8-2016

Aging effects in automated face recognition

Miguel Cedeno Agamez
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Miguel Cedeno Agamez

Entitled
Aging Effects in Automated Face Recognition

For the degree of Master of Science

Is approved by the final examining committee:

Stephen J. Elliott Chair	Stephen J. Elliott
Melissa J. Dark	Melissa J. Dark
Elisa Bertino	Elisa Bertino

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Stephen J. Elliott

Approved by: Eugene H. Spafford 6/10/2016
Head of the Departmental Graduate Program Date

AGING EFFECTS IN AUTOMATED FACE RECOGNITION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Miguel Cedeno Agamez

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2016

Purdue University

West Lafayette, Indiana

First, to God, who has signaled my path through the completion of this goal.

To my mom, Mirta Agamez de Cedeno (R.I.P.) and my dad, Miguel Cedeno Montecer, you have been, and will always be my inspiration. To my grandparents, Natividad de Cedeno (R.I.P.), Miguel Cedeno Barrios, Edilma Lamela and Adalberto Agamez (R.I.P.). You have been my sources of support and encouragement to continue to study until today. You have also taught me the value of the hard work and sacrifice.

To my brothers, Manuel A. Cedeno y Cesar M. Cedeno, for all the arguing and moments of joy. I hope I can serve as a role model for you.

To my partner in crime, my wife, Elba, who has been my study buddy, friend, confident, and my inspiration. It was a long way and you, most than anybody else, know it. I do love you with all my heart, and this triumph is also yours.

Miguel

ACKNOWLEDGEMENTS

To my brother from another mother, Eder, who despite all, had always been there when I more needed him.

To Angelica, the sister I never had, thanks for all. You have been an angel in my life, and you have a special space in my heart.

To Dr. Stephen Elliott, and all the research team in the ICBR for all the help and support. Without it I could not have been able to complete this thesis.

To the Fulbright Program, the US Embassy in Panama, SENACYT and the Panamanian government for all the support. I hope I can give back to my country and serve as a role model for current and future students.

To the CERIAS program and to Purdue University that accepted me and gave me the opportunity of growing my knowledge and skills. So far, it's been the best time of my life, and I will never forget it.

To my CERIAS fellows, thanks for all your help and the support during all those long nights of study.

To the Purdue Fulbright Association, and the current and previous board members and advisors. You have made my stay at Purdue more enjoyable, and I will never forget you.

To the United States of America and the State of Indiana, thanks for being my home for the last two years. I hope that more international students could enjoy similar opportunities like the ones I had, I promise I will come back to visit.

To Christopher, Brian, and Ashrith. You guys are awesome, do not ever change. I know you all have a bright future ahead.

To the Curnett family, thanks for all. You have been so kind to me, and I will never forget you. I consider you as my family.

To my friend Servio, thanks for making me part of your family and for all the support during my time at Purdue.

To my fellow Panamanian Fulbrighters Edgardo, Joaquin, and Joseph. I do not have words to thank you for all your support, advice, and bad jokes. I hope to see you back at home soon.

To Scott, Sandra, and the Vana family. Thank you, thank you so much for making me part of your beautiful family.

To the ILI in Massachusetts and its staff. It was my first academic institution in the United States, and I am so proud to say I was an ILI student. A special thank to Amy Ben-Ezra for all the support and also for opening your house to me during my stay in Massachusetts.

To my Turkish sister, Berna, we did have a good time in Northampton. I promise that Elba and I will visit you in Turkey.

To all my friends at ILI. Thanks, thanks for becoming my family! I will never forget you guys!

To the Wardlaw family. You made me feel at home from the very beginning. You have a special place in my heart.

To Mrs. Vilma, my mother-in-law. I have no words to thank you. I cannot pay you back to what you have done for me, and I will always be in debt with you. Thanks for helping and supporting Elba while I was abroad.

To Ingrid and her family, my dear friends. Thanks for all the support and the words of encouragement.

To my uncle Adalberto and Alex and my aunts Ariadne and Analiz, thanks for the support and all your advice. I hope I have been a good nephew, and I will always do my best to do so.

To my uncle-in-law David, thanks for all the support and for helping Elba out while I was abroad.

To Papatito and his family. I don't know what to say. You had been always there when I needed you. I will be eternally grateful for all the support and moments of joy we have shared. Thank you, thank you so much.

To all my professors and staff at the Universidad Tecnológica de Panamá. You were key towards the completion of this goal. Thanks for all the shared knowledge and guidance during my time as an undergrad student.

To Francisco Martinez and his family, thanks my Puerto Rican brother, thanks for all!

To LTC. Kevin Huggins and his family. Thanks, thanks for giving me the opportunity that pave the way towards the completion of this goal.

To LTC. Ronald Dodge and the United States Military Academy, thanks for opening your doors to me. It made a huge difference in my life.

To Angel, Ariel and Ylenia, my college friends. Thanks for your help, support and for all the moments we have enjoyed.

To my friends at DELL, thanks for all your support. I miss you guys!

To Laura, my boss, thanks for all the help and support. You gave me the opportunity to grow and develop myself. Your help was key towards the consecution of this goal.

Finally, to all the persons that have helped me. I would have to write a book to mention every single person who has helped me along the way. For all of those people, thank you so much.

Boiler Up! Hail, Purdue!

Miguel

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Cedeno Agamez, Miguel. M.S., Purdue University, August 2016. Aging Effects in Automated Face Recognition. Major Professor: Stephen J. Elliott.

The main objective of this work was to analyze the effects of aging in the automated face recognition process.

A dataset was used to perform experiments and obtain indicators to measure the impact of aging. To compare the effects of aging the dataset was segmented based on the age difference between the subjects' face images. The image quality metrics were also part of the analysis performed in this study.

The results of the experiments shown that the higher the gap between the images, the higher the error rates. These were the expected results and it is consistent with other experiments performed in the past. The False Rejection Rate (FRR) was measured at 1%, 0.1%, and 0.01% False Acceptance Rate (FAR) obtaining the similar output as the gap between the images increased.

CHAPTER 1. INTRODUCTION

Face recognition is arguably one of the most performed tasks by human beings on a daily basis. Consequently, to identify a person, we rely on the natural ability to remember other people's faces. Similarly, automated face recognition tries to replicate this ability using algorithms and devices to capture and analyze the different features of the face. This biometric modality already has many limitations due to variations in pose, lighting, expression and age (Park, Tong, & Jain, 2010). The variations that impact the structure of the face have been a subject of study, yet the aging process has not received much attention. Such a topic has become into an interesting subject for research (Zeng, Ling, Latecki, Fitzhugh, & Guo, 2012). The ability of an algorithm to successfully recognize a face will reduce the amount of false positives and false negatives. Some studies have shown that a person recognizes faces 78% of the time compared to only 38% when using computer algorithms (Zeng et al., 2012). These results specifically relate to recognizing faces while people age. There are different scenarios where face recognition plays an important role in the authentication process, and it can be impacted by age. Typical applications of this modality include border control, human-computer-interaction, voter registration, medical records and advanced video surveillance (Zhao, Chellappa, Phillips, & Rosenfeld, 2003).

High throughput environments such as immigration areas demand high performance in face recognition for the authentication process. Skilled and experienced border guards are most likely to perform this task (Gohringer, 2012). The human accuracy of face recognition ability is around 78%. It is also worth to mention that the human beings' ability (the operator) to recognize faces declines with age (Lamont, Stewart-Williams, & Podd, 2005). Researchers have found significant age-related decrements in participants as young as 50, but found that the largest decrements occurred over the age of 70 (Crook & Larrabee, 1992). This indicates that memory decline in face recognition is not linear but accelerates after the age of 70 and is consistent with evidence suggesting that memory in general deteriorates more rapidly in those over 70 years of age than in those a decade or so younger (Parkin, 1993). Moreover, when people are exposed to a high number of faces to recognize, the ability to complete such a task becomes more demanding (Lamont et al., 2005).

Studying how the effects of the aging of faces impact the automated face recognition process will help to avoid some of the limitations. First, computers are better and more reliable than humans when it comes to complete repetitive tasks. Moreover, computers can be exposed to vast amounts of data without the concern of the memory load that inherently impacts the human performance. There are critical processes that rely on the successful recognition of a face such as in border control. Therefore, more information on how to minimize the effects of aging will help to boost the performance of the algorithms and the operators. At the same time, studying, and finding measures to diminish the effects of aging will increase the security levels of the processes using automated face recognition for authorization and authentication purposes.

1.1 Statement of the problem

This study examines the problem of automated face recognition performance of people's faces while they are aging.

1.2 Significance of the problem

The effects of aging have a significant impact on the automated face recognition process, leading to the increase of the False Rejection Rate (FRR) and the False Acceptance Rate (FAR). These rates could negatively affect operations in places with a high volume of users (e.g. airports) causing security issues and unexpected delays. Understanding the impact caused by aging on the performance of face recognition algorithms is critical, turning aging effects on an interesting object of study. The study also addresses image quality issues that could also impact the performance and accuracy of face recognition algorithms. Face recognition is a traditional method for authorization and authentication purposes (Esme & Sankur, 2010). The aging effects are one of the intrinsic problems of face recognition and have attracted a lot of the researchers' attention (Ling, Soatto, Ramanathan, & Jacobs, 2007) (Lamont et al., 2005). Research of aging effects on automated face recognition has benefited from the evolution of different techniques. These techniques have allow algorithms to deal with aging effects on the recognition process and, at the same time it has improved the overall performance (Gohringer, 2012). In spite of all the development over the last years, there are factors that still affect the performance and the accuracy of the recognition process while people age. These factors include the age of the person in a picture, the age gaps within the capture of two different pictures, the gender, and the race (Zeng et al., 2012).

Some approaches have been taken to study the effects of aging on the automated face recognition process. Li, Park, and Jain used a 3D aging model that models the shape and texture separately at different ages using shape and texture pattern space. Generative approaches have also been the subject of study. Ramantath developed a shape transformation model formulated as a physically-based parametric muscle model that captures the subtle deformations that facial features undergo with age (Ramanathan & Chellappa, 2008). Most of the studies focuses on aging models. The desired output of these studies are models that can predict and create an updated version of a subject's face. Then, the outputted image is used as input for the automated face recognition process. Ramanathan and Chellappa used growth parameters and Park et al. used age simulation as the base to build and implement their models. One of the few studies of the performance of algorithms in terms of False Rejection Rate (FRR) and False Acceptance Rate (FAR) was performed by NIST (Grother, Quinn, & Phillips, 2010). According to Gohringer, the NIST study has demonstrated an improvement in the FRR every four years at a 0.001% FAR (Gohringer, 2012). In the same study, for photos taken eight years apart, the top vendor performed at an FRR of 0.05% at a 0.1% FAR. The NIST's study only included seven vendors (four companies and three universities). There were no additional studies on aging, which tested algorithms performance in terms of FRR and FAR. There were no studies dealing with both, the parameters above (FRR and FAR) and how the aging effects affects them.

1.3 Purpose of the study

The first part of the study examines the effects of aging in the automated face recognition process for an *n* amount of pictures from a single subject for all the subjects in the set, considering the FRR at a 1%, 0.1%, and 0.01% FAR.

The second part of the study will examine the effect of aging on the automated face recognition process for a pair of images at the same FRR and FAR levels as in the first part of the study.

A similar study was performed in 2006 over the Morph Database. However, the FAR and the FRR were not taken into account. The study also used a smaller sample size (Ricanek & Tesafaye, 2006).

1.3.1 Research questions

The study addresses the following research questions:

1. Is there any difference at 1%, 0.1%, and 0.01% FAR in terms of FRR across years?
2. Is there any difference in image quality among the subsets?

1.4 Definitions of terms

<u>Biometric System:</u> “System for the purpose of the automated recognition of individuals based on their behavioral and biological characteristics” (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 3).

<u>Detection Error Trade-off (DET) curves:</u> "DET curves can show the relationship between the false reject rate (FRR), and the false accept rate (FAR)" (Johnson, Tan, & Schuckers, 2010, p. 2).

<u>Equal Error Rate (ERR):</u> "The point at which FRR equals FAR". (Johnson et al., 2010, p. 3).

<u>Failure to Acquire (FTA):</u> "The expected proportion of transactions for which the system is unable to capture or locate an image or signal of sufficient quality." (Mansfield & Wayman, 2002, p. 6)

<u>False Accept Rate (FAR):</u> "Proportion of verification transactions with wrongful claims of identity that are incorrectly confirmed" (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 80).

<u>False Rejection Rate (FRR):</u> "Proportion of verification transactions with truthful claims of identity that are incorrectly confirmed" (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 80).

<u>Impostor scores:</u> "is the resulting value when two images that belong to different persons are compared" (He et al., 2010, p. 1795).

<u>Performance:</u> "an assessment of the TAR, TRR, FAR, FRR, FTA, and FTE of a biometric system" (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 81).

<u>Quality:</u> "The degree to which a biometric sample fulfills its specified requirements for its targeted application quality" (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 81).

<u>Sample:</u> "an image, signal, or pattern based interpretation of a physical human feature used for identification or verification using biometric techniques" (ISO/IEC JTC1 SC37 Working Group 1, 2007, p. 81).

True speaker (genuine) scores: "is the resulting value when two images that belong to the same person are compared" (He et al., 2010, p. 1795).

1.5 Assumptions

The following assumptions are part of this study:

- The software and algorithms used during the study did not change.
- The capture process of the images in the dataset used variable environmental conditions.

1.6 Limitations

- The sample dataset does not represent the population.
- The analysis only considers age bins due to a lack of biographical information such as race or ethnicity.

1.7 Delimitations

- MORPH (non-commercial), a publicly available longitudinal face database, was the only dataset used.
- Measuring the impact of other parameters such as pose or occlusions is outside the scope of this study.
- The Quality Scores were computed using the Aware Pre-Face Profile Analysis Tool.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Introduction

The literature review examines the history of face recognition, industry drivers, algorithms, face recognition modalities, aspects impacting face recognition and recent work on aging impacts over the face recognition process.

2.2 Face recognition history

The face is the part of the head from the forehead to the chin, and from one ear to the other. The face also provides identity as an individual human (Moore, Dalley, & Agur, 2013). The basic shape of the face is determined by the structure of the skull bones and other soft tissues such as fat, and skin (Moore et al., 2013). Face recognition consists of the analysis of features of the face or the relative distance between specific points on the face (Kindt, 2013). The result of this analysis is then compared against a stored template to complete the automated recognition process.

Even though recognizing faces is a task that human beings perform on a daily basis, automated face recognition is a relatively new concept. The first semi-automated face recognition system came to light in the 1960's. The system required that the operator locates facial features such as ears, and nose, and calculate the distances to a reference point, and then compare it with the reference data (NSTC Subcommittee on Biometrics, 2006). The correct identification, and extraction of the facial features is the main goal of the automated face recognition process. Different vendors use different methods for automated face recognition. However, all focus on the measurement of key features of the face (Woodward, Horn, Gatune, & Thomas, 2003).

Several approaches to automate the face recognition process were used. In the 1970's, to automate the face recognition process, a computer model for file search, and retrieval used 22 specific subjective markers. These markers included the nose length, and upper-lip thickness (Goldstein, Harmon, & Lesk, 1971). The Eigenface approach based the face recognition on a small set of images features. The features try to approximate the set of known face images, and do not have to correspond to intuitive notions of facial parts, and features (Turk & Pentland, 1991). Elastic Bunch Graph Matching is another of the approaches used. This one consists in the use of Gabor wavelet transform, where labeled graphs represent the face. Then, new faces are extracted by an elastic matching process, and compared through a similarity function (Wiskott, Fellous, Kuiger, & Von Der Malsburg, 1997). The Gabor wavelet transform allows to extract both the time, and frequency information from a given signal. The Gabor wavelet transform approach is broadly used in several face recognition methods (Chao, 2011).

Automated face recognition technology captured the attention of the public due to several implementations. In 2013, the Boston Police Department teamed up with IBM to test IBM's Smart Surveillance System, and Intelligent Video Analytics. The trial consisted of capturing the face of each of the concertgoers who walked through the doors at the 2013 Boston Calling music festival (Macri, 2014). This trial consisted of capturing details as skin color, height, and clothing that could help in a possible forensic identification (Macri, 2014). In the Boston Calling music festival, the captured data was not matched against a database.

Most recently, in February 2015 NEC announced the deployment of their face recognition technology for the Surat City Police in India (NEC, 2015). The system will enable the City Police to hook up into live CCTV cameras in critical locations to perform the identification of possible criminals. According to the media announcements, the live feeds will be matched against prerecorded videos, still face images, and sketches of the suspects (Planet Biometrics, 2015).

In another example, a pedophile was captured 19 years after of being convicted of sexual assault. The FBI used a Next Generation Identification (NGI) system built by Lockheed Martin to track him down. The system has used facial recognition to capture biometric data of his face. A match was spotted amongst driving license photos, and the FBI tracked down the subject to his current location (Russon, 2015).

The above examples show how face recognition can help governments, corporations, and individuals to add new capabilities for entertainment, government, and commercial services. Privacy has been of concern since the invention of the photography; the issues are becoming more critical as new technology is being widely used (Li & Jain, 2011).

But, automated face recognition is not only used for law enforcement, and surveillance purposes. Moreover, and due to the decreasing price of desktop, and embedded computing systems, face recognition has created interest in applications such as multimedia management, human-computer interaction, and biometric authentication (Li & Jain, 2011). The technology has received a lot of attention, and has also advanced during the last years. However, there is still a need for highly reliable and accurate systems.

The industry and governments' growth will continue, and new, and different problems will arise. These problems will require special attention by the automated face recognition technologies (NSTC Subcommittee on Biometrics, 2006).

2.3 Industry drivers

The market for the biometric technology will grow in the next decade. In fact, and according to The Global Government Biometric Systems Market 2015-2025 face recognition accounts for 29.8% of the biometrics market. The biometric industry is worth $4.4 billion, and the Compound Annual Growth Rate (CAGR) will increase to 8.7% to peak at $10.2 billion by 2025 (Strategic Defence Intelligence, 2015).

Strategic Defence Intelligence report divides the market into five categories of biometric systems: fingerprint recognition, facial recognition, iris/retinal recognition, signature recognition, and other technologies. Table 2.1 shows the market share by biometric system:

Table 2.1**.**

Market share by biometric system 2015

Modality	Market share percentage
Fingerprint Recognition	34.9%
Face Recognition	29.8%
Iris/Retinal Recognition	14.9%
Signature Recognition	7.8%
Other	12.7%

There are different elements that can help in the development of the facial recognition technology. First, the existence of an international standard for travel documents (ICAO). The ICAO includes face recognition as one of the possible forms of biometrics to be used. Second, the fact that facial recognition is a mature biometric modality with on-going progress on its algorithms. Third, the availability of high-resolution cameras that allow to take better pictures that make the recognition process to be more precise. These factors keep facial recognition as a competitive player in the biometrics industry (Morpho, 2010).

The International Standards Organization (ISO) has put in place the ISO/IEC 19794 (ISO, 2014), which focus on the interoperable exchange of biometric data. ISO also approved the ISO/IEC 19795 (ISO, 2012). This standard focused on the tests to specifically address absolute performance, sufficiency, and the interoperability available from biometric data formatted. ISO/IEC 19795 is concerned with the evaluation of biometric systems giving recommendations, and requirements to conduct performace evaluation through the enrollment, verification and identification steps (ISO, 2012). The existence of these standards is helping the development and growth of the biometric modalities, including facial recognition.

Although projections and standardization, are promising for face recognition, the modality still has to face intrinsic hurdles. By nature, face recognition is known for being difficult than other techniques such as fingerprint, and iris recognition. In contrast to the face features, the iris features are considered permanent and stable. John Daugman, the originator of the first patented iris recognition algorithm, made this assertion (Browning & Orlans, 2014). However, this assumption has been challenged recently. Researchers have found a clear, and consistent evidence that the aging effects in the iris template are noticeable at one year and that increases with increasing time lapse. Also, for a state-of-the-art iris matcher, and three years of time lapse, at a decision threshold corresponding to a one in two million false match rate, a 153% increase in the false non-match rate was observed (Fenker & Bowyer, 2012).

Also, the face recognition capture process has advantages over the same process in fingerprint, and iris recognition. Face recognition capture process is more natural, and less intrusive. Capture devices distant from the subject can capture facial images in a covert manner (Li & Jain, 2011). But, face recognition has some hurdles added by the effects of the aging. In fact, fingerprints, and irises do not vary as the face does. Contrary to the face, the iris has a natural protection within the body. The eyelids and cornea prevent the wear, and tear of the physical surface, and loss of moisture experienced by fingerprints and the skin, and soft tissue on the face (Browning & Orlans, 2014). In the case of fingerprints, his appearance of a person's print depends on age, grease, and cut or worn fingers due to occupation or lifestyle (Majekodunmi & Idachaba, 2011). Aging effects introduce more complexity than iris, and fingerprint recognition modalities to the recognition process.

2.4 Biometric systems

Five major modules compose a Biometric System: the data collection module, the transmission module, the data storage module, the signal processing module, and the decision module (Mansfield & Wayman, 2002). The face matching process takes place in the signal processing module.

In face recognition, the detection of the face is the most important step in the process (Zhang & Zhang, 2010). If the face cannot be detected, the face recognition process will not continue. When it comes to aging, feature extraction, template creation, and matching are arguably the most important sub-processes over the automated facial recognition process. There is also research that analyzes the impact of factors such as pose, lighting, and expression and how these affect the matching process performance. The analysis of the effects of factors such as pose or lighting is out of the scope of this study.

2.5 Face recognition algorithms

There are several algorithms used in the automated face recognition process. Algorithms fall into two big categories: image-based algorithms, and video-based algorithms (Grgic & Delac, 2008). Based on the purpose of this study, only image-based algorithms will be addressed. There are several image-based algorithms. Principal Components Analysis

(PCA), Linear Discriminant Analysis (LDA), and Elastic Bunch Graph Matching (EBGM) are the most studied in the face recognition field (NSTC Subcommittee on Biometrics, 2006).

Automated face recognition is a pattern recognition problem, which is hard to solve. The higher the dimension of space, the more computation power is required to find a match. A reduction technique is required to project the problem onto a lower dimension space (Zhang & Lu, 2013).

2.5.1 Principal Components Analysis (PCA)

This approach transforms faces into a small set of essential characteristics, so called, the eigenfaces (Slavković & Jevtić, 2012). When an image has to be tested, the eigenface is calculated, and the output is compared to the entire database. The eigenface approach is considered by many to be the first working facial recognition technology. This approach served as the basis for one of the top commercial face recognition technology products (Lata et al., 2009).

2.5.2 Linear Discriminant Analysis (LDA)

LDA can separate different objects in Fisher faces, which is the output of this method. The face contains several distinctive features, and the goal of LDA is to reduce the number of features to a manageable number before the classification. The main objective of using LDA is to maximize the ratio of the between-class variance, and the within-class variance (Gavrilova & Monwar, 2013). LDA is an alternative for PCA. Experiments have shown that LDA has a better performance and outperforms PCA when the number of samples per class is small (Martínez & Kak, 2001).

2.5.3 Elastic Bunch Graph Matching (EBGM)

This method is used to transform images into templates that will serve as input for the matching process. EBGM locates the nodes at the fiducial points, and edges between those nodes as well as the definition of the correspondences between nodes of different poses. The matching process then consists of extracting the model graphs from the images in the gallery, and the image graphs from the proof image. Then these image graphs are compared with the models, and the one with the highest similarity score is selected (Wiskott et al., 1997).

2.6 Aspects impacting face recognition accuracy

Humans naturally use faces to recognize people, but sometimes it is hard to recognize people whose face has changed for one reason or another. The same happens during the automated face recognition process. There are several factors, and aspects that can make the face recognition process more complicated, especially when algorithms have to be used to perform the task.

Although this study concentrates on the effects that aging has on the automated recognition process it is important to highlight other factors that can impact the overall process. Some of the factors mentioned below can be taken care of under controlled conditions such as a laboratory experiment with a controlled image capture process. But, sometimes the data capture is done in the wild, and some of these factors cannot be controlled. Moreover, there is a potential impact on the recognition process that has to be taken into account while developing new algorithms and techniques. Considering the potential impact will help to design better, and more accurate methods.

Also, the quality distribution could be considered a negative factor for the matching step. Moreover, the quality metrics and the genuine, and impostor scores can be impacted by the lack of quality of the images. For this study, we will assume that the distribution of the quality of the pictures in the database is similar to that of the pictures taken in the wild. This assumption will help us to generate scenarios with different, and varying conditions that could be present during the capture process, and, therefore, impact the matching process.

The following sections depict some of the more common aspects that can potentially impact the face recognition process.

2.6.1 Pose variations

The pose or rotation of the head at the moment of the image capture can impact the recognition process. For face recognition systems deployed in wild environments such as airports, and cities, more reliable algorithms are required. These algorithms should be able to correct pose variations encountered in these environments. Like other problems of face recognition, variations in the pose are tremendously important in many applications (Shah, Sharif, Raza, & Azeem, 2014). Several approaches are being developed to overcome the pose variation in face recognition. For example, one approach uses a three-module implementation that first corrects the angle of the image, then extracts the features. Finally, this approach matches the obtained face images using statistical models (Cament, Galdames, Bowyer, & Perez, 2015).

Figure 2.1. Pose Variations in MORPH database (University of North Carolina Wilmington, 2015)

2.6.2 Environmental

Background type, and lighting are amongst the most important environmental factors that can impact the face recognition process. In a real setting there are huge differences between different illumination settings. Also, it is difficult to detect the face in complex background settings. The presence of these factors result in incorrect results (Li et al., 2014).

2.6.2.1 Background type

The background type variation can be limited if the image capture process takes place in a controlled setting. However, if the image takes place in the wild, algorithms have to be able to find the face in a complex background setting. Recent work has achieved a face detection rate of 95.51% in complex background settings (Xin Wang & Pan, 2014). A successful detection process will then allow the face feature extraction module to work with a better input for the identification of the subject(s). Another approach is using the skin color, and texture to identify the face in complex background settings such as the ones find in nature (Xiaohua Wang, Zhang, & Yao, 2011).

2.6.2.2 Illumination

Illumination variation is another factor that not only limits the performance of the face recognition process but the performance of the matching process. For example, traditional feature-based face detectors, tend to fail under severe illumination variations such as heavy shadows and overexposure (Han, Shan, Chen, & Gao, 2013). As with background type, illumination variation can be limited in controlled environments. The angle between the light source, and the face can generate shadows that mask face features. The matter in question will diminish the ability of an algorithm to make a good identification decision. Moreover, even images from the same subject can have large differences. One of the causes of these differences can be due to the variation of illumination conditions (Zhang, Wang, Zhu, Liu, & Chen, 2015).

Figure 2.2. Illumination variations in MORPH database (University of North Carolina Wilmington, 2015)

2.6.3 Occlusions

Occlusions are objects that impede algorithms to extract the face features successfully. Internal occlusions are some intrinsic biological element of the face such as the mustache, the beard, and even the hair. On the other hand, external occlusions are all the items that a person wears and that obscure some facial trait. This category includes items such as hats, sunglasses, and turbans. Occluded parts on the face images usually degrade the recognition performance, and thus robust algorithms are necessary (Oh, Lee, & Lee, 2008). A research team used a framework based on a machine learning algorithm to detect the disguised (occluded) face and to determine the type of occlusion. The framework successfully detected caps, glasses, and masks (Li et al., 2014).

Figure 2.3. Illumination variations in MORPH database (University of North Carolina Wilmington, 2015)

2.6.4 Esthetical

Esthetical factors can also affect the face recognition identification process. Procedures such as plastic surgery aim to correct or restore the appearance or functionality of visible parts of the human body, including the face (De Marsico, Nappi, Riccio, & Wechsler, 2011). Therefore, the markers, and the templates used for the face recognition process will be also impacted. These changes can occur due to of accidents or regular plastic surgery procedures. On the other hand, plastic surgery can also be misused to conceal a person identity with the intent to commit fraud or evade law enforcement. Face recognition after plastic surgery can lead to rejection of genuine users or acceptance of impostors (Singh, Vatsa, & Noore, 2009). There are other esthetical factors such as bruising or blows that can temporarily change the structure of the markers in the face. Some examples are shown in Figure 2.4.

Figure 2.4. Esthetical variations in MORPH database (University of North Carolina Wilmington, 2015)

2.6.5 Aging

Aging is the biological effect that all body parts (including the face) undergo while people are growing. Aging is the main topic of this research.

The whole body suffers the effects of the aging process, but arguably the face is the part of the body where these effects have more impact and are more noticeable. The face has three basic elements: skin, soft tissue and the underlying bone structure that provides the shape to the face (Friedman, 2005). There are major forces that contribute to the facial aging including gravity, skeletal remodeling, subcutaneous fat redistribution and loss, hormonal imbalance, chronic solar exposure, and smoking (Coleman & Grover, 2006). All the age-related changes, and factors mentioned, impact the performance of the automated face recognition process (Albert, Sethuram, & Ricanek, 2011).

Figure 2.5. Aging variations in MORPH Database (University of North Carolina Wilmington, 2015)

There is a need for robust algorithms, and systems that can deal with the effects of aging over the automated face recognition process.

These age-related face changes go against one of the features that physical traits have to have to be valid as a biometric identifier: permanence. When a biometric identifier is permanent, it means it does not change over (some) time (Kindt, 2013). Besides persistence, there are several features that a biometric identifier has to fulfill: universality, uniqueness, collectability, performance, acceptability and circumvention (Zhang & Lu, 2013). Compared to other biometric identifiers, the face score is lower in terms of permanence.

Aging effects have become an interesting research topic. Some approaches have been proposed to offset the effects of aging on the face recognition process. Frequently update the face's template through a data capture process is one these proposed approaches. However, template updates can be performed if all subjects with their faces in a database are regularly available, and willing to provide up to date face images (Lanitis, 2009). The following section addressed some of these approaches, and other recent work in aging.

2.7 Recent work on aging

Recently, aging effects over face recognition has become an interesting research topic. Researchers have come up with different approaches, and methods to deal with the aging effects that impact the performance, and accuracy of the face recognition process.

There are three different research activities around to the human aging: 1) Age Invariant Face Recognition, 2) Age Estimation, and 3) Modeling/Simulating aging process (Esme & Sankur, 2010). Table 2.2 shows some recent work in aging.

Table 2.2.

Recent Work in Aging

Study	Author	Sample
A Discriminative Model for Age Invariant Face Recognition	Li, Park & Jain (2011)	20,000 images
Age Transformation for Improving Face Recognition Performance	Singh, et al. (2007)	1,578 images
Ageing effect on face recognition	Ng, Hon & Lee (2007)	46 subjects
Automatic Representation of Adult Aging in Facial Images	Patterson, et al. (2006)	1724 images
Age-Invariant Face Recognition	Park et al. (2010)	1002 images

2.7.1 Age invariant face recognition

Age invariant methods to overcome the aging effects of the recognition process can be categorized in two main blocks: generative, and non-generative. The generative approach applies computer-generated aging models as an input for the recognition process. On the other hand, the non-generative approach derives an age-invariant signature from the subject's face (Esme & Sankur, 2010).

2.7.1.1 Generative

Generative approaches generate models based on some input and try to recreate the aging process. This approach uses landmarks, and associated growth parameters. One of the drawbacks of this approach is that only takes into account the geometric growth, and does not include other facial attributes such as texture (Esme & Sankur, 2010). This approach is related to the modeling, and simulation of the aging process.

2.7.1.2 Non-Generative

One of the non-generative approaches used the gradient orientation technique to build the difference between a pair of passport images taken at different years (Ling et al., 2007). Once the difference is calculated, it is used in conjunction with an SVM (Support Vector Machine) for the verification task. An SVM is supervised learning model with associated learning algorithms that analyze data, and recognize patterns (Press, 2007).

In another approach, the matching is performed based on the coherency of feature drifts (Biswas, Aggarwal, Ramanathan, & Chellappa, 2008). The study claimed that if the two images belong to the same subject, the drifts in features follow a coherent pattern that may not be the case if the images belong to different subjects. Finally, a comparison between the coherency score, and the threshold takes place to make the matching decision.

2.7.2 Age estimation

Humans learn to recognize thousands of faces (Lamont et al., 2005), and can also estimate the age of another person solely based on the face. Though, this estimation might not be correct (Zeng et al., 2012). The ability of humans to estimate others' age is still an interesting research topic.

Researchers have also evaluated the motivations behind age estimation (Lanitis, Draganova, & Christodoulou, 2004). Some applications of age estimation are:

- Age-specific human computer interaction: if a computer algorithm can determine the subject age, the computer environment, and preferences can be adjusted based on the age of the subject.
- Age-based indexing of face images: for photo albums, user can retrieve their own images based on a specific age-range.
- Development of automatic age progression systems: research in this area can help to identify missing people by predicting the current facial appearance of a person based on the age.
- Understanding the process of age perception by humans: automatic age estimation can help to understand how people perceive other person's age.

2.7.3 Modeling/Simulating aging process

The successful simulation of the aging process could fill-in the gaps in face recognition algorithms when it comes to deal with the effects of aging. Instead of using training sets (Park et al., 2010), a single image could be used to simulate the aging process starting from the picture, and ending at the desired age. This approach could be used to update existing databases automatically or to predict the current face structure of missing persons (Esme & Sankur, 2010).

CHAPTER 3. PROCEDURES AND DATA COLLECTION

Procedures and data collection will be discussed in this chapter. The topics for discussion will include the high level research design, the criteria for the dataset selection, and the description of the performed experiment. Figure 3.1 shows the research design schema:

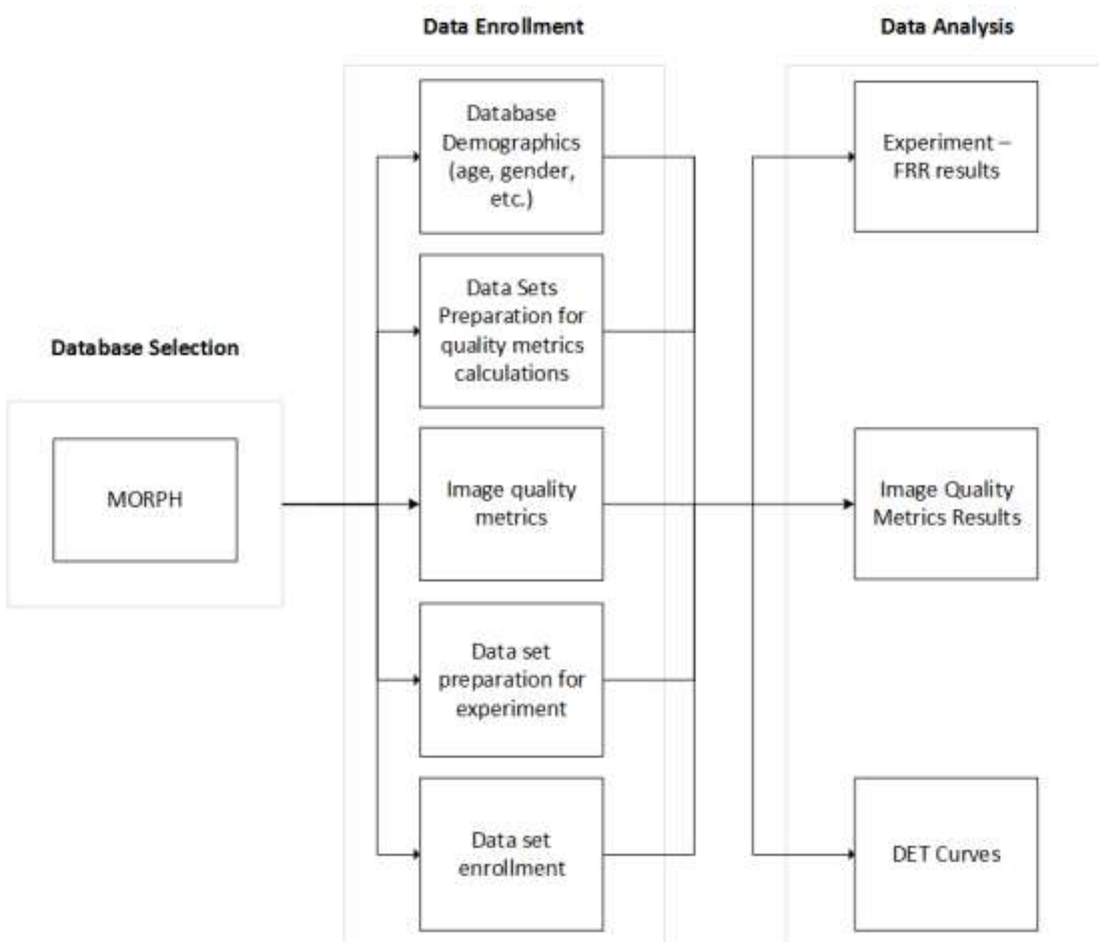

Figure 3.1. High-level research design

3.1 Variables of interest

The variables to be studied include the False Rejection Rate (FRR) of both experiments. The resulting FRR of the first experiment, which will be called $\mathbf{FRR}_{\textit{all images}}$, will be analyzed to determine the impact of the age difference in the recognition process. The same will be done in the second experiment. The difference, is that in the second experiment, the resulting measure of FRR consists only on a pair of images. These pairs have a specific age difference (1, 2, 3, …, *n*) and will be called $\mathbf{FRR}_{\textit{image pair}}$. The sets for the second experiment consists on image pairs, of the same age difference, of all the subjects that match the criteria in the dataset. The $\mathbf{FRR}_{\textit{image pair}}$ results will then be compared to determine if there is any impact on the face recognition process due to the age difference of the subjects in the dataset.

Demographic information related to the study will also be collected. The demographic identifiers to be collected are the subject age, and the age difference between pictures. Information regarding to gender will also be available in chapter 4.

3.2 Sample selection

The study will use an existing aging database, MORPH (Ricanek & Tesafaye, 2006). The database contains 55,134 images of 13,000 individuals collected over five years. The non-commercial version of the MORPH database was used for this study. This version is specifically designed for academic research purposes.

The sample was selected using the following criteria and assessments:

a. Images per subject: in order to make a successful match, at least two images from the same subject are required. For this purpose the images metadata were extracted from the filename, and stored in a database. The metadata contains subject id, age and gender. Once all the metadata information was uploaded to the database, a set of queries were run in order to select the subjects that met this criteria. These queries also compare the age difference between the first, and the subsequent images from each subject. If the age difference is zero, the script tossed out the image from the study. Also, if the subject has only one image, the subject is not considered for the study. The complete analysis of the images per subject will be found in chapter 4.

b. Image Quality Metrics: after the data clean up performed in part a, the image quality metrics were calculated for the complete dataset and uploaded to the database. Another set of queries were run to remove the images whose quality metrics couldn't be extracted by the Face Profile Analysis Tool. After this assessment, the same queries used to calculate the images per subject were run again. This time, the queries used the resulting dataset after two clean-ups, images per subjects, and quality metrics. The result of these queries is the resulting dataset for the study.

The complete analysis and the tools used will be described in chapter 4.

3.3 Data collection

For this study, the data to be collected can be divided into three categories: Demographic data, Image Quality Metrics, and Matching Results.

3.3.1 Demographics

The relevant demographic information for this study is the subject age, and the age difference between among the face images from each subject. Although gender is not part of this study, some gender demographic information will also be included. The demographic detailes of the dataset will be addressed in chapter 4.

3.3.2 Image quality metrics

The Image Quality Metrics were extracted using the Face Profile Analysis Tool available at the ICBR. The quality metrics include indicators such as: background type, degree of blur, eye axis angle, and many more. A complete list of the Image Quality Metrics used for this study is available in Appendix A.

3.3.3 Matching results

The matching results will be pulled using MegaMatcher v 9.0. Once the matching results are obtained, the genuine and impostor scores were used to build the DET curves for the analysis.

3.4 External validity

External validity examines whether or not an observed causal relationship should be generalized to and accross different measures, person, settings, and times (Calder, Phillips, & Tybout, 1982). Based on this definition, potential threats for the external validity of these study are listed below.

3.4.1 Dataset

A subset of the MORPH database is the dataset used in this study. Using this database could represent a potential threats for external validity. Specifically regarding to the quality of the images, there was no previous knowledge of the quality of the whole dataset. An analysis of the quality of the dataset will be performed after the quality metrics are pulled.

3.5 Internal validity

Internal validity addresses whether or not an observed covaritaion should be considered a causal relationship (Calder et al., 1982). Based on the given defininiton, two internal validity threats were identified: selection bias effects and mortality.

3.5.1 Selection bias

Datasets to study the effects of the aging process over the automated face recognition are not easily available. The MORPH database was selected as the main data container for this study. There was no previous knowledge regarding to the quality of the images on the dataset or details of the capture process used to build the dataset.

The only variable of study is the effect of aging in the automated face recognition process, therefore, other variables' distributions such as gender, pose or illumination variation were not taken into consideration for this study.

3.5.2 Mortality

As there was no previous knowledge regarding the quality of the images, there is the probability that some of them are low-quality images. Consequently, a quality metrics analysis was performed to discard these low-quality images. The results of the analysis will be shown in CHAPTER 4.

3.6 Data analysis techniques

The performance of the matching process were measured using:

- Detection Error Trade-off (DET) curves at 1%, 0.1%, and 0.01% FAR.

The database was evaluated using:

- Quality Metrics
- Failure to Extract results
- Filename structure from the images in the dataset

CHAPTER 4. RESULTS AND ANALYSIS

Results of the experiments are divided up into Demographics, All Images and Image Pair sections. Each experiment uses the same population as explained in the following sections. The only difference is the amount of images used for each experiment. The All Images experiment, uses all the images of all the subjects for the analysis. On the other hand, the Image Pair experiment only uses the oldest, and the most recent images of the subjects that overlaps among the dataset. That being said, the subject has to be present in all of the different subsets. The subsets will be divided based on the age difference (gap) between images.

4.1 Data collection

Results of the data collection process include a description of the dataset used, and a description of the images that were considered for the final analysis.

4.1.1 Dataset demographics

The dataset used for this work was MORPH, which is a novel longitudinal face corpus that offers images of a large number of diverse subjects over a period of many years (Ricanek & Tesafaye (2006).

For MORPH's academic version there was no metadata available in the downloaded file. The demographics details were extracted from the file name with a script that was developed in PHP, and the results were stored in a MySQL database for the analysis. These details were limited to: user ID, picture ID, gender, and age. This is an example of the naming convention of the images in the dataset: 009055_0M54.jpg. Some of the fields also had to be normalized. For example, both the user ID and some picture ID's had leading zeros. These zeros were removed for data consistency purposes and in order to run queries based on a single data format per column.

Table 4.1, Table 4.2, Table 4.3, and Table 4.4 show the summary of the original dataset structure. These metric were extracted through queries over the data stored in the MySQL used for the analysis.

Table 4.1.

Number of facial images by gender (full dataset)

	Total
Males	46,645
Females	8,489
Total	55,134

Table 4.2.

Number of facial images per subject (full dataset)

	1	2	3	4	5+	Total
Males	372	2,349	3,606	1,976	3,155	11,458
Females	85	479	712	351	532	2,159
Total	457	2,828	4,318	2,327	3,687	13,617

Table 4.3.

Number of facial images by gender and decade-of-life (full dataset)

	< 20	20 – 29	30 – 39	40 – 49	50 +	Total
Males	6,638	14,016	12,447	10,062	3,482	46,645
Females	831	2,309	2,910	1,988	451	8,489
Total	7,469	16,325	15,357	12,050	3,933	55,134

Table 4.4.

Summary statistics (full dataset)

Age range	16 – 77
Age median	33
Average number of images per individual	4
Images per individual range	1 – 53

4.1.2 Image quality

The whole dataset was evaluated in terms of quality metrics. This exercise served as a starting point for the study, and also allowed to remove the images with too low quality for the study.

The complete set of results are in the Appendix B. In general, the metrics behaved as expected. In the metrics with high variability, a normal distribution was found. In the metrics with low variability, skewed distributions were found. For example, the metric of degree to which eyes are closed is expected to be left skewed. This because is not expected that the subjects to close the eyes during the capture. On the other hand, metrics such as eye axis angle has more variability and it presents a normal distribution.

This expected behavior, in terms of distributions, made the selection process easily as the images could be selected based on the established criteria with no additional concerns. It allowed to build 4 datasets for the study.

4.1.3 Data clean-up

The dataset contained low-quality-images for which metrics would be hard or even impossible to extract. Before performing a final analysis the dataset was cleaned up using queries over the dataset, and using a Face Profile Analysis Tool available at the ICBR.

The Face Analysis Profile Tool outputs CSV files with the images quality metrics using the ISO Frontal profile for the analysis. The tool also outputs a report detailing the images which metrics were extracted. Then these reports were uploaded to the database and compared against the full dataset. The subjects with images which metrics couldn't been extracted were removed from the dataset.

A total of 601 subjects were removed from the study due to image quality issues, and issues with the subject's data itself. Besides the failure to extract, three other issues were identified within the subject's subset:

- Random order of the images: some of the images in the subject subset were not sequentially ordered. Therefore, some of the pictures seems to not be congruent with the ages found in the image metadata. It was decided to remove the subjects with this issues from the dataset because it makes it difficult to define whether the supposed age of the picture was the real one.

- Subject subsets with other subject images: there were subsets with images from other subjects.
- Subjects with a single image: there were subjects with only one image in the dataset. This could be attributed to the fact that the user only did one visit.

Subsets with these three issues were also removed from the study. Table 4.5 shows the amount of subjects removed and retained for the study.

Table 4.5.

Removed and retained subjects

	Total
Removed	601
Retained	13,016
Total	13,617

Once the subjects with issues were removed from the original dataset, the final dataset was built. In some cases, there were some subjects with multiple images from the same age. The dataset was build using the demographic data gathered from the picture filename, and a PHP script that selects the image with the maximum image ID for any age in the subject's data subset. For example, if a subject has two pictures at the age of 36 with ID 0 and 1 respectively, the PHP script only takes the one with the ID 1, provided that it is not the first image in the subject's data subset.

Table 4.6, Table 4.7, Table 4.8, and Table 4.9 show the structure of the final dataset to be used for the analysis.

Table 4.6.

Number of facial images by gender (final dataset)

	Total
Male	21,592
Female	3,979
Total	25,571

Table 4.7.

Number of facial images per subject (final dataset)

	2	3	4	5+	Total
Males	5,284	2,798	533	97	8,712
Females	1,039	482	84	23	1,628
Total	6,323	3,280	617	120	10,340

Table 4.8.

Number of facial images by gender and decade-of-life (final dataset)

	< 20	20 – 29	30 – 39	40 – 49	50 +	Total
Males	2,772	6,490	6,139	4,631	1,560	21,592
Females	386	1,114	1,333	940	206	3,979
Total	3,158	7,604	7,472	5,571	1,766	25,571

Table 4.9.

Summary statistics (final dataset)

Age range	16 – 77
Age median	33
Average number of images per individual	2
Images per individual range	2 – 7

4.2 Matching process

Once the dataset was cleaned up, the next step was to match the pictures in the different datasets. The matching process consists of two main steps: the enrollment and the matching.

Both steps were performed using MegaMatcher v 9.0 MegaMatcher is a commercially available technology for large-scales Automatic Biometric Identifications Systems (Neurothecnology, 2016).

4.2.1 Enrollment

The enrollment process consisted in creating the datasets subsets for the matching process. For this study several datasets were created. Table 4.10 shows the datasets created for the experiments in this study.

Table 4.10.

Datasets

Dataset Name	Experiment	Total Pictures
Full dataset	1	14,000
Age difference 1	2	194
Age difference 2	2	194
Age difference 3	2	194
Age difference 4	2	194

Even though the dataset after the final cleanup had 25,571, all the images were not used for the matching process. In order to measure the matching performance for the second experiment, the same subjects needed to be present in all the age difference based subsets.

There were only 97 subjects (2 images per subject) that were present in all the subsets for the age difference matching experiment.

4.2.2 Matching

The matching process consisted in matching the templates generated in the enrollment process. In order to measure the matching process performance, each template is matched against all the other templates in the dataset.

The output of this process are the genuine scores and the impostor scores. The genuine scores are obtained comparing two images that belong to the same person, while the impostor scores are the result of comparing two images that belong to different people (He et al., 2010).

The genuine and impostor scores are used to create the Detection Error Trade-off (EDT) charts used for the analysis of the matching performance.

4.3 Data analysis

The analysis of this study relied on the Detection Error Trade-off (EDT) curves. The DET curves can show the relationship between the false reject rate (FRR), the percentage of genuine users that are rejected, and false accept rate (FAR), the percentage of imposters that are accepted (Johnson et al., 2010) at different levels.

These charts also estimate the equal error rate (EER), or the point at which the false match rate is equal to the false non-match rate. The DET curves were chosen as a suitable graphical representation, and were plotted using the Bio-Metrics Performance Metrics Software by Oxford Wave Research Bio-Metrics (OxfordWaveResearch, 2015).

The application uses the genuine, and the impostor scores as input. The format of the input is a .dat file. The output is a graphical representation through the DET curves.

4.4 All-images experiment

For the all picture experiment, all the images of all the subjects in the dataset were used during the matching process. It did not matter the amount of pictures from each subject, the intention was to gain a general overview from the dataset in terms of matching performance.

Figure 4.1 shows the resulting DET curve for the whole dataset and Table 4.11 shows the summary of the FAR and FRR from the whole dataset. The EER, where the FAR and the FRR are equal, is 4.3641%, which means that per every 100 subjects, the algorithm will falsely accept or reject between four and five subjects. In settings requiring high security this measure could be considered as too high.

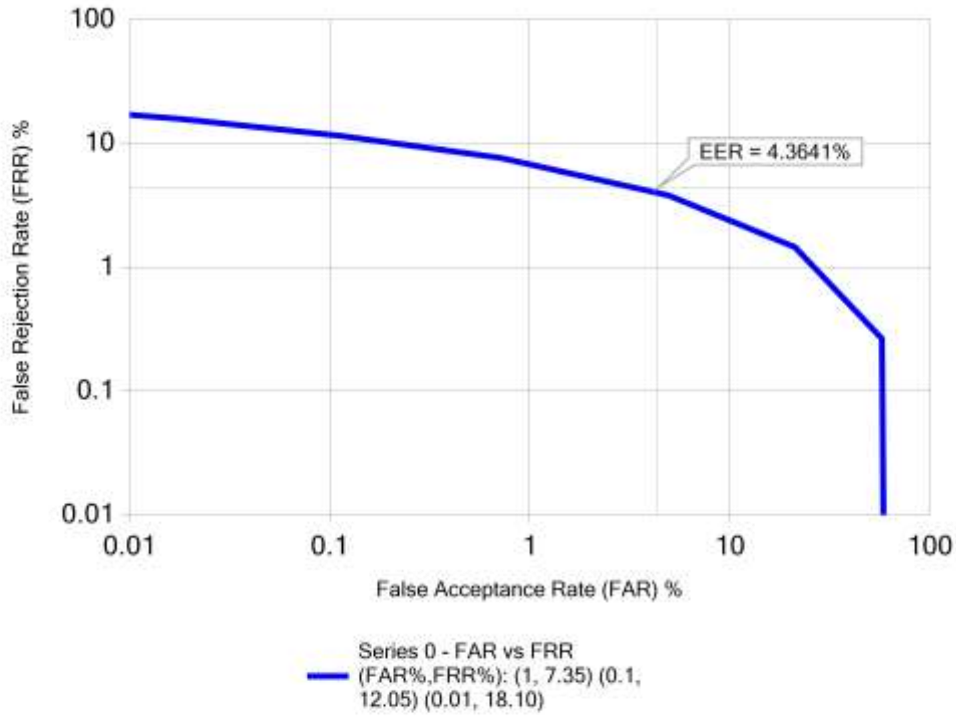

Figure 4.1. Whole dataset EDT curve

Table 4.11.

Full Dataset FAR vs FRR

FAR	FRR
1%	7.35%
0.1%	12.05%
0.01%	18.10%

The above results will be used as a comparison measure between the experiments.

4.5 Image-pairs experiment

For the image pairs experiment only the subjects that has images in all the datasets were selected for the matching process. In total there were ninety-seven subjects who has images in datasets one through four. Each dataset number represents the distance (gap), in years, between the pictures. For example, dataset one has ninety-seven pairs of images with one-year difference between them. Dataset two has two years' difference. The same happens for dataset three and four as well.

The following sections will show the results of the analysis performed over the resulting datasets. The results, as in the first experiment, consisted in the EDT curve, and a table showing the comparison between the FAR and the FRR at FAR levels of 1%, 0.1% and 0.01%.

The explanation of the second experiment results, in terms of FAR and FRR for each sub-experiment, will be the same. Instead of explaining each of them separately, the whole results will be discussed in the section following to the results.

4.5.1 One-year gap

This experiment consisted on ninety-seven subjects whose images (pair) were taken with one year of difference (gap). Figure 4.2, and Table 4.12 show the results of this experiment.

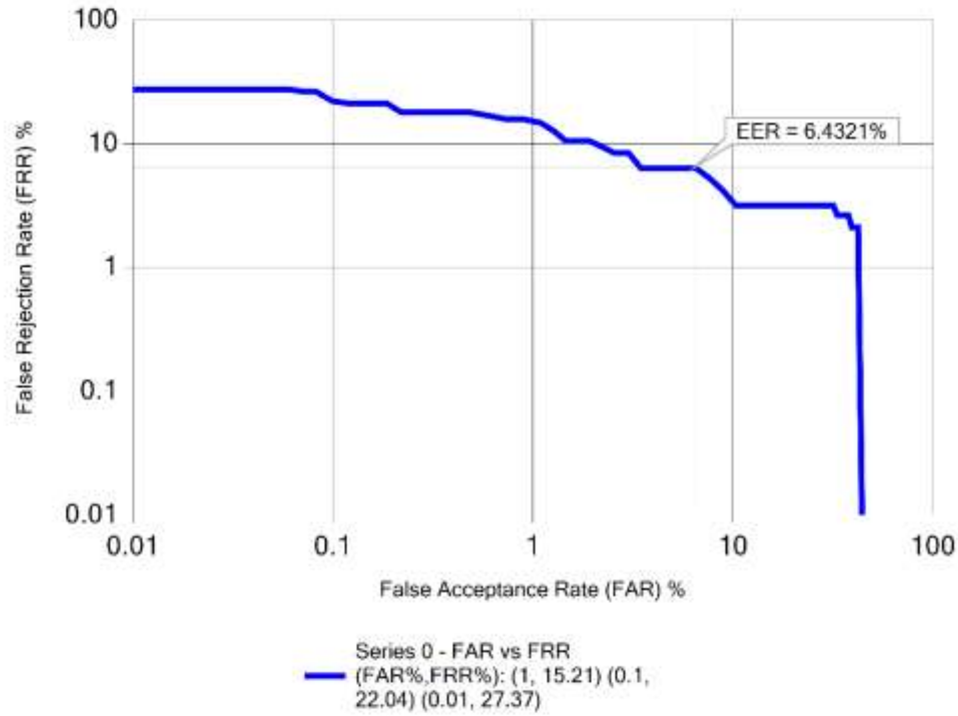

Figure 4.2. Dataset one EDT curve

Table 4.12.

Dataset one FAR vs FRR

FAR	FRR
1%	15.21%
0.1%	22.04%
0.01%	27.37%

4.5.2 Two-year gap

This experiment consisted on ninety-seven subjects whose images (pair) were taken with two years of difference (gap). Figure 4.3, and Table 4.13 show the results of this experiment.

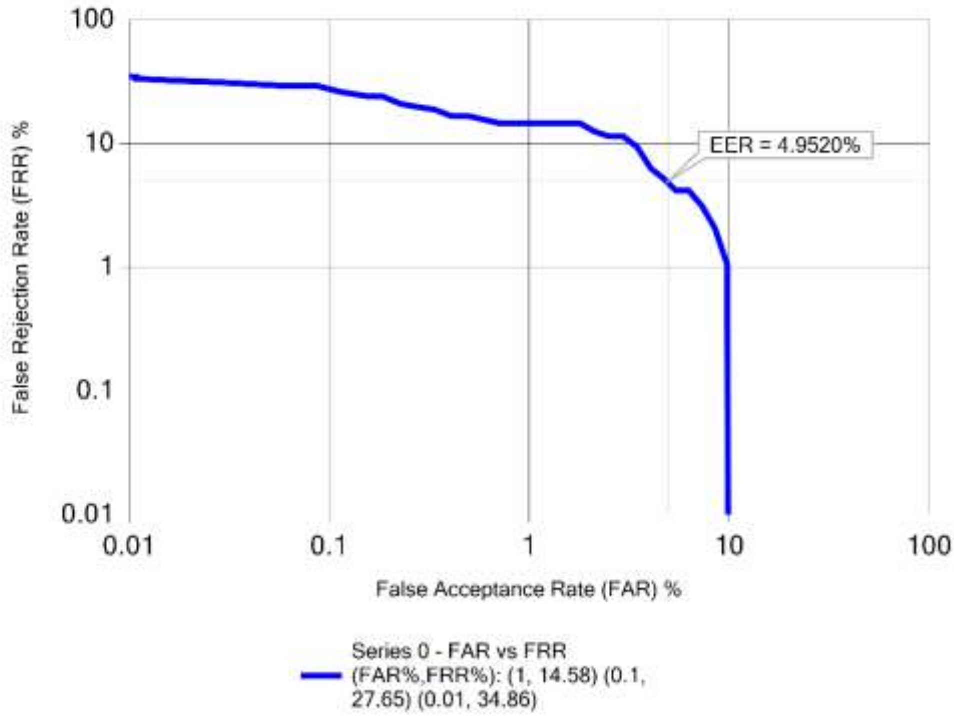

Figure 4.3. Dataset two EDT curve

Table 4.13.

Dataset two FAR vs FRR

FAR	FRR
1%	14.58%
0.1%	27.65%
0.01%	34.86%

4.5.3 Three-year gap

This experiment consisted on ninety-seven subjects whose images (pair) were taken with three years of difference (gap). Figure 4.4, and Table 4.14 show the results of this experiment.

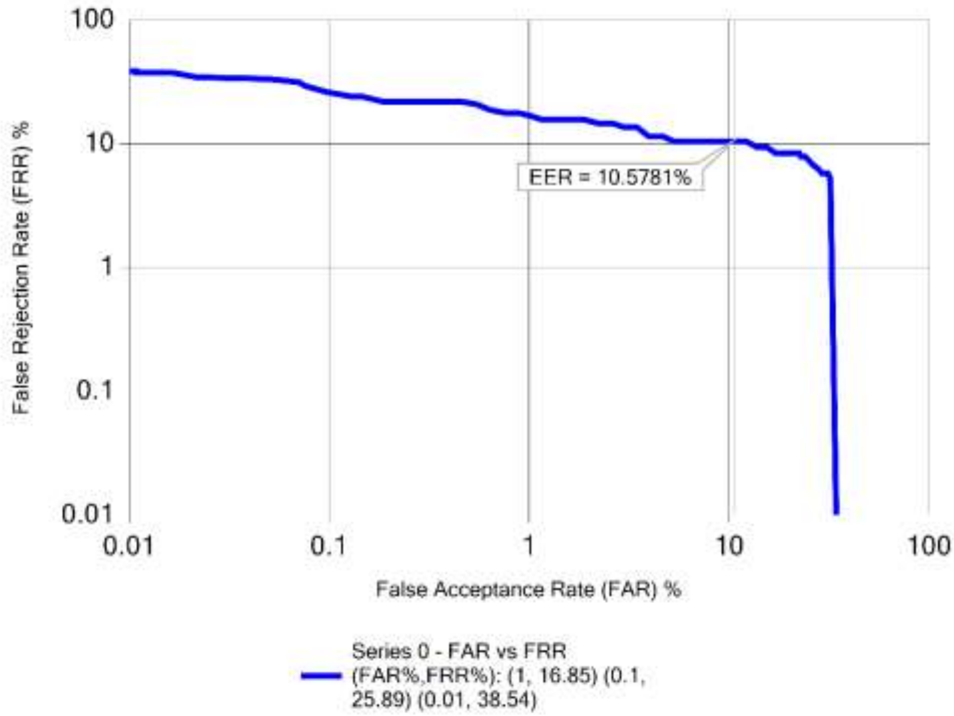

Figure 4.4. Dataset three EDT curve

Table 4.14.

Dataset three FAR vs FRR

FAR	FRR
1%	16.85%
0.1%	25.89%
0.01%	38.54%

4.5.4 Four-year gap

This experiment consisted on ninety-seven subjects whose images (pair) were taken with four years of difference (gap). Figure 4.5, and Table 4.15 show the results of this experiment.

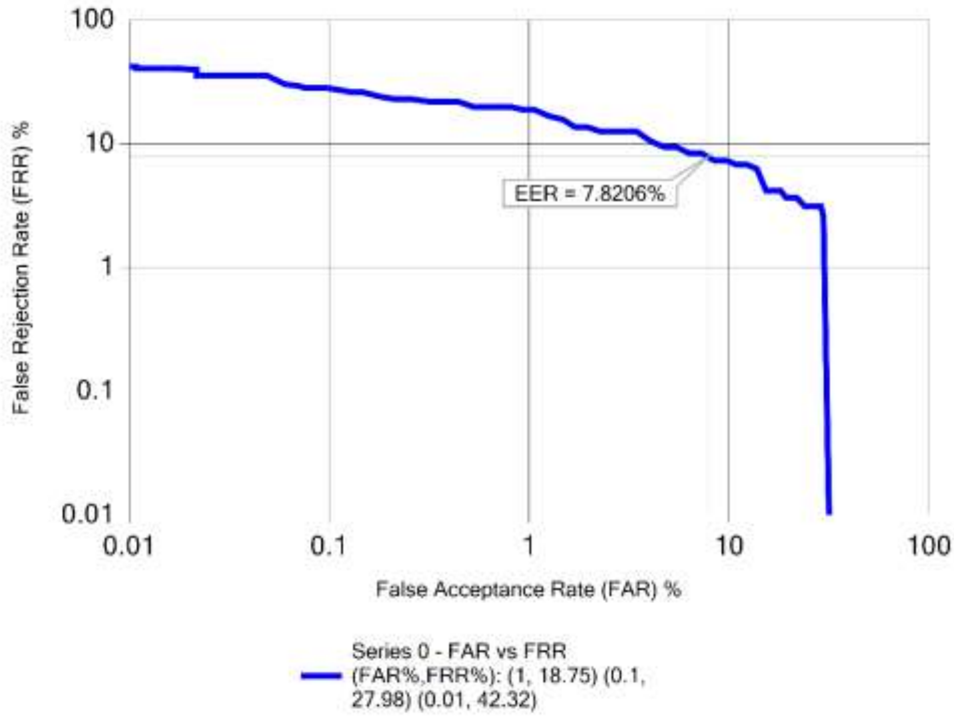

Figure 4.5. Dataset four EDT curve

Table 4.15.

Dataset four FAR vs FRR

FAR	FRR
1%	18.75%
0.1%	27.98%
0.01%	42.32%

4.6 Summary of experiments

Matching performance was compared across all the datasets. Table 4.16 shows the FRR values for all the datasets at 1%, 0.1% and 0.01% FAR. All these values were extracted from the DET curves, presented in the previous sections, from each of the datasets.

Table 4.16.

All datasets: FAR vs FRR

FAR	FRR				
	Baseline	1	2	3	4
1%	7.35%	15.21%	14.58%	16.85%	18.75%
0.1%	12.05%	22.04%	27.65%	25.89%	27.98%
0.01%	18.10%	27.37%	34.86%	38.54%	42.32%

Table 4.17.

All datasets EER

ERR				
Baseline	1	2	3	4
4.3641%	6.4321%	4.9520%	10.5781%	7.8206%

The results shown that the best matching performance was obtained with datasets one, and two. There is a slightly difference between both datasets, in terms of FRR. Dataset two performed better that dataset one at 1% FAR with 14.58% FRR. Dataset one performed slightly better at 0.1% and 0.01% FAR with 22.04% and 27.37% FRR respectively.

The results are the same when comparing them against the ones obtained from the full dataset. Research work has shown that the performance decreases when the gap between the pictures increases (Gohringer, 2012). The same happened in the experiments performed in this research work. The performance steadily decreases while the gap between the images was increased. It is worth to note that the higher the value of the FRR, the lower the performance.

The similarity of the results obtained with dataset one, and two could be attributed to the gap between the pictures and to other factors, such as gender, and ethnicity that were not taken into account for this research.

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

The purpose of this study was to understand the impact of aging in the automated face recognition process. The results of this work lead us to some conclusions and recommendations for future research work.

5.1 Conclusions

The results of the study suggest that the algorithm performed better on datasets one, and two. Datasets one, and two contained images with gaps of one, and two years respectively. The FRR values were quite similar between these datasets. FAR. Dataset one performed better that dataset two at 0.1% FAR with 22.04% FRR compared to the 27.65% FRR from dataset two. Dataset two performed slightly better at 1% FAR with 14.58% FRR. This, compared against 15.21% FRR of dataset one at 1% FAR.

FRR values grew consistently as the gap between the images increased besides the values of datasets one, and two previously mentioned. It can be concluded that the bigger the gap between the images the lower the performance of the algorithm. The study shows that the results of the aging process impacts the performance of the algorithm. This conclusion is reached based on the resulting FAR values for each dataset. Figure 5.1 shows that the results are also congruent with the results of similar studies (Gohringer, 2012).

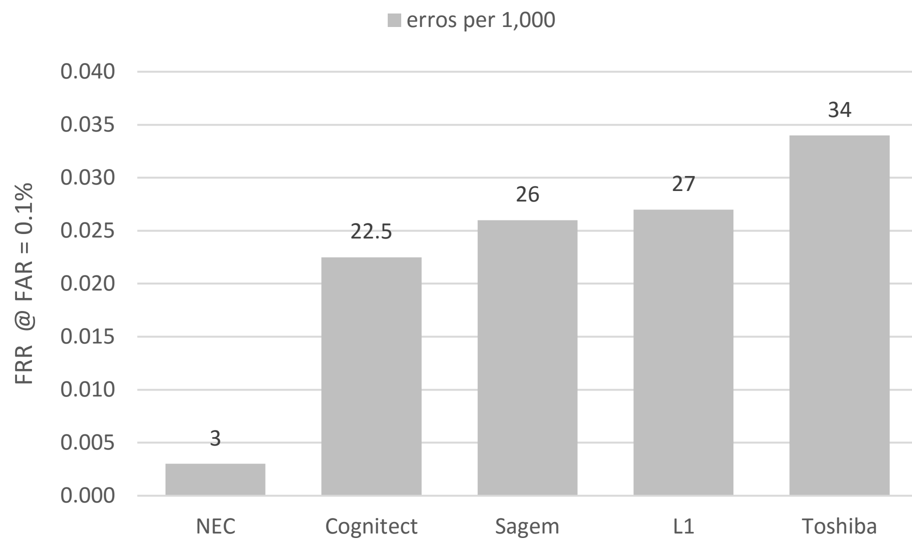

Figure 5.1. FRR results from NIST MBE 2010 (Gohringer, 2012)

5.2 Recommendations for future research

Effects of aging in face recognition are not the only factor impacting the face recognition process. There are other areas that could be explored for future research:

- Demographical study: the effects of aging could be measured based on the ethnicity and/or gender of the population.
- Use a larger dataset: the amount of images for the second experiment was small. With more datasets, with larger gaps between images, in depth results could be gotten.
- Algorithm vs. Operator performance: an experiment to measure the difference in the recognition process performance between the algorithm, and a human operator could be performed. This could also be used to measure factors impacting the operator's performance.

- Low vs. high quality images: compare algorithm performance between low, and high quality aging datasets. The dataset used for this study had variable conditions and low quality images.

LIST OF REFERENCES

LIST OF REFERENCES

Albert, M., Sethuram, A., & Ricanek, K. (2011). *Implications of adult facial aging on biometrics*. INTECH Open Access Publisher.

Biswas, S., Aggarwal, G., Ramanathan, N., & Chellappa, R. (2008). A Non-generative Approach for Face Recognition Across Aging. In *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems* (pp. 1–6). IEEE. doi:10.1109/BTAS.2008.4699331

Browning, K., & Orlans, N. (2014). Biometric Aging.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research*, 240–244.

Cament, L. A., Galdames, F. J., Bowyer, K. W., & Perez, C. A. (2015). Face recognition under pose variation with local Gabor features enhanced by Active Shape and Statistical Models. *Pattern Recognition*, *48*(11), 3371–3384. doi:10.1016/j.patcog.2015.05.017

Chao, W. (2011). Gabor wavelet transform and its application.

Coleman, S. R., & Grover, R. (2006). The anatomy of the aging face: volume loss and changes in 3-dimensional topography. *Aesthetic Surgery Journal*, *26*(1 Supplement), S4–S9.

Crook, T. H., & Larrabee, G. J. (1992). Changes in facial recognition memory across the adult life span. *Journal of Gerontology*, *47*(3), P138–P141.

De Marsico, M., Nappi, M., Riccio, D., & Wechsler, H. (2011). Robust face recognition after plastic surgery using local region analysis. In *Image Analysis and Recognition* (pp. 191–200). Springer.

Esme, B., & Sankur, B. (2010). Effects of Aging over Facial Feature Analysis and Face Recognition, (March), 1–4.

Fenker, S. P., & Bowyer, K. W. (2012). Analysis of template aging in iris biometrics. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on* (pp. 45–51).

Friedman, O. (2005). Changes associated with the aging face. *Facial Plastic Surgery Clinics of North America*, *13*(3), 371–380.

Gavrilova, M. L., & Monwar, M. (2013). *Multimodal biometrics and intelligent image processing for security systems*. IGI Global.

Gohringer, C. (2012). The application of face recognition in airports. *Biometric Technology Today*, *2012*(7), 5–9. doi:10.1016/S0969-4765(12)70148-0

Goldstein, A. J., Harmon, L. D., & Lesk, A. B. (1971). Identification of human faces. *Proceedings of the IEEE*, *59*(5), 748–760.

Grgic, M., & Delac, K. (2008). Face Recognition Algorithms. Retrieved March 5, 2015, from http://www.face-rec.org/algorithms/

Grother, P. J., Quinn, G. W., & Phillips, P. J. (2010). Report on the evaluation of 2D still-image face recognition algorithms. *NIST Interagency Report*, *7709*, 106.

Han, H., Shan, S., Chen, X., & Gao, W. (2013). A comparative study on illumination preprocessing in face recognition. *Pattern Recognition*, *46*(6), 1691–1699.

He, M., Horng, S.-J., Fan, P., Run, R.-S., Chen, R.-J., Lai, J.-L., … Sentosa, K. O. (2010). Performance evaluation of score level fusion in multimodal biometric systems. *Pattern Recognition*, *43*(5), 1789–1800.

ISO. (2012). ISO/IEC 19795 Information technology - Biometric performance testing and reporting.

ISO. (2014). ISO/IEC 19794-7 Information technology - Biometric data interchange formats.

ISO/IEC JTC1 SC37 Working Group 1. (2007). ISO/IEC JTC1 SC37 Standing Document 2 (SD 2 – Version 5), Harmonized Biometric Vocabulary.

Johnson, P. A., Tan, B., & Schuckers, S. (2010). Multimodal fusion vulnerability to non-zero effort (spoof) imposters. In *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on* (pp. 1–5).

Kindt, E. J. (2013). *Privacy and Data Protection Issues of Biometric Applications*. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-7522-0

Lamont, A. C., Stewart-Williams, S., & Podd, J. (2005). Face recognition and aging: effects of target age and memory load. *Memory & Cognition*, *33*(6), 1017–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16496722

Lanitis, A. (2009). Facial Biometric Templates and Aging: Problems and Challenges for Artificial Intelligence. In *AIAI Workshops* (pp. 142–149).

Lanitis, A., Draganova, C., & Christodoulou, C. (2004). Comparing different classifiers for automatic age estimation. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *34*(1), 621–628.

Lata, Y. V., Kiran, C., Tungathurthi, B., Ram, H., Rao, M., Govardhan, A., & Reddy, L. P. (2009). Facial recognition using eigenfaces by PCA.

Li, & Jain, A. (2011). *Handbook of face recognition* (2nd ed.). London ; New York : Springer.

Li, Li, B., Xu, Y., Lu, K., Yan, K., & Fei, L. (2014). Disguised face detection and recognition under the complex background. In *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2014 IEEE Symposium on* (pp. 87–93).

Ling, H., Soatto, S., Ramanathan, N., & Jacobs, D. W. (2007). A study of face recognition as people age. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on* (pp. 1–8).

Macri, G. (2014). Boston PD Facial Recognition Test Recorded Faces At Festival. *The Daily Caller*. Retrieved April 13, 2015, from http://dailycaller.com/2014/08/18/boston-pd-tested-facial-recognition-software-by-recording-every-face-at-local-music-festivals/

Majekodunmi, T. O., & Idachaba, F. E. (2011). A Review of the Fingerprint, Speaker Recognition, Face Recognition and Iris Recognition Based Biometric Identification Technologies.

Mansfield, A. J., & Wayman, J. L. (2002). *Best practices in testing and reporting performance of biometric devices*. Centre for Mathematics and Scientific Computing, National Physical Laboratory Teddington, Middlesex, UK.

Martínez, A. M., & Kak, A. C. (2001). Pca versus lda. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(2), 228–233.

Moore, K. L., Dalley, A. F., & Agur, A. M. R. (2013). *Clinically oriented anatomy*. Lippincott Williams & Wilkins.

Morpho. (2010). *Automatic Facial Recognition: A Review*.

NEC. (2015). NEC provides Face Recognition Technology to Surat City Police. *Press Releases*. Retrieved July 30, 2015, from http://www.nec.com/en/press/201502/global_20150225_03.html

Neurothecnology. (2016). SDK for large-scale AFIS or multi-biometric face-voice-iris-palm-fingerprint systems. Retrieved March 20, 2016, from http://www.neurotechnology.com/megamatcher.html

NSTC Subcommittee on Biometrics. (2006). Biometrics “Foundation Documents.” Retrieved February 28, 2015, from http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/biometric-center-of-excellence/modalities/facial-recognition

Oh, H. J., Lee, K. M., & Lee, S. U. (2008). Occlusion invariant face recognition using selective local non-negative matrix factorization basis images. *Image and Vision Computing*, *26*(11), 1515–1523.

OxfordWaveResearch. (2015). Bio-Metrics Performance Metrics Software. Retrieved April 23, 2014, from http://www.oxfordwaveresearch.com/j2/products/bio-metrics-1-1-performance-metrics-software-2

Park, U., Tong, Y., & Jain, A. K. (2010). Age-invariant face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *32*(5), 947–954.

Parkin, A. J. (1993). *Memory: Phenomena, experiment and theory*. Psychology Press.

Planet Biometrics. (2015). Indian city goes live with "crime-busting" face recognition surveillance. Retrieved July 30, 2015, from http://www.planetbiometrics.com/article-details/i/3293/desc/indian-city-goes-live-with-crime-busting-face-recognition-surveillance/

Press, W. H. (2007). *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press.

Ramanathan, N., & Chellappa, R. (2008). Modeling shape and textural variations in aging faces. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on* (pp. 1–8).

Ricanek, K., & Tesafaye, T. (2006). Morph: A longitudinal image database of normal adult age-progression. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on* (pp. 341–345).

Russon, M.-A. (2015). FBI $1bn facial recognition system caught a paedophile 19 years after he went on the run. *International Business Times*. Retrieved December 9, 2015, from http://www.ibtimes.co.uk/fbi-1bn-facial-recognition-system-caught-paedophile-19-years-after-he-went-run-1509954

Shah, J. H., Sharif, M., Raza, M., & Azeem, A. (2014). Face recognition across pose variation and the 3S problem. *Turkish Journal of Electrical Engineering and Computer Science*, *22*(6), 1423–1436.

Singh, R., Vatsa, M., & Noore, A. (2009). Effect of plastic surgery on face recognition: A preliminary study. In *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on* (pp. 72–77).

Slavković, M., & Jevtić, D. (2012). Face recognition using eigenface approach. *Serbian Journal of Electrical Engineering*, *9*(1), 121–130.

Strategic Defence Intelligence. (2015). *The Global Government Biometric Systems Market 2015-2025*. Retrieved from http://www.marketresearchreports.biz/analysis-details/the-global-government-biometric-systems-market-2015-2025

Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on* (pp. 586–591).

University of North Carolina Wilmington. (2015). MORPH Database. Retrieved December 10, 2015, from https://ebill.uncw.edu/C20231_ustores/web/store_main.jsp?STOREID=4

Wang, X., & Pan, H. (2014). Research on Fast Face Detection Algorithm in complex background. In *Advanced Materials Research* (Vol. 1030, pp. 1779–1782).

Wang, X., Zhang, X., & Yao, J. (2011). Skin color detection under complex background. In *Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference on* (pp. 1985–1988).

Wiskott, L., Fellous, J.-M., Kuiger, N., & Von Der Malsburg, C. (1997). Face recognition by elastic bunch graph matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *19*(7), 775–779.

Woodward, J. D., Horn, C., Gatune, J., & Thomas, A. (2003). *Biometrics: A look at facial recognition*.

Zeng, J., Ling, H., Latecki, L. J., Fitzhugh, S., & Guo, G. (2012). Analysis of Facial Images across Age Progression by Humans. *ISRN Machine Vision*, *12*, 7. Retrieved from http://www.hindawi.com/journals/isrn.machine.vision/2012/505974/cta/

Zhang, C., & Zhang, Z. (2010). *A survey of recent advances in face detection*.

Zhang, & Lu, G. (2013). 3D Biometrics Technologies and Systems. In *3D Biometrics* (pp. 19–33). Springer.

Zhang, Wang, L., Zhu, Q., Liu, Z., & Chen, Y. (2015). Noise modeling and representation based classification methods for face recognition. *Neurocomputing*, *148*, 420–429.

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, *35*(4), 399–458.

APPENDICES

Appendix A Metric's Definitions and Scores

For the purposes of this study, we use the Face Profile Analysis Tool available at the ICBR. The software outputs the metrics shown in Table A.2. The definitions in Table A.2 are similar to the ones in the ANSI INCTIS 385-2004, and the ISO/IEC 19794-5:20005. The individual ranges for each metric are shown in Table A.1.

Table A.1.

Face metric's scoring

Metric Name	Scoring
Facial Dynamic Range	1 to 8
% Background Uniformity	0 to 100%
% Background Gray	0 to 100%
% Facial Brightness	0 to 100%
% Facial Saturation	0 to 100%
Brightness Score	1 to 5
Eye Contrast	1 to 5
Degree of Background Clutter	0 to 5
Degree of Blur	0 to 5
Background Type	Simple or Complex
Eye Separation	Variable Integer
Eye Axis Angle	Variable degree
Eye Axis Location Ratio	0 to 1
Centerline Location Ratio	0 to 1
Image-Width to Head-Width Ratio	Variable Integer
Head-Height to Image-Height Ratio	Variable Integer
Height to Width Ratio	Variable Integer

Table A.2.

Metrics Scoring and Definitions

Metric Name	Scoring
Facial Dynamic Range	Facial Dynamic Range indicates the number of bits in the dynamic range of the facial region of the input image. A minimum of 7 is required.
% Background Uniformity	Percent Background Uniformity reflects the variation of color throughout the background of the image. Values can be in the range 0 to 100%. Optimal is 100%
% Background Gray	Percent Background Gray reflects the level of fray in the background. Values can be in the range 0 to 100%. Optimal is usually 18%.
% Facial Brightness	Percent Facial Brightness is the average luminance of the facial region as a percent.
% Facial Saturation	Percent Facial Saturation is the percent fraction of pixels saturated in the facial region.
Brightness Score	Brightness indicates how well the dynamic range is centered in the facial region of the image. A valid value for this score will be an integer in the range of 1 to 5. Ideally this value should be greater than or equal to 3. Values below 3 indicate that the facial region may be too dark. A special value of 0 applies to facial images that have too much saturated black.
Eye Contrast	Eye Contrast indicates how well the dynamic range is spread in the eye regions of the image. The contrast value will be an integer in the range of 1 to 5. A score of 3 or higher is adequate (the higher the better). A score of 2 or less is inadequate.
Degree of Background Clutter	Degree of Clutter indicates how much background clutter occurs in the image. Scores are in the range 0 to 5. With 0 indicating no background clutter and 5 indicating no background clutter and 5 indicating a high degree of background clutter.

Table A.2 Continued

Metrics Scoring and Definitions

Metric Name	Scoring
Degree of Blur	Degree of Blur Indicates how much focus and/or motion blur is present in the image. Scores are in the ranges 0 to 5.
Background Type	Simple or Complex.
Eye Axis Angle	Eye axis angle is the slope of the eye-axis measured in degrees clockwise (positive) from the horizontal.
Eye Axis Location Ratio	Eye Axis Location Ratio is the location of the eyes axis as a fraction of the image height up from the bottom.
Centerline Location Ratio	Centerline Location Ratio is the location of the centerline as a fraction of the image width measured from the left side of the image.
Image-Width to Head-Width Ratio	Image Width to Head Width Ratio is the ratio of image width to head width.
Head-Height to Image-Height Ratio	Head Height to Image Height Ratio is the ratio of the head height to image height.
Height to Width Ratio	Height to Width is the ratio of image height to image width.

Appendix B Image Quality Metrics Charts

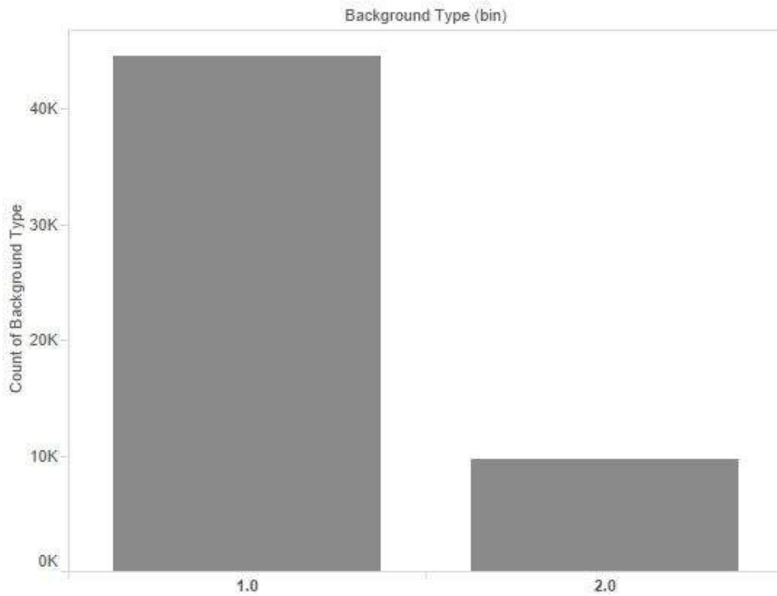

Figure B.1.1. Background Type

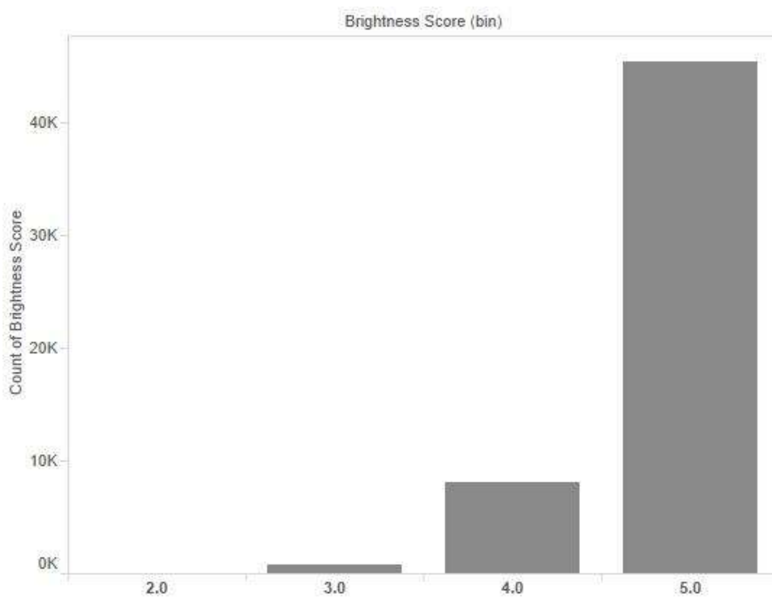

Figure B.1.2. Brightness Score

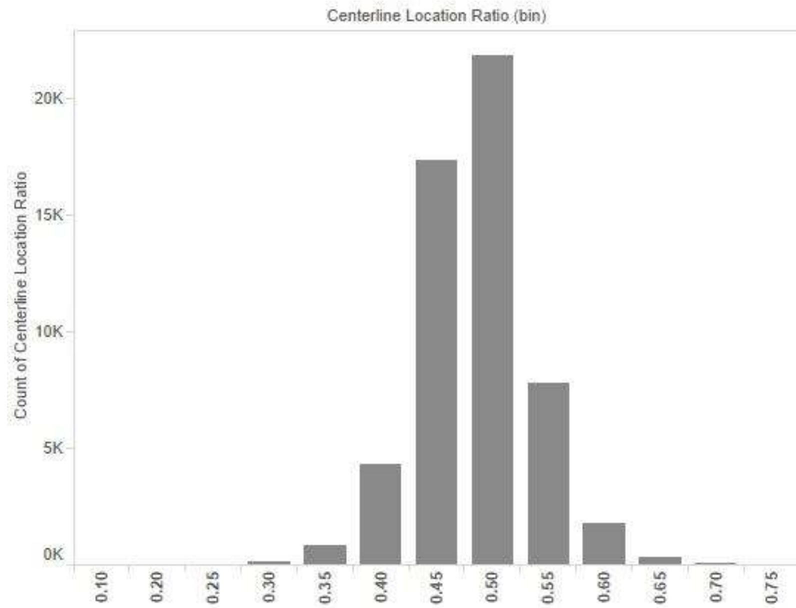

Figure B.1.3. Centerline Location Ratio

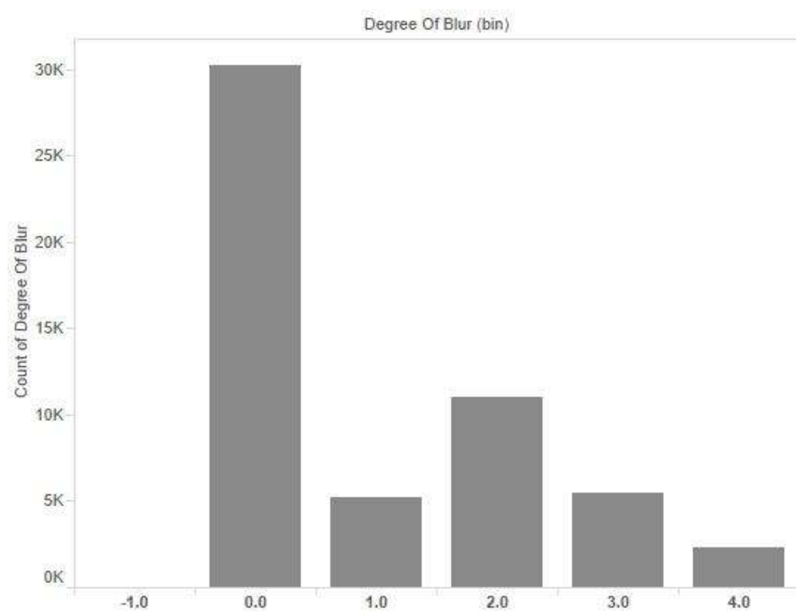

Figure B.1.4. Degree of Blur

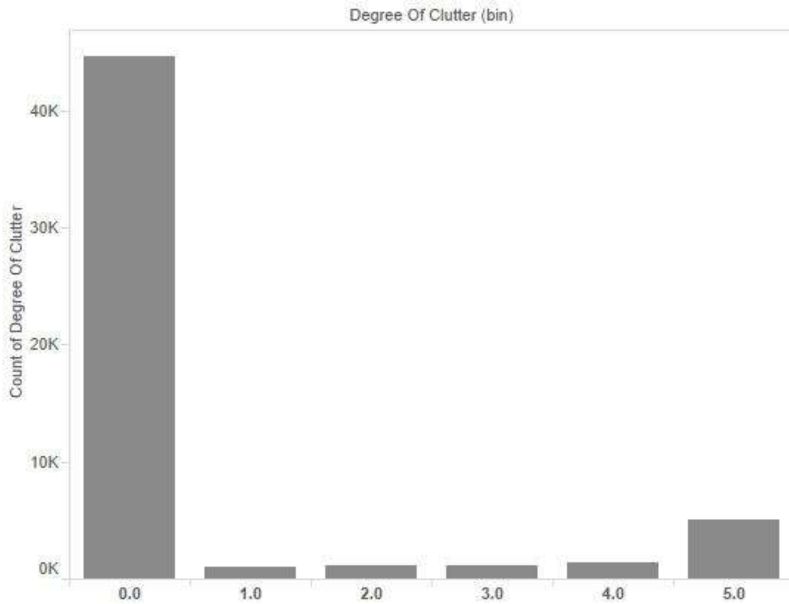

Figure B.1.5. Degree of Clutter

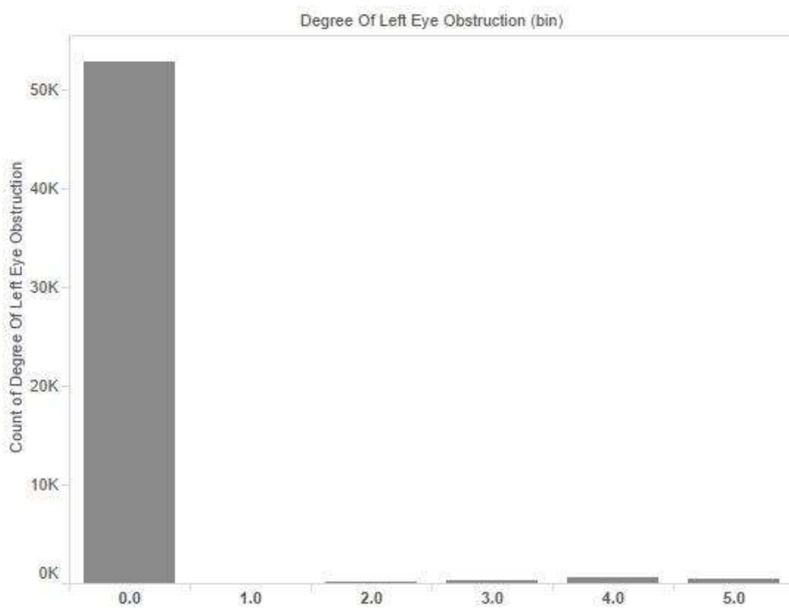

Figure B.1.6. Degree of Left Eye Obstruction

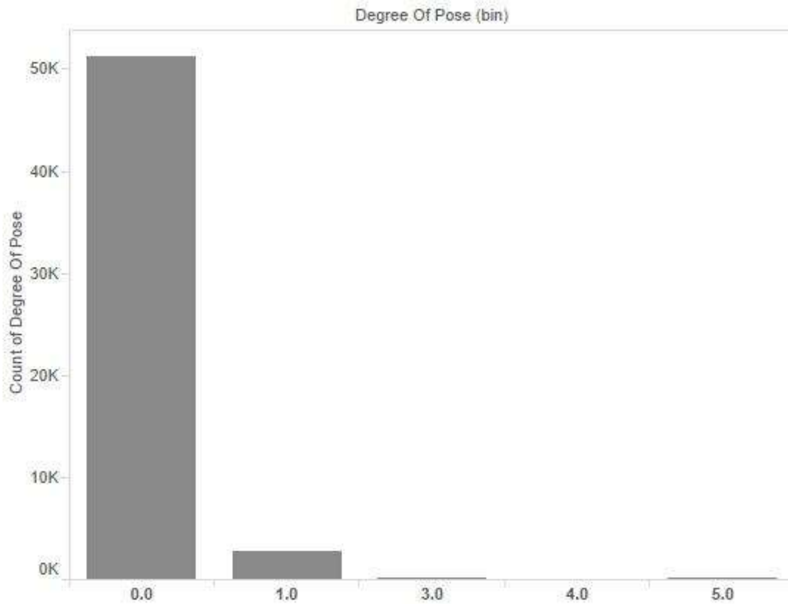

Figure B.1.7. Degree of Pose

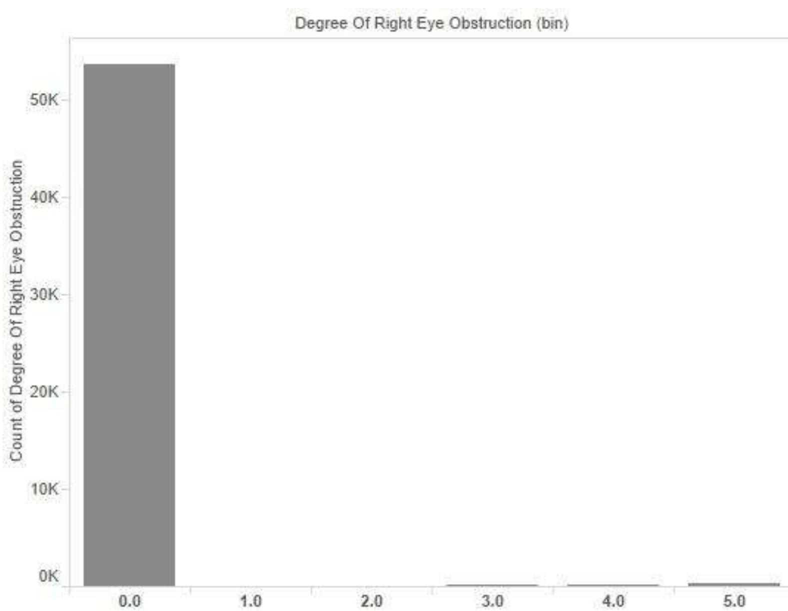

Figure B.1.8. Degree of Right Eye Obstruction

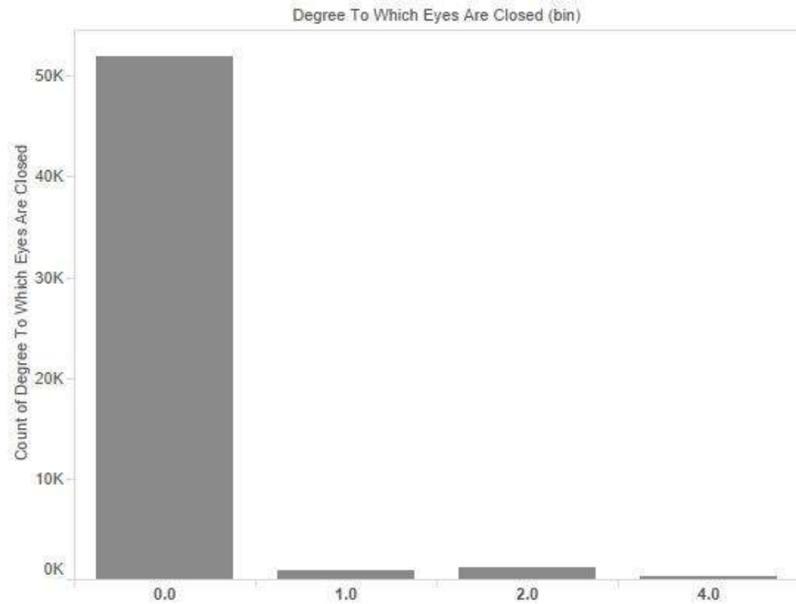

Figure B.1.9. Degree to Which Eyes are Closed

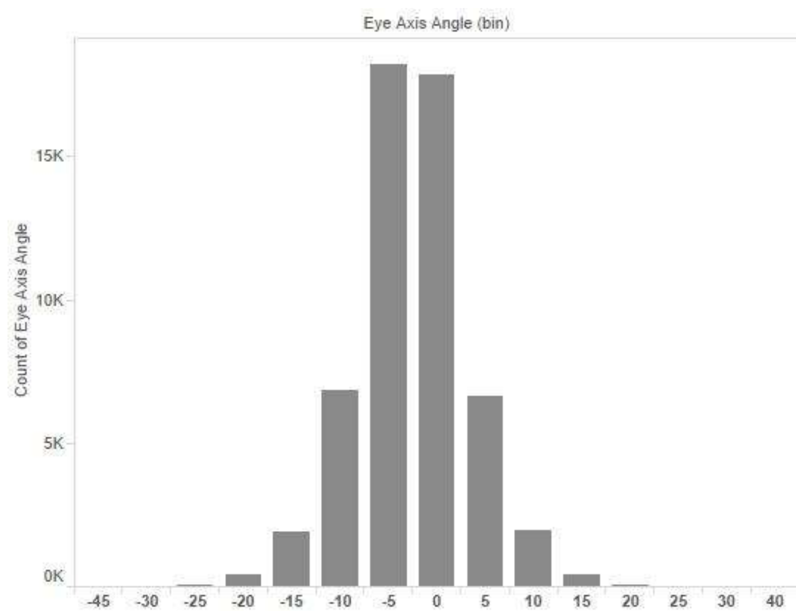

Figure B.1.10. Eye Axis Angle

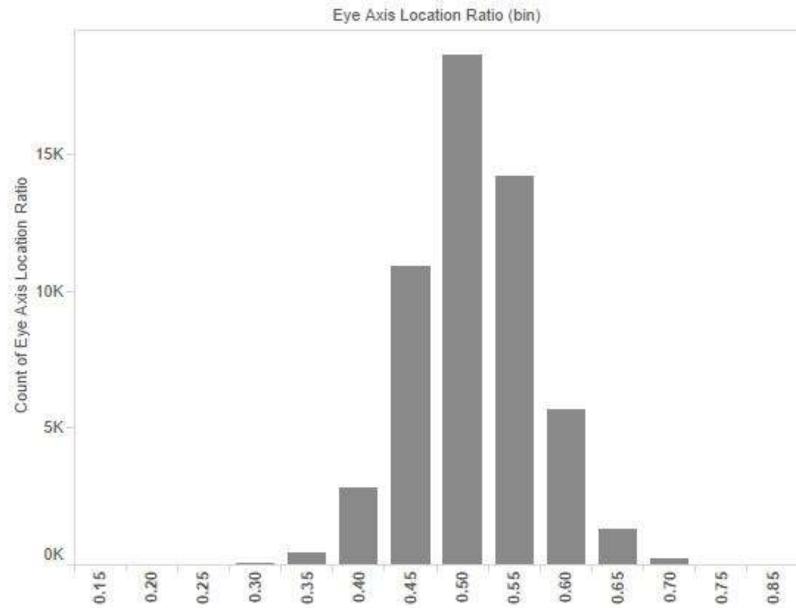

Figure B.1.11. Eyes Axis Location Ratio

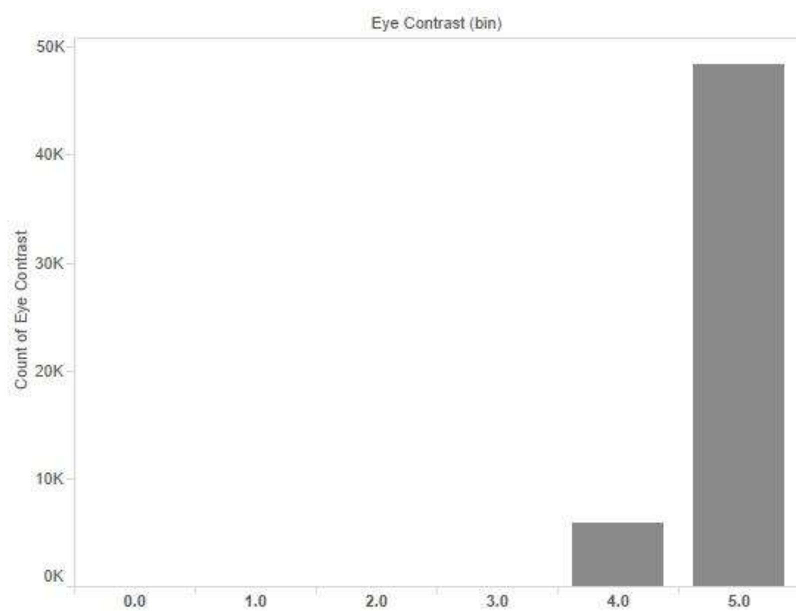

Figure B.1.12. Eye Contrast

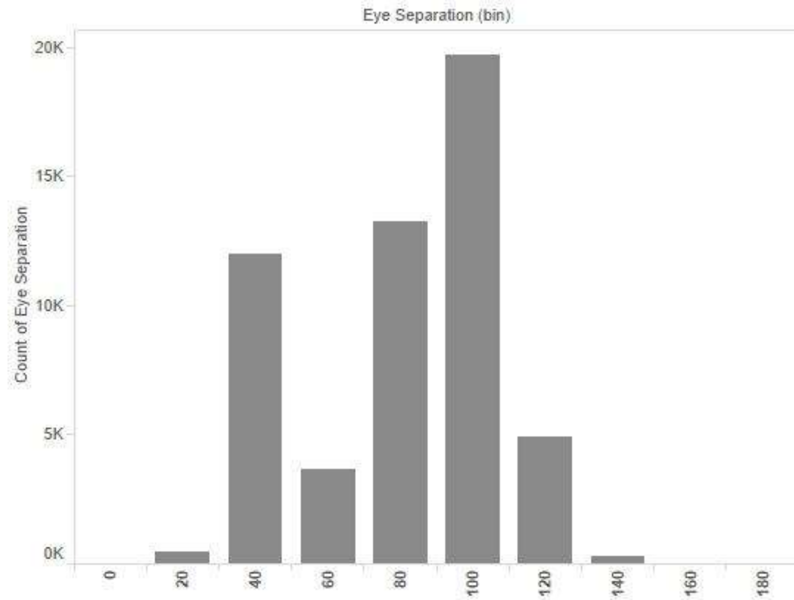

Figure B.1.13. Eye Separation

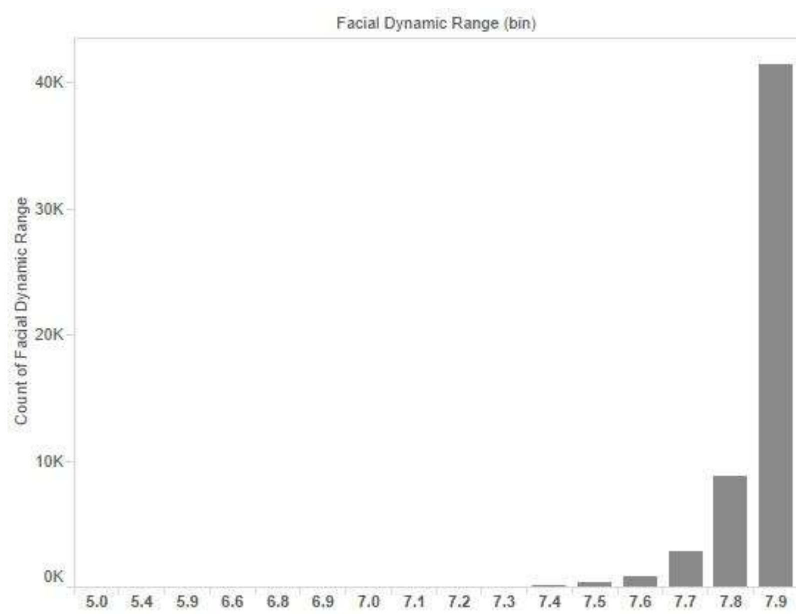

Figure B.1.14. Facial Dynamic Range

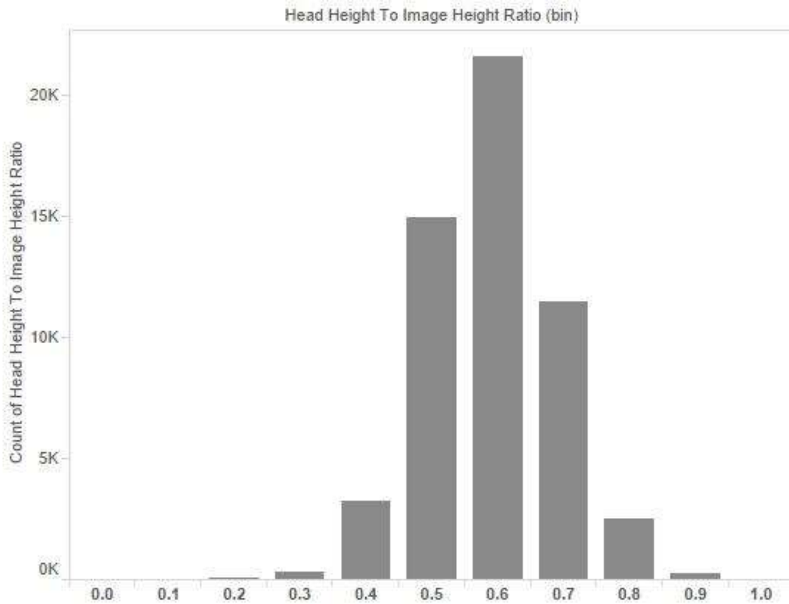

Figure B.1.15. Head Height to Image Height Ratio

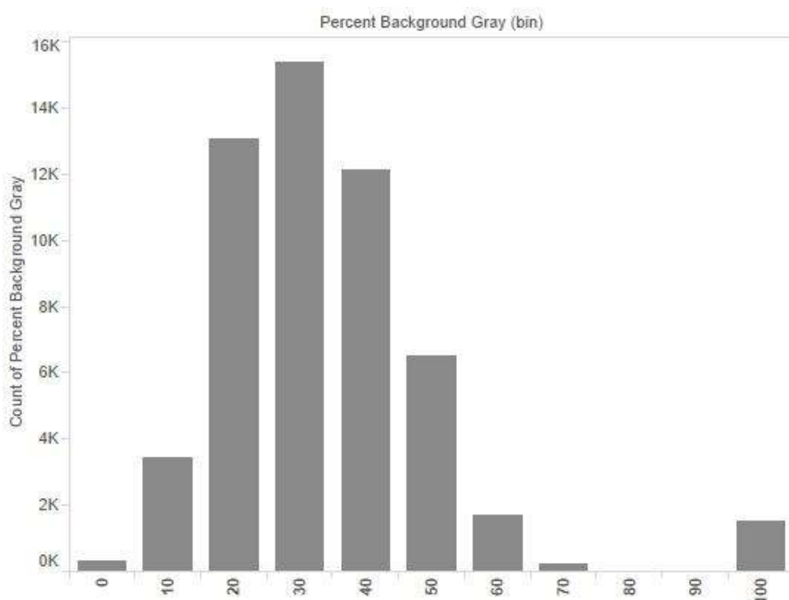

Figure B.1.16. Percent Background Gray

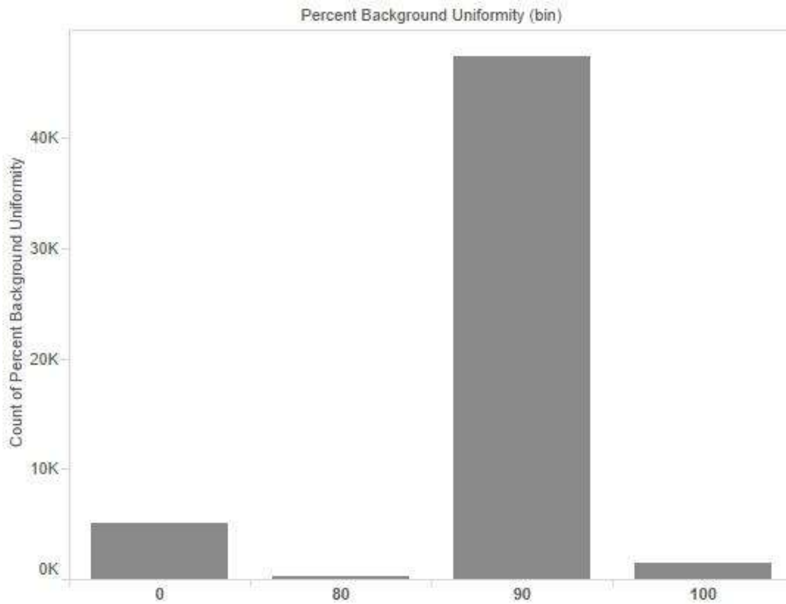

Figure B.1.17. Percent Background Uniformity

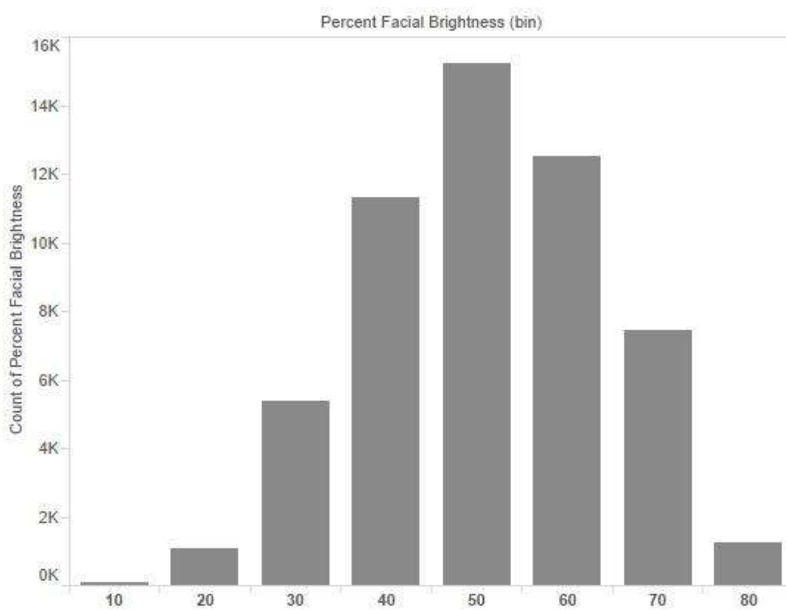

Figure B.1.18. Percent Facial Brightness

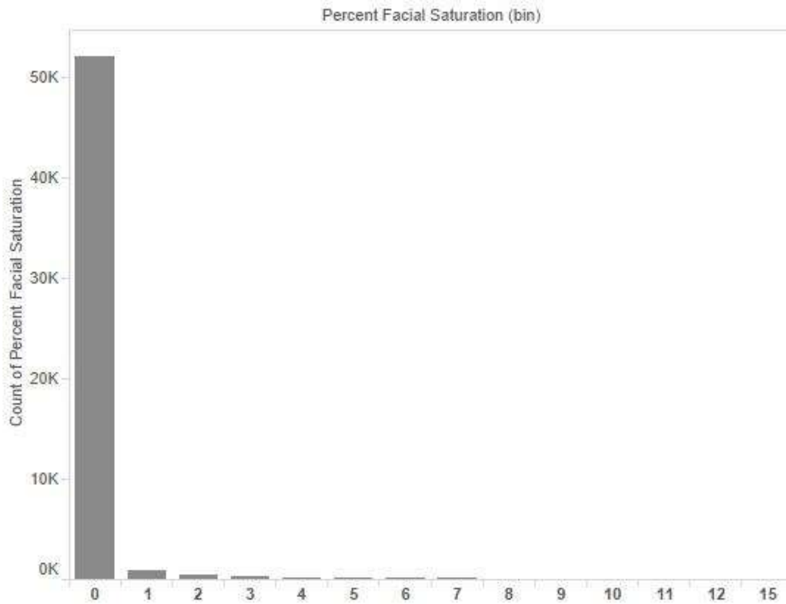

Figure B.1.19. Percent Facial Saturation

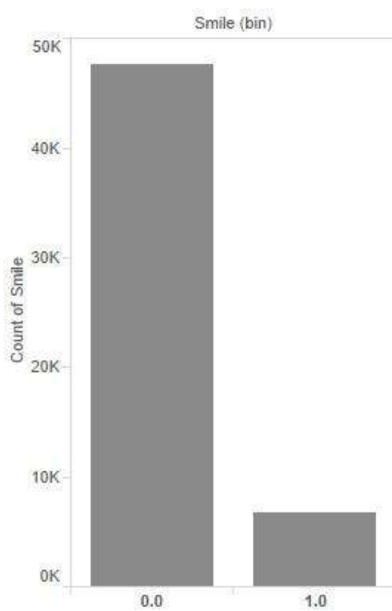

Figure B.1.20. Smile

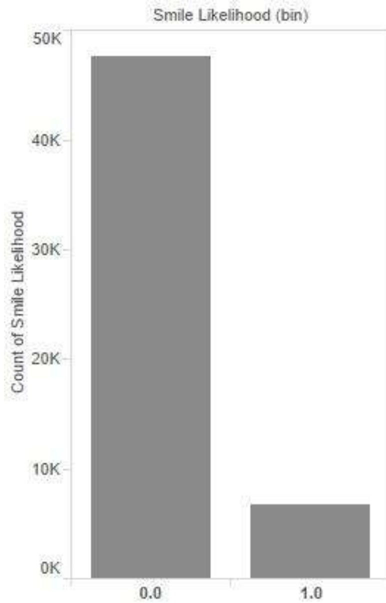

Figure B.1.21. Smile Likelihood