

# **Gender Identification from 3D Facial Surface Model**

*Hao Wang*

Master of Science  
Informatics  
School of Informatics  
University of Edinburgh  
2019

# Abstract

This dissertation introduces a method for 3D facial gender identification using conformal mapping. Conformal mapping is valuable in applications as it preserves the orientation and angles in local neighbor areas. There has been research on implementing conformal mapping to surface classification. However, there has been no research exploring the influence of conformal mapping on 3D identification problems on demographic attributes of human beings in the computer vision domain, which inspires the core framework of this project.

The main contribution of this dissertation is introducing conformal mapping to 3D facial gender identification and analyzing the performance of this novel system. Given a 3D point cloud, we wish to map the 3D model to 2D space conformally, and then extract and select features for classification. This dissertation explores the performance produced by two binary classifiers, an RBF-SVM classifier, and a k-NN classifier, on data with and without geometric registration. The results of the evaluation are competitive comparing to previous methods, and also indicate that the 3D approach with conformal mapping in this project is robust to the poses of 3D models. The evaluation of performance in this dissertation also reflects a fluctuation due to the limitation of the dataset used in this project, which inspires further evaluation using more massive datasets in future work. The performance also supports the value to research in 3D computer vision problems with geometric information extracted from conformal mapping.

## Acknowledgements

I have imagined the finale of this year for many times but still cannot calm down. I was lucky to meet such a lot of kind friends, tutors, and professors. All of you made who I am and what I have achieved now, and decorated my beautiful life in Edinburgh.

First and foremost, I express my gratitude to my supervisor, **Prof. Bob Fisher**. **Bob**: I would not have accomplished this project without your help. Your professionalism not only equipped my knowledge for this project but also promoted my enthusiasm for Computer Vision. Thank you for being a friend to me. I will never forget your kindness and humor, which made me feel relaxed when we had meetings. And I will never forget the pizza party held in your house as well, which was indeed a relief during the stressful time. Most of all, thank you for your consistent humility and your unshakeable patience. I have been a layman to 3D problems, but I have learned so, so much.

Then, I would like to thank all the professors and tutors I met. I appreciate your devotion to guide me to explore unknown knowledge for me. Moreover, I admire your knowledge and your enthusiasm in a specific domain, which has reinforced my conviction that nothing would be done correctly without enthusiasm. In particular, I appreciate Gary P. T. Choi and Dr. Carmelo Mineo, who provided generous support for conformal mapping and boundary points detection in my dissertation, respectively.

I would like to thank all my friends who I met here in Edinburgh. They were the people who I spent most of my time with, and they were the people who gave me the most direct support generously in the past year. **Shengwen**: thank you for being a friend, a brother to me. It's not easy to have a friend to discuss something technical even in a casual conversation. **Zihan**: thank you for asking me to help you with coding and cooking. My skills got improved a lot, as well. **Wenjun**: thank you for being my company for such a long time. I enjoyed the time we studied, cooked, and played together. **Xiaochu**: thank you for always cheering me up throughout the toughest time. Maybe I would have achieved another 80% if you had stayed in the library when you finished all your exams but I still had one. **Wei**: thank you for your consistent kindness. Although you could not provide me any academic support, your kindness and smile would always make my day. **Sunny and Taro**: I will never forget the time we spent together in the restaurants, bars, classrooms, and kitchen, indoors and outdoors. I appreciate your knowledge and your generous help on my study. I hope that I can be like you being erudite in my beloved academic areas. **Kiyoon and Boniface**: I never

thought that I could achieve a mark over 80% here before. Thank you for all your effort in our MLP project. It was an enjoyable experience to work with both of you. I have learned a lot from your hard work and dedication, from which I will benefit in all my life. Many thanks to **Zili, Sijin, Xiaojun, Shiqin, Xiaoyang, Zhaoliang, Ruochen, Hui, Lin, Runze, Yadi, Shihui, Guangliang, Callum, and Jack**. I know that you could provide little help to my dissertation technically, but I am happy and honored to have you as my friends. I always felt relief from stressful work after chatting with you.

The paper is too short to express my appreciation to all my friends. I hope that you will all have a happy life and all goes as smoothly as possible. I will never forget you no matter where we are, and hope will see you again.

Last but not least, I thank my parents for their unconditional and unshakeable love and support. Thank you, Dad, for being a mentor to me. Although we have chosen different subjects to study, I definitely need your experience. Thank you, mom, for your meticulous care. I forgot to make some phone calls when I was busy. I will make my schedule more organized to squeeze some time to make a call periodically. I promise. All in all, I love you, both of you.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>5</b>
2.1	Literature Review . . . . .	5
2.1.1	2D-based system . . . . .	6
2.1.2	3D-based system . . . . .	7
2.1.3	Hybrid (3D + 2D) system . . . . .	9
2.2	Dataset . . . . .	10
2.3	Packages . . . . .	10
<b>3</b>	<b>Methodology</b>	<b>12</b>
3.1	Data Preprocessing . . . . .	13
3.2	Conformal Mapping . . . . .	15
3.3	Feature Extraction . . . . .	19
3.4	Feature Selection . . . . .	23
3.5	Classification . . . . .	24
<b>4</b>	<b>Experimental Results</b>	<b>25</b>
4.1	Conformal mapping . . . . .	25
4.2	Classification . . . . .	27
<b>5</b>	<b>Conclusions</b>	<b>33</b>
5.1	Summary of achievements . . . . .	33
5.2	Future work . . . . .	34
	<b>Bibliography</b>	<b>35</b>

# **Chapter 1**

## **Introduction**

Identification of demographic attributes of human beings, such as gender, ethnicity, and age, based on computer vision techniques, has captured increasing attention from both academia and industry in recent years. Gender is one of the fundamental and principal demographic characteristics that should be classified initially, which can not only enhance the subsequent evaluation of gender-dependent biological identity [1], for example, age, but also promote more sophisticated identification, for example, face recognition. It also acts an critical role in technologies related to human interaction in various potential scenarios in real life [2], including Human Computer Interaction (HCI), video surveillance, multimedia retrieval, anonymous customized advertisement and biometrics [3, 4, 5, 6].

While human beings can recognize gender effectively and naturally [4, 7], gender is still non-trivial to be identified by the computer automatically. Although there are rich gender cues embedded in our appearance, voice, behavior and other biometrics, such as iris, fingerprint, and palm print, it is more naturally accessible for a human being to identify gender from facial appearance [8]. Therefore, there has been extensive research on gender identification from the facial data in the computer vision domain in the past decades [2].

The origin of gender identification with a facial model can be traced back to the early 1990s [9], where an automated technique was proposed to recognize gender from two-dimensional (2D) face images. From then on, extensive research has focused on 2D-based facial gender identification approaches. Moreover, this is still one of the most popular subjects in the computer vision domain today due to not only the numerous applications that are facilitated by gender identification but also the broader availability of 2D face images and the lower expense of 2D sensors. Researchers have

proposed several solutions previously based on raw images, texture features, and geometric features [3]. Also, there are deep learning-based techniques been researched recently [10, 11].

Although significant achievement has been made on 2D facial gender identification, both pose and illumination are still two obstacles that could severely restrict the performance [12, 13]. For example, face poses may be not aligned to the proper frontal position due to personal reasons, or if the camera is not adjusted correctly, and the light is dim, then both of which could deteriorate the identification performance. Three-dimensional (3D) face models could ameliorate the classification performance with the embedded geometric information under the conditions where 2D techniques are restricted. There is no projection from 3D space to 2D space, which improves the robustness of 3D facial gender identification to the variation of poses (face or camera) and illumination. Moreover, the booming of 3D sensor technology reduces the cost of 3D data acquirement, which further enables the accessibility and development of the research on 3D facial gender identification [14]. There is also research on gender identification with the combination of 2D and 3D information since the fusion of 2D, and 3D features could generally improve robustness and accuracy [15], and it may reflect how human beings classify gender more naturally [16].

We present the general framework for facial gender identification in Figure 1.1. Gender identification with 3D models is more complicated as it is challenging to find patterns in 3D data due to the lack of pixel information in 3D data, such as color or infrared intensity information. Therefore, 3D identification approaches contain additional processes to extract patterns embedded in 3D data in the feature extraction stages.

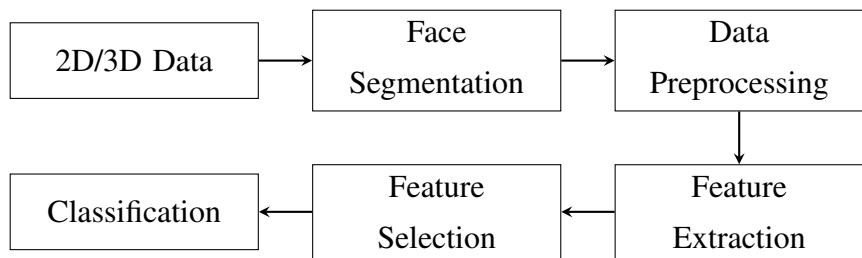


Figure 1.1: General framework for facial gender identification

We proposed a novel method for gender identification from 3D facial surface models in this project. In this algorithm, we introduced conformal mapping initially to map the 3D face data to 2D space conformally, and then classify with the combination of

features extracted from 3D models and 2D conformal-mapped data, including heights of facial feature points, distances between any two of the selected feature points, and ratios of any two of the distances. We present an example of a 2D intensity image, a 3D point cloud, and a 2D conformal map image of a sample in Figure 1.2.

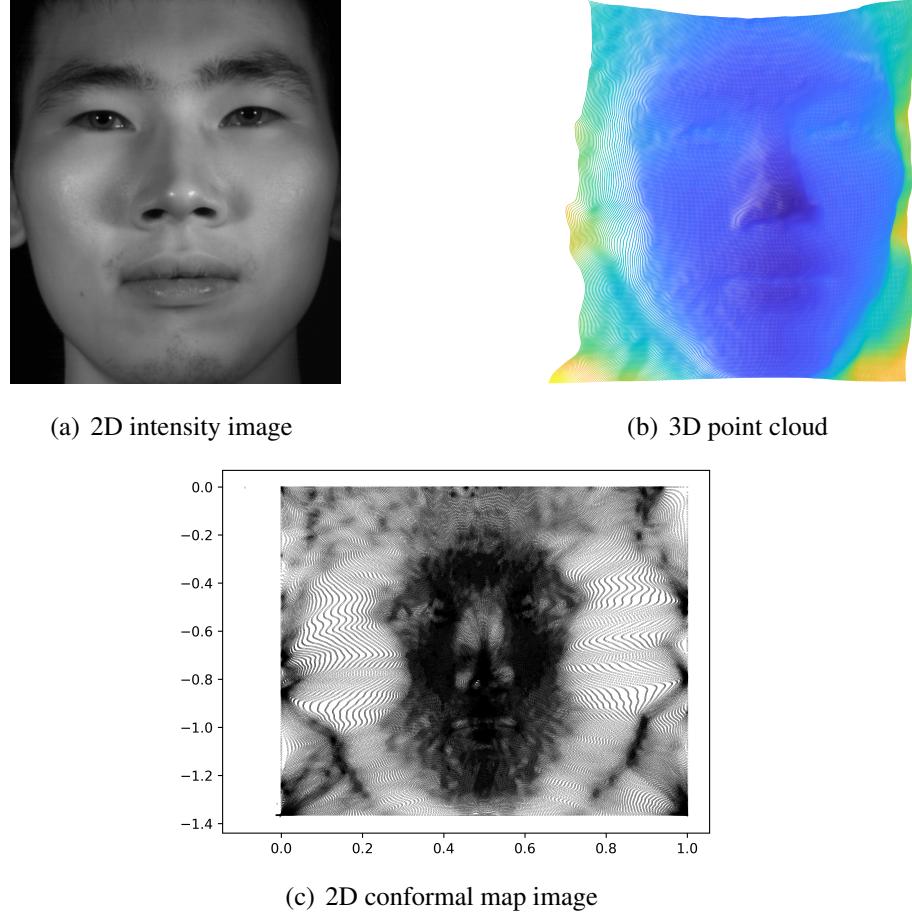


Figure 1.2: Example of 2D intensity image, 3D point cloud and 2D conformal map image of a sample used in this project.

Conformal mapping was initially introduced for surface classification, arguing that “Conformal mapping is refiner than topological equivalent class and coarser than isometric equivalent class, making it suitable for practical classification purposes” [17]. Our research provided an alternative solution to extract 2D features from the corresponding conformal map of a 3D head rather than the projected 2D image. We achieved an accuracy of 88.33% on 57 samples without registration and normalization in data preprocessing, which was competitive to the best approaches proposed previously, and indicated that identification with 3D models and conformal mapping could produce results robust to the variation of poses of 3D models. To the best of

our knowledge, this project is the first work to explore the influence of introducing conformal mapping to 3D facial gender identification problems.

The rest of this dissertation is organized as follows. Chapter 2 introduces the background of this project, including the review of relevant literature, the dataset, and tools employed in this project. Chapter 3 explains the detailed methodology of this project. Experiments and results, as well as discussion, are presented in Chapter 4. Chapter 5 concludes the project and discusses possible future work.

# **Chapter 2**

## **Background**

### **2.1 Literature Review**

Research on facial gender identification has been vastly developed since it emerged in recent decades. As one of the most fundamental demographic features from biometric information, gender identification plays an indispensable role in numerous applications, such as HCI, surveillance, information retrieval, and intelligent marketing. Although human beings can differentiate between male and female naturally and effectively [4, 7], it is non-trivial for a computer to discriminate without human interventions. Various methods for automated gender identification have been proposed, which can be classified into three categories in general, i.e., 2D-based, 3D-based, and hybrid methods.

2D facial gender identification has been studied most extensively due to the easy availability of 2D face data, such as color and infrared intensity images. The global and local features embedded in 2D data can facilitate the classification. However, 2D facial gender identification can be easily affected by the variation of face or camera poses or the illumination conditions. 3D facial gender identification could not only break the fundamental limitations of 2D methods [14] but also promote the identification accuracy in the circumstances that are challenging for the 2D system with the 3D geometric information. Also, 3D-based methods have attracted more research recently benefitting from the booming of 3D sensing technology. Besides, researchers have proposed hybrid technologies considering features extracted from both 3D and 2D images, which can generally be more robust and accurate than considering only either of the single modality [15]. We review these areas in more detail below.

### 2.1.1 2D-based system

There has been extensive research on 2D-based facial gender recognition since the first automated gender classification approach proposed in the early 1990s [9]. The widespread use of 2D sensors, such as digital cameras and infrared detectors, supports the availability of 2D color or intensity images, which facilitates the research on 2D facial gender identification. There are global and local features in 2D images that can be useful for gender identification, including color or intensity information from a color or intensity image, local texture or 2D shape variations of gender-discriminative regions. There are three categories of 2D-based gender identification approaches in general corresponding to the above global and local features, i.e., raw image-based, texture feature-based, and geometry feature-based approaches [3].

**Raw image-based approaches** process the entire raw frontal face images as the original data generally. There are also various dimensionality reduction techniques, such as down-sampling and subspaces transformation, implemented on original data typically before feeding in the classifier to reduce the computational complexity and memory requirement. This category of approaches could be traced back to the early 1990s when back-propagation SexNet [9] was proposed. Two neural networks were introduced for image compression and gender identification, respectively. Researchers also implemented normalization and dimensionality reduction to obtain input data with the size of  $30 \times 30$  pixels and the dimension of 40 from a set of 90 images (45 male and 45 female participants). Their experiment achieved an error rate of 8.1% on average, which displayed an improvement compared to that of 11.6% from psychophysical studies where five humans were asked to identify. There was other research exploring gender classification with raw images using different classifiers. For example, a support vector machine (SVM) was introduced to achieve an error rate of 3.38% on a dataset with 1755 samples (1044 males and 711 females) [18].

**Texture feature-based approaches** extract local features from texture variation of specific gender-discriminative regions. This kind of technique displays strength compared to raw image-based approaches with its low computational complexity. Researchers investigated classification with different texture features extracted from 2D images; for example, Haar-like features [19], Gabor features [20] and achieved accuracy of 91.6% on a database with 1240 samples. There was also research on promoting the performance of classification by improving the classifier. A Min-Max Modular SVM ( $M^3$ -SVM) was proposed to classify Gabor features, and the result was promoted

by around 6% higher than SVM on a dataset of 12912 frontal images with various degree view face, expressions, and other variance [21]. Also, researchers proposed to achieve demographic attribute classification with local binary patterns (LBP) and obtained promising results [22].

**Geometry feature-based approaches** consider 2D shape changes, such as angles, distances, and areas of the specific facial region, of different genders based on the 2D location of a set of facial landmarks, such as the nose tip, the inner and outer corners of eyes and mouth, and the chin. Some approaches claim that face shape can also contribute to gender identification significantly. Researchers have achieved a correct classification rate of 79% with sixteen selected geometry features [23]. Also, research on gender identification claimed that around 85% of 406 geometry features contain significant gender-discriminative information [24].

Besides, researchers have also introduced techniques based on deep learning, for example, deep convolutional neural networks (CNN), to 2D facial gender identification tasks recently to learn representations for human attributes, including skin color, hairstyle, expression, and others [11, 25, 26] for classification.

### 2.1.2 3D-based system

There has been enormous research on 2D facial gender identification, mainly based on texture features. However, 3D geometrical features also embed gender information according to anatomical studies [3] and other pioneering research [27, 28]. Also, the 3D structure of a head can promote the effectiveness of gender identification compared to methods based on intensity information [29]. Due to the booming of 3D sensing technology, it is more convenient to obtain data containing 3D shape information, which has boosted the research on 3D facial gender identification techniques recently. The application of 3D face models can overcome the internal weakness of 2D counterparts as the real facial geometrical information can be captured by the 3D sensor without projection to 2D space so that the classification would be enhanced to be more robust to the variation of face or camera pose or illumination [2]. Moreover, some 3D-based gender identification approaches delivered comparable or higher performance [3, 30].

The research on 3D facial gender identification can be traced back to the late 1990s when researchers proposed introducing Principal Component Analysis (PCA) to extract features from 130 face models represented by 3D coordinates [29]. They reduced the dimensionality of data to a 17-dimensional subspace and reported a correction re-

sult with a peak rate of 96.9%, which improved the performance of 2D methods on corresponding grayscale images (93.8%). Moreover, they argued that the fusion of features of both modalities could promote the performance further to 97.7%. Also, researchers proposed a geometry feature-based algorithm, in which several face regions, such as the forehead, eyebrow, cheek, and nose, were manually selected [31]. Features for gender identification consisted of the volume of selected regions compared to the whole face and ratio of surface area. They implemented a non-linear SVM classifier on the GavabDB dataset and achieved an error rate of 17.44% on average. Another method for gender classification with 3D human faces considered only facial feature points, such as face contour, eyes, eyebrows, nose, and mouth [32]. There were 23 facial feature points collected in this research for calculation on 3D point coordinates, Euclidean distances between any two points, the ratio of any two distances and angles between any two line segments any two points as the endpoints. These four types of statistics comprised the essential feature space of a 3D human face. A peak gender classification rate of 94% was achieved with the Matcher Weighting fusion method.

Researchers have also studied the realization of 3D facial gender identification with other various feature extraction techniques. An attempt was proposed to consider an individual's face as a deformed version of a "standard" human face, and classify gender considering the difference between both of these two models [33]. Instead of explicitly extracting facial features, a face recognizer was proposed considering the non-negative distance between two subjects to classify gender. Haar wavelets were used together with the steerable pyramid transform for the decomposition of an individual's facial structure to facilitate the comparison. This wavelet approaches achieved an accuracy of around 94% for gender identification on a public dataset, Face Recognition Grand Challenge version 2.0 (FRGCv2) dataset [34]. Researchers also proposed to extract radial and iso-level curves centered at the nose tip of a 3D face model [35]. Then these selected curves were analyzed by an existing shape analysis framework to compute similarities between a query face and templates of different genders. This research achieved the best classification rate of 84.98% with Adaboost on a subset of FRGCv2. Besides, a Correspondence Vector (CV) was introduced to measure the distance between a candidate face and a reference face [36], which achieved the best classification rate of 89.2% classified by an SVM with a linear kernel. Also, a fusion-based method was proposed to combine the similarity scores of five self-defined facial regions, i.e., the internal face, the upper region of the face, the lower region of the face, the left eye and the nose, as the final classification result [37]. They implemented an SVM

classifier to achieve an accuracy of 94.3% on 945 face models of a mixed database.

These previous researches supported the claim that the additional features contained in 3D face models could promote the performance of gender identification effectively. Nevertheless, although 3D facial gender identification has displayed distinguished achievement, there are still unsettled issues, for example, the precise positioning of salient gender discriminating features on 3D facial surfaces, that deserve research to enhance the performance further [7, 38, 39].

### 2.1.3 Hybrid (3D + 2D) system

Previous research elaborated that human beings might classify gender considering both 2D and 3D features, and their interrelationships [16]. This argument resulted in a significant trend for identification problem with face models to consider both 2D texture and 3D shape features jointly. Another factor promoting the development of hybrid approaches is that both 3D face models and corresponding 2D images are produced together using most of the current 3D imaging systems. Researchers claimed that compared to single modality-involved approaches, the fusion of 3D and 2D clues could generally promote a more robust and accurate classification system [15].

Pioneering research attempted to consider multiple modalities where the integration of similarity measurements of 3D shape and 2D texture improved the accuracy of gender identification with an SVM classifier [40]. Although this research was intuitive, there were several downsides, such as using pixel information of both facial intensity and depth images directly and considering the entire face region equally. Another research introduced a fusion method combining 3D shape and 2D texture information [30]. In this work, researchers extracted Dense Scalar Field (DSF) features from 3D mesh and LBP feature from range and gray images, and achieved a correct rate of 93.27% on FRGCv2 classified using a Random Forest. There was also an attempt introducing the shape index as another feature representing the curvature of the 3D model and then processed together with LBP [41]. Researchers later proposed a different approach based on local circular patterns (LCP) aiming at representing the variance between different genders or ethnicities comprehensively [3]. Introducing LCP improved the discriminative performance and robustness to noise by implementing a clustering-based quantization instead of a binary one compared to local binary patterns (LBP) and its variants [3].

Although 3D features can overcome some limitations that depend only on 2D fea-

tures, there are certain disadvantages to implementing 3D facial gender detection. For example, 3D approaches require expensive scanners, and introduce high computational complexity and memory requirement. Researchers proposed an approach with facial needle-maps, a 2.5D facial surface normal-based representation, extracted from 2D images, which contained 3D facial geometric information from a fixed viewpoint [42]. A weighted Principal Geodesic Analysis (PGA) was proposed to extract features, which were useful for gender discrimination, from 2.5D facial needle-maps, and achieved an accuracy of 97% on the 200 facial models in the Max-Planck face database [43].

## 2.2 Dataset

We used Speech-driven 3D Facial Motion Dataset (S3DFM) to evaluate the algorithm proposed in this project.

The Speech-driven 3D Facial Motion Dataset (S3DFM) was produced for research on speech-driven 3D facial dynamics originally, which describe dynamic 3D human faces while speaking [44]. The original data was captured by a binocular stereo vision video system from DI4D Ltd where two intensity cameras were employed to form a high-frame-rate 3D video sensor. There are two parts contained in this dataset overall. One is Speaking with Frontal Pose (S3DFM-FP) with 770 samples from 77 participants, and the other is Speaking with Varying Pose (S3DFM-VP) with 260 samples from 26 participants. There are ten video sequences of 500 frames for each participant repeating the same sentence. A 3D depth sequence and a corresponding pixel-wise aligned 2D intensity sequence were recorded in each video sequence with resolutions of  $600 \times 600$  points for each. All participants have been labeled considering their ages, genders, ethnicities, and mother-tongues.

We adopted S3DFM-FP (27 female and 50 male participants) in this project and selected the fifth frame of the first video sequence to form the dataset for the project processes. There are four examples displayed in Figure 2.1. Images in the first row are the 3D depth images, and those in the second row are the corresponding aligned 2D intensity images.

## 2.3 Packages

This project was deployed using mainly Python3 and MATLAB R2019a, and we employed several relevant packages for different procedures. In general, we implemented

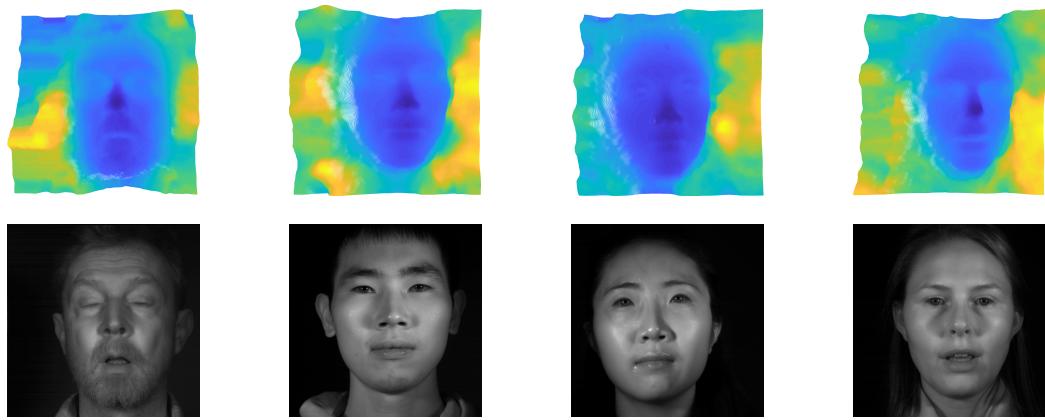


Figure 2.1: Example images from 4 participants (2 males and 2 females) in S3DFM-FP. The first row: 3D depth images; the second row: aligned 2D intensity images.

data preprocessing and conformal mapping with MATLAB R2019a, and achieved classification and its relevant processing with Python3.

We used the Lidar and Point Cloud Processing from Computer Vision Toolbox in MATLAB R2019a initially to preprocess the original point cloud data. For example, we adopted `pcregistericp` for point cloud registration by the Iterative Closest Point (ICP) algorithm. We adopted previous research on the realization of the conformal mapping on disk-type point cloud [45], which required not only the 3D point cloud data but also the indices of the boundary points of the data. Therefore, boundary in MATLAB was adapted to detect the boundary points of a point cloud. Conformal mapping and the corresponding evaluation of this operation could be produced automatically by adapting this software.

We employed scikit-learn [46] for feature selection, model selection, and classification in general. Scikit-learn is an open-source Python module that has been employed extensively in machine learning problems. There are many state-of-the-art algorithms in scikit-learn covering a wide range of machine learning techniques and its relevant processes, which are convenient to build a classification framework based on machine learning. For example, we used Recursive feature elimination with cross-validation (RFECV) from `sklearn.feature_selection.cross_validate` from `sklearn.model_selection` and Support Vector Machine (SVM), `sklearn.svm`.

Besides, we used `shape_predictor_68_face_landmarks.dat` in Dlib for facial landmarks detection. Dlib is a modern C++ toolkit that has been widely used in both industry and academia to solve practical problems in various domains including robotics, embedded or mobile systems, and high performance computing [47].

# Chapter 3

## Methodology

In general, the framework of the 3D facial gender identification system employed in this project can be represented as the combination of several procedures, as shown in Figure 3.1.

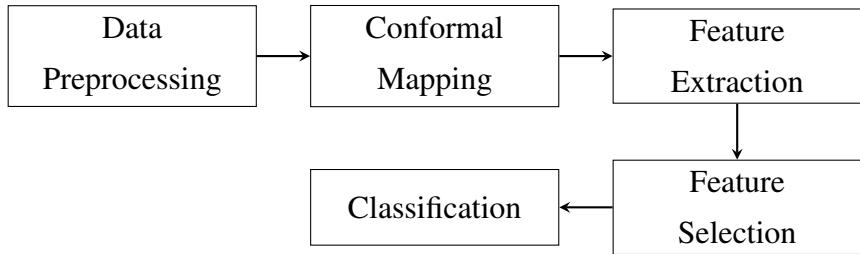


Figure 3.1: The workflow of the principal model in this project.

Some particular steps are employed commonly in most of the other identification problems in the computer vision domain, including object detection, preprocessing, feature extraction, feature selection, and classification [14, 48]. In this project, the 3D face model was the object of interest, which was isolated from a noisy external background initially with the help of landmark detection. Then, the 3D model was mapped to 2D space conformally, and a height map was constructed to represent the features of the original 3D data. Also, the 2D coordinates of facial landmarks on the 2D mapped image were extracted to produce other features, such as Euclidean distances between points and ratio of distances. Feature selection was employed before classification to select the most discriminative group of features. Both computation and memory expense can be reduced to prevent the problem of the curse of dimensionality. In the end, a binary classifier was adopted as this project focuses on whether a 3D facial model represents a male or female. A non-linear support vector machine (SVM) with

the radial basis function (RBF) kernel was employed initially considering its popularity in previous research, and we also evaluated the system with a k-nearest neighbors algorithm (k-NN).

### 3.1 Data Preprocessing

The original 3D facial model cannot be adopted as the inputs of the following procedures directly since there are not only human faces, but also many outer regions with distracting features and other inevitable noisy background captured during data acquiring procedure, as shown in Figure 3.2(a), Figure 3.2(b), Figure 3.2(e), and Figure 3.2(f). These regions, such as neck area, ear, and hair, can lead to deterioration of performance of the following data processing procedures and classification. Although these features could be helpful for identification in the human scenario, the variation of these external features could mislead 3D facial gender identification algorithm, and therefore, these regions should be removed initially. We present two examples of original data in the S3DFM dataset, and the corresponding cropped models in Figure 3.2. The first two columns contain the original 2D and 3D data, respectively, and the last two columns contain the corresponding cropped facial model.

The initial step of data preprocessing in this project was facial region detection. The original 3D facial model was preprocessed to isolate the facial region out from the outside area with the help from the explicitly identifiable facial part, for example, nose. The facial region was segmented by an auxiliary bounding box determined by the position of the nose tip, chin and boundary points of the left and right cheek, i.e., the highlighted blue and red dots in Figure 3.3(c), respectively. Moreover, with the segmentation of the facial region, the problem of gender identification is decoupled from face detection, and the influence of the location of the face in the original data can be reduced. For example, the location of the face in the original data is different as shown in the pair of example, Figure 3.2(a) and Figure 3.2(e), Figure 3.2(b) and Figure 3.2(f). The cropped face was adjusted to be located at the center of the image after segmentation, as shown in the pair of example, Figure 3.2(c) and Figure 3.2(g), Figure 3.2(d) and Figure 3.2(h), and the possible influence of face location was eliminated.

Considering that corresponding aligned 2D intensity images were also contained in the S3DFM dataset together with 3D data, we introduced an approach for 2D facial feature point detection [49] to locate these fiducial points selected for locating the auxiliary bounding box mentioned above. `get_frontal_face_detector()` from Dlib was

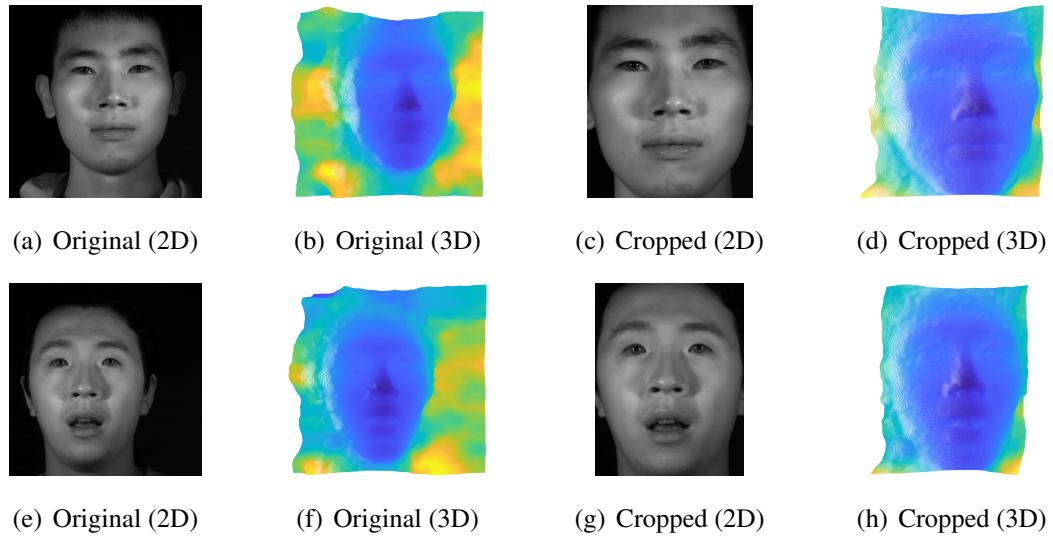


Figure 3.2: Example of original data and cropped data. We present original 2D grayscale images and the corresponding 3D models in the first two columns, respectively, and the corresponding cropped model in the last two column. We also present two examples in the first and second row, respectively, to display that there is variance of face location in the original data, and this variance could be eliminated after face segmentation.

employed as facial landmark detector, and shape predictor `68_face_landmar.dat` from Dlib was implemented as predictor to locate 68 facial feature points. The detector was designed by the combination of a linear classifier, an image pyramid, and a sliding window detection scheme, considering the classic Histogram of Oriented Gradients (HOG) feature [49]. The predictor was provided within Dlib and was pre-trained on the iBUG 300-W face landmark dataset [50]. It was convenient to detect facial landmarks automatically with the cooperation of the detector and predictor on 2D color or intensity images. The result of detection was recorded as a list of coordinates of these facial feature points. The approach used in this project can detect 68 facial landmarks, including the contour of eyes, eyebrows, nose, mouth, and chin. After detection on 2D images, we mapped these landmarks to the corresponding 3D model with the help from the correspondence between 2D and 3D data, and then segmented the 3D facial regions. There is an example of the result of 2D facial feature points detection, as shown in Figure 3.3(a), and a mapped result of its corresponding 3D data, as shown in Figure 3.3(b).

The auxiliary bounding box in this project was defined as a rectangle whose left,

right, and the lower boundary was determined by the boundary point of the left cheek, the right cheek, and chin, respectively. We highlight these three points in the example with red dots, as shown in Figure 3.3(c). Considering that the facial landmarks detected in this project did not reflect the upper boundary of a face model correctly, we defined the upper boundary of the bounding box aided by the location of the nose tip, which is highlighted with a blue dot in Figure 3.3(c). We assumed that the distance between the upper boundary of the box and nose tip was equal to that between the lower boundary of the box and nose tip for most samples. The upper boundary of the auxiliary bounding box would be modified to the boundary of the image to prevent overflowing if the face in the original image was located by the boundary of the original image. We highlight the auxiliary bounding box in the example with a yellow rectangle, as shown in Figure 3.3(c).

Once the cropped facial data is obtained, there were other preprocessing procedures followed before feeding it as input to the feature extraction and classification stages. The geometric alignment was implemented to the cropped model to restrict the possible error caused by the variation of face pose, which was achieved by adapting MATLAB function `pcregistericp`. Downsampling could be necessary to reduce the requirement of computing resources and storage capacity. Besides, rescaling or normalizing the pixel values, for example, to zero mean and unit variance, was also considered to be applied to the facial data. These procedures were designed to minimize the sensitivity of the system to variations and other distracting features and to make the feature extractor and classifier more effective. There are also corresponding functions in MATLAB that are convenient to be used for the operations mentioned above.

## 3.2 Conformal Mapping

We introduced conformal mapping to this project, with which a 3D facial model can be mapped to 2D space conformally. To the best of our knowledge, this project is the first work that introduces conformal mapping to 3D facial gender identification. We display an example of the mapping result in Figure 3.4.

With the development of 3D sensing technology, it is increasingly more convenient to obtain 3D models for computer vision tasks, which has promoted the development of the 3D approach for solving computer vision problems. According to anatomical studies, 3D geometrical information is essential for gender identification [3]. More-

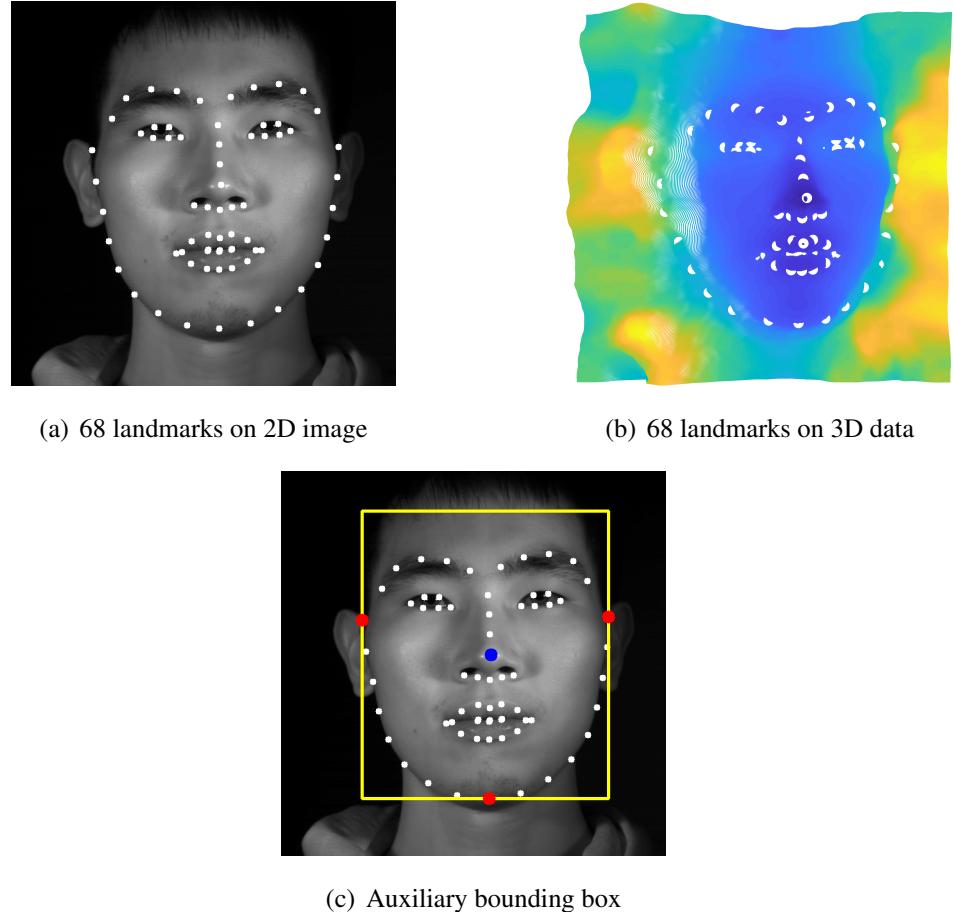


Figure 3.3: Example of 68 facial feature points detection and auxiliary bounding box. 68 facial feature points are highlighted by white dots. The boundary points of left cheek, right cheek and chin selected for locating boundary of the auxiliary bounding box are highlighted by red dots in (c). The nose tip is highlighted by a blue dot in (c). The auxiliary bounding box is highlighted by a yellow rectangle in (c).

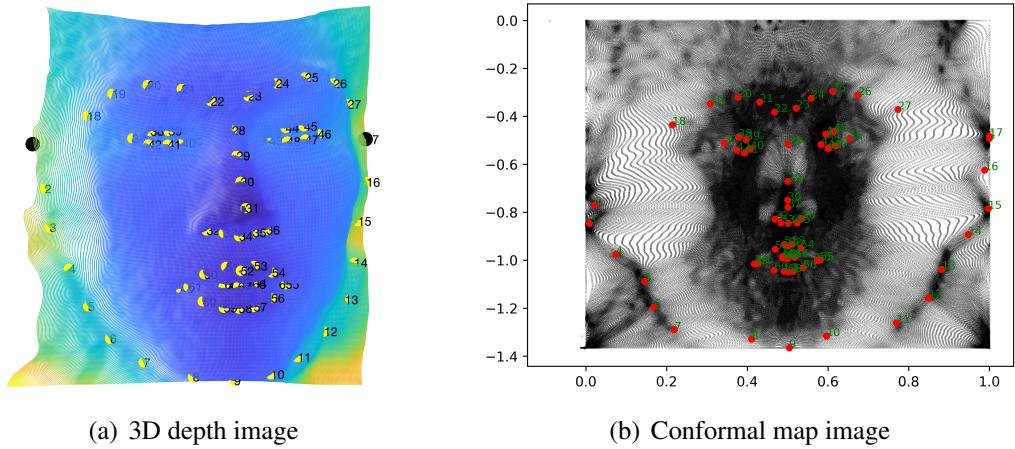


Figure 3.4: Example of conformal mapping result. The 68 facial landmarks are highlighted by yellow and red dots in (a) and (b), respectively. The point 1 ( $P_1$ ) and point 17 ( $P_{17}$ ) are highlighted by the larger black dots. These two points were used to define the centroid of a head ( $P_c$ ).

over, previous research has claimed the strength of introducing 3D models that the classification could be more productive with a 3D model than with image intensity information [29]. Also, the real facial shape can be captured in 3D facial models without a 2D projection, which enhances the robustness of 3D classification to the variation of illumination conditions and poses of face or camera [2]. However, processing 3D data requires high computational complexity and memory cost typically, which inspired the research on hybrid (3D + 2D) systems.

Conformal mapping has been used for surface classification for years and has been proved that “conformal equivalent class is refiner than topological equivalent class and coarser than isometric equivalent class” [17], which inspired the principal framework of this project. Due to the preservation of orientation and angles in local neighbor areas, conformal mapping has been introduced to various complex analysis problems that exhibit inconvenient geometries. There are also practical applications where conformal mapping plays a significant role, such as the Mercator projection and the stereographic projection in cartography, which are similar to this project as the core is to map a 3D model to 2D space.

There are programs published previously for the realization of conformal mapping on 3D data in the form of point cloud and triangular meshes, covering disk-type, genus-zero and non-zero-genus data [51, 52, 45, 53], which are convenient to be adapted by other projects.

The S3DFM dataset contains both 3D point cloud models and corresponding aligned 2D intensity images. Each of the 3D point cloud models consists of  $600 \times 600$  points represented by 3D Cartesian coordinates  $(x, y, z)$ . Thus, we adapted previous work on conformal mapping for disk-type point clouds [45] to achieve this procedure automatically. The software package for this algorithm is provided on its author's personal website, which is a MATLAB program requiring both 3D Cartesian coordinates of the points in the point cloud and the indices of the boundary points of the point cloud in anti-clockwise orientation as the inputs. To locate these boundary points, we adapted boundary in MATLAB considering the map of points on  $x$ - $y$  plane. This MATLAB function returned the indices of boundary points of a point cloud anti-clockwise. We display an example of the boundary points, which are highlighted by black “o”s, in Figure 3.5 with front view, top view and left view, respectively.

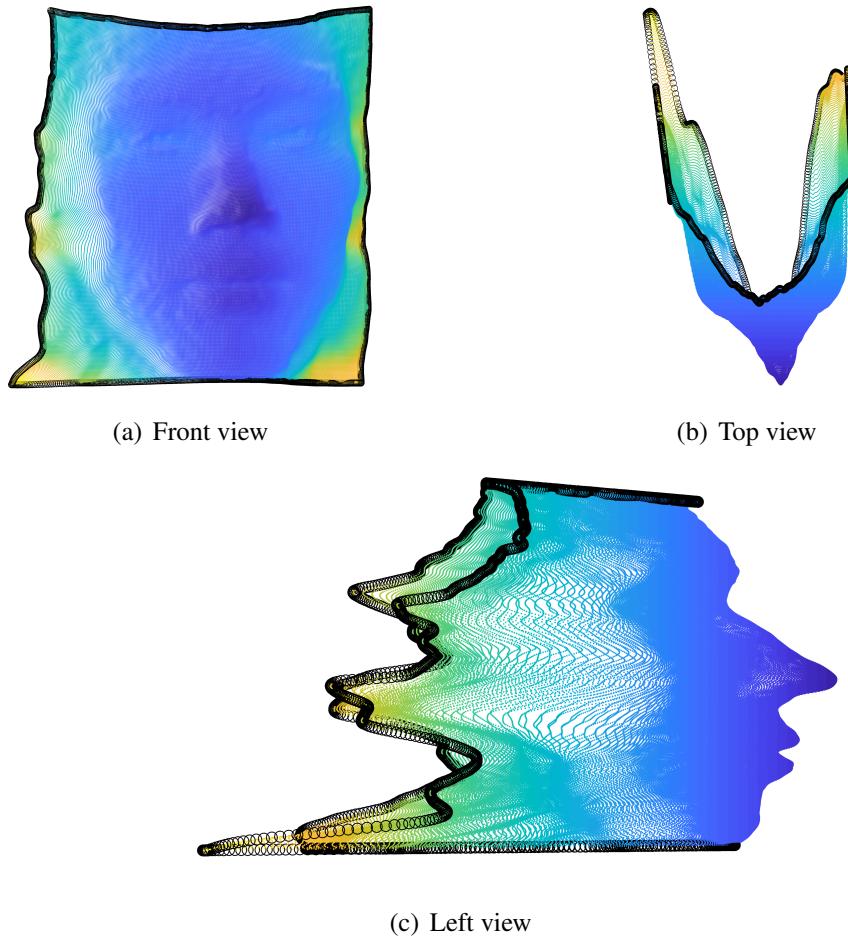


Figure 3.5: Example of boundary points of a point cloud. The boundary points are highlighted by black “o”s.

We implemented conformal mapping on 3D disk-type point cloud with or with-

out registration to a template 3D model, and with or without normalization on three axes respectively in advance. However, due to the restriction of Delaunay triangulation, there was various quantity of samples that could be processed by this conformal mapping software for each setting. We list the number of samples that were mapped successfully in Table 3.1.

REGISTRATION	NORMALIZATION	NUM_S
YES	YES	62
No	No	57

Table 3.1: The quantity of samples that were mapped successfully (NUM\_S) under each setting.

The free software processed the 3D point cloud data and calculated the corresponding conformal map in MATLAB automatically, and returned the map result consisting of 2D coordinates of points in the conformal map, which was convenient for use in the following procedures.

### 3.3 Feature Extraction

Identification of 3D models is more complicated than that of 2D images as there is no texture information, for example, the color information of a pixel, but only spatial information, for example, the coordinates of points in a point cloud, contained in the original 3D data. 2D height maps were generated initially to represent the feature of the 3D facial models in this project. For point cloud data, each point in a 3D model was allocated with a height, which was the Euclidean distance between the corresponding point on the 3D facial surface model and the centroid of the head. Since the datasets used in this project contain only a frontal face, we assumed that the centroid of a head ( $P_c$ ) is located at the midpoint of the line connecting the ears. Considering the 68 facial feature points detected in advance, we defined  $P_c$  as the midpoint of the line connecting point 1 ( $P_1$ ) and point 17 ( $P_{17}$ ), as shown in Figure 3.4(a). Assuming the coordinates of  $P_c$ ,  $P_1$  and  $P_{17}$  are  $(x_c, y_c, z_c)$ ,  $(x_1, y_1, z_1)$  and  $(x_{17}, y_{17}, z_{17})$ , respectively, then  $x_c = \frac{1}{2} \times (x_1 + x_{17})$ ,  $y_c = \frac{1}{2} \times (y_1 + y_{17})$  and  $z_c = \frac{1}{2} \times (z_1 + z_{17})$ . We present an example of all height vectors in an example of 3D models in Figure 3.6. The length of each height vector was the corresponding height to be calculated. The 68 height

features are the first 68 features used for gender classification.

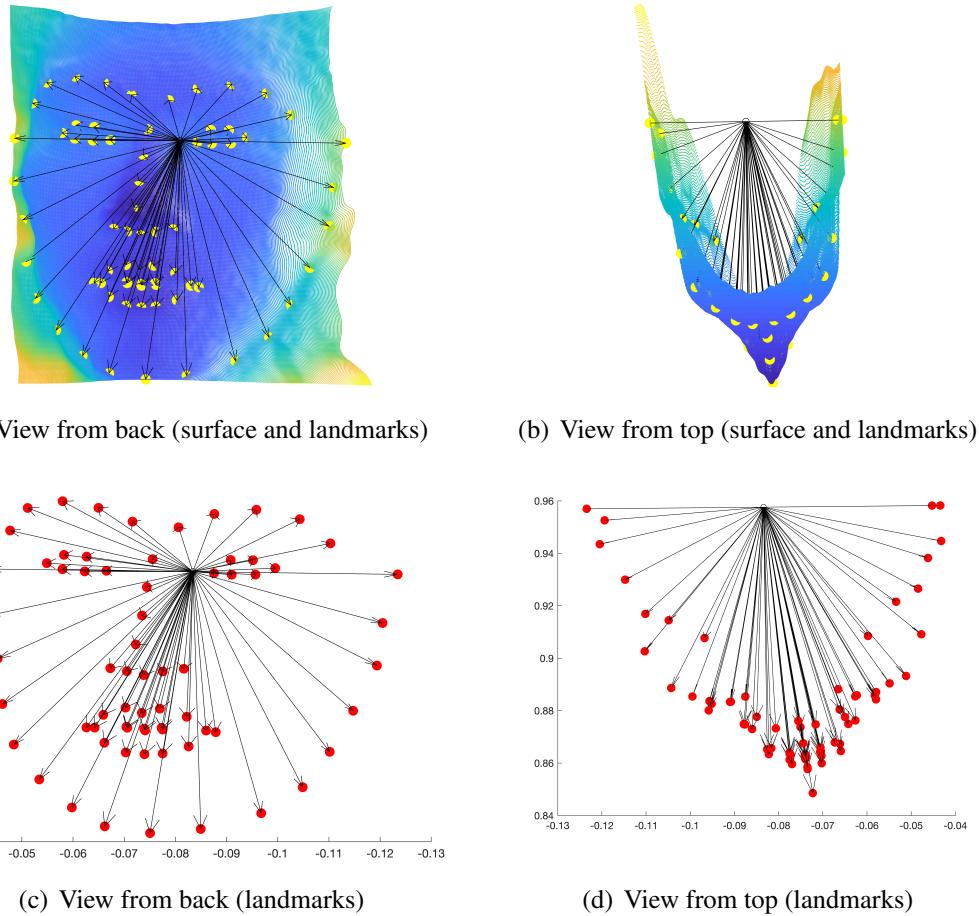


Figure 3.6: Example of height vectors. The top row contains 3D models with both facial surfaces and 68 facial landmarks, where these landmarks are highlighted by yellow dots. The bottom row contains 3D models with only 68 facial landmarks to make it more explicit, where these landmarks are highlighted by red dots. The centroid of a head ( $P_c$ ) is highlighted by a black “o”, which is the origin of all height vectors (black).

We also considered extracting features from the 2D conformal mapped data to feed in classifier together with height features, including Euclidean distance between selected facial feature points and the ratio of any two distances. A representative subset consisting of 28 points was selected initially from the 68 facial feature points on the 2D map images, as shown in Figure 3.7, including the corner of eyebrows, eyes, mouth, cheeks, nose tip, and chin. The indices of these 28 selected points were 2, 4, 7, 9, 11, 14, 16, 18, 20, 22, 23, 25, 27, 28, 31, 32, 34, 36, 37, 40, 43, 46, 49, 52, 55, 58, 63, and 67 with respect to the 68 facial landmarks (from 1 to 68). Then, we defined the point-pairs for the Euclidean distance calculation as listed in Table 3.2. We calculated

Euclidean distances concerning these point-pairs and then calculated the ratio between any two of these distances. We present the distance to be calculated in this project in Figure 3.7 represented by line segments in black.

GROUP	POINT-PAIRS FOR EULIDEAN DISTANCE CALCULATION
A	(1,7), (2,6), (3,5), (8,9), (8,10), (9,10), (10,11), (11,12), (11,13), (12,13), (14,20), (14,21), (16,17), (16,18), (17,18), (19,20), (20,21), (21,22), (23,24), (23,25), (23,26), (23,27), (23,28), (24,25), (25,26), (25,27), (25,28)
B	(1,2), (2,3), (5,6), (6,7), (9,19), (9,20), (12,21), (12,22), (14,16), (14,17), (14,18), (24,27), (26,28)
C	(4,1), (4,2), (4,3), (4,5), (4,6), (4,7), (4,8), (4,9), (4,10), (4,11), (4,12), (4,13), (4,14), (4,15), (4,16), (4,17), (4,18), (4,19), (4,20), (4,21), (4,22), (4,23), (4,24), (4,25), (4,26), (4,27), (4,28)
D	(15,1), (15,2), (15,3), (15,5), (15,6), (15,7), (15,8), (15,9), (15,10), (15,11), (15,12), (15,13), (15,14), (15,16), (15,17), (15,18), (15,19), (15,20), (15,21), (15,22), (15,23), (15,24), (15,25), (15,26), (15,27), (15,28)

Table 3.2: Four groups of point-pairs for Euclidean distance calculation. Group A: horizontal distances; Group B: vertical distances; Group C: distances including the chin; Group D: distances including the nose tip.

Considering the computational complexity, memory cost, and the volume of datasets, we implemented dimensionality reduction before feeding the features to the classifier. There are plenty of dimensionality reduction techniques adopted widely to accomplish feature extraction, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA). Besides, it was argued that the performance of classification could be enhanced with a fusion of features from various methods in some cases [48]. We adapted `sklearn.decomposition.PCA` to discover the appropriate number of components initially, and then explored how PCA would influence the performance of classification.

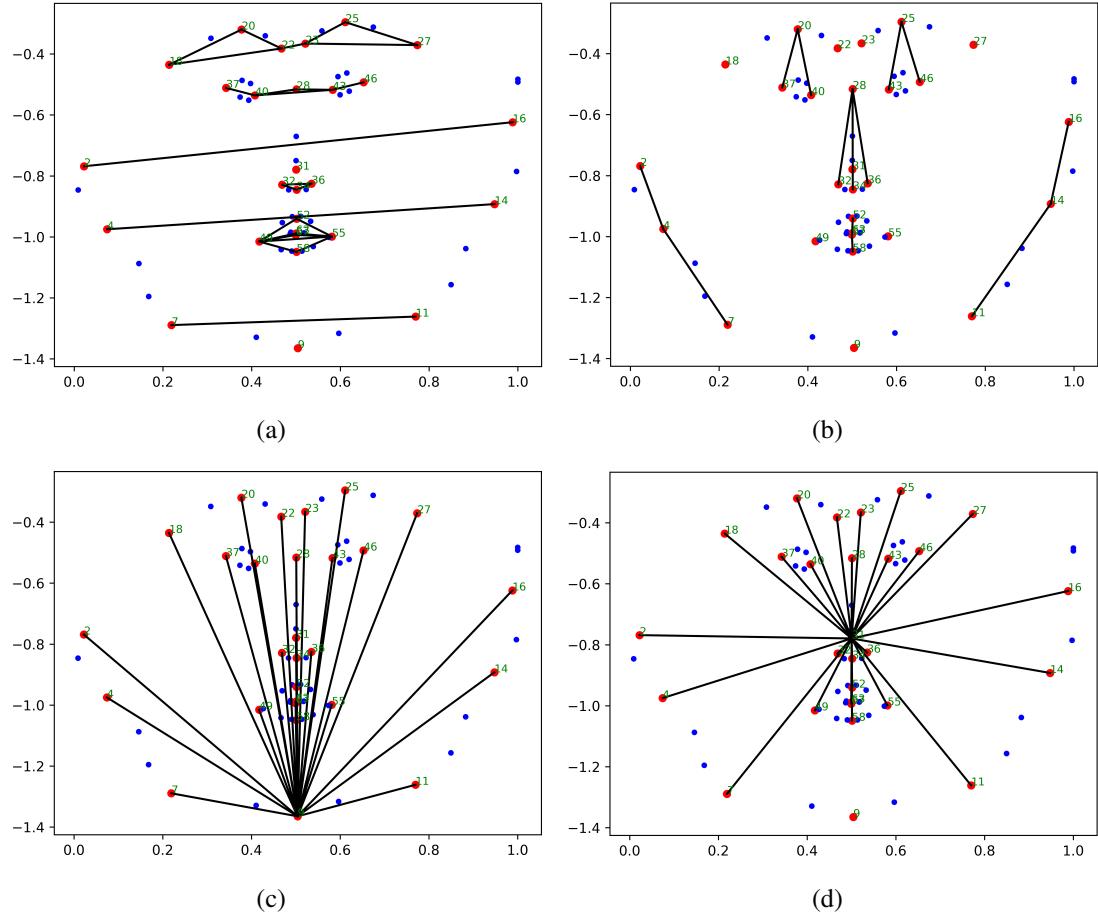


Figure 3.7: Example of Euclidean distances to be calculated with selected facial landmarks. The selected 28 facial landmarks are highlighted by the larger red dots with indices, while other landmarks are smaller blue dots without indices. The length of the black line segments are the Euclidean distances to be calculated as the distance features of 2D conformal mapped data.

### 3.4 Feature Selection

Selection of the most discriminative group of features is implemented in machine learning projects commonly to release the computational intensity and increase the robustness of the system when there are a large number of features in some situations. Representative descriptors of the data will be identified and selected as the inputs of classifiers.

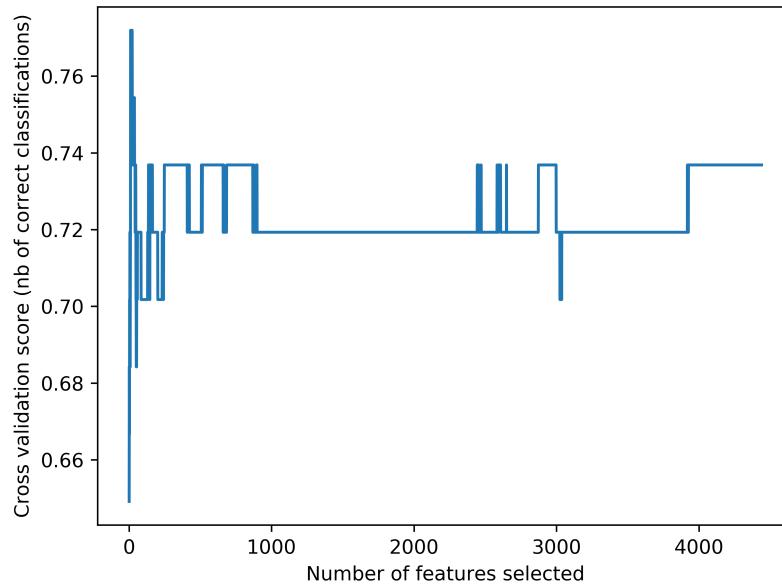


Figure 3.8: Example of evaluation of classification performance versus various number of features selected by RFECV. The  $x$ -axis indicates the number of features, and the  $y$ -axis indicates the grid score of cross validation.

In this project, there were inevitably a different number of feature points contained in the cropped 3D face data due to the various sizes of human heads. To facilitate feature selection and classification, we selected some facial regions manually in advance, including eyes, nose, mouth, and chin, to constitute the feature vector with equal length for each face model. These selected regions were represented by the 68 facial landmarks detected in the previous procedure. The primary feature vector consisted of height features, distance features, and ratio features obtained in previous stages. There were 4439 features in total, including 68 height features, 93 distance features, and 4278 ratio features, which increased both computational complexity and risk of the curse of dimensionality. We employed *recursive feature elimination with cross validation* (RFECV) to select further the optimal subset of features. It was convenient to achieve this procedure by adapting `sklearn.feature_selection.RFECV`. This func-

tion processed the feature matrix with respect to the label vector and searched for the optimal subset greedily concerning the classification performance. This function could tune the number of features selected with cross-validation, and return the indicator on whether the particular feature was selected automatically, with which we extracted the subset of features from the primary feature matrix. We employed an SVM as the estimator in RFECV and adopted stratified 3-fold cross-validation to select the subset. There was one feature to remove in each iteration. We display an example of evaluation of the classification performance versus the various number of features selected during the operation of RFECV in Figure 3.8.

## 3.5 Classification

Binary classifiers were introduced to this project as we only considered whether a 3D facial model represented a male or female. There have been plenty of classifiers employed in gender identification, including SVM, Adaboost, neural networks, k-NN, and the Bayesian classifier [48]. Besides, classification accuracy may be promoted with the fusion of various classifiers [54]. Based on our review of previous research, SVM is the predominant classifier employed in gender classification, followed by boosting approaches such as Adaboost [48]. Moreover, usually a non-linear kernel is adopted in the classifiers to achieve higher accuracy, for example, non-linear SVM with the radial basis function (RBF) kernel is applied widely in previous work [48]. We started with a non-linear SVM with the RBF kernel, and also compared the performance with that produced by a k-NN classifier.

Model selection is equally critical in solving machine learning problems. We adopted k-fold cross-validation to evaluate performance produced under different hyperparameter settings. Moreover, to obtain an unbiased evaluation of the final classification model, we adapted stratified separation for the train and test sets considering the ratio of gender in the different datasets. All these related operations could be implemented conveniently by adapting corresponding functions in `scikit-learn`.

# Chapter 4

## Experimental Results

### 4.1 Conformal mapping

Conformal mapping was one of the core procedures in this project whose performance would influence the outcome of classification since height features, distance features and ratio features were extracted based on the 2D conformal-mapped images. We adapted free software [45] to achieve conformal mapping in this project. This software processed 3D point cloud data and generated the corresponding 2D mapped data, and produced the evaluation of the mapping operation automatically. To be more specific, the mapping operation in this project was quasi-conformal mapping as the frontal facial surface was recorded as disk-type point cloud data. The Beltrami coefficient was introduced as a measure of the non-conformality, whose norm indicated the degree of conformality distortion.

We present some examples of the results of conformal mapping in Figure 4.1. The first column contains the original 3D point clouds with facial landmarks highlighted by yellow dots and their indices, and the second column contains the corresponding 2D mapped images with facial landmarks highlighted by red dots. The performance of conformal mapping was evaluated by the degree of conformality distortion. The evaluation on particular samples was displayed in the last column, where the  $x$ -axis represents the norm of the Beltrami coefficient, and the  $y$ -axis represents the number of vertices belonging to the specific norm.

The Beltrami coefficient of an ideal quasi-conformal mapping has a constant norm. It can be observed that the histogram in the middle row highly concentrates at one value, which indicates that parameterization on this sample achieved uniform conformality distortion, while the histogram in the top row is more loose, which reflects a

weaker but acceptable parameterization. The histogram in the bottom row presents a much flatter distribution, which indicates a much weaker conformal mapping that can be observed in Figure 4.1(h). The more detailed analysis of the performance of conformal mapping algorithms can be found in related research.

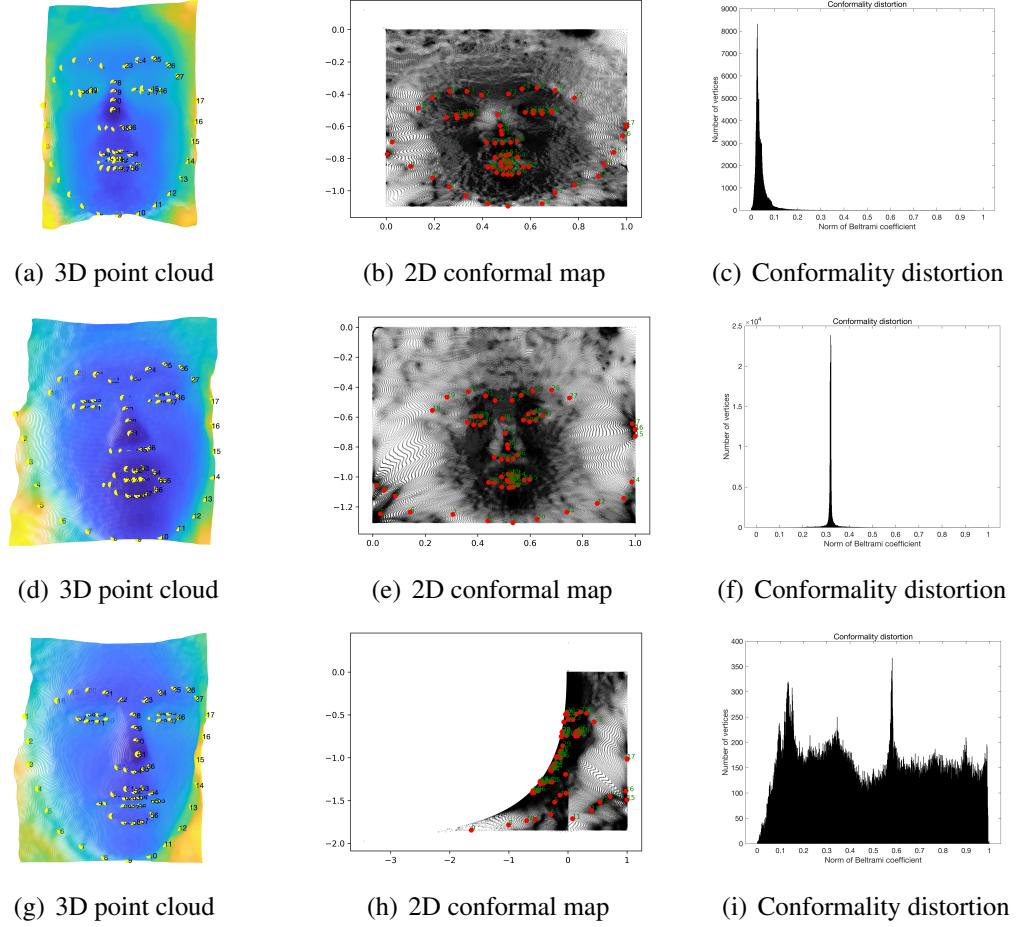


Figure 4.1: Example of conformal mapping results and the corresponding evaluation of the conformality distortion. Top row: moderate mapping result; middle row: good mapping result; bottom row: weak mapping result.

Due to the limitation of the conformal mapping algorithm and distribution of points in 3D point clouds, we obtained a 2D mapped model generated from a subset of original 3D point clouds, as we mentioned in Chapter 3. Moreover, some 3D point clouds were parameterized with a weak conformal mapping, for example, Figure 4.1(h), which reduced further the number of premium 2D mapped data for feature extraction and classification. A larger dataset can be adopted to produce more acceptable 2D mapped data in the future to reduce the impact on classification due to a small dataset.

## 4.2 Classification

An RBF-SVM classifier and a k-NN classifier were adopted in this project, both of which were trained based on k-fold cross-validation. We describe the detail of the experiments as follows.

---

**Algorithm 1:** The algorithm of one single experiment in this project.

---

**Result:** Average accuracy of 10 tests.

- 1 Generate feature matrix  $X$ , label vector  $y$ ;
  - 2 Shuffle the samples in  $X$  randomly;
  - 3 Implement stratified split on  $X$  to generate train sets ( $X_{train}, y_{train}$ ) and test sets ( $X_{test}, y_{test}$ ),  $n\_splits = 10$ ;
  - 4 Select feature considering  $X_{train}$  using Recursive Feature Elimination with 10-fold cross-validation;
  - 5 **for**  $i = 1$  **to** 10 **do**
  - 6     Train model using  $X_{train}(i), y_{train}(i)$  with 10-fold cross-validation;
  - 7     Test model using  $X_{test}(i), y_{test}(i)$ ;
  - 8 **end**
  - 9 Calculate the average performance of these ten tests.
- 

We implemented the stratified split on the full set of data to generate training and test set initially, keeping the ratio of the two labels in training and test set the same as that in full set of data. Due to the limited quantity of samples, we split the full set of data ten times and evaluated the system with ten prediction results on the test set. The data was shuffled randomly before splitting in advance, and separated by adapting `StratifiedShuffleSplit` from `scikit-learn`. We pre-defined the size of test set as 10% of full set of data. Then, we adapted recursive feature elimination with cross-validation (RFECV) from `scikit-learn` on the training set to select further the optimal subset of features, and extracted these features from both training and test set to create new training and test set with selected features, respectively. For RFECV, we used a linear-SVM as its estimator, and set the function with 10-fold cross-validation, considering one feature in one iteration, i.e.  $step = 1$ . We also standardized the selected training and test set, respectively, after feature selection by RFECV. Then, we implemented a non-linear SVM with the RBF kernel and a k-NN as classifiers in this project by adapting the corresponding functions from `scikit-learn`, respectively. The hyperparameter settings of these two classifiers were selected by adapting `GridSearchCV`.

from scikit-learn, which could evaluate various combinations among specified hyperparameters exhaustively for a specific estimator with cross-validation, and return the optimal hyperparameter settings automatically. For SVM classifier used in this project, we selected RBF kernel. The kernel coefficient was 0.01, and the penalty parameter of the error term was 10. The other parameters remained default settings of the function. For k-NN classifier, number of neighbors to consider was 3, and the other hyperparameters remained the default settings of the function. Each classifier was then trained with stratified 10-fold cross-validation and tested on separate test set. We present the framework of this one single experiment in Algorithm 1. There were ten re-shuffling and splitting iterations using `StratifiedShuffleSplit` by default. Thus, ten pairs of train and test set were created randomly, and we adapted our program to train and evaluate the model using these ten pairs automatically. The final evaluation of this one single experiment was the average accuracy of these ten tests.

	1	2	3	4	5
ACC (%)	83.33	100.00	100.00	66.67	83.33
NUM_F	7	7	15	3505	12
	6	7	8	9	10
ACC (%)	83.33	100.00	66.67	100.00	100.00
NUM_F	10	19	2204	3	11

Table 4.1: 10 accuracy results (ACC) and corresponding number of features (NUM\_F) selected by RFECV in one single experiment with RBF-SVM.

We listed ten results from one single experiment with an RBF-SVM classifier in Table 4.1, which reflects that there was explicit fluctuation among the results of each test. S3DFM is a relatively small dataset that could lead to the phenomenon that there were different trends among data embedded in training and test set after dataset separation. It could be also noticed that the quantity of features in the optimal subset selected by RFECV varies in different tests. As cross-validation was implemented in this stage, selection on features could also be affected due to the small dataset, which inspires future work using a more massive dataset. We also noticed from Table 4.1 that the classification was more accurate and stable with less than twenty selected features. It was argued in previous research that around 89% of the difference among 406 features could be explained by 18-20 features [24]. Although the results in Table 4.1 cannot in-

dicate the range of optimal number of features to be selected, it deserves more research in the future to coarsely selected a subset of features, automatically or manually, before implementing standard feature selection algorithms, for example, RFECV, which might make feature selection and classification more stable and accurate. Besides, it was possible that Test 4 and Test 8 triggered off the curse of dimensionality considering the small quantity of dataset and the massive quantity of features.

We also repeated the experiment described in Algorithm 1 for ten times, and calculated the average accuracy (ACC) and the corresponding standard deviation. We listed the best average performance, as well as the medium, minimum, and maximum accuracy among all iterations under different settings in Table 4.2.

It can be observed that classifying with an RBF-SVM classifier on data without registration and normalization in advance produced the best accuracy of 88.33%. Moreover, the outcome produced with an RBF-SVM classifier was better than that with a k-NN classifier generally, considering medium due to its robustness.

CLA	REG	NOR	NUM_S	ACC (%)	MED (%)	MIN (%)	MAX (%)
SVM	YES	YES	62	86.67 ± 15.30	91.65	66.67	100.00
	NO	No	57	88.33 ± 13.71	91.65	66.67	100.00
k-NN	YES	YES	62	81.67 ± 18.33	83.33	50.00	100.00
	NO	No	57	83.33 ± 17.56	83.33	50.00	100.00

Table 4.2: Best average performance (ACC) among ten iterations under different settings of registration (REG) and normalization (NOR) in data preprocessing and classifier (CLA). We also list the standard deviation, median (MED), minimum (MIN) and maximum (MAX) accuracy among the ten results.

Some previous research suggested implementing geometric alignment of the facial model in the data preprocessing procedures to ensure the appropriate location of specific facial landmarks. However, researchers also claimed that previous automatic and manual alignment algorithms gave limited benefit to the performance of gender identification [55] while significant processing time was required due to the computational complexity [19]. Our results present the same trend that the variance between performance of the systems with or without alignment in advance was limited. The outcome of the system without registration was slightly better than that with registration generally, which indicates the possibility that the classifier could be more robust.

to face misalignments [56]. Also, we can explain this limited variance with knowledge of cartography. We implemented conformal mapping to map the 3D point cloud data to 2D space. The mapped results would be robust generally regardless of the poses of 3D models as we stretched and flattened the original 3D data to 2D space. Hence, the distance and ratio features extracted in later stages would be robust. Nonetheless, it deserves more experiments with more massive datasets to explore this argument.

	1	2	3	4	5
HEIGHT (%)	0.00	0.00	0.00	1.34	0.00
DISTANCE (%)	0.00	0.00	0.00	2.17	0.00
RATIO (%)	100.00	100.00	100.00	96.49	100.00
	6	7	8	9	10
HEIGHT (%)	0.00	0.00	1.45	0.00	0.00
DISTANCE (%)	0.00	0.00	3.09	0.00	9.09
RATIO (%)	100.00	100.00	95.46	100.00	90.90

Table 4.3: The percentage of height features (Height), distance features (Distance), and ratio features (Ratio) selected in each of the ten tests. These ten tests corresponds to the ten tests listed in Table 4.1.

We calculated the percentage of each group of features considering the quantity of features selected in ten tests, respectively, corresponding to Table 4.1, to explore which part of features were more predominant for gender identification in this project. The percentage results are listed in Table 4.3 There were three groups of features extracted from both 3D data and 2D conformal-mapped data, including 68 height features (Height), 93 distance features (Distance), and 4278 ratio features (Ratio). The outcomes of feature selection reflect that ratio features were predominant in the selected features, which indicates that ratio features contain more gender discriminative information. The ratio features, as a relative measure, could be more robust for gender classification considering the variation among size of human heads.

Moreover, the identification system designed in this project achieved competitive performance comparing to the best approaches proposed previously, as compared in Table 4.4. The conformal mapping can promote the performance of gender identification remarkably as although a small dataset was used to train and test the system in this project, the outcome was competitive to those methods trained and tested with a much

larger dataset, which indicates the value of researching the identification problem with more geometry knowledge. Previous research claimed that a human can identify the gender with photos of adult faces with hair concealed by the accuracy of 96% [57], which indicates that there is still possible improvement for the algorithm proposed in this project to promote the performance.

Besides, there were still other aspects that could be considered to enhance the performance of this algorithm. The limited number of samples processed in this project could lead to the fluctuation of accuracy. Larger datasets could be introduced in the future, for example, the Face Recognition Grand Challenge dataset (FRGCv2), to explore the influence of dataset and the more robust performance. There are also texture features embedded in 2D conformal-mapped images, which could be extracted further to discuss their influence. In addition, it was argued that human facial shape would change while growing up [60]. We did not consider these age-related changes in this project, which could be explored in future work.

METHOD	DATASET (NUM_S)	FEATURES	CLASSIFIERS	FOLD	ACC (%)
[3]	FRGCv2 (3676)	LCP (2D+3D)	ADABOOST	10	95.50
[31]	GAVABDB (61)	GEOMETRY FEATURES (3D)	SVM (RBF)	5	82.56 ± 0.92
[33]	FRGCv2 (3675)	WAVELETS (3D)	SVM (POLYNOMIAL)	10	93.50
[37]	FRGCv1 + SELF DATA (945)	CUVATURE-BASED SHAPE INDEX (3D)	SVM (RBF)	5	94.30
[39]	FRGCv2 (466)	3D FEATURES (3D)	RANDOM FOREST	10	95.28
	FRGCv2 (4007)				93.61
[41]	HFB (192)	3D COORDINATES (2D+3D)	SVM (RBF)	5	94.80 ± 5.30
	FRGCv2 (4007)				93.70 ± 2.00
[58]	FRGCv2 (466)	FACIAL CURVES (3D)	ADABOOST	10	86.05
[59]	FRGCv2 (466)	LANDMARK DISTANCES (3D)	LDA	10	97.05
	FRGCv2 (4007)				96.12
THIS PROJECT	S3DFM-FP (57)	HEIGHT + LANDMARK DISTANCES (3D)	SVM (RBF)	10	88.33 ± 9.17
			k-NN		85.01 ± 10.54

Table 4.4: Comparison among approaches.

# **Chapter 5**

## **Conclusions**

### **5.1 Summary of achievements**

We explored 3D facial gender identification with conformal mapping in this project and achieved a competitive performance on 57 samples selected from the S3DFM dataset comparing to the best approaches proposed before. We implemented conformal mapping to map the 3D data to 2D space initially, and extracted features from both 3D models and 2D conformal-mapped data, including height features from 3D models, distance and ratio features from 2D mapped data. The original feature matrix contained 4439 features, including 68 height features, 93 distance features, and 4278 ratio features. We adapted recursive feature elimination with cross-validation to select the optimal subset of features before feeding the feature matrix to the classifier. The results of the feature selection indicated that ratio features contained more gender discriminative information in the data used in this project. An RBF-SVM classifier and a k-NN ( $k = 3$ ) classifier were adapted in classification procedure and trained with 10-fold cross-validation in this project. The results of classification presented that the RBF-SVM classifier produced better results than k-NN. Moreover, the outcome of the recognition system in this project displayed the limited difference between the performance produced with data with and without registration, which indicated that it was robust to variation of poses of face to identify gender with 2D data mapped conformally from a 3D facial model. However, due to the limited amount of data used in this project, there was explicit fluctuation among the performance of each test, which inspired the possible experiments on more massive datasets in future work to generate more robust results.

## 5.2 Future work

The original purpose of this project was to discover how conformal mapping would influence gender identification on 3D skull models, which was changed since access to the skull dataset was not granted. We could design a new project on 3D skull models referring to this project in the future. Similar to the dataset used in this project, skull datasets typically contain 3D skull models represented by point clouds or triangular meshes, which can be adopted to generate corresponding conformal maps by adapting related free software mentioned in Chapter 3. There may be a few differences among operations in data preprocessing, for example, generating the required inputs for specific software. The other procedures in this project can be adapted to new datasets to extract similar features. Then the optimal subset of features can be selected and classified following a framework similar to that in this project.

The result of this project was competitive to some existing 3D approaches for gender identification on 3D facial models. However, there is explicit fluctuation among the performance of classification, which could be affected by the small number of data in S3DFM. The classification system proposed in this project could be evaluated using other more massive datasets, for example, the Face Recognition Grand Challenge dataset (FRGCv2), to generate more robust outcomes. Besides, there were facial features neglected in this project, which was argued to be useful for gender identification, for example, the vertical width of eyes and mouth. Due to samples of S3DFM were recorded while participants repeating a sentence, there are inevitable nictation and motivation of mouth captured in various frames. We selected the same frame from each video sequence, which resulted in that eyes were closed, or mouth was opened in some samples. We could select those frames with eyes opened, and mouth closed manually in future work so that more probable gender-discriminative features could be selected. Moreover, the limited difference between results on data with and without registration inspired that the introduction of conformal mapping could enhance the robustness of 3D facial gender identification to variation of poses of 3D models to some extent, which indicates the value on researching computer vision problems in 3D domain with exploration of embedded geometry information.

# Bibliography

- [1] Hao Wang. Gender identification from 3d skull models, 2019.
- [2] Baiqiang Xia. Which facial expressions can reveal your gender? a study with 3d faces. *arXiv preprint arXiv:1805.00371*, 2018.
- [3] Di Huang, Huaxiong Ding, Chen Wang, Yunhong Wang, Guangpeng Zhang, and Liming Chen. Local circular patterns for multi-modal facial gender and ethnicity classification. *Image and Vision Computing*, 32(12):1181–1193, 2014.
- [4] Hu Han, Charles Otto, Xiaoming Liu, and Anil K Jain. Demographic estimation from face images: Human vs. machine performance. *IEEE transactions on pattern analysis and machine intelligence*, 37(6):1148–1161, 2014.
- [5] Antitza Dantcheva, Carmelo Velardo, Angela D’angelo, and Jean-Luc Dugelay. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, 51(2):739–777, 2011.
- [6] Yunlian Sun, Man Zhang, Zhenan Sun, and Tieniu Tan. Demographic analysis from biometric data: Achievements, challenges, and new frontiers. *IEEE transactions on pattern analysis and machine intelligence*, 40(2):332–351, 2017.
- [7] Syed Zulqarnain Gilani, Kathleen Rooney, Faisal Shafait, Mark Walters, and Ajmal Mian. Geometric facial gender scoring: Objectivity of perception. *PloS one*, 9(6):e99483, 2014.
- [8] Rama Chellappa, Charles L Wilson, Saad Sirohey, et al. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705–740, 1995.
- [9] Beatrice A Golomb, David T Lawrence, and Terrence J Sejnowski. Sexnet: A neural network identifies sex from human faces. In *NIPS*, volume 1, page 2, 1990.

- [10] Kaipeng Zhang, Lianzhi Tan, Zhifeng Li, and Yu Qiao. Gender and smile classification using deep convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 34–38, 2016.
- [11] Rajeev Ranjan, Vishal M Patel, and Rama Chellappa. Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):121–135, 2017.
- [12] Chenghua Xu, Yunhong Wang, Tieniu Tan, and Long Quan. Depth vs. intensity: Which is more important for face recognition? In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 1, pages 342–345. IEEE, 2004.
- [13] Kevin W Bowyer, Kyong Chang, and Patrick Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. *Computer vision and image understanding*, 101(1):1–15, 2006.
- [14] Song Zhou and Sheng Xiao. 3d face recognition: a survey. *Human-centric Computing and Information Sciences*, 8(1):35, Nov 2018.
- [15] D. Huang, W. Ben Soltana, M. Ardabilian, Y. Wang, and L. Chen. Textured 3d face recognition using biological vision-based facial representation and optimized weighted sum fusion. In *CVPR 2011 WORKSHOPS*, pages 1–8, June 2011.
- [16] Vicki Bruce, A Mike Burton, Elias Hanna, Pat Healey, Oli Mason, Anne Coombes, Rick Fright, and Alf Linney. Sex discrimination: how do we tell the difference between male and female faces? *perception*, 22(2):131–152, 1993.
- [17] Xianfeng Gu and Shing-Tung Yau. Surface classification using conformal structures, 2003.
- [18] Baback Moghaddam and Ming-Hsuan Yang. Gender classification with support vector machines. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pages 306–311. IEEE, 2000.

- [19] Gregory Shakhnarovich, Paul A Viola, and Baback Moghaddam. A unified learning framework for real time face detection and classification. In *Proceedings of Fifth IEEE international conference on automatic face gesture recognition*, pages 16–23. IEEE, 2002.
- [20] Shihong Lao and Masato Kawade. Vision-based face understanding technologies and their applications. In *Chinese Conference on Biometric Recognition*, pages 339–348. Springer, 2004.
- [21] Hui-Cheng Lian, Bao-Liang Lu, Erina Takikawa, and Satoshi Hosoi. Gender recognition using a min-max modular support vector machine. In *International Conference on Natural Computation*, pages 438–441. Springer, 2005.
- [22] Zhiguang Yang and Haizhou Ai. Demographic classification with local binary patterns. In *International Conference on Biometrics*, pages 464–473. Springer, 2007.
- [23] Brunelli Poggio, R Brunelli, and T Poggio. Hyperbf networks for gender classification. 1992.
- [24] Ashok Samal, Vanitha Subramani, and David Marx. Analysis of sexual dimorphism in human face. *Journal of Visual Communication and Image Representation*, 18(6):453–463, 2007.
- [25] Neeraj Kumar, Peter Belhumeur, and Shree Nayar. Facetracer: A search engine for large collections of images with faces. In *European conference on computer vision*, pages 340–353. Springer, 2008.
- [26] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of the IEEE international conference on computer vision*, pages 3730–3738, 2015.
- [27] Vicki Bruce, Tim Valentine, and Alan Baddeley. The basis of the 3/4 view advantage in face recognition. *Applied cognitive psychology*, 1(2):109–120, 1987.
- [28] Arnulf BA Graf and Felix A Wichmann. Gender classification of human faces. In *International Workshop on Biologically Motivated Computer Vision*, pages 491–500. Springer, 2002.

- [29] Alice J O’toole, Thomas Vetter, Nikolaus F Troje, and Heinrich H Bülfhoff. Sex classification is better with three-dimensional head structure than with image intensity information. *Perception*, 26(1):75–84, 1997.
- [30] Baiqiang Xia, Boulbaba Ben Amor, Di Huang, Mohamed Daoudi, Yunhong Wang, and Hassen Drira. Enhancing gender classification by combining 3d and 2d face modalities. In *21st European Signal Processing Conference (EUSIPCO 2013)*, pages 1–5. IEEE, 2013.
- [31] Xia Han, Hassan Ugail, and Ian Palmer. Gender classification based on 3d face geometry features using svm. In *2009 International Conference on CyberWorlds*, pages 114–118. IEEE, 2009.
- [32] Yuan Hu, Li Lu, Jingqi Yan, Zhi Liu, and Pengfei Shi. Sexual dimorphism analysis and gender classification in 3d human face. *IEICE transactions on information and systems*, 93(9):2643–2646, 2010.
- [33] George Toderici, Sean M O’malley, George Passalis, Theoharis Theoharis, and Ioannis A Kakadiaris. Ethnicity-and gender-based subject retrieval using 3-d face-recognition techniques. *International Journal of Computer Vision*, 89(2-3):382–391, 2010.
- [34] P Jonathon Phillips, Patrick J Flynn, Todd Scruggs, Kevin W Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. Overview of the face recognition grand challenge. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05)*, volume 1, pages 947–954. IEEE, 2005.
- [35] Lahoucine Ballihi, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Driss Aboutajdine. Geometric based 3d facial gender classification. In *2012 5th International Symposium on Communications, Control and Signal Processing*, pages 1–5. IEEE, 2012.
- [36] Ryan Tokola, Aravind Mikkilineni, and Christopher Boehnen. 3d face analysis for demographic biometrics. In *2015 International Conference on Biometrics (ICB)*, pages 201–207. IEEE, 2015.
- [37] Yuan Hu, Jingqi Yan, and Pengfei Shi. A fusion-based method for 3d facial gender classification. In *2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, volume 5, pages 369–372. IEEE, 2010.

- [38] Baiqiang Xia, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, and La-houcine Ballihi. Combining face averageness and symmetry for 3d-based gender classification. *Pattern Recognition*, 48(3):746–758, 2015.
- [39] Baiqiang Xia, Boulbaba Ben Amor, and Mohamed Daoudi. Joint gender, ethnicity and age estimation from 3d faces: an experimental illustration of their correlations. *Image and Vision Computing*, 64:90–102, 2017.
- [40] Xiaoguang Lu, Hong Chen, and Anil K Jain. Multimodal facial gender and ethnicity identification. In *International Conference on Biometrics*, pages 554–561. Springer, 2006.
- [41] Xiaolong Wang and Chandra Kambhamettu. Gender classification of depth images based on shape and texture analysis. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 1077–1080. IEEE, 2013.
- [42] Jing Wu, William AP Smith, and Edwin R Hancock. Gender discriminating models from facial surface normals. *Pattern Recognition*, 44(12):2871–2886, 2011.
- [43] Jing Wu, William AP Smith, Edwin R Hancock, and Michal Kawulok. Extracting gender discriminating features from facial needle-maps. In *2009 16th IEEE International Conference on Image Processing (ICIP)*, pages 2449–2452. IEEE, 2009.
- [44] Jie Zhang and Robert Fisher. 3d visual passcode: Speech-driven 3d facial dynamics for biometrics. *Signal Processing*, 2 2019.
- [45] Ting Wei Meng, Gary Pui-Tung Choi, and Lok Ming Lui. Tempo: feature-endowed teichmuller extremal mappings of point clouds. *SIAM Journal on Imaging Sciences*, 9(4):1922–1962, 2016.
- [46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [47] Chiara Galdi, Lara Younes, Christine Guillemot, and Jean-Luc Dugelay. A new framework for optimal facial landmark localization on light-field images. In *2018*

- IEEE Visual Communications and Image Processing (VCIP)*, pages 1–4. IEEE, 2018.
- [48] Choon-Boon Ng, Yong-Haur Tay, and Bok-Min Goi. A review of facial gender recognition. *Pattern Analysis and Applications*, 18(4):739–755, 2015.
  - [49] Vahid Kazemi and Josephine Sullivan. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1867–1874, 2014.
  - [50] Christos Sagonas, Epameinondas Antonakos, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 300 faces in-the-wild challenge: Database and results. *Image and vision computing*, 47:3–18, 2016.
  - [51] Pui Tung Choi and Lok Ming Lui. Fast disk conformal parameterization of simply-connected open surfaces. *Journal of Scientific Computing*, 65(3):1065–1090, 2015.
  - [52] Pui Tung Choi, Ka Chun Lam, and Lok Ming Lui. Flash: Fast landmark aligned spherical harmonic parameterization for genus-0 closed brain surfaces. *SIAM Journal on Imaging Sciences*, 8(1):67–94, 2015.
  - [53] Gary Pui-Tung Choi, Kin Tat Ho, and Lok Ming Lui. Spherical conformal parameterization of genus-0 point clouds for meshing. *SIAM Journal on Imaging Sciences*, 9(4):1582–1618, 2016.
  - [54] Erno Mäkinen and Roope Raisamo. An experimental comparison of gender classification methods. *pattern recognition letters*, 29(10):1544–1556, 2008.
  - [55] Erno Makinen and Roope Raisamo. Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):541–547, 2008.
  - [56] Michael Mayo and Edmond Zhang. Improving face gender classification by adding deliberately misaligned faces to the training data. In *2008 23rd International Conference Image and Vision Computing New Zealand*, pages 1–5. IEEE, 2008.
  - [57] A Mike Burton, Vicki Bruce, and Neal Dench. What’s the difference between men and women? evidence from facial measurement. *Perception*, 22(2):153–176, 1993.

- [58] Lahoucine Ballihi, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Driss Aboutajdine. Boosting 3-d-geometric features for efficient face recognition and gender classification. *IEEE Transactions on Information Forensics and Security*, 7(6):1766–1779, 2012.
- [59] Syed Zulqarnain Gilani, Faisal Shafait, and Ajmal Mian. Biologically significant facial landmarks: How significant are they for gender classification? In *2013 International conference on digital image computing: techniques and applications (DICTA)*, pages 1–8. IEEE, 2013.
- [60] DJJ Farnell, J Galloway, AI Zhurov, and S Richmond. Multilevel models of age-related changes in facial shape in adolescents. In *Annual Conference on Medical Image Understanding and Analysis*, 2019.