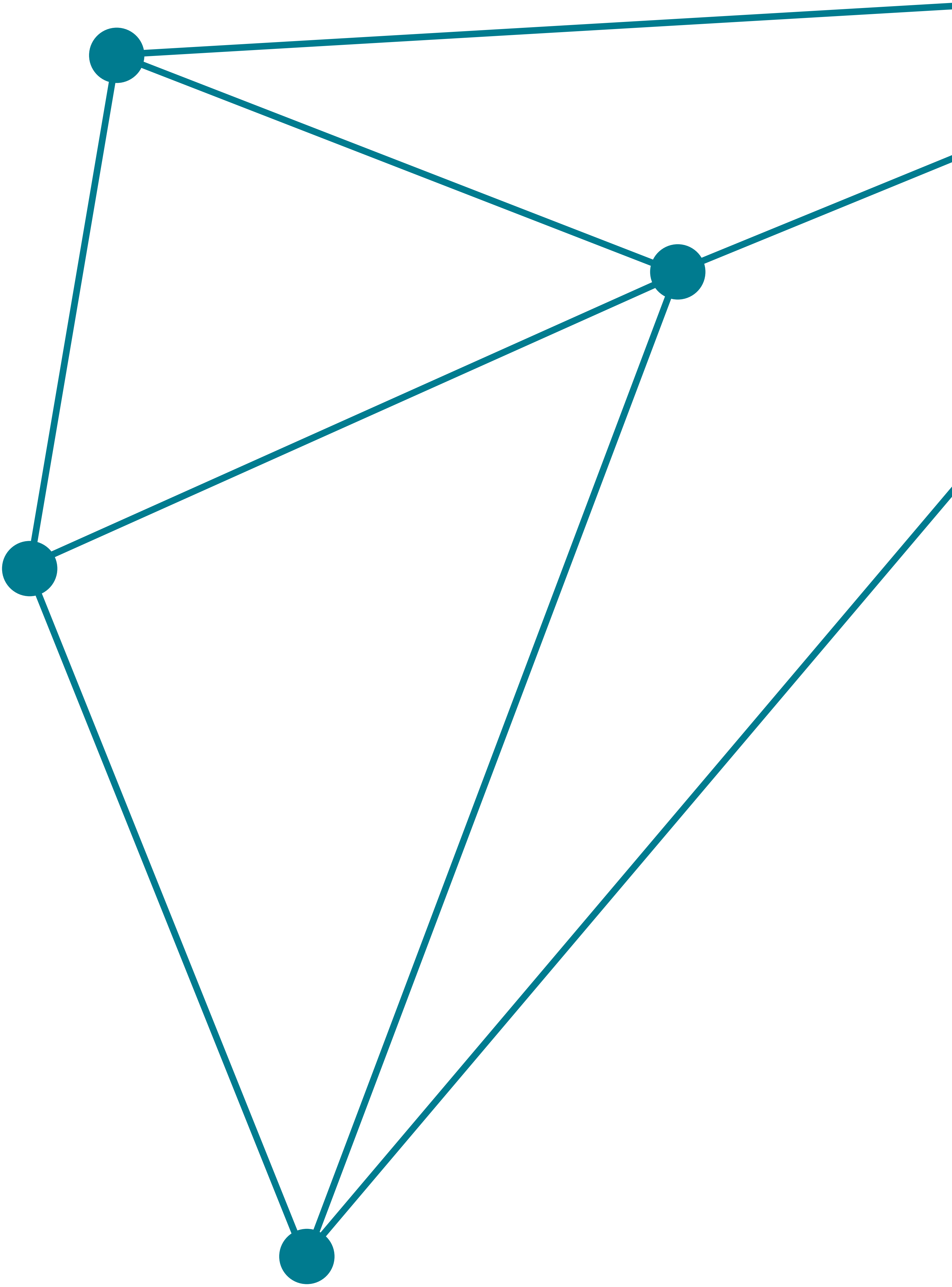
Figure 9.11: The scenario of shared-decision making when implementing the tool

Scenario 1



Scenario 2

Scenario 3



Chapter 10

The concept: Fraos

This chapter shows the concept called Fraos, which stands for facial recognition after orthognathic surgery. It exists out of a communication tool and a facial recognition system. The communication tool is the tool which is used during the consultation. The doctor can use the tool to communicate if the patient is expected to be unrecognizable after the orthognathic surgery. This communication tool is personalised per patient and the facial recognition system is the system behind this communication tool. This system analyses two patient pictures, whereof one is manipulated and the landmarks are displaced according to the type of surgery the patient will undergo. This chapter shows how the communication tool and facial recognition system work. The next chapter, chapter 11, will validate the communication tool and chapter 12 will validate the facial recognition tool. The whole design process can be found in appendices 21 till 27.

The concept can be seen as something the VUmc can use as an inspiration for the future. It can be seen as a mean to make research outcomes concrete and it can be seen as a goal for the VUmc to apply the research they are planning to do in practice. Before being able to apply this concept in the VUmc, more research needs to be done. Therefore, a roadmap to implementation has been created in chapter 13.1.

10.1 The communication tool

Figure 10.1 and 10.2 show the design of the front and back of the flyer, which is the communication tool. The real format of the flyer is in A4, which can be found in appendix 25.

Emotional recovery is provoked by the design of the flyer by using the theory of psychological adaptive nodes (appendix 27). The design of this flyer is almost completely based on research and is designed for both introvert as extrovert patient profiles (appendix 24). The market research, communication research and research about emotional recovery can be found in appendix 22, 26 and 27.

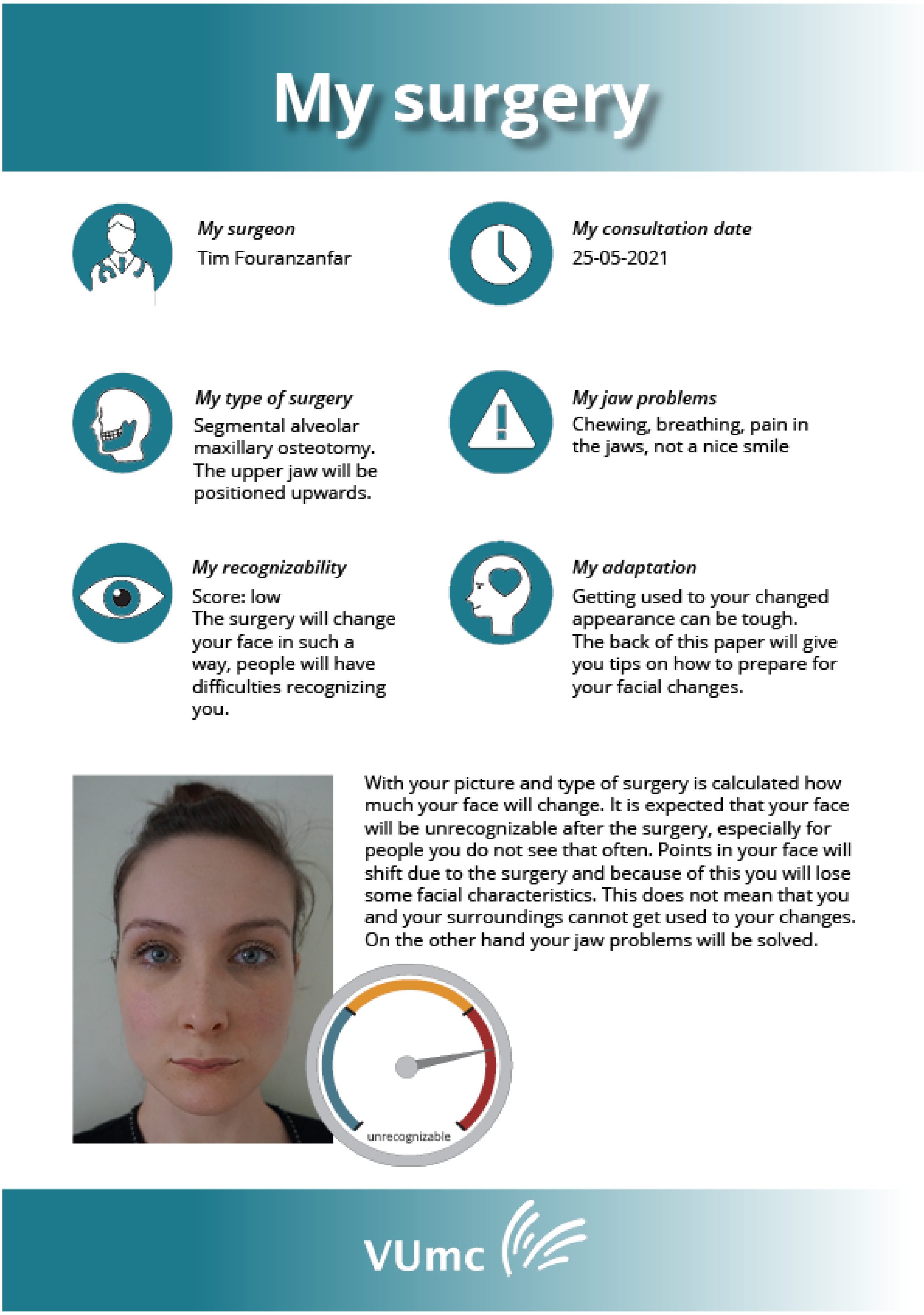


Figure 10.1: The front of the Communication Tool

All the aspects of the design are carefully selected and argued. The most important design decisions and arguments are shown in figures 10.4 and 10.5. These figures also explain the function of all the design aspects presented in the flyer. Chapter 11 validates the design.

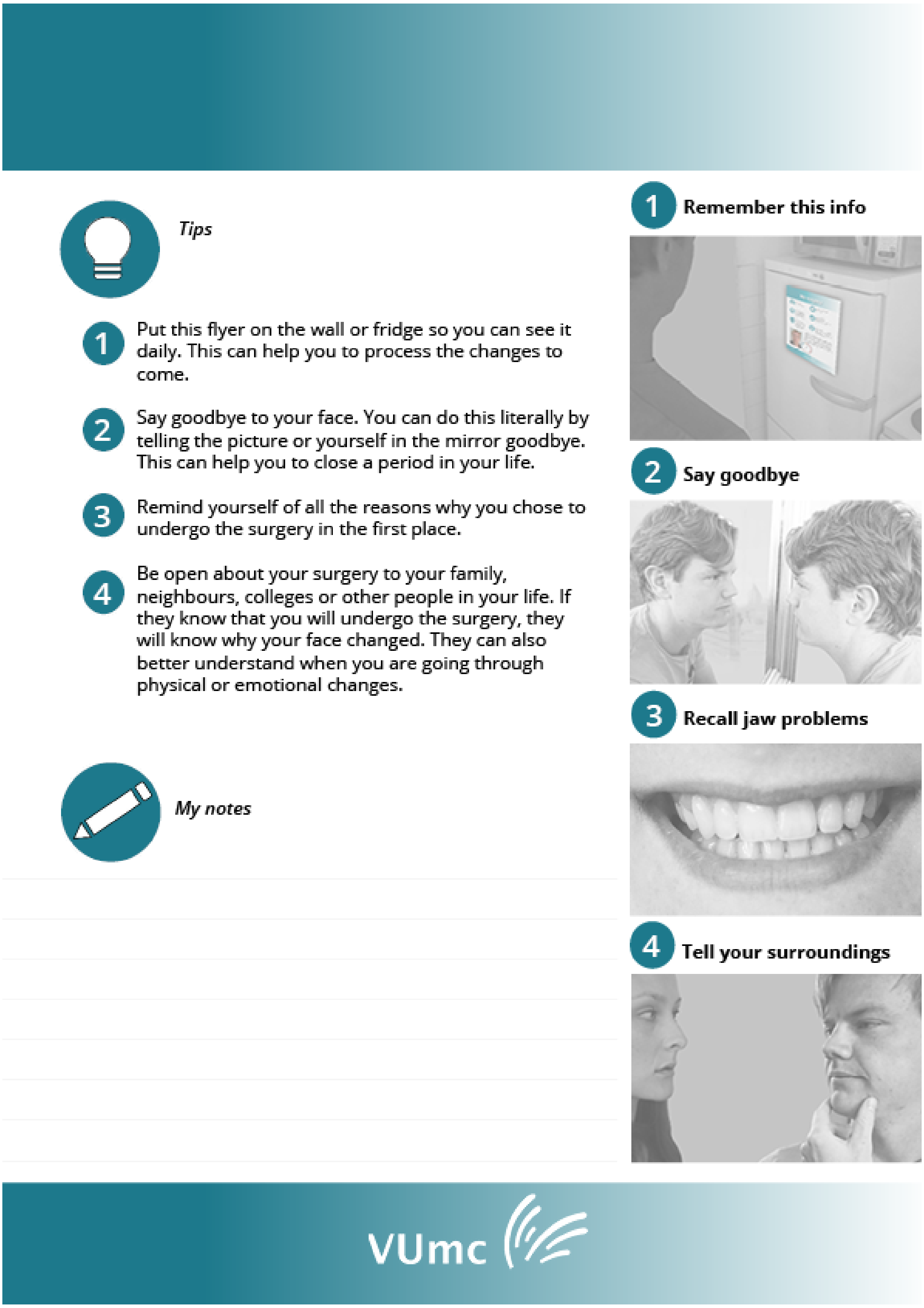


Figure 10.2: The back of the Communication Tool

Icons are used to find the subjects fast, to remember information, for the illiterate, it looks positive and it improves the appearance. This will influence the patient emotionally by feeling positive and confident (Interaction Design Foundation, 2002).

The word ‘My’ is used through the whole flyer to underscore that the information is **personal**.

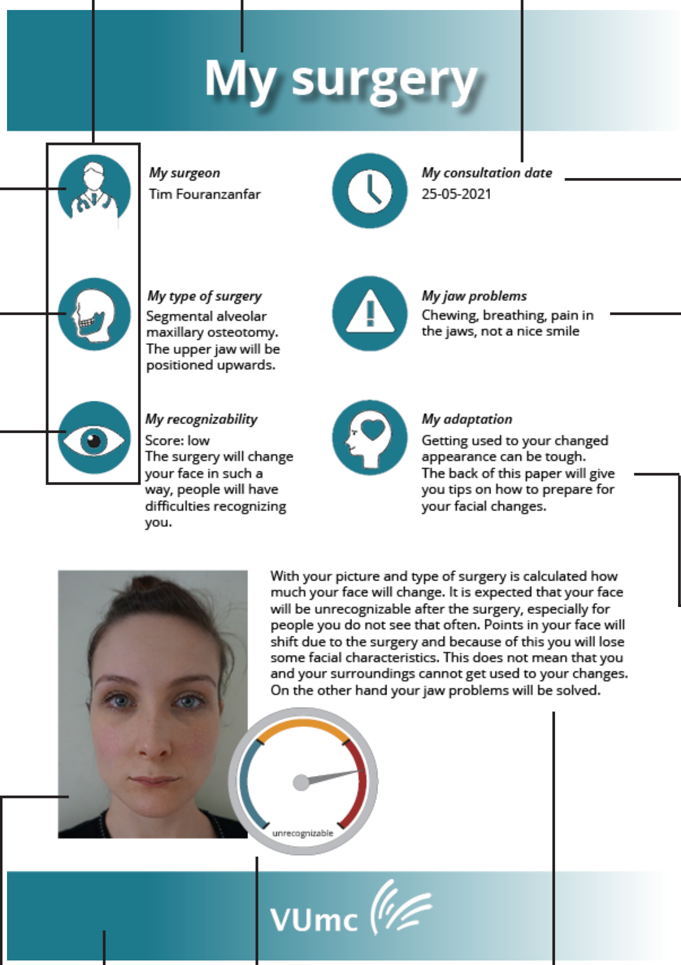
Open Sans has been chosen as **font** since it is the best to read when printed. It is sans serif which makes it easier to read (Vistaprint, 2018). The size of the font is 12 pt, which is a comfortable reading size for most people (Flyerzone, 2020).

The name of the surgeon as a **cue** for patients to inform themselves. This is a second adaptive skill (Lehti, 2016).

The medical term of the surgery as a **cue** for patients to inform themselves and an explanation of what it means. This is a second adaptive skill (Lehti, 2016).

The score and a short description of the **consequences**. For the doctor to explain fast to the patient.

The picture of the patient is shown which is used to calculate the similarity score. It gives the patient a feeling the information is **personal**. It can help to process information which is an adaptive skill (Lehti, 2016). It corresponds to tip 1 and 2.



Practical information so the patient can always recall when the consultation has been.

To **remember** the patient why he chose to undergo the surgery. It can help to process information which is an adaptive skill (Lehti, 2016). Corresponds with tip 3.

Empathize with the patient. Referring to the back of the flyer. Provokes an emotional reaction of the patient which activates behaviour towards the surgery (Interaction Design Foundation, 2002)

This **colour** has been chosen because it gives a feeling of calmness and trust but also depth, expertise and stability (Color psychology, 2020)(Bruens, 2011, pp. 66-67). Patients do respond best when colours correspond with the intended results of the treatment (Bruens, 2011, pp. 66-67).

This **facial recognition clock** visualises the similarity score to make it understandable for patients.

This longer story gives an explanation of the outcome of the analysis, the consequences concerning recognition and jaw problems. It is a **summary** of the things the doctor should say during the consultation.

Figure 10.4: Design decisions of the front page

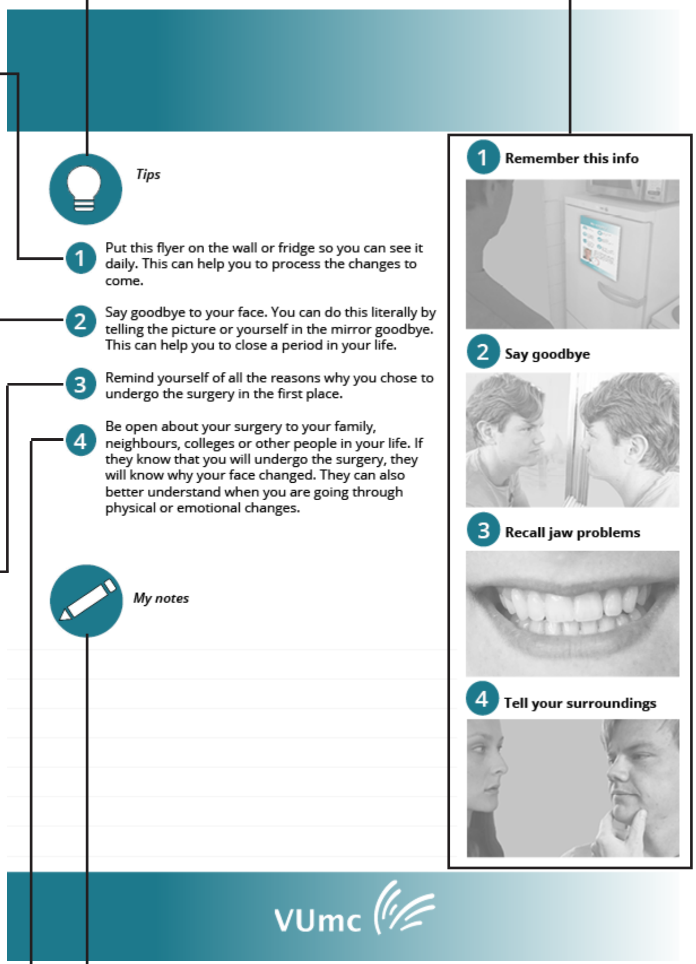
Tips are given to support emotional recovery when needed. Not every patient needs (all) the tips, depending on the mental state.

The tips on the right support the tips on the left by using a key-word for describing the tip and by showing pictures of how to execute the tips. This can help **remembering** the information and it can help illiterate patients.

This tip **confronts** the patient with the results of facial recognition on the front page. It stimulates direct problem solving which is an adaptive skill (Lehti, 2016).

This tip **confronts** the patient with their own face and the changes which are about to come. It stimulates direct problem solving which is an adaptive skill (Lehti, 2016).

This tip helps the patient to put things in **perspective** by remembering the original reason he wanted to undergo the surgery. It stimulates information processing which is an adaptive skill (Lehti, 2016).



This tip helps the patient be **transparent** to the surroundings about the surgery. This can help to support the patient and to accept the facial changes. It stimulates mobilizing the social network. Which is an adaptive skill (Lehti, 2016).

The patient can use this space to write down any thoughts during emotional recovery. This stimulates **information processing** which is an adaptive skill (Lehti, 2016).

Figure 10.5: Design decisions of the back page

Since the facial recognition score shown in the facial recognition clock can still be too abstract for the patient to understand the meaning and because the user test in chapter 11 showed that patients are interested in how their appearance will be after the surgery, an overview of four clocks with patients who belong to that clock regarding the similarity score can be given to the patients. This can give the patients an insight into how much they can change after the surgery. It would be even better to show a picture to the patient which is equal to the surgery the patient will undergo since he can confuse the type of surgery with his own. However, since so many combinations of surgeries are possible, this is too time-consuming to create in this project.

Chapter 10.2.2 explains more about the clocks and it shows two more clocks with higher similarity scores than the ones showed in figure 10.6. However, since the similarity scores of these two clocks did not exist in the 75 patient pictures used in the computer recognition test (chapter 7), these are not represented in figure 10.6. However, it might be possible that patients at the VUmc do have high similarity scores matching the two clocks, and therefore, figure 10.6 could be optimized when further research has been done at the VUmc.

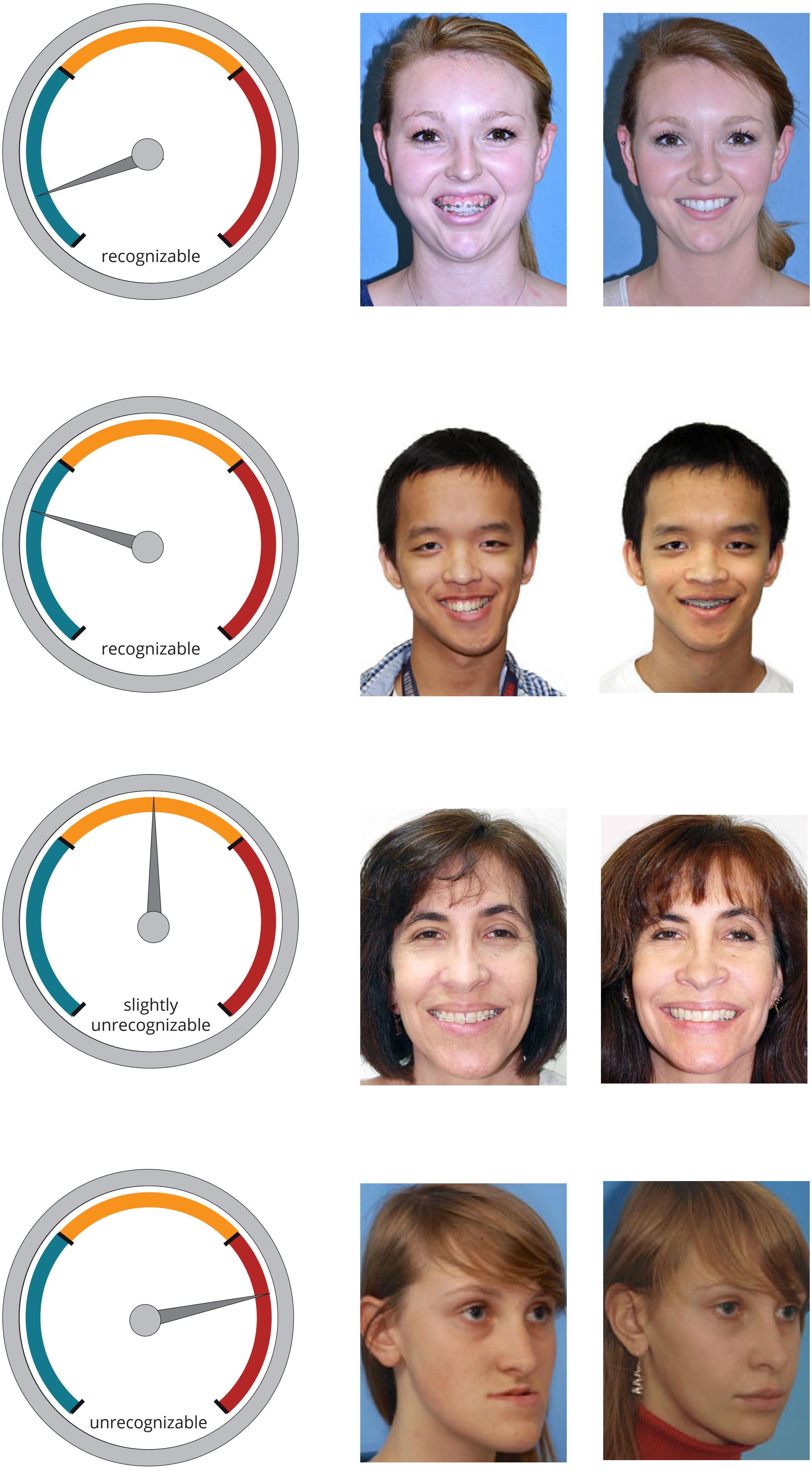


Figure 10.6: Facial recognition clocks belonging to patients which can be shown to the patients to make the clock more concrete. (Reference pictures: Dr. Richard W. Joseph, 2019; Drantipov, 2020; Facial Sculpture Clinic, 2020)

10.2 The facial recognition system

This chapter will explain how the facial recognition system works and how it can be developed. The facial recognition system is the system working together with the communication tool (chapter 10.1). The communication tool is the representation of all the information generated by the facial recognition system. The facial recognition system is what generates the similarity scores for all patients and it translates it to the communication tool. It is integrated into the software the VUmc uses currently, called Epic (Electronic Privacy Information Center).

The system is not visible for the ones using it, as the surgeon, but it uses the technology which is unmissable. This project only comes up with a concept. The development of the tool can be done in the future. Before it can be implemented, there are still things which should be done. What should be done and in what order can be found in the roadmap in chapter 13.1.

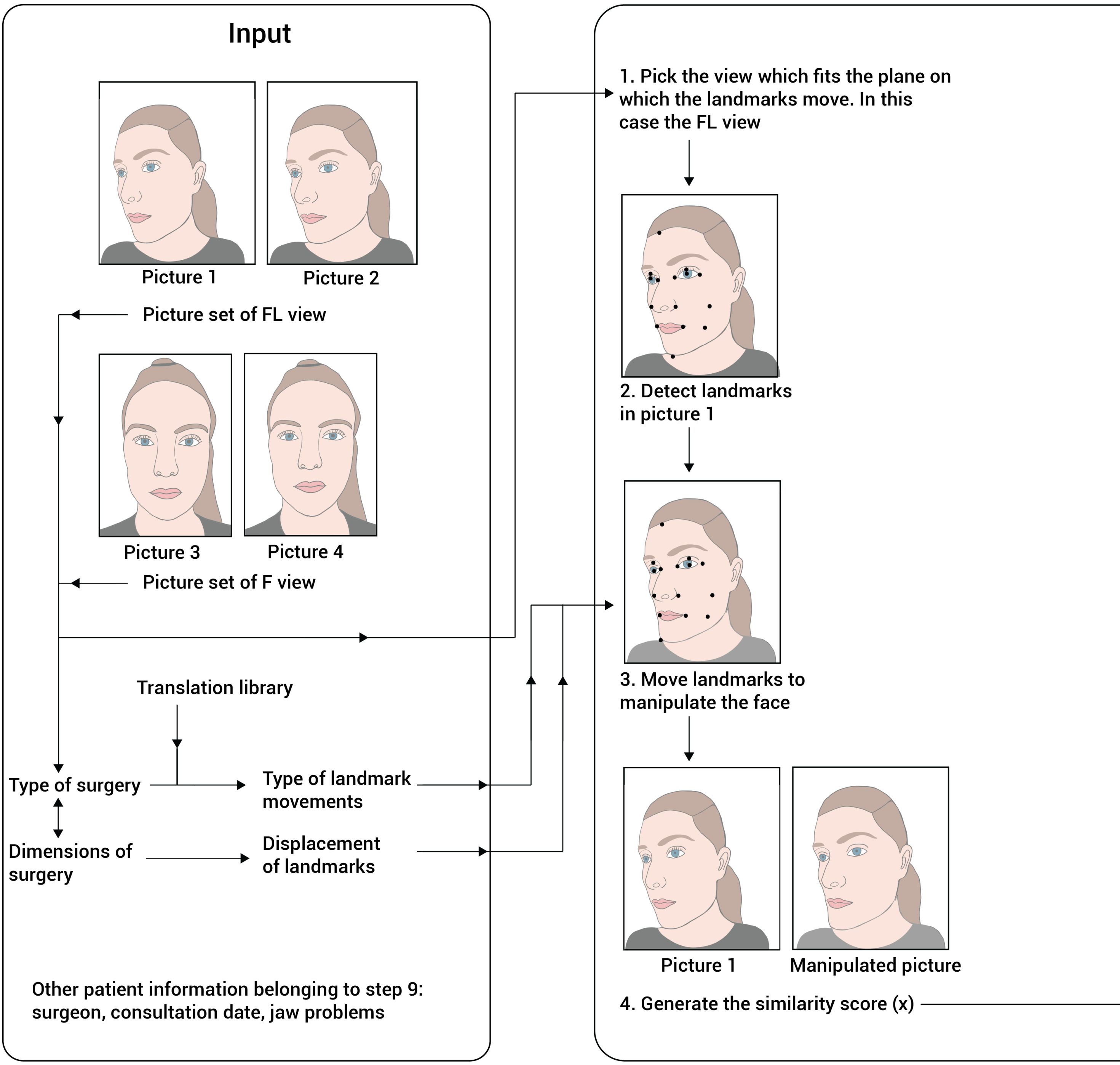


Figure 10.7: Flowchart of the facial recognition system

10.2.1 The procedure

To explain what the facial recognition system should do before the flyer can be printed, a flowchart is created (figure 10.7). This flowchart shows the input, work of the system and the output. The input is the part which is partly covered by the patient information present in Epic and by the pictures the doctor should upload. The type of surgery should be translated to information the system can use to perform the manipulation on the face. Therefore, the 'translation library' in the input can be seen as a database which combines the type of surgery with the type of landmark movements belonging to the specific type of surgery (figure 10.8). In this project, only regions have been researched and not the specific type of surgery as shown in figure 10.9. However, future research should take the type of surgery into account (chapter 10.9). The 'dimensions of surgery' is the term which is shown to the user, e.g. the surgeon, in Epic. This is the same as 'displacement of landmarks' but it is called differently since the user might not know what displacement of landmark means. The user should make sure when using the system, all the needed input is present in Epic.

System

5. Define recognizability:

Recognizable x < 0.43
Slightly Unrecognizable 0.43 ≥ x < 0.45
Unrecognizable x ≥ 0.45

6. Find the facial recognition clock matching x

7. Activate information belonging to type of recognizability

Unrecognizable: Appendix 25, page 112

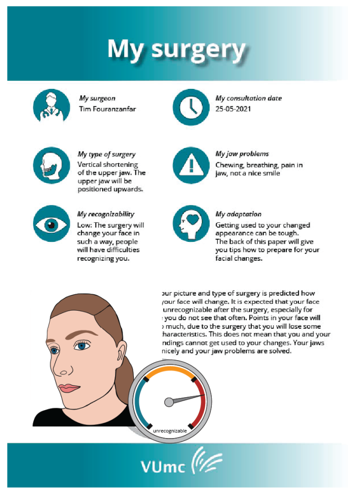
Slightly recognizable: Appendix 25, page 114

Recognizable: Appendix 25, page 116

8. Put the patient information from Epic in the flyer: surgeon, consultation date, jaw problems, picture.

9. Save the flyer in Epic

Output



The flyer specified to the patient

The landmarks which should be used for manipulating the face are the ones from figure 10.10 since these describe the landmarks which move during orthognathic surgeries. However, to generate the similarity score, the 68 landmarks of figure 10.11 can be used since these landmarks are only used to identify the face as face. Also, the method used to generate the similarity score (chapter 3.3 and 7), is fast and relatively easy to use.

However, when a software engineer is creating the software needed for this system, it could be possible other methods are better to use. The 68 landmark system is mostly used in computer science, and therefore it is expected it will be used in the software as well. Although, other landmark systems may suffice also.

Figure 10.9: Area of the surgery of Le fort I, II and III (Royal medical services, 2017)

Type of surgery	**Landmark movement**	**Landmarks to move**	**Movement on plane**
Le fort I	1(aL)	YR, YL, Ls	L
Le fort I	1(aS)	YR, YL, Ls	L
Le fort II	1(aL)	YR, YL, AIR, AIL, Sn, Ls	L
Le fort II	1(aS)	YR, YL, AIR, AIL, Sn, Ls	L
BSSO	2(aL, bL, d)	ZR, ZL, ChR, ChL, Me, Li	FL

Figure 10.8: An example of some surgeries in the translation library

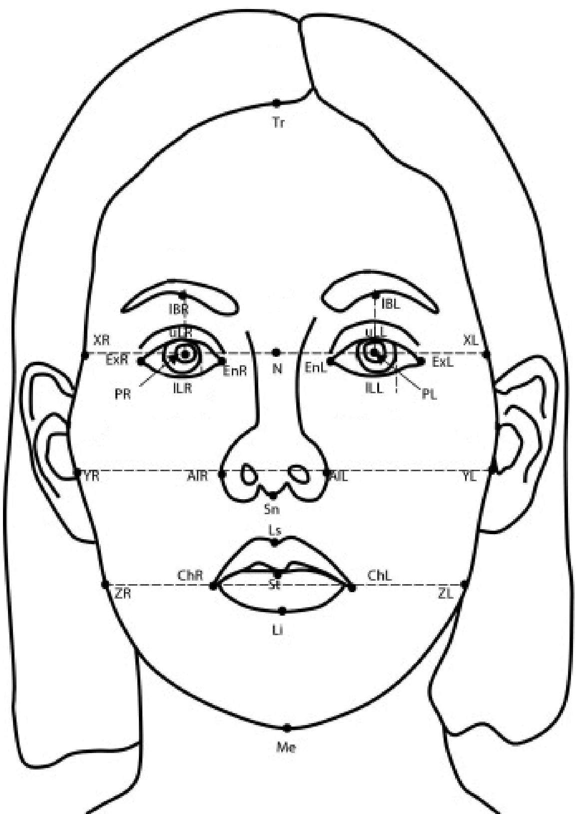


Figure 10.10: Recap of the frontal landmarks (Kiekens et al., 2008).

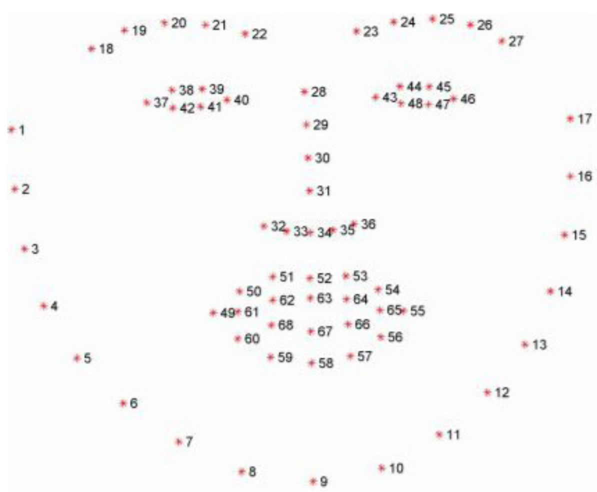


Figure 10.11: The amount of landmarks used the most in literature and software is 68 (Amato et al., 2018)

10.2.2 Feasibility of development

The facial recognition system should be produced to be able to complete the steps listed in the system of figure 10.7 and before it can be implemented. Mark that before it can be implemented, more research should be done in facial recognition after orthognathic surgery (chapter 13.1). An example of how development can be done is shown here. However, since the concept will be implemented in the future, and since further research might change the concept, this development is only an example of how it could be done. The things which should be done to develop the system, seem feasible at the moment because the technologies which are needed for each step already exist. Only integrating all of these steps is something which needs to be done. The details of how it can be programmed are not shown, since the VUmc should hire a software developer to do this. The steps the system should undertake are shown, with a description of how it could be developed.

Integration of the steps

For developing the facial recognition system by integrating all the steps, software should be developed. This software should be integrated into Epic, the software the VUmc uses right now for saving patient data (chapter 10.2.3). It is researched if it would be possible to implement software in Epic. Fortunately, this has already been done before at the Multiple Scleroses (MS) department in the VUmc (de Jong, 2016). They integrated specific MS software in Epic. Therefore, this can also be done for the facial recognition system. The VUmc should consult the developers who integrated the MS software in Epic and redo it the way they performed it.

Step 1: Pick the view which fits the plane on which the landmarks move.

This step is linked to the type of surgery which is used to select the type of view. The system should get the 'movement on plane' information from the translation library (figure 10.8), which is a database. This database should be created together with physicians from the VUmc, to use the medical names the VUmc specifically uses for the 'type of surgery' column since this can differ per hospital.

Step 2: Detect landmarks in the chosen view.

The landmarks of Kiekens et al., (2008) are used to manipulate the patient's face in the picture. This is different than what happens in the dlib library (chapter 3.3), where the landmarks of Amato et al., (2018) are used to detect the face. The landmarks of Kiekens et al., (2018) need to be trained by manually labelling images (appendix 4.3).

Step 3: Move landmarks to manipulate the face.

Face manipulation by using landmarks is already done e.g. in applications like Instagram (Pysource, 2019)(figure 10.12). The system should manipulate the landmarks according to the landmarks belonging to the selected surgery from the translation library (figure 10.8). Multiple methods can be used for augmentation of the face. Instagram's method is described now but, in the future when another method seems more fitting, this could be applied as well. Instagram's method of detecting the face is similar to the method used in this project (chapter 3.3) by using the Histogram of Oriented Gradients. After, the SVM (Supported Vector Machine) method is used to detect the facial landmarks.

After this is done, the augmentation of the face is done, called the mask. The location of the landmarks where the augmentation is happening is located. The type of augmentations is saved in a library. Which one Instagram uses is unknown, but it is comparable to libraries like OpenCV and Dlib (Erminesoft, 2017). A new library should be created, containing the manipulations belonging to the landmark movements from the translation library (figure 10.8). Also, the scale of the picture should be defined so the system can manipulate according to the dimensions of the surgery which the surgeon puts in Epic.

Figure 10.12: An example of augmation of a face in Instagram (Gumbinner, 2020)

Step 4: Generate the similarity score
The code needed to achieve this is already written (appendix 15). However, it should be integrated into the facial recognition system.

Step 5: Define recognizability
The system should define where the similarity score (x) lies:
Recognizable $x < 0.43$
Slightly Unrecognizable $0.43 \geq x < 0.45$
Unrecognizable $x \geq 0.45$
The boundaries are defined by the thresholds defined in chapter 8. The lower threshold of 0.43 is variable and can be lowered depending on future research on familiarity. Linking a recognizability score to the similarity score should be a simple translation.

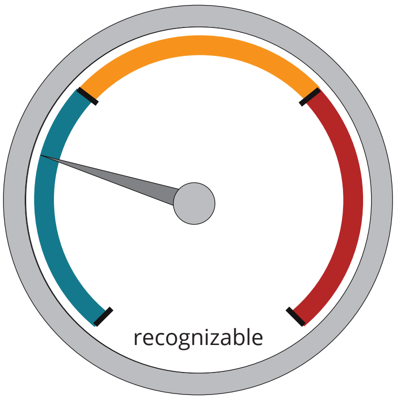
Step 6: Put x and the recognizability in the facial recognition clock
The system should generate the clock based on the outcomes of steps 5 and 6. It is easiest to create some clocks belonging to a range of recognizability scores as shown in figure 10.13. In this example is chosen to create two clocks in the range ‘recognizable’, because it gives meaning if almost nothing is changing in your face or when it is closer to ‘slightly recognizable’, which means there might be enough facial change to have difficulties recognizing the patient for people who are not that familiar, also depending on the context the patient is in. There is one clock in ‘slightly recognizable’ because it is such a small range. The ‘unrecognizable’ has three ranges because this is the outcome with the most impact, the score can have meaning for the patient by telling how unrecognizable the patient will become. However, The range of the clocks of $0.6 \leq x < 0.8$ and $x \geq 0.8$, did not exist in the patient data of chapter 7. However, it could be possible this range does exist at the VUmc, and therefore these two clocks belonging to the range are designed. If the similarity score never exceeds 0.8, and the difference between 0.6 and 0.8 is big when looking to patient experiences, the ranges could be improved by making them smaller, The system could link the score to the belonging clock and, therefore, the system does not have to generate a personal clock for each separate patient, which makes it easier for the software engineer to write for.

Step 7: Activate information belonging to the type of recognizability
This is about the story belonging to the recognizability next to the patient's picture, explaining more about the meaning of the score. It works the same as the facial recognition clocks in step 7. Three stories are predefined and linked to the outcomes of ‘recognizable’, ‘slightly recognizable’ and ‘unrecognizable’. The stories can be found in appendix 25.

Step 8: Put the patient information from Epic in the flyer: surgeon, consultation date, jaw problems, picture.
This information is already stored in Epic and the system should get this information from Epic. After, the retrieved information should be placed in fixed spots in the flyer.

Step 9: Save the flyer in Epic
The last step of the system is to save the flyer in Epic where after the surgeon can print it for the patient to take home. The system should save it on the part of the Epic the VUmc uses as well as in ‘My Chart’, which is the platform the patients use to retrieve information (chapter 10.2.4). It is expected this is an easy step as well.

Facial Recognition Clock when x< 0.3



Facial recognition clock when 0.3 ≤ x < 0.43

Facial recognition clock when 0.43 ≤ x < 0.45

Facial recognition clock when 0.45 ≤ x < 0.6

Facial recognition clock when 0.6 ≤ x < 0.8
This score was never reached with the 75 patient pictures of chapter 7

Facial recognition clock when x ≥ 0.8. This score was never reached with the 75 patient pictures of chapter 7.

Figure 10.13: An example of a set facial recogntion clocks

10.2.3 Integration in Epic

It is practical if the software which should be produced for the tool is linked to the current software they use at the orthognathic department to store patient data. This way all information is consistent and human errors of selecting the information by hand are minimized. The software the VUmc uses currently is called Epic Systems (Amsterdam UMC, 2019; VUmc / VU University Medical Center Amsterdam, 2017; Ribbers et al., 2018). The software is used at the VUmc by the physicians to write and see medical reports of patients. It is also linked to MyChart, consisting of a platform and app, where patients can log in to see their medical report (figures 10.16). The big reason that the VUmc implemented Epic is to improve information supply to the patient which, therefore, fits the tool. VUmc worked in 2016, while implementing Epic, on storage of imagery (VUmc / VU University Medical Center Amsterdam, 2016). Unfortunately, the VUmc is not clear on how they store their imagery, if Epic is used to store pictures as well and the regulations regarding privacy. More about regulations can be found in chapter 13.1.

The interface of Epic exists of multiple dashboards. The interface can be changed to an own environment of which the user desires it to look (BJC HealthCare, 2017). How the interface looks exactly is different per hospital, but also depends on what extensions the hospital bought (Wagenberg, 2020). Unfortunately, it is not known how the interface of Epic in the VUmc is designed. Therefore, only an example of how facial recognition can be accessed via Epic is shown in figure 10.14. It shows an example of how the interface can look. The position and type of information about the tool can differ when it is applied in Epic at the VUmc.

10.2.4 Integration in MyChart

The communication tool will be printed and given to the patient during the consultation. However, when a patient wants to print it again because he lost the paper or wants to give one to his family, it is practical if he can download it from MyChart (figure 10.15). In MyChart, the head called 'Dossier' (File) shows a subhead called 'Medische tools' (Medical tools) (figure 10.16). It is convenient if the tool would be accessible under the head of 'Medische tools'. The user test (chapter 11) showed that questions pop-up for most patients when they go home from the consultation. Since the communication tool is taken home to process the information more and to increase emotional recovery, it could happen that during processing information, new questions arise. The patients can ask their questions to the treatment team in MyChart under the head 'Berichten' (Messages) (figure 10.17).

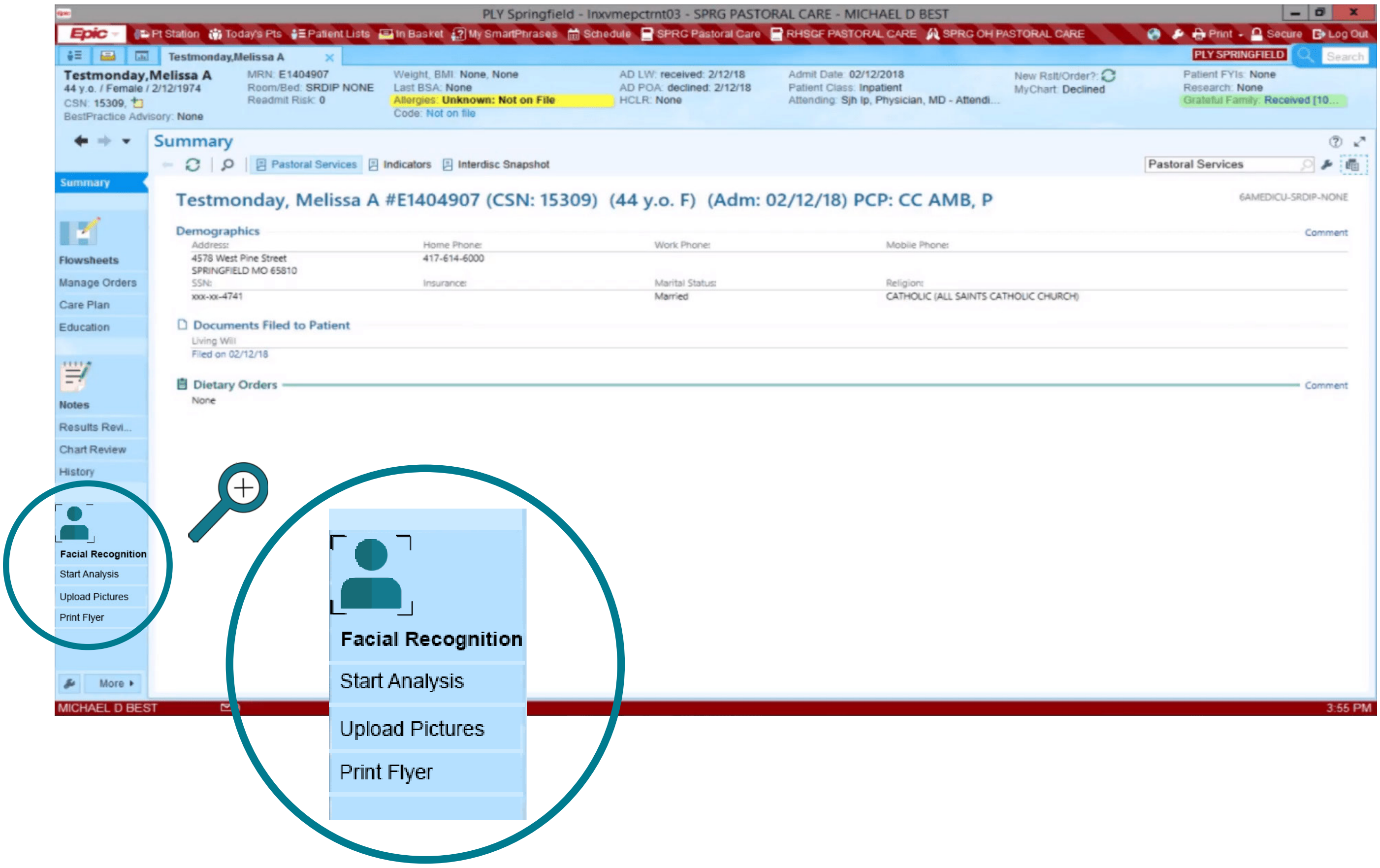


Figure 10.14: An example of integration in Epic with a magnification of the facial recognition. (Software Connect, 2020, adjused image)

Figure 10.15: MyChart, available in Google Play (2020) for all devices

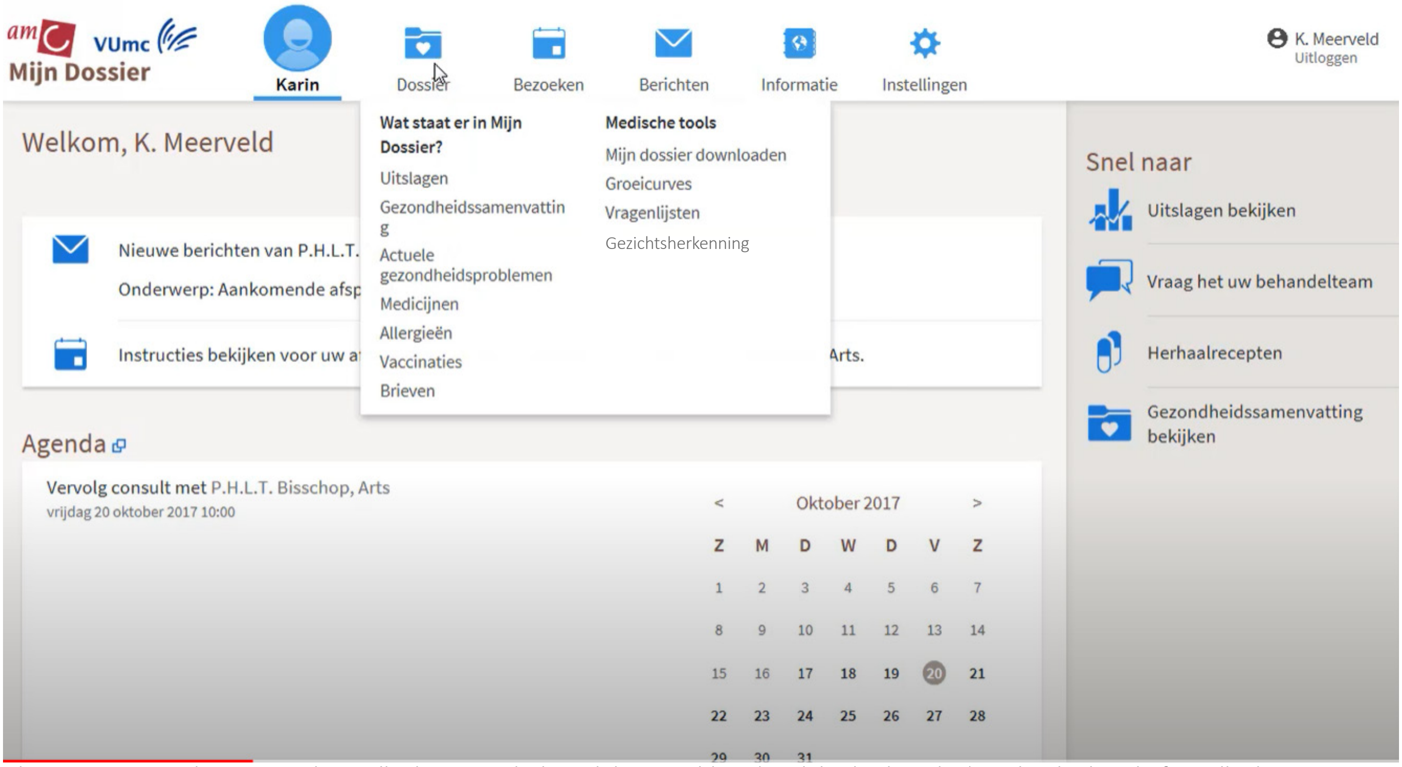


Figure 10.16: MyChart, open de medical report. Find 'Facial recognition' (Gezichtsherkenning) under the head of 'Medical tools' (Medische tools)(Amsterdam UMC, 2017, adjused image)

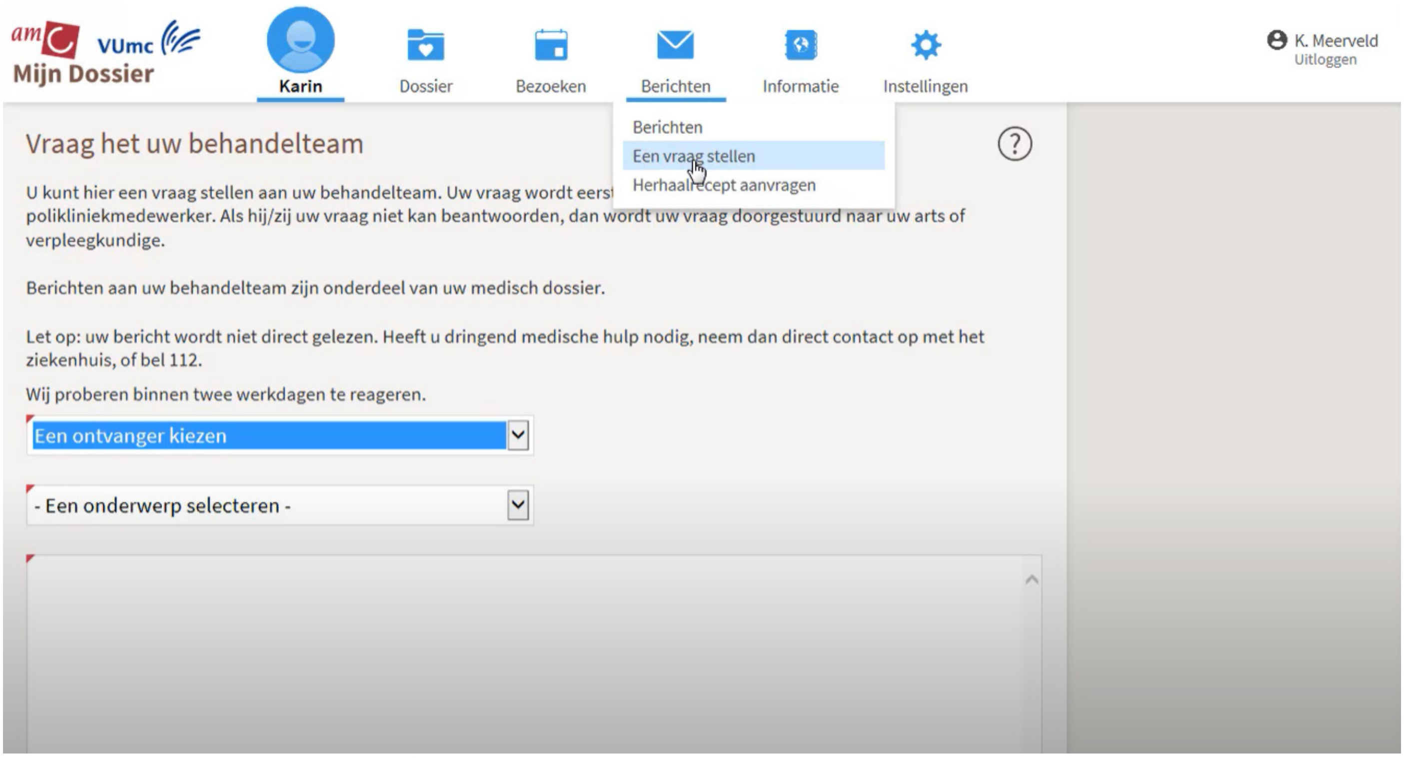
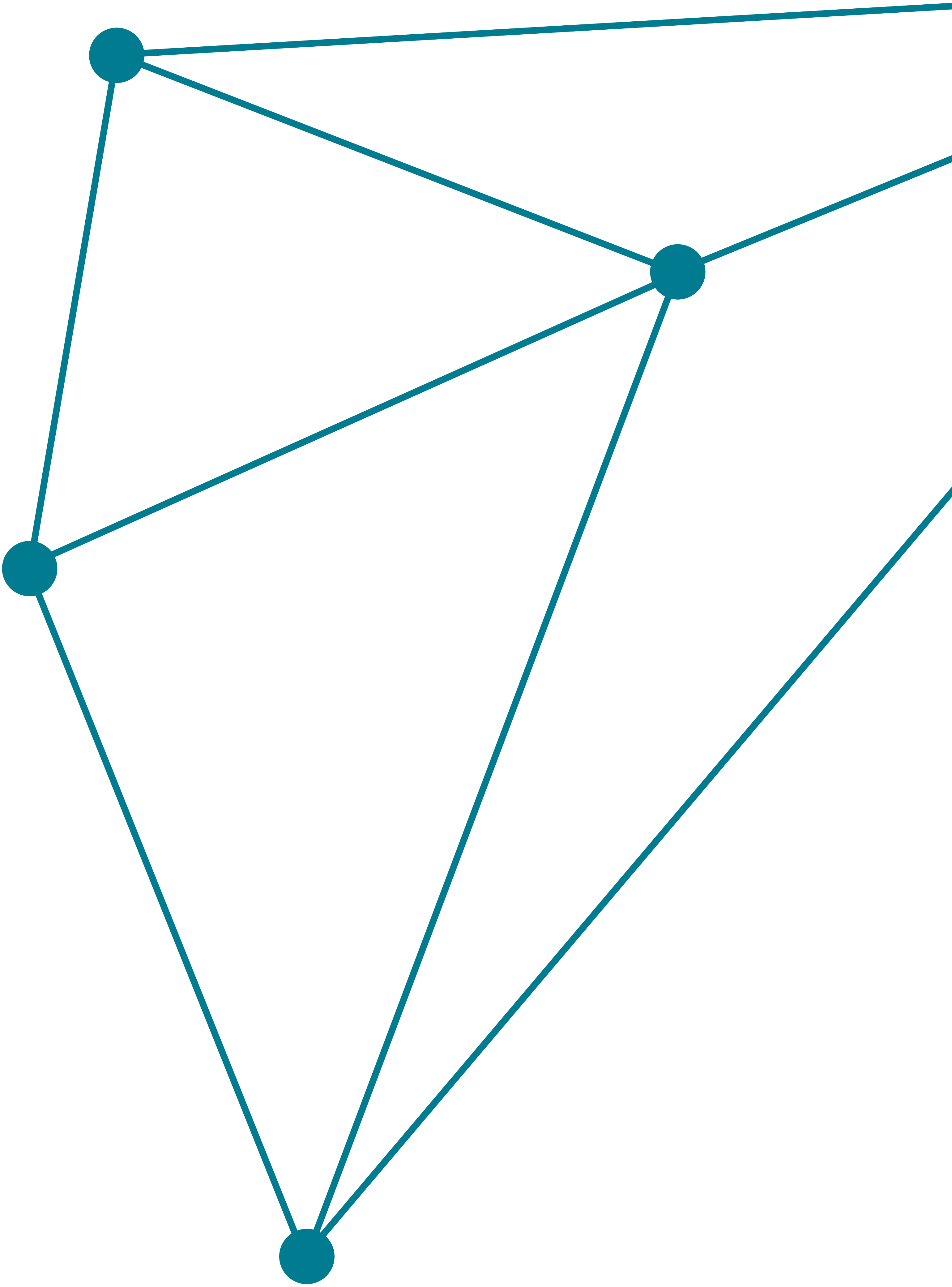


Figure 10.17: MyChart, ask a question (Amsterdam UMC, 2017)



Chapter 11

Validation communication tool

A user test has been done to validate the flyer. The test was done with eight participants from different ages, professions and patient profiles. The user test existed out of four parts, an explanation of the goal of the test and background of the project, patient profiling, a role-play and an interview. The complete test guideline can be found in appendix 28. This chapter shows shortly how the test is performed and what the results are. The flyer was perceived well and some small textual changes have been made to the flyer after the test. The responses of all participants can be found in appendix 29.

Research objective
What do people think of the flyer?

Research questions
1. What is the response when hearing about becoming unrecognizable?
2. What is the balance of visuals and text?
3. How clear are the written explanations?
4. How clear are the visuals?
5. How is the amount of information perceived?
6. Is any information unnecessary or missing?
7. What is thought of the layout?
8. Is the facial recognition clock clear?
9. How is the tone of the flyer perceived?
10. What is the overall thought of this type of communication?

11.1 Method

1 Participants
Eight participants participated whereof four women (age 22, 49, 54, 78) and four men (age 20, 22, 52, 62) with all different professions. Four of them were estimated as introvert patients, two as extrovert and two between intro- and extrovert. All participants could remember the last time they went to see a doctor.

2 Stimuli
The flyer was printed as it should be in the VUmc, on one A4 and with the designed front and back. It was printed in Dutch (appendix 25, (page 118 - 119), since the participants were Dutch.

3 Apparatus
The answers of the interviewees were written down with pen and paper. To guide the test, a test guideline has been printed (appendix 28), to guide the session in a good direction.

4 Procedure
First, a pilot has been done to test the test setup where after it has been improved. The session of the user test started with an explanation of the goal of the test and some questions to estimate the patient profile. After a role-play was done where the interviewer was playing the surgeon and the interviewee was playing the patient. Before this role-play was explained that the interviewee plays the patient on the picture of the flyer and that he is undergoing a certain surgery. The role play was simulated as how the consultation of the VUmc is currently going (appendix 26). During the role-play, it was asked twice if there were any questions. The answers of the participants were written down. After the role-play, questions were asked about the flyer.

5 Measures
The complete response of the participants are shown in appendix 29.

11.2 Results

Most of the participants always go home while still having questions about the consultation. They all thought the way of communicating with a flyer, where after taking it home, could be an improvement for the shared-decision making, information transfer but also for remembering and processing the information from the consultation.

The consultation role play
During the role-play where the consultation was simulated, the interviewer playing the surgeon explained the results of the facial recognition analysis and the emotional recovery while pointing to the flyer. For all the participants who were playing the patient, everything was clear. Three participants had worries about facial recognition and one was not sure if he would continue with the surgery if he would change that much. These participants were all older than 49. The male participants in their early 20's together with the 62-year-old participant, were completely literal about the surgery and they had a 'what must be done should be done' attitude, not worrying about their facial changes and changes in facial recognition. Three participants, whereof two females with the age of 49 and 54 and one male with the age of 52, asked if they would look better after the surgery where after the interviewer answered based on only the functional results of the surgery: Your face will be more in proportion and your mouth will look healthier. But what people perceive as beautiful is different per person. If their appearance improves is, therefore, something which is a concern. The three participants who were concerned about facial recognition, whereof two females with the age of 54 and 78 and one male with the age of 52, understood when empathizing with the patients that when jaw problems were present combined with some dissatisfaction about the appearance, the surgery had to be done. The other five patients were more rational: if there is a problem it should be solved and the facial recognition is just something which you should accept.

The flyer

In general, all participants liked the flyer. They thought it looked professional, structured and it gave a good overview of all the information said by the surgeon (interviewer). The icons and images make it look nice, but also make it easier to find back information. One male participant with the age of 20 had difficulties reading, and also for him, the flyer was easy and fast to read. Four participants look on the internet for information about a medical problem or treatment before the surgery. One participant noticed that the name of the surgeon could be used to look for information about the expertise. The participant was an extrovert patient, so this behaviour matches the research about this patient profile (appendix 23). Extrovert participants gave more comments about the role play and flyer than introvert participants. Two participants commented on the tips. According to them, the tips will not address each patient and it could happen that some tips will be applied by the patient and some tips will not. But according to them, it is good that they are called 'tips', because the patients can choose if the tips will be used or not. The style of writing was rated and perceived well. One participant, whose profession was related to Dutch writing styles, was more critical about the writing and send an email after the test with some points of improvement. The opinion of the colour of the pictures differed. Two participants thought the pictures with the tips were too flat and the original colour would look nicer and more clear while one participant liked the flat colour of the images with the tips, he thought it fit the colour scheme and it made the overall look calmer then what it would be with the original colour. He also thought it could look nicer to change the patient picture on the front page to a flat colour as well. Since these opinions were personal and most of the participants liked it, nothing was changed about the colours of the pictures. The blue-green colour scheme was perceived well. They thought it fitted the medical purpose and it looked calm, which was the intention for the design as well (appendix 26). One participant, with knowledge in graphical design, thought the format of the A4 paper looked cheap and could be more luxurious. The interviewer explained that the results of the analysis of the patient picture and printing the flyer is done right before the consultation so the A4 paper is chosen so the hospital does not need any extra recourses and it can be done fast. However, the participant thought the surgery costs a lot of money, so they could also spend money to create the flyer with high-quality paper, folded with a more volumized feeling and to print it at a printing company. However, this participant used to create luxurious flyers during his graphical design profession so he was critical about this. The other participants thought it looked all good. The facial recognition clock to represent the score was understood by all participants. One participant thought that the red colour in the clock relating to 'unrecognizable' could be perceived as bad or negative. But she also thought it could make patients aware that something was really happening with their face. Two patients had a comment about the name of the score which was described as 'Low' under the head 'my facial recognition'. According to them, the meaning could be clearer if it was written down like 'Score: Low.' This has been improved in the flyer.

11.3 Discussion

Of the participants, one female participant of 54 years old, did undergo orthognathic surgery in the past. The role play was done for the participants to empathize with the patient. However, this is not the same as actually being a patient and getting a result which is about your personal changes in facial recognition. It is expected that orthognathic patients react differently when they get the flyer about their own face. Therefore, this flyer should be tested with real orthognathic patients to test how they react on it and if it supports emotional recovery. Three participants already found becoming unrecognizable because of an orthognathic surgery quite shocking, which would probably be even more shocking for real patients.

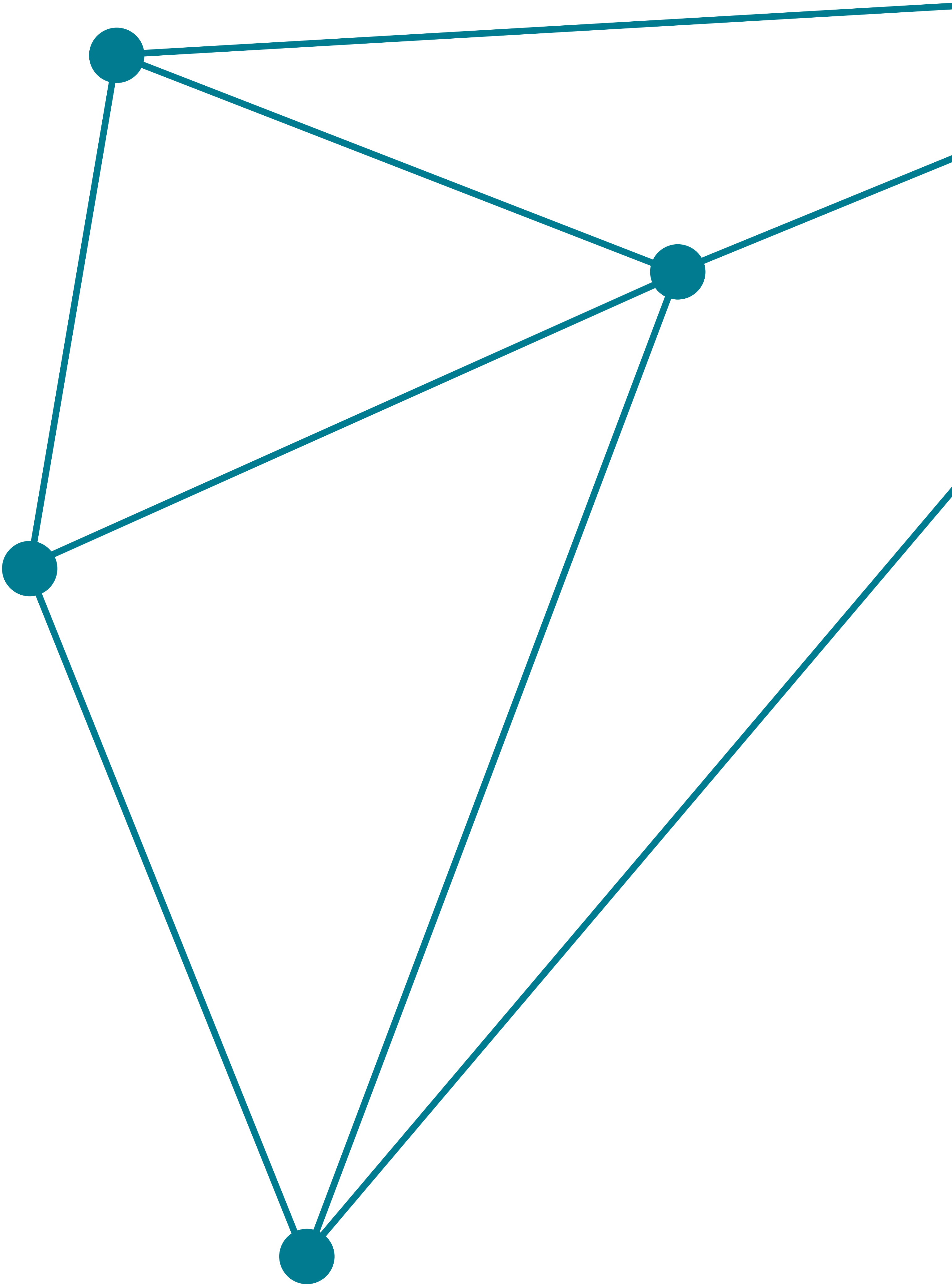
The ages, professions and patient profiles were distributed which shows a complete result of the understanding and opinion of a variety of people. However, the 'supported' and 'unsupported' patient profiles which can be combined with 'introvert' and 'extrovert' (appendix 23) are not tested in this test although they do influence emotional recovery. It is expected that the flyer will both give extra support to the 'supported' profile, but this profile might not need it, and to the 'unsupported' profile, from whom it is expected they need the extra support. The flyer should be tested on both of these patient profiles as well, to test if extra support outside the flyer is needed, especially for the unsupported profiles. The images and icons are designed to remember the information better and for patients who have difficulties reading or are illiterate. None of the participants was illiterate, so if this works for this type of patients is unknown. However, the test did show that the flyer is easy to read for people who do not read that easy, it is attractive for people to read and the visuals improve reminding information.

11.4 Conclusion

The flyer was perceived well by the participants and only a few comments were given. The opinions on colour choices did differ between participants. Since this is a personal opinion, nothing will be changed regarding colour. Some writing style optimizations will be made to the design, but the layout will not be changed. The layout and colours of the flyer were chosen with the use of research and the psychological model called 'the theory of psychological adaptive modes' was used to design for the emotional recovery. In theory, it all should have worked. According to this test, it also works in reality. However, the participants were not orthognathic patients and a test with patients should be done to discover if the tool really works for patients. At the end of the project, the facial recognition clock has been changed to the one in chapter 10, this had nothing to do with the user test but was done to improve patient experiences (appendix 24).

Key-points from this chapter

- The flyer was perceived well and only a few comments were given
- The flyer has been optimized after the user test
- The design decision based on research did provoke the reactions which were desired and designed for
- Younger people (early 20's) are less worried about facial recognition changes than other people (the age of 49+)
- 3 of the 8 participants asked how their appearance would be after the surgery and were worried about facial recognition.
- A test with real orthognathic patients should be done to see if the information is presented clearly, how it is perceived and how it contributes to emotional recovery



Chapter 12

Validation facial recognition system

Two tests are done to validate the functioning of the facial recognition system. Without functioning of the facial recognition system, the communication tool does not work as well. A dimension test has been done, proving that dimensions influence facial recognition and that facial recognition depends on the patient's face and type of surgery. Also, a test has been done to prove that predicting similarity scores before an orthognathic surgery can be done. This test also shows a difference in similarity scores and patient expectations which are caused by an inaccurate human threshold and age difference. First, the setup of the dimension test is shown whereafter the setup of the proof of predictability.

12.1 Dimension test

For the design of the facial recognition system has been chosen to take the dimensions into account. The reason for this is because dimensions of the surgery do have an influence on the amount of displacement of landmarks and therefore, it influences facial recognition. A dimension test has been done to prove this. The communication tool and facial recognition system are both build upon the fact that facial recognition is dependent on the face of the patient and on the type of surgery, which was the conclusion of the human and computer recognition test (chapter 8). This dimension test shows this fact by comparing manipulations of landmark movements and by showing a comparison between the similarity scores of an equal manipulation performed on two different people.

Research objective
In what way do the dimensions influence the similarity score?

Research questions
1. What dimensions cause a similarity score passing the human or computer threshold?

2. Do the directions of landmark movements influence the results?

3. Do the results differ when another face is used?

Hypothesis
The bigger the dimensions are and therefore the landmark displacement, the higher the similarity score. This is based on the knowledge that when landmarks displace more, the face distance the code calculates will be bigger which results in a higher similarity score.

12.1.1 Method

1 Stimuli
One picture in frontal view is taken and manipulated for combinations 1b, 1c, 2b and 2c in four different dimensions (figure 12.1). Another picture in frontal view from the same person, person 1, has been taken which is used as the picture to compare with (figure 12.2). A picture from person 1 is made in frontal-lateral view so the landmark movements of 1a and 2a movements could be manipulated as well. Another picture from person 1 in frontal-lateral view has been made and has been used as the picture to compare with (figure 12.3). Two pictures in frontal view have been made from another person, person 2 (figure 12.7). One of these pictures is manipulated for combination 1c in the four different dimensions. This manipulated picture is compared to the other picture which has not been used for manipulation. Person 2 has been compared with person 1.

2 Apparatus
The test has been done in the Spyder software, with the code (appendix 15) written in the programming language of Python. The dlib and face_recognition libraries have been used (chapter 3.3).

3 Procedure
The procedure is further examined in appendix 29. The picture has been manipulated manually with the use of the Warp tool in Photoshop (appendix 11), the same as done in the human recognition test (chapter 6). The picture has been manipulated according to the landmark movements from the classification system (chapter 4). A grid was placed in the picture, on the face, to create dimensions in the picture which would be the same during each manipulation. The chosen dimensions were in millimetres. There has been decided on dimensions by measuring the face at the maxillary plane, the possible dimensions are estimated from this and are shown in figure 12.1. All the manipulated pictures can be seen in appendix 30.
Another picture from the same person has been made in a similar position as the frontal image. This picture is compared to the manipulated pictures with the code. This is done because the pre- and post-surgical pictures of the patients from the data (chapter 7) are also not made on the same day or in the exact same angle. It is, therefore, closer to reality. Also, the two pictures from the same person without any manipulations have been compared

Manipulation step	1	2	3	4
Dimensions of landmark movements [mm]	5	10	15	20

Figure 12.1: The manipulation steps

to see if the faces were seen as similar (figure 12.2 and 12.3). Secondly, two pictures from another face, made on another day, have been taken (figure 12.7). One of them has been manipulated with the 1c-down landmark movements in all four different dimensions. After, the manipulated pictures have been compared with the other not manipulated picture of this person to compare the results between the two faces. Also, the two pictures from the same person without any manipulations have been compared to see if the faces were seen as similar.

12.1.2 Results

The two frontal pictures were compared without any manipulations and score a similarity of 0.35 and the frontal-lateral pictures score a similarity of 0.26. This is positive since this means the faces are seen as similar and these pictures could, therefore, be used for the analysis. The dimensions do have an influence on the similarity scores for either manipulation in the mandible and maxilla (figure 12.4, 12.5 and 12.6). The bigger the dimension, the higher the similarity score is. This is meeting the hypotheses. How much the score increases depends on the manipulation.

Mandible

The 2c-shortening manipulation, which is vertical shortening of the mandible, and the 1c-lengthening manipulation which is vertical lengthening of the maxilla, already achieves the threshold of 0.45 at the lowest dimension of 5mm. This meets the results of the computer and human recognition test (chapter 8). It does not automatically mean that if the displacement is increased it will succeed the threshold line, since the 1c-lengthening manipulation did not exceed the threshold at any displacement. Moving the mandible forwards (1a) did not succeed the threshold line. The 2b-narrowing manipulation does achieve the threshold at a lower displacement of 10mm than the 2b-widening manipulation which achieves the threshold at 15mm. Also, the increase of the two lines differs.

Maxilla

The manipulations 2b-narrowing and 2c-lengthening manipulations, which is narrowing and vertical lengthening of the mandible, have the same increase when the displacement goes up. But the 2c-lengthening manipulation already exceeds the threshold line at the lowest displacement while the 2b-narrowing exceeds it at a displacement of 15mm. The 2b-widening and 2c-shortening manipulations do have almost the same similarity scores. Moving the maxilla forwards (2a) did not succeed the threshold line.

Comparison of two faces

The two pictures of the other person have been compared as well, without any manipulations, and has a similarity score of 0.36 which means this picture could also be used for the analysis.
Figure 12.8 shows that the same manipulation with the same displacement does give other results. Person 1 is the person used for the previous analysis and person 2 is the new person. Both lines have the same slope between 5mm and 10mm. Where the slope increases linear for person 2, between 5 and 15mm, it is not for person 1. Also, when the first displacement of 5mm is done, person 1 is on the computer threshold line while person 2 is not, so person 1 is unrecognizable while person 2 is still recognizable. Therefore, it can be concluded that the similarity scores do increase when the displacement increases. However, the exact scores and the slope is dependent on the face. Therefore, the similarity score should be predicted per patient.

Figure 12.2: On the left, the F picture used for the manipulations. On the right, the F picture used for comparing the similarity score. The similarity score of these two pictures is 0.35.

Figure 12.3: On the left, the FL picture used for the manipulations. On the right, the FL picture used for comparing the similarity score. The similarity score of these two pictures is 0.26.

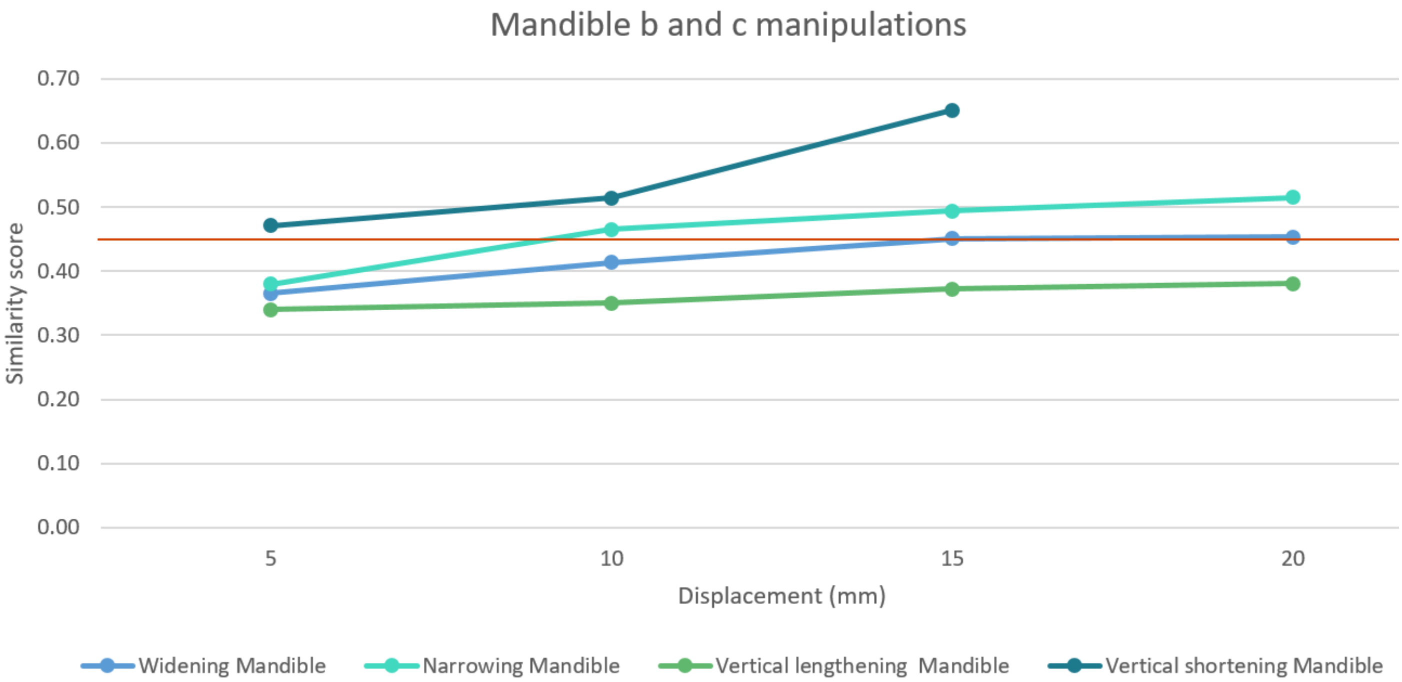


Figure 12.4: Manipulations b and c on the mandible

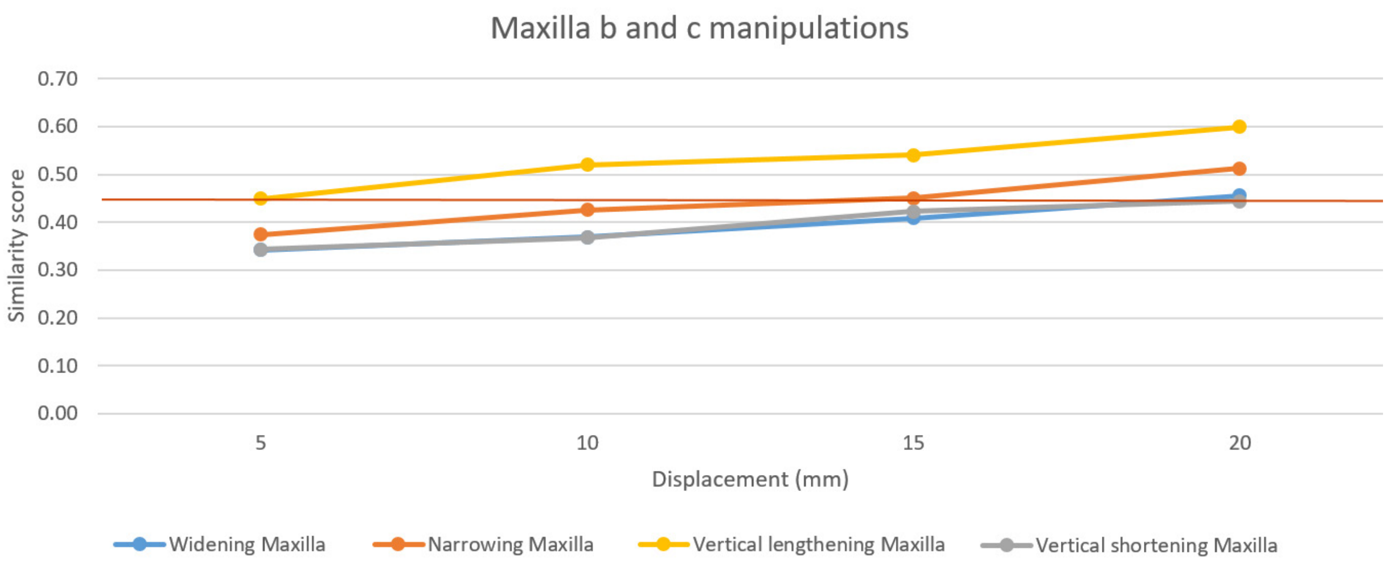


Figure 12.5: Manipulations b and c on the maxilla

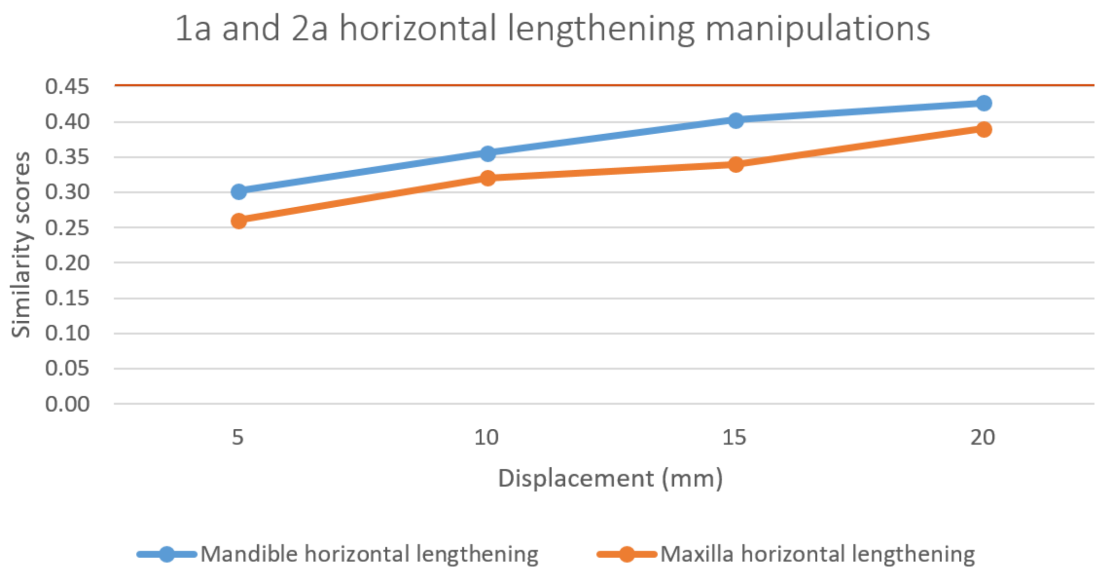


Figure 12.6: Manipulations on the mandible and maxilla

Figure 12.7: Pictures of another person. On the left, the F picture used for the manipulation. On the right, the F picture used for comparing the similarity score. The similarity score of these two pictures is 0.36.

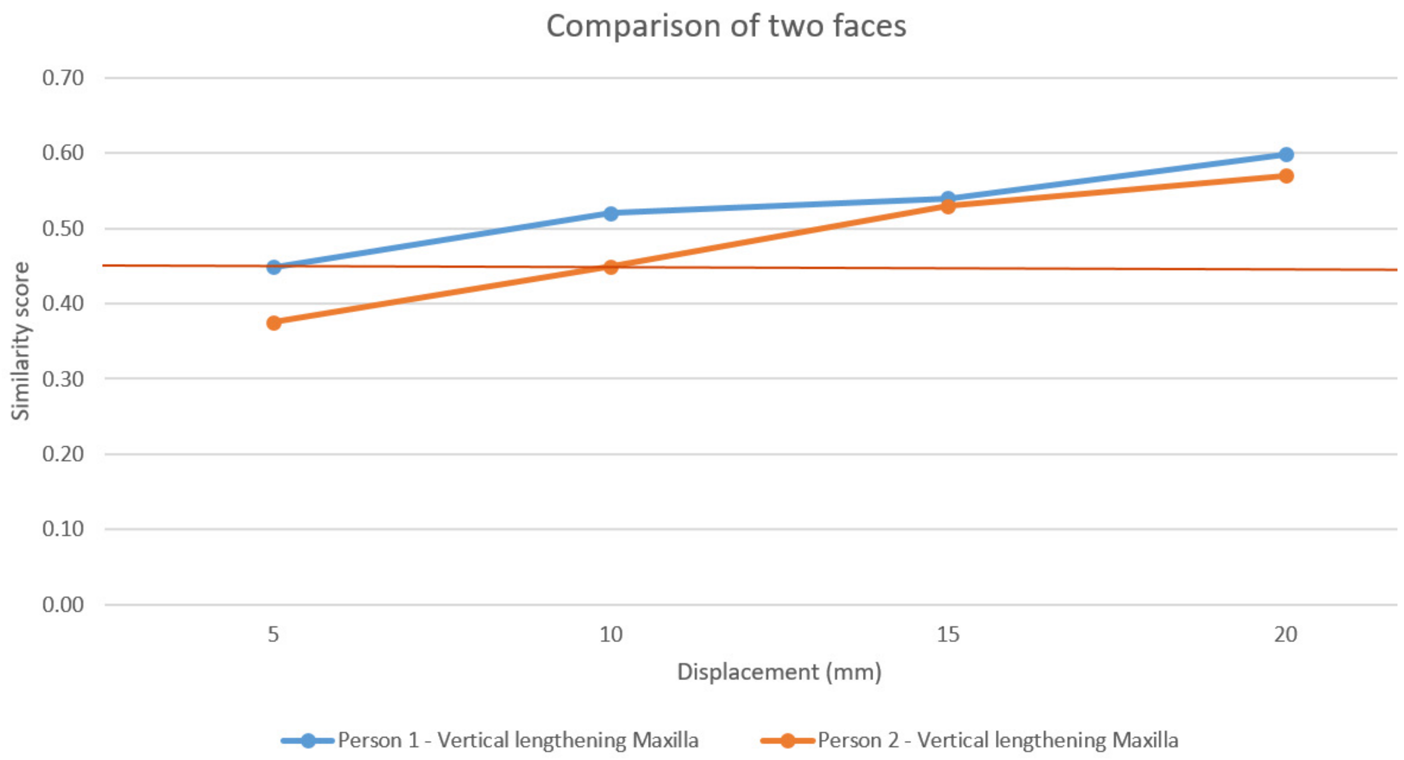


Figure 12.8: Comparison between two people with both the same manipulation

12.1.3 Discussion

When the dimensions were getting higher, the face did look less realistic. Therefore, it is expected that the displacements of 15mm and 20mm with only one type of landmark movement are rare in reality. However, when it is combined with another type of landmark movement, the planes in the face will align more instead of getting holes or bumps which was the case with higher displacements. Only single landmark movements were tested in this test. More research in how manipulating multiple landmark movements can be done best should be done because yet it is not know how this can be done in an accurate way.

The prediction of how the face will change will always be a prediction and is it expected that it will not predict the exact same score. When predictions are done on patients, the manipulated picture should therefore never been shown to them, since it can give them wrong expectations about how they can look. However, this research shows that the similarity score is depending on the face and surgery, it can, therefore, give an indication of how much the face will change related to facial recognition.

The manipulations from the frontal-lateral view could have been less accurate than the ones in the frontal view since both the frontal as the lateral plane could be seen, an estimation had to be made how the face would move in 3D. Also, creating the manipulations cost ca. two days. Since it is time-consuming, the tool should be able to manipulate the picture automatically. How this can be done is described in chapter 10.2.2.

From this test is not clear if the predicted similarity scores of patients meet the real similarity score of the pre- and post-surgical picture. Therefore, a test has been done in chapter 12.2 which proves the method works.

The angle of the face in the picture does influence the similarity score. When the facial recognition system is applied in practice, a controlled environment should point the faces in the same direction. Also, person 2 had some hair in front of his face. Since person 2 had a lower similarity score in comparison with person 1, it is expected this did only influence the results of this test a little. However, a controlled environment in the VUmc should remove hair from the face to decrease errors.

12.1.4 Conclusion

The dimensions do influence the similarity score, the bigger the landmark displacement, the higher the similarity score is. This is also meeting the hypothesis before the test. How much the similarity score increases with the displacement depends on the type of manipulation and, therefore, it depends on the type of surgery. It can be concluded that the directions of landmark movements do influence the similarity scores and therefore all the landmark movement annotations should also be divided into directions (e.g. 1b-narrowing and 1b-widening instead of only 1b).

The similarity scores can differ per face, therefore each patient should get an own similarity score. Not all type of landmark movements have been analyzed because manually manipulating the pictures is time-consuming. This test did prove, among other things, that landmark movements in both the frontal view as frontal-lateral view can be manipulated by using dimensions.

12.2 Proof of predictability

The concept predicts the similarity score of patients before they undergo the surgery. Unfortunately, this type of application has never been done before. This chapter shows a test on two patients who went through orthognathic surgery and compares the predicted similarity scores to the real ones from the pre- and post-surgical pictures. This test shows that predicted similarity scores by using manipulated pictures can indeed match the real similarity score. Also, it shows that the experiences of these two patients do not match the results of facial recognition and this is expected to be caused by the familiarity of the human threshold and the age on which the surgery was performed.

Research objective
How well do the predicted similarity scores match the real similarity scores and the patient experiences?

12.2.1 Background information patients

Patient 1
This patient had to undergo two separate surgeries, the first done in 2006 and the second in 2008 on the age of 43. The first one was widening the maxilla with a SARME surgery (surgically assisted rapid maxillary expansion)(Zuyderland Medisch Centrum, 2018). The maxilla was made weaker with the use of cuts in the bone. A distractor was placed in the bone of the maxilla and she had to turn the distractor to slowly widen the maxilla which created a gap between the anterior teeth 8 and 9 (appendix 7). Pictures have been found from two years before the first surgery and from after the first surgery when the maxilla was already widened with the distractor. Also, a picture was found from after the second surgery, which rotated the maxilla and horizontally lengthened the mandible. Unfortunately, no picture was found between the first and second surgery. The experience of the patient was that she was happy with her appearance after the surgery but it hurt her that she was not recognized by some distant acquaintances she had not seen in a few years. An interview with patient 1 can be found in appendix 2.

Patient 2
The surgery was done in 2013 at the age of 18 years old. A set of pictures has been used from 2012 where the patient was 17 years old. The patient's maxilla was rotated and the mandible was lengthened horizontally. This happened during one surgery. The patient did not have problems with her appearance before the surgery but she did like the appearance of her teeth and jaw more after the surgery. She had the feeling she did not change that much and she was recognized by everyone.

12.2.2 Method

1 Stimuli
Pictures of two patients who went through orthognathic surgeries have been used for the analysis.

Patient 1 went through two surgeries, widening the maxilla performed in 2006 and rotating the maxilla with horizontal lengthening of the mandible in 2008. From this patient, two pictures were found in her personal archives before the first surgery, made on the same day in 2004, one picture from between the surgeries made in the hospital and one picture after both surgeries made in the hospital.

Patient 2 went trough one surgery in 2013, rotating the maxilla and horizontal lengthening of the mandible. From this patient, five pictures were found in her personal archives. In the first test, two pictures before the surgery made on the same day in 2011 and one after the surgery in 2014 were used. In the second test, one picture before the surgery in 2010 and one after the surgery in 2020 were used.

2 Apparatus
The test has been done in the Spyder software, with the code (appendix 15) written in the programming language of Python. The dlib and face_recognition libraries have been used (chapter 3.3).

3 Procedure
First, for both patients the two pictures before the surgery, the similarity score was calculated. just like done in the dimension test (chapter 12.1). For patient 1 the similarity score was 0.26. For patient 2 the similarity score was 0.36. Therefore, these pictures could be used for the analysis.

Patient 1
One of the pictures from 2004 has been manipulated twice, according to the landmark movements belonging to the first surgery, with a displacement of 5 and 10mm. The similarity score has been calculated between the manipulated and non-manipulated picture from before the first surgery. The similarity score also has been calculated between the non-manipulated picture from before the first surgery and the real post-surgical picture after the first surgery. Also, the similarity score of the non-manipulated picture from before the first surgery and the real post-surgical picture after the second surgery has been calculated and this has been compared with the patient's experiences.

Patient 2
One of the pictures from 2011 has been manipulated according to the landmark movements belonging to the surgery. The similarity score has been calculated between the manipulated and non-manipulated picture from before surgery and is compared with the similarity score between the non-manipulated picture from before the surgery and the real post-surgical picture after the first surgery.

The similarity score between one picture from before the surgery in 2010 and from after the surgery in 2020 has been compared to the predicted similarity score as well to prove that multiple picture sets could get the same result.

12.2.3 Results
Figure 12.9 and 12.10 show all the pictures with belonging similarity scores used in this test. The results are separated into patient 1 and 2.

Patient 1 (figure 12.9)
The predicted similarity score of both 5mm as 10mm lie close to the similarity score of the non-manipulated pre- and post-surgical pictures. It is unknown what the dimensions of the surgery were and what the result would be if the corresponding dimension would be applied. However, the landmark movements of the manipulated picture lie close to the real surgery according to the similarity scores. It is expected that the optimal dimension lies between 5 and 10mm. By finding this optimal value, the prediction of the similarity score can be improved. In the future, more research in optimizing the range in dimensions and if manipulations of the real dimensions correspond with the similarity score of the
pre- and post-surgical picture should be done.

The comparison of the picture before the first surgery has been made with a picture after the second surgery where the similarity score was 0.42 (last pictures of figure 10.9). This means that the score lies close to the human threshold but does not achieve it. This does not meet the experiences of the patient. Although she was recognized by most people she did not see for years did not recognize her. The human thresholds calculated from the human recognition test in chapter 6 were based on a group of people who were familiar to each other and they saw each other regularly. A distinction between the amount of familiarity has not been made. Further research should calculate a new human threshold for people who do not have a relationship and only see each other once in a few years. It is expected that this threshold will lie below the 0.42 score of this patient.

Patient 2 (figure 12.10)
The predicted similarity score of 0.54 does lie close to the actual similarity score of 0.52, which means the patient is unrecognizable. This means the score can be predicted on teenagers above the age of 14 although it is unknown what the influence of age and age difference is. However, when this is done in the hospital, the pictures will be made close to the surgery, which will make the age difference not bigger than two years. It is therefore expected it will not be a problem in the hospital when the surgery is performed at an age of 18 years old. However, when patients are younger this could be a problem since the face changes more during childhood than during adulthood and this should be further examined when the tool will be implemented.

The similarity score did not meet the experience of the patient, she had the feeling she did not really change while the similarity score passes the computer threshold spaciously. It is expected that the similarity score lies higher than the patient's experience because the patient was still a teenager. The facial changes caused by age could, therefore, have influenced the similarity score more than the surgery did.

A picture from 2010 before the surgery and 2020 after the surgery are compared, to see if the similarity score is not based on probability. The surgery was done in 2012 so there is even more age difference between the pictures. The similarity score of these two pictures is 0.53. This is almost the same as the similarity score of 0.52. This means that the similarity score can be trustworthy and can be the same although different pictures are used.

Figure 12.9: Comparison between pictures of patient 1 with its description on the right.

These pictures of patient 1 are both made in 2004 on the same day. The similarity score of these pictures was 0.26. This means they are seen as similar and can be used for the rest of the analysis.

The right picture has been manipulated, the maxilla has been widened 5mm. The similarity score is 0.31.

The right picture has been manipulated, the maxilla has been widened 10mm. The similarity score is 0.34.

The right picture is the real post-surgical picture after the first surgery and after the patient widened the distractor. The similarity score is 0.33.

The right picture is the real post-surgical picture after the second surgery. The similarity score is 0.42.

Figure 12.10: Comparison between pictures of patient 2 with its description on the right.

These pictures of patient 2 are both made in 2011 on the same day. The similarity score of these pictures was 0.36. This means they are seen as similar and can be used for the rest of the analysis.

The right picture has been manipulated, the maxilla has been rotated 5mm and the mandible horizontally lengthened 5mm. This is the only performed dimension because of the difficulty of the manipulation. The similarity score is 0.54.

The right picture has been manipulated, the maxilla has been widened 10mm. The similarity score is 0.52.

The left picture from 2010 before the surgery and the right from 2020 after the surgery are compared. The similarity score is 0.53.

12.2.4 Discussion

Only two patients have been analysed in this test. Therefore, it is unknown if the predicted similarity scores matching the real similarity scores are based on probability. More patients should be analysed in the future to find out if predicting the similarity score is statistically significant or that the manipulation technique needs to be optimized.

The pictures which the hospital should take from the patients should be in the same setting, meaning the patient should look to a certain point putting the head of the patient in the same angle. Since this test was done on patients where most pictures were chosen from their personal life, it was impossible to find pictures with the same quality, distance to the face, position and facial expression. A controlled setting can improve the outcomes.

The hair of the patients in the pictures sometimes covered part of their face. Since the method used uses geometry to generate the similarity scores, this could have influenced the results. The hair of patient 1 always covered the same part of the face in all pictures so it is expected the similarity scores are higher in all of the comparisons of the patient pictures and the results of patient 1 are, therefore, more consistent than of patient 2.

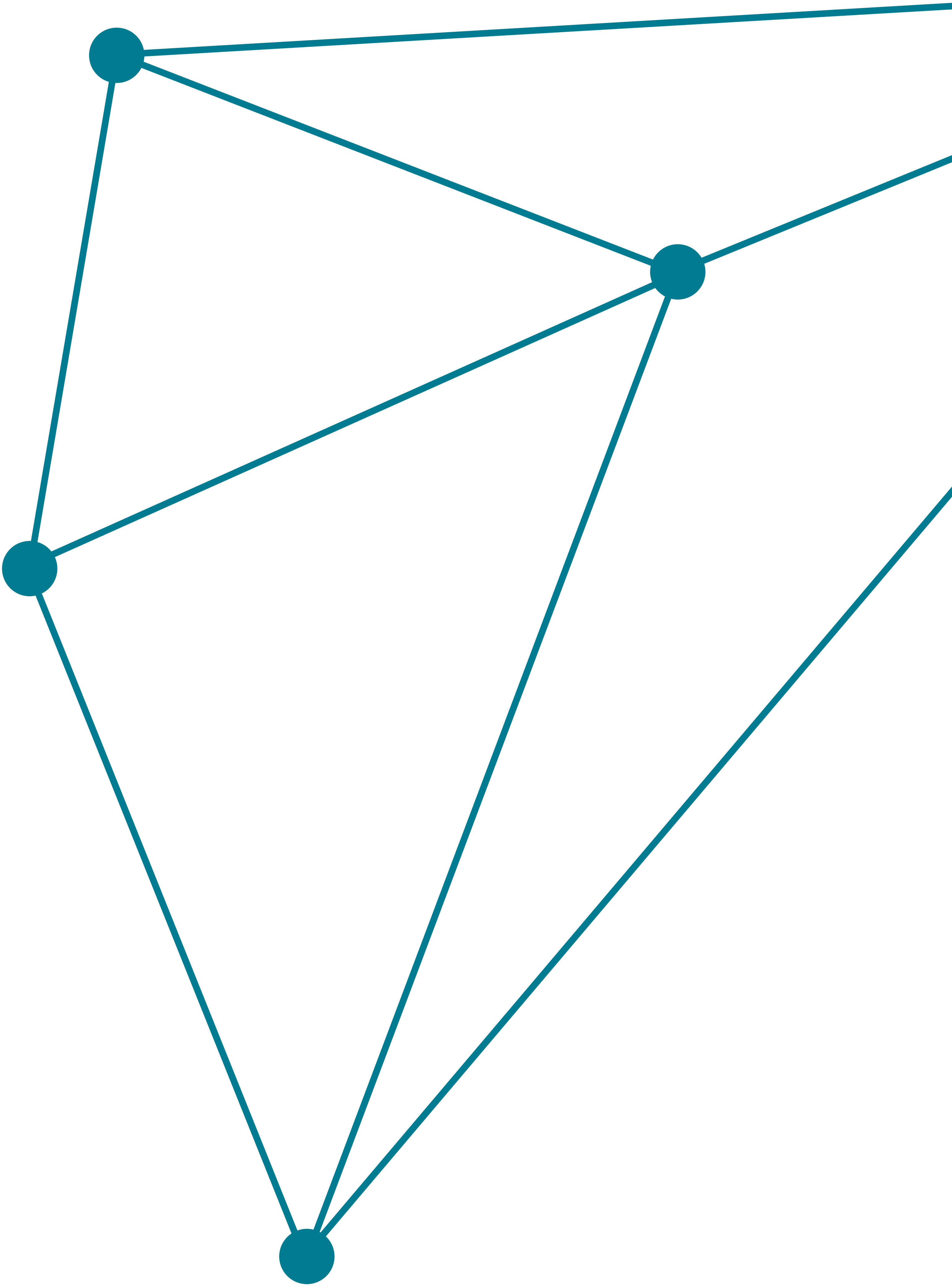
It was hard to manipulate the picture of patient 2 because both surgeries changed the face on all planes of the face. In the hospital, pictures in frontal-lateral position should be made so manipulations can be created in all directions. In the future, 3D scans could be used to create 3D manipulations to easen manipulation and to optimize the similarity score even more. Also, since two surgeries were performed two different manipulations had to be done. The way this manipulation is performed is estimated since there has been done research on only single manipulations in this project. In the future, researching multiple manipulations should be done as well, to find out how the landmarks are moving when multiple surgeries are done. This can improve the predicted similarity scores.

12.2.5 Conclusion

The most important founding, which is the foundation for the whole concept, is that the similarity score can be predicted. However, only two patients are analysed and more patients should be analysed in the future to find out if the prediction method is statistically significant.
The patient experiences of both patients do not match the similarity scores. For patient 1, it is expected that this caused because the human threshold should be lower for people who are less familiar then tested in the human test (chapter 6). Therefore, a distinction needs to be made in the amount of familiarity. For patient 2, it is expected that this is caused because the age difference has not been taken into account in the similarity score. Therefore, the high similarity score could have been caused by facial changes caused by age difference instead of the surgery.

Key-points from this chapter

- Dimensions do influence facial recognition
- Facial recognition changes caused by an orthognathic surgery depends on the patient's face. Therefore, facial recognition changes should be tailored towards patients and each patient needs an own similarity score
- Facial recognition changes caused by an orthognathic surgery depends on the type(s) of surgeries performed
- Similarity scores can be predicted
- More research should be done in the effect of the amount of familiarity and facial recognition. Another human threshold could evolve from this
- Age difference could probably higher the similarity score. Resulting in not meeting patient experiences



Chapter 13

Future implementation

Before the concept can be implemented at the orthognathic department in the VUmc, some steps still have to be made. The first step in researching facial recognition after orthognathic surgery is done in this project and the concept shows an idea of how this research could be implemented in practice in the future. However, to be sure the concept is feasible, more research should be done and the concept should be optimized. This chapter will, therefore, give a recommendation to the VUmc about how they can continue with this project and how they can eventually, implement the concept. A roadmap is shown, which the VUmc can follow. This includes a testing plan, the regulations and development plan. Also, an overview of what the investment costs and benefits of the concept would be, are shown. Finally, recommendations for the VUmc are described.

13.1 Roadmap to implementation

If the VUmc is going to implement the concept, steps should be undertaken before this can be done. Therefore, a roadmap to implementation is shown step by step. Research in regulations has been done in appendix 33 and an exploratory description of the research gaps which are filled and still open can be found in appendix 1. The roadmap on pages 135 till 137 shows a short description of the things which could be done by the VUmc.

Figure 13.1 shows the gaps which were present before this project, defined within the scope, which are now filled in. It also shows the gaps which are present after this project. These gaps are estimated and probably more gaps will pop-up when future research is done. The future gaps can be filled up when the roadmap to implementation is followed.

Gaps before project
2.1 Regions
2.2 Familiarity
2.3 Combinations
2.4 Identic face
2.5 Displacements
2.6 Predictability
2.7 Communication

Gaps after project
2.8 Prediction all surgeries
2.9 Manipulate rotations
2.10 Complex surgeries familiar group
2.11 Specific area
2.12 Type of familiarity
2.13 Age
2.14 Expectations vs. similarity score
2.15 Experience Fraos
2.16 VUmc Data
2.17 3D pictures
2.18 Texture, expression, quality
2.19 Influence orthodontist
2.20 Surgery protocols
2.21 Patient evaluations

Figure 13.1: Defined gaps before and estimated gaps after the project

1. Research the gaps

This project covers only part of the knowledge in the topic.

Similarity score

If **prediction** of the similarity score is trustworthy should be examined on patients in the VUmc.

When reshaping the jaw, the **local skin texture** around this may also change which can influence facial recognition.

The influence of **age** on the similarity score. Puberty influences facial changes, therefore it might be possible that the concept cannot be used for the age < 18.

The influence of the **orthodontist** and wearing braces on the similarity score is still unknown.

Influence of **quality and illumination** of picture, angle of the face, expression on the accuracy of the similarity score.

VUmc data

A difference in surgical techniques between the used 75 patient pictures and the techniques the VUmc uses in their hospital could have influenced results. The VUmc could analyse their patient pictures with the same method as used in this project to optimize the results.

Familiarity

The **threshold** of recognizability of a patient can be further researched to find a threshold per type of familiarity, e.g. for people who are far acquaintances and therefore not very familiar.

Landmarks

Higher accuracy of predicting similarity scores can be achieved when more research is done in how the landmarks move per surgery. In this research, only estimations are made.

A distinction in manipulating a **specific area** of a certain region belonging to specific surgeries should be further explored. This could have a different influence on landmark movements and the similarity score.

Manipulations

In what way the manipulations can be done best is unclear. The manipulations in this project have been done manually which could have lead to an inconsistency in the way of manipulation.

Multiple surgeries, rotational surgeries and surgeries on the lateral plane are difficult to manipulate because it is an in-depth movement, which is hard to do in 2D. Further exploration in **3D images** could solve this problem.

Patients

Evaluations of patients should be integrated into future research and when implementation of the concept starts. A distinction between ages and types of surgeries should be made when evaluating since this could influence perceptions.

Optimize concept accordingly

2. Prepare and execute a test

Before the concept can be developed, tests should be done

Regulations

A **test protocol** should be written for the METc committee at the VUmc when patients are involved in the research called WMO (Scientific Human-related Research).

The **most important regulation** when developing and implementing the concept is the Norm HRP.3 ME3 part of article 12 of the WMO, which means the requirements include that sponsors protect the privacy and confidentiality of the subject data.

The **AVG regulations** apply when pictures of people who are recognizable are made and saved online.

Exchange of patient data is already done in the VUmc by using Epic, which should follow the EGiz (Gedragscode Elektronische Gegevensuitwisseling in de Zorg) and NEN 7510:2011 and 7513. They should make sure the exchange of patient pictures also follows these regulations when doing tests.

Test group

Find a test group of **patients** including a variety of patient profiles and differences in surgeries.

Find a test group of **physicians** to get feedback from the users of the concept.

Develop a pilot plan

Decide on the goal by determining things like the planning, time available, approach, scale and budget.

Manually simulate the operation of the concept. Create a personalised flyer per patient, this is the prototype of the concept.

Execute the test

Test the prototype of the concept during the **consultation**.

Test if the **workflow** of the facial recognition system part of the concept is integrated enough in the workflow of the VUmc.

Evaluate if and how the concept contributes to **emotional recovery** of patients.

Evaluate **experiences** of patients and surgeons.

Optimize concept accordingly

3. Develop the concept

Develop the final concept and train surgeons to use it

Hire a software engineer

The software engineer should develop the software of the **facial recognition system**.

When the software is developed it should be **integrated into Epic**. He can contact the MS (multiple sclerose) department of the VUmc where they already integrated software Epic, so their method can be redone.

The concept could be integrated into **the brochure** the VUmc hands out during the current consultations, to create a consistent lay-out.

Train surgeons

Provide resources for the surgeons to find more about the facial recognition research to give them trust in the concept.

Explain to the surgeons how they can use the tool most effectively during their consultation.

Show a simulation to surgeons of how the software can be used in Epic and show it can be done fast and easy.

Parallel activities

Things which can be done during implementation of the concept

Improve surgical techniques

When more research is done in facial recognition after orthognathic surgeries, it could be possible to change surgical techniques to make patients more recognizable. Other research, e.g. in soft-tissue should also be done to see how movements of soft-tissue and therefore appearance, can be predicted per patient.

Expand to other departments

The research of facial recognition could be expanded to other departments in the VUmc such as the plastic surgery department or transgender surgeries. The tool might be interesting to implement in other departments as well.

13.2 Investments

For implementing the concept investments are needed. In this section, the costs and benefits are described. The costs exclude all the steps from the roadmap (chapter 13.2) which still need to be done before the concept can be implemented.

13.2.2 Costs

The biggest investment will be the **development costs** of the facial recognition system. The system should be implemented in Epic, which already has been done before with other software at the VUmc (chapter 10.2). The costs include collecting feedback, improving the concept and executing pilots.

Sporadically investments are needed for covering the **time the physicians need** to create the database and structural investments are needed to cover the time the physicians need to take pictures from the patients.

It is possible the **consultation will take longer** when the concept is implemented because the surgeon has more information to share and the patient might have more questions. This can increase the costs as well and might also increase work pressure.

13.2.3 Benefits

The patients will be more satisfied with their surgery because their **expectations meet the outcome.** This will decrease complains towards surgeons, resulting in a decrease of work pressure.

The concept will help the patients to **recover emotionally**. The patients and their close acquaintances will be supported in accepting the changes in the patient's face. This will decrease the need for psychological care during the lifetime of the patients.

The **surgeon will be able to explain** how much the patient's face will change which they cannot do without the concept. It can spare time informing the patients with concrete information.

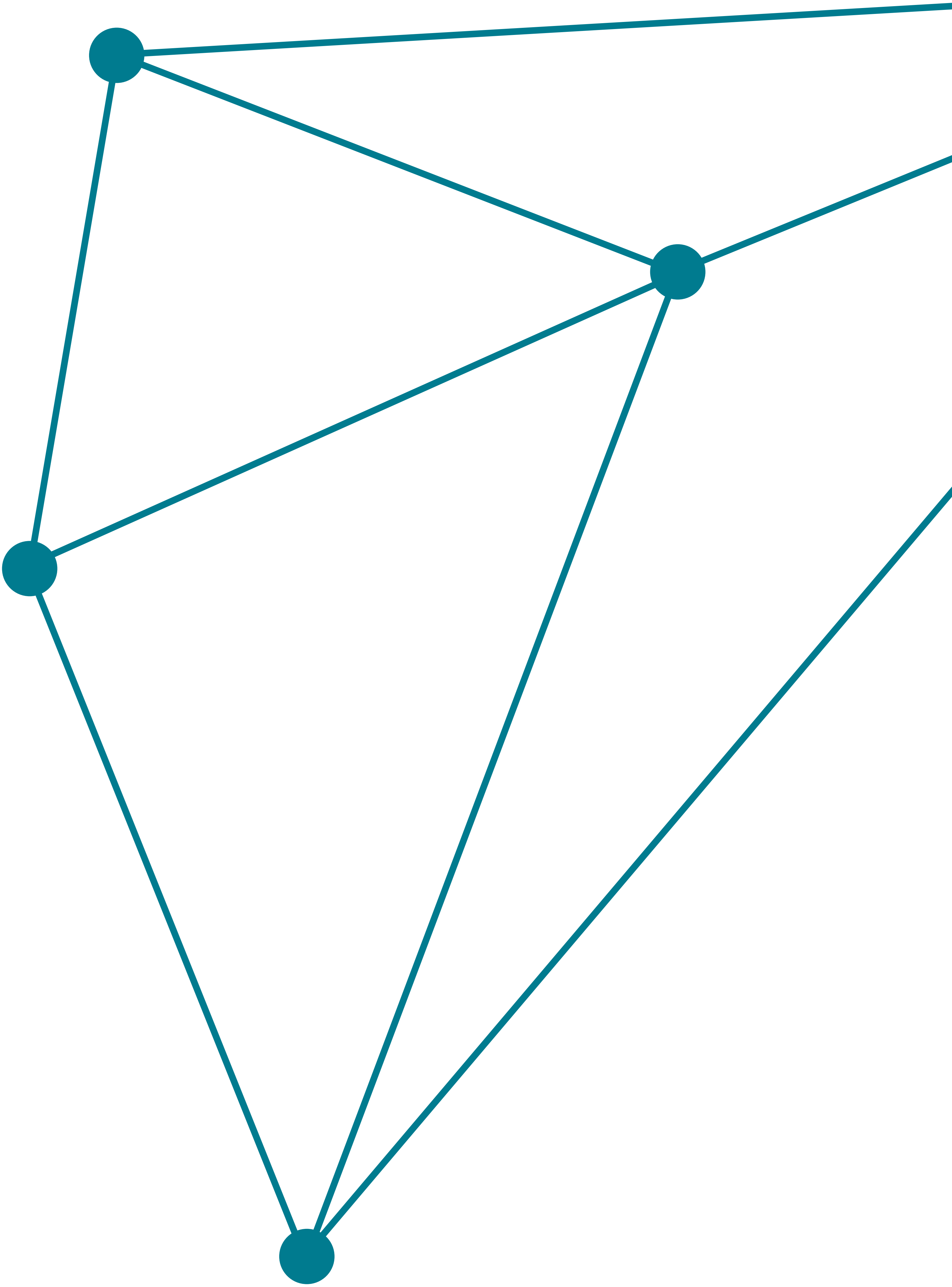
13.3 Recommendations

The most important thing for the VUmc to do is to follow the roadmap to implementation. This will both expand the research and, finally, implement the concept. However, some more detailed recommendations are described here which the VUmc could take into account when continuing with the research and implementation.

Rating own patient pictures
People could rate patient pictures manually on similarity in order to get a human-computer ratio which is based on specific orthognathic surgeries. This is an easier method for the VUmc to perform than the human recognition test because they can pick random people for rating the patients instead of selecting a familiar group. However, this method has a problem as well. According to the research of Bindemann et al. (2012), facial matching tasks of identifying unfamiliar faces have a high error rate. Also, chapter 3.2 describes that processing unfamiliar and familiar faces are different in human brains. The human recognition test has been done at familiar faces to simulate the context of the patient. However, what could be done is finding a familiar-unfamiliar ratio, which allows the VUmc to perform this method after all.

Emotional recovery
According to the defined patient profiles (appendix 23). Fraos has only been designed for extrovert and introvert patients and not for the supported or unsupported patients. Supported patients do not need extra support in emotional recovery, however, unsupported patients might need extra support. Ideas have been created for this (appendix 21), but have not been implemented. Further evaluation of Fraos should point out if it increases emotional recovery enough or if extra care is needed for unsupported patients. This does not have to be implemented in Fraos per se, but it could also be provided via other types of care (e.g. a patient organisation, psychologist). Also, how emotional recovery is going before implementing Fraos should be tested to create an image of how Fraos can best be optimized.

Integration
It is recommended to use different views of physicians and patients to improve research and Fraos. Integration of the interests of all stakeholders, the system of the VUmc and the knowledge from multiple fields will provide the most complete image and will lead to the most success in implementing Fraos.



Chapter 14

General discussion

This chapter shows a reflection about how the implementation of the project in the sensemaking model creates an overview of the accomplishments and the iterations between topics. Hereafter, the current and future impact of the project is described with a summary of the impact. Last, the limitations of the project are described.

14.1 Reflection

After finishing the concept, the sensemaking model of Aselmaa (2017) has been used to create an overview of the knowledge obtained during the project and the iterations which were done (figure 14.1). All the tasks done and its connections made more 'sense' after creating this overview and might make the process and foundings more clear for future researchers. The exploration of all the keywords belonging to the numbers can be found in appendix 1. Overall, all the information which has been found is linked to each other. The link between orthognathic surgeries and facial recognition was missing before this project and has now been created.

When looking at the gaps before the project which are now filled, it is clear that a start has been made on the topic and this project resulted in some interesting findings. However, during the project more and more questions popped-up, keeping it challenging to maintain the boundaries around the project and to prevent wandering off.

The gaps which are present after the project (figure 14.2) are based on the questions which popped up during the project and more question will probably pop-up in future research since still a lot in this field can be discovered. Because of the initial project structure from chapter 2, it made it easier to follow the plan. Also, because of new findings during the project, flexibility was required since the initial ideas did not always meet the realistic way to go. Therefore, the planning was adjusted every week (appendix 35 and 36). The initial plan for the concept which was formulated in the project brief (appendix 37), a tool to optimize orthognathic surgeries, seemed too ambitious regarding the complexity of the topic and did not fit in the time-frame. However, in the end, a concept focusing more on the patient's experiences was created.

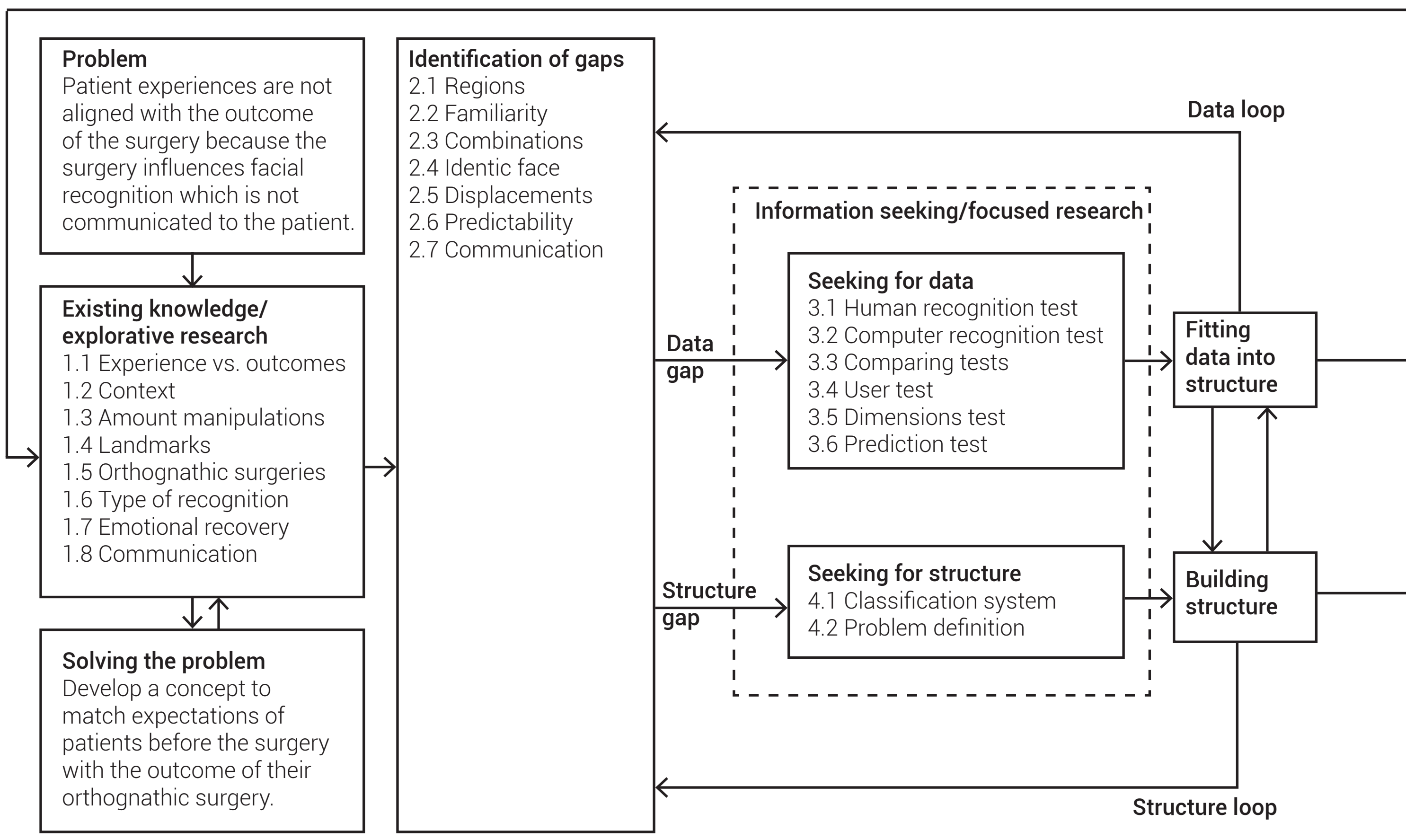


Figure 14.1: Implementation of the model of the cognitive process and mechanisms of individual sensemaking

Gaps before project
2.1 Regions
2.2 Familiarity
2.3 Combinations
2.4 Identic face
2.5 Displacements
2.6 Predictability
2.7 Communication

Gaps after project
2.8 Prediction all surgeries
2.9 Manipulate rotations
2.10 Complex surgeries familiar group
2.11 Specific area
2.12 Type of familiarity
2.13 Age
2.14 Expectations vs. similarity score
2.15 Experience Fraos
2.16 VUmc Data
2.17 3D pictures
2.18 Smile and score
2.19 Influence orthodontist
2.20 Surgery protocols
2.21 Patient evaluations

Figure 14.2: Defined gaps before and estimated gaps after the project

Outcomes updated knowledge

Accretion:
Instantiated structure
6.1 Final version of Fraos
6.2 Roadmap for future implementation
6.3 Recommendations and discussion

Tuning:
Adapted structure
5.1 The communication part of Fraos.
5.2 Facial recognition system of Fraos.

Re-structuring:
New structure
Three versions of 4.1
Two versions of 4.2

Iterative pattern
- 3.1 resulted from 1.2, 1.3, 1.4, 1.5 and 1.6
- 3.2 resulted from 1.3, 1.4 and 1.5,
- 3.3 resulted from 3.1 and 3.2
- 3.4 resulted from 1.7, 1.8 and 5.1
- 3.5 resulted from 1.4 and 5.2
- 3.6 resulted from 3.1 and 5.2
- 4.1 resulted from 1.4 and 4.5
- 4.2 resulted from all existing knowledge and 3.1, 3.2 and 3.3

14.2 Impact of the project

This project did have a certain impact (figure 1543). The expected current and future impact have been described.

14.2.1 Current impact

Orthognathic surgeries influencing facial recognition is an existing problem at the orthognathic department of the VUmc which T. Fouranzanfar, an orthognathic surgeon, addressed multiple times for a long time. However, no one picked it up. This project proved the importance of the subject, and it proved that changes in facial recognition after orthognathic surgeries are real. The research section touched this problem and showed more insights into how the problem could be researched and what happens when this is researched. This can be a motivator for researchers to continue working on the topic. The concept shows how the VUmc could translate their research into something more concrete, which has an effect on the care of patients. This could motivate them to follow the roadmap to implementation. This project could also open doors at other departments at the VUmc, especially since they are specialised in facial reconstruction. Facial recognition can play a role in all type of facial surgeries. The project could be a set-up for further research and improvement of care regarding facial recognition.

14.2.2 Future impact

When more research is done, the concept could be implemented which will improve matching the patient experiences with the outcome of the surgery. In the future, the research could lead to a change in surgical techniques, leading to more recognizable patients after the surgery. Technologies in facial recognition are emerging, so future research in the topic might be easier to carry out and what happens with facial recognition due to orthognathic surgeries might be explored in detail. Overall, this project is a start for improving care regarding facial recognition in the future.

"Now, there is proof that orthognathic surgeries do influence facial recognition"

"The research in the topic opened a new conversation and can provoke further research"

"The concept shows how research can be translated to something concrete which improves care"

Figure 14.3: Summary of the impact of the project

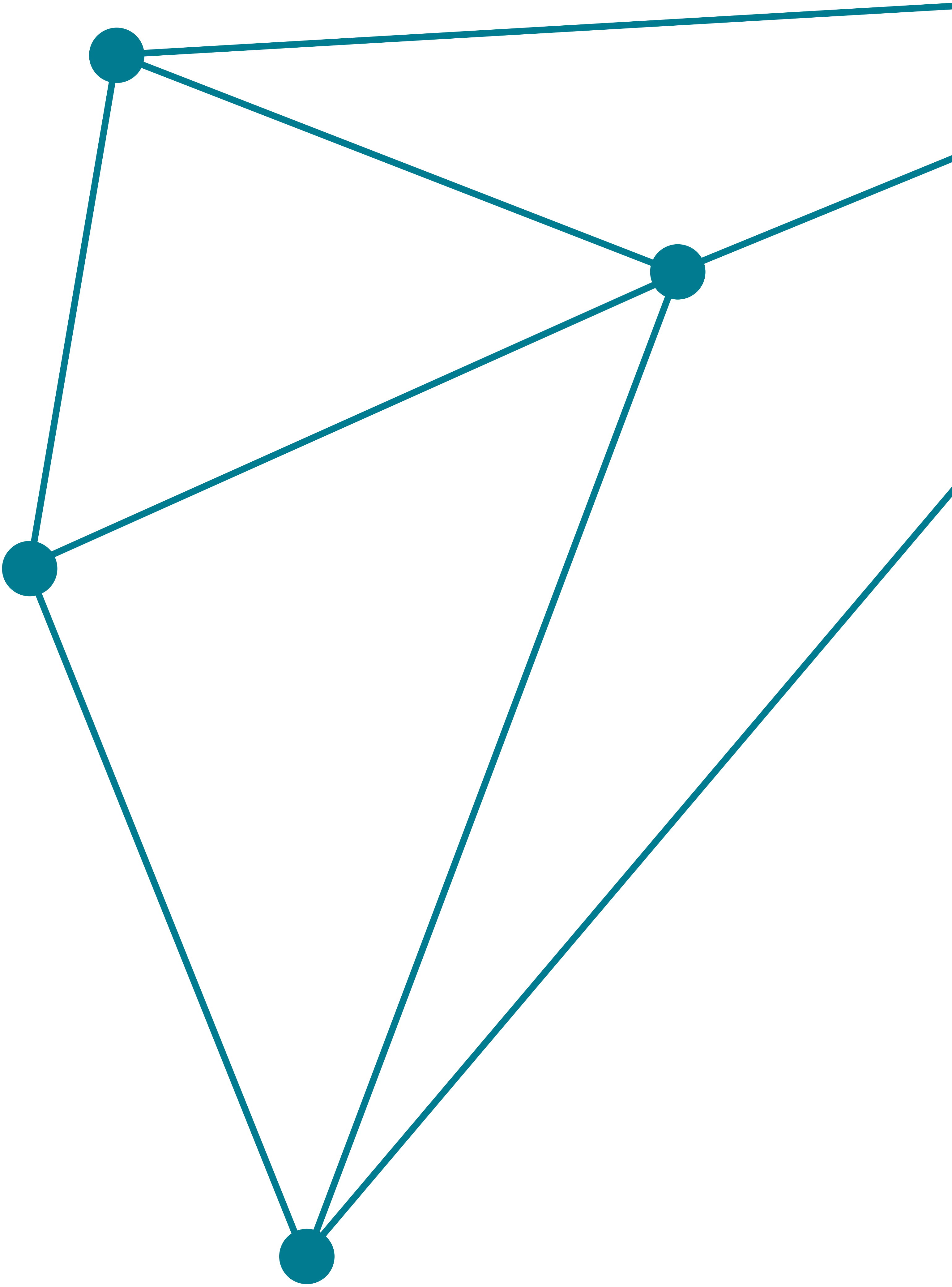
14.3 Limitations of the project

In general, the biggest limitation was that the whole project has been done alone behind a desk resulting in less communication which required more creativity to achieve results. Also, a limitation was the complexity of the topic and the relatively little amount of time. Therefore, not all questions needed to implement the concept were answered.

The concept has not been evaluated by physicians at the VUmc. The comment received from T. Fouranzanfar (personal communication, June 8, 2020) was when the first draft of the flyer part of the concept was done. He liked the concept and thought it would be great, but more steps should be done before it could be implemented. At this moment, the predictability test was not done and he was questioning if predicting the similarity score would be possible. After all these things were sorted out, there has not been any evaluation with physicians anymore. Therefore, it is unknown what they think of Fraos.

Also, within this project it was not possible to test with patients, therefore it is unknown what they think about Fraos. The user test showed that younger participants were flexible and they do not seem to worry about facial recognition. Therefore it could be possible that part of the patients do not need Fraos since they do not have problems with changes in facial recognition. The way of communicating via a flyer was appreciated by the participants of the user test, but it is unknown how patients respond to it. Also, it might be possible that patients are not comfortable if pictures of them are made. Because of all of these gaps, it is unclear at the moment if the implementation of Fraos is a good idea or that it should be further optimized.

It was not possible during this project to do any field research at the VUmc. It would have been valuable to observe some consultations to be able to integrate Fraos in the system of the VUmc. Now, it is only integrated with the use of online information, whereof only a little information was available.



References

Ahdid, R., Safi, S., & Manaut, B. (2017). Euclidean and geodesic distance between a facial feature points in two-dimensional face recognition system. *International Arab Journal of Information Technology*.

AHRQ. (2017, July 1). Tools to Help Patients Communicate Their Needs. Retrieved 10 June 2020, from https://www.ahrq.gov/cahps/quality-improvement/improvement-guide/6-strategies-for-improving/communication/strategy6htools.html

Amato, G., Gennaro, C., Falchi, F., & Vairo, C. (2018). A Comparison of Face Verification with Facial Landmarks and Deep Features. In *MMedia 2018*. Retrieved from https://www.researchgate.net/publication/338048224_A_Comparison_of_Face_Verification_with_Facial_Landmarks_and_Deep_Features

Amsterdam UMC. (2019). https://mijnamcvumcdossier.nl/. Retrieved 11 July 2020, from https://mijnamcvumcdossier.nl/MyChart-PRD/Authentication/Login?

Amsterdam UMC. (2017, October 25). MyChart instructiefilmpje [Video file]. Retrieved from https://www.youtube.com/watch?v=SNUmVTgLKCg&list=PL5n5f65T4tspW8Gq6NA98GnULg1hj4_KA&index=26&t=0s

Amsterdam UMC. (2020, March 9). Patiëntenfolders mondziekten, kaak- en aangezichtschirurgie. Retrieved 11 June 2020, from https://www.vumc.nl/zorg/expertisecentra-en-specialismen/mondziekten-kaak-en-aangezichtschirurgie/informatie-voor-patienten-mondziekten-kaak-en-aangezichtschirurgie/patientenfolders-mondziekten-kaak-en-aangezichtschirurgie.htm\

Anand, B., & Shah, P. (2016, January 1). Face Recognition using SURF Features and SVM Classifier. Retrieved 1 April 2020, from https://www.semanticscholar.org/paper/Face-Recognition-using-SURF-Features-and-SVM-Anand/0b798016118aa778355a4e78b3900b05c4dbf67f

Aselmaa, A. (2017). Incorporating sensemaking perspective in design: Supporting physicians during thecontouring tasks in radiotherapy. https://doi.org/10.4233/uuid:a9ff69cd-e5a5-451e-a6c8-0681b633927d

Atkinson, A. P., & Adolphs, R. (2011). The neuropsychology of face perception: Beyond simple dissociations and functional selectivity. In *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2010.0349

Barton, J. J. S., & Corrow, S. L. (2016). Recognizing and identifying people: A neuropsychological review. In *Cortex*. https://doi.org/10.1016/j.cortex.2015.11.023

Bellucci, C. C., & Kapp-Simon, K. A. (2007). Psychological Considerations in Orthognathic Surgery. In *Clinics in Plastic Surgery*. https://doi.org/10.1016/j.cps.2007.04.004

BJC HealthCare. (2017, December 11). WUSM Epic Vlog Episode 8: POD 1 Go-Live of Epic, Inbasket, Laptops, and Training [Video file]. Retrieved from https://www.youtube.com/watch?v=_UhdVhG0pkA&t=251s

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0003022

Bonanthaya, K., & Anantanarayanan, P. (2013). Unfavourable outcomes in orthognathic surgery. In *Indian Journal of Plastic Surgery*. https://doi.org/10.4103/0970-0358.118592

Bruce, V., Hancock, P. J. B., & Burton, A. M. (1998). Comparisons between human and computer recognition of faces. Proceedings - *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, FG 1998. https://doi.org/10.1109/AFGR.1998.670983

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bruens, G. (2011). *Form Color Anatomy*. Zaltbommel, Netherlands: Van Haren Publishing.
Color psychology. (2020). Teal Color. Retrieved 23 June 2020, from https://www.colorpsychology.org/teal/

Cadogan, J., & Bennun, I. (2011). Face value: An exploration of the psychological impact of orthognathic surgery. *British Journal of Oral and Maxillofacial Surgery*. https://doi.org/10.1016/j.bjoms.2010.07.006

Çeliktutan, O., Ulukaya, S., & Sankur, B. (2013). A comparative study of face landmarking techniques. In *Eurasip Journal on Image and Video Processing*. https://doi.org/10.1186/1687-5281-2013-13

Chellappa, R., Sinha, P., & Phillips, P. J. (2010). Face recognition by computers and humans. *Computer*. https://doi.org/10.1109/MC.2010.37

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/s13423-018-1559-x

Cunningham, S. J., Gilthorpe, M. S., & Hunt, N. P. (2000). Are orthognathic patients different? *European Journal of Orthodontics*. https://doi.org/10.1093/ejo/22.2.195

Dahiru, T. (2011). P-Value, a true test of statistical significance? a cautionary note. *Annals of Ibadan Postgraduate Medicine*. https://doi.org/10.4314/aipm.v6i1.64038

Dorresteijn, M. (2015, October 26). AMC stapt over op nieuw epd. Retrieved 11 July 2020, from https://www.zorgvisie.nl/amc-stapt-over-op-nieuwe-epd-2707812w/

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*. https://doi.org/10.1068/p080431

Epic. (2020, June 18). Privacy Policies | Epic. Retrieved 30 July 2020, from https://www.epic.com/about/privacypolicies

Erminesoft. (2017, May 22). Technology Behind Snapchat and Instagram Face Filters. Retrieved 29 July 2020, from https://erminesoft.com/how-snapchat-instagram-face-filters-work/

Facial Sculpture Clinic . (2020, January 1). Facial Sculpture Clinic . Dr Joël Defrancq, Jaw surgery and implant reconstructive surgery. Retrieved 2 April 2020, from https://facialsculptureclinic.com/en/

Farley, P. (2019, April 26). Face detection and attributes concepts. *Azure Cognitive Services*. Retrieved 18 March 2020, from https://docs.microsoft.com/bs-latn-ba/azure/cognitive-services/face/concepts/face-detection

Fagertun, J., Harder, S., Rosengren, A., Moeller, C., Werge, T., Paulsen, R. R., & Hansen, T. F. (2014). 3D facial landmarks: Inter-operator variability of manual annotation. *BMC Medical Imaging*. https://doi.org/10.1186/1471-2342-14-35

Flyerzone. (2020). Welk font gebruik ik voor mijn brochure? - Flyerzone. Retrieved 25 June 2020, from https://www.flyerzone.nl/blog/view/welk-font-gebruik-ik-voor-mijn-brochure

Fuentes-Hurtado, F., Diego-Mas, J. A., Naranjo, V., & Alcañiz, M. (2019). Automatic classification of human facial features based on their appearance. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0211314

Gainotti, G. (2014). The neuropsychology of familiar person recognition from face and voice. *Psychologica Belgica*. https://doi.org/10.5334/pb.at

Gateno, J., Alfi, D., Xia, J. J., & Teichgraeber, J. F. (2015). A Geometric Classification of Jaw Deformities. Journal of Oral and Maxillofacial Surgery : *Official Journal of the American Association of Oral and Maxillofacial Surgeons*. https://doi.org/10.1016/j.joms.2015.05.019

Gateno, J., & Xia, J. J. (2020). Contemporary Surgical Planning for jaw surgeries. *Houston Methodist Leading Medicine*. Retrieved from https://issuu.com/instituteforacademicmedicine/docs/contemporarysurgicalplanningforjawd_66b39cfcbb7674?e=8044025/37817576

Geitgey, A. (2017, January 1). Face Recognition — Face Recognition 1.2.3 documentation. Retrieved 6 May 2020, from https://face-recognition.readthedocs.io/en/latest/readme.html

Geitgey, A. (2018, November 7). Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning. Retrieved 5 May 2020, from https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The Perception of a Face Is No More Than the Sum of Its Parts. *Psychological Science*. https://doi.org/10.1177/0956797611427407

Google Play. (2020, June 10). MyChart - Apps op Google Play. Retrieved 14 July 2020, from https://play.google.com/store/apps/details?id=epic.mychart.android&hl=nl

Groene Hart Ziekenhuis. (n.d.). EPD. Retrieved 11 July 2020, from https://www.ghz.nl/voorbereiden-op/uw-patientendossier/epd/

Groeneveld, B. S., Melles, M., Vehmeijer, S., Mathijssen, N., & Goossens, R. (2019). Tailored information technology in healthcare: Methodology of a case study using a web application in total hip arthroplasty. In *T. P. Cotrim, F. Serranheira, P. Sousa, S. Hignett, S. Albolino, & R. Tartaglia (Eds.), Health and Social Care Systems of the Future: Demographic Changes, Digital Age and Human Factors - Proceedings of the Healthcare Ergonomics and Patient Safety,* HEPS 2019 (Vol. 1012, pp. 334-341). (Advances in Intelligent Systems and Computing; Vol. 1012). Springer Verlag. https://doi.org/10.1007/978-3-030-24067-7_38

Groeneveld, B., Melles, M., Vehmeijer, S., Mathijssen, N., van Dijk, L., & Goossens, R. (2019). Tailored Patient Experiences. https://doi.org/10.1007/978-3-319-96098-2_26

Gumbinner, L. (2020, February 25). *I*nstagram trick: How to find all those cool Instagram Story filters everyone else seems to have. Retrieved 29 July 2020, from https://coolmomtech.com/2020/01/how-to-find-save-new-instagram-story-filters/

Ha, J. F., & Longnecker, N. (2010). Doctor-patient communication: A review. In *Ochsner Journal.*

Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition*. https://doi.org/10.3758/BF03197270

HealthManagement.org. (2020, June 8). Grand Plans for Potential COVID-19 Vaccine Production. Retrieved 10 June 2020, from https://healthmanagement.org/c/it/news/how-to-choose-clinical-communication-tools

HIPAA Journal. (2020, February 26). Communication Tools in Nursing. Retrieved 10 June 2020, from https://www.hipaajournal.com/communication-tools-in-nursing/

Homsi, N. (n.d.). Distraction osteogenesis. Retrieved 6 April 2020, from https://surgeryreference.aofoundation.org/cmf/sequela/mandible/mandibular-facial-asymmetry/distraction-osteogenesis?searchurl=%2fSearchResults#placement-of-the-distractor

Kak, F.S., Mahmood Mustafa, F., & Valente, P. (2018). A Review of Person Recognition Based on Face Model. https://doi.org/10.23918/iec2018.01

Interaction Design Foundation. (2002). What is Emotional Design? Retrieved 17 June 2020, from https://www.interaction-design.org/literature/topics/emotional-design

ITACA-TSB. (2012). Proceedings IV Workshop on Technology for Healthcare and Healthy Lifestyle. Retrieved from https://books.google.nl/books?id=KHIXggo_mN0C&pg=PA2000&lpg=PA2000&dq=communication+tools+healthcare&source=bl&ots=4-nPE95ysr&sig=ACfU3U3lECduPu1R5ksnqmtty4Jz_0F6A&hl=en&sa=X&ved=2ahUKEwi49sSeuffpAhWQ2qQKHc0vDlM4FBDoATAHegQIChAB#v=onepage&q=communication%20tools%20healthcare&f=false

Kashani, H. (2016, August 31). Osteotomies in Orthognathic Surgery | IntechOpen. Retrieved 29 June 2020, from https://www.intechopen.com/books/a-textbook-of-advanced-oral-and-maxillofacial-surgery-volume-3/osteotomies-in-orthognathic-surgery

Kessels, R. P. C. (2003). Patients' memory for medical information. In *Journal of the Royal Society of Medicine.* https://doi.org/10.1258/jrsm.96.5.219

Keshtgar, S., Keshtgar, A., Mistry, P., & Shakib, K. (2019). Assessing facial recognition after orthognathic surgery at automated border controls in airports. *British Journal of Oral and Maxillofacial Surgery*. https://doi.org/10.1016/j.bjoms.2018.12.021

Kiekens, R. M. A., Kuijpers-Jagtman, A. M., van 't Hof, M. A., van 't Hof, B. E., Straatman, H., & Maltha, J. C. (2008). Facial esthetics in adolescents and its relationship to "ideal" ratios and angles. *American Journal of Orthodontics and Dentofacial Orthopedics*. https://doi.org/10.1016/j.ajodo.2007.07.013

King, D.E. Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research 10, pp. 1755-1758, 2009

Kirmayer, L. J., Groleau, D., Guzder, J., Blake, C., & Jarvis, E. (2003). Cultural consultation: A model of mental health service for multicultural societies. In *Canadian Journal of Psychiatry.* https://doi.org/10.1177/070674370304800302

KNMG. (2019, March 29). Behandelingsovereenkomst (WGBO). Retrieved 30 July 2020, from https://www.knmg.nl/advies-richtlijnen/dossiers/behandelingsovereenkomst-wgbo.htm

Nederlandse Vereniging van Journalisten. (2018, April 16). AVG en Fotografie. Retrieved 30 July 2020, from https://www.nvj.nl/privacy/avg/avg-en-fotografie#:%7E:text=Bijna%20alles%20wat%20je%20met,direct%20of%20indirect%20identificeerbaar%20is.

Lee, C., Tanikawa, C., Lim, J., & Yamashiro, T. (2019, June 7). Deep Learning based Cephalometric Landmark Identification using Landmark-dependent Multi-scale Patches. ArXiv, abs/1906.02961.

Lehti, J. (2016). Theory of psychological adaptive modes. *Medical Hypotheses*. https://doi.org/10.1016/j.mehy.2016.03.003

Levinson, W., Hudak, P., & Tricco, A. C. (2013). A systematic review of surgeon-patient communication: Strengths and opportunities for improvement. In *Patient Education and Counseling.* https://doi.org/10.1016/j.pec.2013.03.023

Liu, X., Shan, S., & Chen, X. (2013). Face recognition after plastic surgery: A comprehensive study. *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-37444-9_44

Mayo Clinic. (2018, January 3). Jaw surgery - Mayo Clinic. Retrieved 14 April 2020, from https://www.mayoclinic.org/tests-procedures/jaw-surgery/about/pac-20384990

Mishra, M. V., Likitlersuang, J., B Wilmer, J., Cohan, S., Germine, L., & DeGutis, J. M. (2019). Gender Differences in Familiar Face Recognition and the Influence of Sociocultural Gender Inequality. *Scientific Reports*. https://doi.org/10.1038/s41598-019-54074-5

Mostard, W. (2019). Combining Visual and Contextual Information for Fraudulent Online Store Classification. Retrieved from https://www.ai.rug.nl/~mwiering/Thesis_Wouter_Mostard.pdf

Nishimura, M., Doyle, J., Humphreys, K., & Behrmann, M. (2010). Probing the face-space of individuals with prosopagnosia. Neuropsychologia. https://doi.org/10.1016/j.neuropsychologia.2010.03.007

Opdebeeck Orthodontie. (2020). Onze praktijk in tijden van Corona. Retrieved 11 July 2020, from https://www.opdebeeckorthodontie.be/nl/corona

Papesh, M. H., & Goldinger, S. D. (2010). A multidimensional scaling analysis of own- and cross-race face spaces. Cognition. https://doi.org/10.1016/j.cognition.2010.05.001

Phillips, P. J., & O'Toole, A. J. (2014). Comparison of human and computer performance across face recognition experiments. In *Image and Vision Computing*. https://doi.org/10.1016/j.imavis.2013.12.002

Pysource. (2019, March 12). Face landmarks detection - Opencv with Python. Retrieved 29 July 2020, from https://www.youtube.com/watch?v=MrRGVOhARYY&list=PL5n5f65T4tspW8Gq6NA98GnULg1hj4_KA&index=11&t=585s

Ribbers, T., V.d. Linden, J., He, Y., & Cai, Z. (2018). ACD: Design for value based healthcare (Master Project). Delft, the Netherlands: TU Delft.

Robson, D. (2014, July 30). Neuroscience: why do we see faces in everyday objects?. Retrieved from https://www.bbc.com/future/article/20140730-why-do-we-see-faces-in-objects

Rosebrock, A. (2020, February 14). Facial landmarks with dlib, OpenCV, and Python. Retrieved 18 March 2020, from https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/

Rosebrock, A. (2020, April 18). Face recognition with OpenCV, Python, and deep learning. Retrieved 6 May 2020, from https://www.pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/

Rosebrock, A. (2020, April 18). Training a custom dlib shape predictor. Retrieved 5 May 2020, from https://www.pyimagesearch.com/2019/12/16/training-a-custom-dlib-shape-predictor/

Royal medical services - JOS. (2017, April 11). Orthognathic surgery [Slides]. Retrieved from https://www.slideshare.net/lubna_aborob/orthognathic-surgery-74866459

Saman, M., Abramowitz, J. M., & Buchbinder, D. (2013). Mandibular osteotomies and distraction osteogenesis evolution and current advances. In *JAMA Facial Plastic Surgery*. https://doi.org/10.1001/jamafacial.2013.44

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., & De Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*. https://doi.org/10.1111/j.0956-7976.2005.01554.x

Schattner, A., Bronstein, A., & Jellin, N. (2006). Information and shared decision-making are top patients' priorities. *BMC Health Services Research*. https://doi.org/10.1186/1472-6963-6-21

Sforza, C., & Ferrario, V. (2006). Soft-tissue facial anthropometry in three dimensions: from anatomical landmarks to digital morphology in research, clinics and forensic anthropology. *J Anthropol Sci*, 84, 97–124.

Singh, R., Vatsa, M., Bhatt, H. S., Bharadwaj, S., Noore, A., & Nooreyezdan, S. S. (2010). Plastic surgery: A new dimension to face recognition. *IEEE Transactions on Information Forensics and Security*. https://doi.org/10.1109/TIFS.2010.2054083

SciShow Psych. (2018, October 22). Why Our Brains Recognize Faces So Easily... or Fail at It. Video file. Retrieved from https://www.youtube.com/watch?v=lwEvw6a1fS8

Software Connect. (2020). Epic | Electronic Medical Records (EHR) Software | 2020 Reviews. Retrieved 12 July 2020, from https://softwareconnect.com/ehr/epic/

Soh, C. L., & Narayanan, V. (2013). Quality of life assessment in patients with dentofacial deformity undergoing orthognathic surgery - A systematic review. *In International Journal of Oral and Maxillofacial Surgery.* https://doi.org/10.1016/j.ijom.2013.03.023

Speechmed. (2017). See, Touch and hear vital healthcare information in multiple languages. Retrieved 10 June 2020, from https://www.speechmed.com/

Top Class Dental. (n.d.). classes of malocclusion . Retrieved 16 April 2020, from https://topclassdental.com.au/service/orthodontics/classes-of-malocclusion/

Valentine, T. (2001). Face-space models of face recognition. *Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges*. https://doi.org/10.1006/jmps.1997.1193

Vistaprint. (2018, January 1). Design Decoded: The Top 12 Easy to Read Fonts. Retrieved 18 June 2020, from https://www.vistaprint.com/hub/design-decoded-top-12-easy-read-fonts?couponAutoload=1&GP=06%2f18%2f2020+04%3a24%3a56&GPS=5705425366&GNF=0&ncid=cjachapxo5mvotuefzqhkt4q&rd=2

VUmc. (2019, September 3). Over METc VUmc: Medisch Ethische Toetsingscommissie. Retrieved 26 June 2020, from https://www.vumc.nl/research/overzicht/medisch-ethische-toetsingscommissie/over-metc-vumc.htm

VUmc. (2019, April 18). Toestemming voor delen medische gegevens buiten Amsterdam UMC. Retrieved 11 July 2020, from https://www.vumc.nl/zorg/rechten-en-plichten/rechten-voor-het-elektronisch-patientendossier/toestemming-voor-delen-medische-gegevens-buiten-amsterdam-umc.htm

VUmc / VU University Medical Center Amsterdam. (2016). Jaarverslag VUmc 2016. Retrieved 11 July 2020, from https://issuu.com/vumc/docs/jaarverslag_vumc_2016_final

VUmc / VU University Medical Center Amsterdam. (2017). Jaarverslag 2017. Retrieved 11 July 2020, from https://issuu.com/vumc

Vulink, N. C. C., Rosenberg, A., Plooij, J. M., Koole, R., Bergé, S. J., & Denys, D. (2008). Body dysmorphic disorder screening in maxillofacial outpatients presenting for orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery*. https://doi.org/10.1016/j.ijom.2008.06.005

Wagenberg, B. (2020, June). Epic EMR software. Retrieved 12 July 2020, from https://www.emrsystems.net/epic-ehr-software/

Wang, N., Gao, X., Tao, D., Yang, H., & Li, X. (2018). Facial feature point detection: A comprehensive survey. Neurocomputing. https://doi.org/10.1016/j.neucom.2017.05.013

Willems, S., De Maesschalck, S., Deveugele, M., Derese, A., & De Maeseneer, J. (2005). Socio-economic status of the patient and doctor-patient communication: Does it make a difference? In *Patient Education and Counseling.* https://doi.org/10.1016/j.pec.2004.02.011

Z1. (n.d.). How much to make an app - Calculator . Retrieved 9 June 2020, from http://howmuchtomakeanapp.com

Zhang, G., & Wang, Y. (2009). Faceprint: Fusion of local features for 3D face recognition. *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-01793-3_41

Zuyderland Medisch Centrum. (2018, November 10). Kaak verbreding. Retrieved 24 June 2020, from https://www.kaakchirurglimburg.nl/kaakstand-correctie/kaak-ver
breding/