**Faculty of Computer Science and Information Technology**

# 3D Face Analysis using Tensor Approach

**Suriani Ab Rahman**

**Doctor of Philosophy**
**2022**

# 3D Face Analysis using Tensor Approach

Suriani Ab Rahman

A thesis submitted

In fulfillment of the requirements for the degree of Doctor of Philosophy

(Computer Science)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2022

## DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

……………………
Signature

Name: Suriani Ab Rahman

Matric No.: 15010037

Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

Date: 7 August 2020

## ACKNOWLEDGEMENT

All praise belongs to Allah. I would like to offer my amazing thanks to a number of people during my PhD journey. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr Jacey-Lynn Minoi who inspired me to become a better PhD student by having a critical thinking skill. I am thankful to her for her guidance, support, advice, and valuable time she spent with me all the way through my PhD.

I am thankful to my second supervisor, Dr Hamimah binti Ujir who offered me valuable suggestions and advices. I have been lucky to have her as my second supervisor.

Thanks to the Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak for providing a conducive academic laboratory. Not forgotten, the Centre for Graduate Studies for the advice and support given to all the postgraduate students.

Special thanks to Marcella, Khairunnisa, Faizol and Faisal, the members of the Knowledge Transfer laboratory, FCSIT who turn out to be great buddies. Their moral support and technical assistance help me a lot.

I acknowledge the financial support from the Malaysian Ministry of Higher Education (MOHE) for giving the MyBrain scholarship. I owe to them for the funding and thank you for the trust which is given to me to complete my PhD.

My deepest thanks to my family who their prayers, love, support and understanding have been a strong driving force throughout my work. Thank you all.

# ABSTRACT

The advancement of multimodal technology has enabled the creation of large face datasets. Multidimensional characteristics such as covariates and multimodal aspects in 2D, 2.5D, and 3D data are included in these datasets. Early studies in face research used matrix-based and vector-based algorithms to represent faces. According to studies, these methods have the potential to prevent the loss of critical and significant data, which could lead to lower recognition performance. The goal of this research is to develop and validate a tensor-based face recognition method that can overcome the drawbacks of matrix-based Principal Component Analysis (PCA). A face dataset consists of faces with a combination of multiple underlying causal factors such as facial expression, expression intensity, angle of view, gender and race, and the bilinear technique alone is incapable of accurately representing the dataset's multidimensionality. PCA's shortcomings can be overcome by using the tensor decomposition approach to separate these distinct variations. Experimental results have shown that the multilinear tensor approach could statistically outperform the bilinear PCA approach in face recognition applications. This study has added to the understanding of the centring strategy used in tensor models. The median projection operator could maximise the variation for each principal axis in the tensor space and the results have shown that in the median-centred strategy, recognition rates for emotional expressions increased from 0.4 percent to 1.4 percent. Only fear expressions have demonstrated similar recognition performance in mean-centred and median-centred experiments. Current face recognition using PCA is unable to distinguish between different types of centring approaches and cluster them. As a result, this research adopted a hybrid combined framework of a tensor model with an ANOVA (Analysis of Variance) model to uncover the substantial effects of within-subject differences (expression types and expression strengths) on recognition performance. Apart from that, experimental results revealed an interaction effect between those two covariates, showing that the effect of expression types on recognition performance is not continuous and is influenced by intensity levels. This new evidence may well be useful in a variety of recognition processes.

**Keywords:** Tensor model, centring, variation, interaction effect

***Analisis Wajah 3D Menggunakan Pendekatan Tensor***

## ***ABSTRAK***

*Kemajuan dalam penggunaan teknologi multimodal membolehkan penghasilan set data wajah dalam kuantiti yang besar. Set data ini terdiri daripada aspek multidimensi seperti kovariat dan juga aspek multimodal yang merangkumi 2D, 2.5D, dan 3D data. Pendekatan awal kajian wajah telah menggunakan kaedah berasaskan matrik dan vektor untuk mewakili muka. Kajian telah menunjukkan bahawa pendekatan ini menyebabkan kemungkinan berlakunya kehilangan beberapa maklumat penting dan bermakna, yang membawa kepada prestasi pengecaman yang lebih rendah. Tujuan kajian ini adalah untuk membina dan menganalisa kaedah pengecaman wajah berasaskan tensor bagi mengatasi kekurangan 'Principal Component Analysis' (PCA) yang berasaskan matrik. Set data wajah terdiri daripada wajah dengan kombinasi beberapa faktor penyebab yang mendasarinya seperti ekspresi wajah, intensiti ekspresi, sudut pandangan, jantina dan bangsa, dan pendekatan bilinear sahaja tidak dapat mewakili multidimensi set data dengan tepat. Pendekatan penguraian tensor mempunyai keupayaan untuk memisahkan sumber variasi yang berbeza ini dengan baik bagi mengatasi kekurangan PCA. Keputusan eksperimen telah menunjukkan bahawa pendekatan multilinear tensor secara statistik dapat mengatasi pendekatan bilinear PCA yang digunakan dalam aplikasi pengecaman wajah. Kajian ini telah memberikan sudut pandangan baru bagi teknik pemusatan untuk sistem pengecaman wajah berasaskan tensor. Keputusan kajian menunjukkan pengoperasi unjuran median berupaya untuk memaksimumkan variasi pada setiap paksi prinsipal dalam ruang tensor, dan kadar pengecaman bagi ekspresi wajah telah meningkat daripada 0.4 peratus sehingga 1.4 peratus. Hanya ekspresi wajah takut menunjukkan prestasi yang sama di antara kaedah pemusatan min dan median. Pengecaman wajah berasaskan model PCA tidak dapat* *menunjukkan perbezaan apabila kaedah pemusatan yang berbeza digunakan. Oleh itu, kajian ini telah mengintegrasikan kaedah hibrid di antara model tensor dan model ANOVA (Analysis of Variance) dalam mendedahkan kesan signifikan yang dimiliki oleh variasi intra-subjek (jenis ekspresi dan tahap ekspresi) terhadap prestasi pengecaman. Selain daripada itu, keputusan eksperimen juga telah menunjukkan bahawa wujud interaksi di antara dua kovariat tersebut yang bermaksud bahawa kesan jenis ekspresi terhadap prestasi pengecaman adalah tidak tetap dan ia bergantung kepada perubahan yang berlaku pada tahap intensiti juga. Pengetahuan ini dapat memberi faedah selanjutnya dalam pelbagai proses pengecaman.*

**Kata kunci:** *Model tensor, pemusatan, variasi, kesan tindakan bersilang (interaksi)*

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ANOVA | Analysis of Variance |
| AUs | Action Units |
| CANDECOMP | Canonical Decomposition |
| FR | Face Recognition |
| FRGC | Face Recognition Grand Challenge |
| HOSVD | Higher-Order Singular Value Decomposition |
| PARAFAC | Parallel Factor Analysis |
| PC | Principal Component |
| PCA | Principal Component Analysis |
| ROC | Receiver Operating Curve |
| SHREC | Shape Retrieval Contest |
| SPSS | Statistical Product and Service Solutions |
| SVD | Singular Value Decomposition |

# CHAPTER 1

# INTRODUCTION

## 1.1 Overview

The face dataset is made up of several underlying causal factors, including expression type, intensity, head pose, lighting condition, age, gender, and race. Interactions among all these factors form the image itself (Vasilescu & Terzopoulos, 2000a). The commonly used Principal Component Analysis (PCA) is a well-established method for projecting datasets from a high-dimensional space to a new lower-dimensional space while maintaining overall variability (Sirovich & Kirby, 1987; Turk & Pentland, 1991).

However, as being mentioned by Vasilescu and Terzopoulos (2002a), PCA is only applicable to a single factor variation or two-way data only. A two-way data refers to a condition in which the first mode represents the subjects (the single variation) and the second mode represents the true measurement taken: 5090 points ($x$, $y$, $z$) for each face surface. In response, a tensor model (Tucker, 1964, 1966) was successfully implemented (Vasilescu & Terzopoulos, 2000b) to separate the causal factors or the face variants in a different mode (multimode or multiway) to overcome the limitations of PCA. This new arrangement is required in order to avoid losing the important information that is significant in understanding the organisation of the data as a whole (Acar & Yener, 2009).

Despite the fact that many tensor-based face recognition systems have been presented over the years, very few studies have been done on the statistical specifics of the multilinear technique employed in applications. There are still problems with data pre-processing, particularly with the centring technique, because an erroneous approach can increase the projection inaccuracy (Johnson & Wichern, 2008). Furthermore, the relationship among the causal factors or face variants should be studied to verify that the overall results are conclusive and credible (Wickremasinghe, 2005; Givens et al., 2013). These issues are frequently attributable to the biological nature of the human face and/or oral maxillofacial surgery (the explanation excludes the oral maxillofacial surgery as it is not part of this research work).

Therefore, this research study proposes to investigate and analyse the importance of centring techniques in the context of bilinear and multilinear approaches, and to identify the extent of the tensor model given the various causative factors in the context of recognition in faces. The novel findings are two-fold: (i) the benefits of using a tensor model to distinguish between different centring approaches, and to increase the total variations of the dataset in a parsimonious space; and (ii) an extended knowledge of the multivariate relationships among the causal factors in the tensor model, which have not been reported in any previous face recognition literature (Smeets et al., 2012; Lu et al., 2019), by using the proposed tensor-based hybrid approach with Analysis of Variance (ANOVA) model to further explain and convince the relationships.

## 1.2 Research Problem

Data pre-processing should be the first step in any multiway analysis. The technique is essential for determining a correct analysis (Rizzi & Vichi, 1995). Centring is one of the approaches. It refers to the process of subtracting a constant term from each element. In other words, centring means the coordinate system is moved to a new reference point, typically the origin. If the projection is performed without a centring procedure, the total variations carried by the principal axes may not represent the true total variations (Kroonenberg, 1989;

Cadima & Jolliffe, 2009; Honeine, 2014; Alexandris et al., 2017). Despite the fact that some research used centring processes, the majority of them used mean centring. Outliers are well known to have an impact on the value. As a result, despite the fact that centring was used, the major axes obtained may not have the maximum variation information.

The presence of a significant degree of variety in human faces is a major difficulty in face recognition studies. The challenge of extracting intrinsic information from a person's face is challenging since within-subject variations may be larger than between-subject variations (Adini et al., 1997; Lu, 2004). Although a tensor model can successfully separate all the variations, dimension reduction for a tensor object in a parsimonious space contributes to the different amounts of total energy (variation) for each mode, which may still lead to a problem in assessing the face variants relationships (Kroonenberg, 1989).

The following subsection are the concerns that are being addressed in this research.

### 1.2.1 Projection Error

The mean statistic tends to maximise the sum of squared distances from each data point to the selected principal axes since the value is easily affected by outliers. As a result, it will increase the projection error and the principal axes with maximum variation may not be obtained. Maximising the variation is important because it represents the important facial features that characterise all the subjects and can be used as an identity cue to differentiate subjects. Therefore, it is important to minimise the squared distances in order to obtain the principal axes with maximum variation in the tensor space.

### 1.2.2 Relationships Among the Face Variants

In face verification system performance, it has been found that some covariates may have shown different trends when other covariates are considered together (Lu et al., 2019), or that there may exist an interaction effect among the covariates themselves. Such relationships may give some impact towards recognition performance. Several studies used the Receiver Operating Curve (ROC) to establish the relationships among the covariates. However, as mentioned in (Bowyer & King, 2019; Lu et al., 2019), ROC tends to give inconsistent results and does not show a clear trend.

## 1.3 Research Objectives

This research aims to demonstrate and validate the advantages of using a tensor approach as a multidimensional array model to replace the classical and commonly used two-dimensional array in analysing facial datasets by measuring the extent of the variances in a parsimonious space. This is especially important with the increase of high-dimensional datasets and spaces. The evaluation of the advantage of the proposed method will be based on the variances (singular values), the total variations accounted for the retained principal axes, as well as the correct recognition rates obtained. Besides that, a reliable statistical test will be used to prove the existence of the claimed interaction effects among the face variants. As such, the objectives of this study can be listed as follows:

- To examine the effect of various centring approaches on minimising the projection error and increasing the total variations,
- To build a centralised tensor-based facial recognition system and evaluate its performance,

- To examine the effect of expression types and expression strengths, towards recognition performance.

The following hypotheses will be examined.

H1: There is a possibility of reducing the projection error using a median projection operator.

H2: There is a possibility of discovering unknown or hidden interactions among the different face variants.

## 1.4 Expected Outcomes

This study introduces the use of a median statistic as a constant term ($b$) in the distance formula, $d_i = (x_i - b)$, to replace the commonly used mean statistic. Since the median statistic is not affected by the outliers, it can minimise the sum of squared distances from each data point to the principal axes. Through this, it will minimise the projection error and maximise the total variations that the principal axes carry (each subspace carries the maximum variation). The results obtained from this study should be able to demonstrate that the proposed median-centred tensor-based face recognition system could increase the recognition rates of correct identification as compared to the uncentred and mean-centred tensor-based face recognition systems, as well as the traditional PCA approach.

Apart from that, because the model addresses within-subject variations rather than between-subject variations, this study should be able to assess within-subject variations using the repeated measure of ANOVA approach. This approach may be able to overcome PCA's limitation that only decompose the variations according to its components, and not between the sources of the covariates. As a result, the proposed approach can be used to test and validate the impact of expression types and expression strengths towards recognition performance, as well as to uncover any potential hidden interaction between those two covariates. This approach provides statistically reliable evidence that defines the face variants relationships at a deeper level.

## 1.5 Scope of Study

A 3D face dataset will be used in this investigation. Since this study focuses on using a tensor-based multilinear technique on a database of faces with expression variations (expression types and expression strengths), it is not possible to compare it to other face datasets. This is due to the fact that the tensor model is created by separating the causal factors or the face variants into different modes. Binghamton University 3D Facial Expression dataset (BU-3DFE) contains six different expressions with four levels of intensity. In order to make an accurate comparison (variability, recognition rates, and relationships among the face variants), the datasets need to have a similar number of expression types and intensity levels as the BU-3DFE dataset. Since no other 3D face dataset with the same face variants is currently available, this study relies on the BU-3DFE dataset. Although another two face datasets are available at hand: the Imperial College London face dataset and the Notre Dame face dataset, these two datasets do not have a similar number of expression types and both were prepared without intensity data.

## 1.6 Summary of Research

The research investigation is summarised in Table 1.1.

**Table 1.1:** Research Summary

| Problem statements | Research questions | Research objectives | Research contributions | Deliverables |
|---|---|---|---|---|
| The principal axes obtained using mean centring may not carry the maximum variation since the mean value is affected by outliers. | Can the mean statistic be replaced by the median statistic? | To examine the effect of centring in minimising the projection error and increasing the total variations. | Introduce a median projection operator which can minimise the projection error and maximise the total variation. | A median-centred tensor-based face recognition system (as an alternative to the traditional mean-centred tensor-based face recognition system). |
| It is difficult to extract the interaction effects among the face variants, due to the difference in the total amount of energy (variation) in each mode of a tensor object. | (i) Are there any possible interactions among the face variants?<br>(ii) If the relationships exist, how to assess or reveal them? | To examine the effect of expression types and expression strengths, towards recognition performance. | Reveal or establish interaction effects that exist among expression types and expression strengths (this extends findings from a previous study by Smeets et al. (2012)). | A hybrid approach using tensor and ANOVA, (which emphasises on variations within a subject) to establish the interaction effects among the face variants (as an alternative to using the core tensor and the ROC). |

## 1.7 Chapter-by-Chapter Summary

The organisation of the remaining chapters of this thesis is as follows:

**Chapter 2** discusses the existing research on centring approaches and the relationships among the face variants. A list of techniques that will be used in this study is also provided.

**Chapter 3** presents a brief summary of the methodology used for this study. It gives an overview of the experimental setup, dataset, and the procedures to be implemented. Most importantly, the core of the thesis is presented, the two main contributions – namely the median centring approach, and establishing the interaction effects that exist among expression types and expression strengths.

**Chapter 4** is essentially about the possible implementation and testing, for the centralising and ANOVA procedures, together with the development of the tensor-based face recognition system. The chapter also investigates the impact of intra-subject variations on recognition performance.

**Chapter 5** concludes the thesis with a summary of contributions to the computer science field generally, and to the face recognition community specifically. In addition, possible extensions for future research are also indicated.