



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

	Page
DECLARATION	i
ACKNOWLEDGEMENT	ii
ABSTRACT	iii
<i>ABSTRAK</i>	v
TABLE OF CONTENTS	vii
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1: INTRODUCTION	1
1.1 Overview	1
1.2 Research Problem	2
1.2.1 Projection Error	3
1.2.2 Relationships Among the Face Variants	4
1.3 Research Objectives	4
1.4 Expected Outcomes	5
1.5 Scope of Study	6
1.6 Summary of Research	6
1.7 Chapter-by-Chapter Summary	8

CHAPTER 2: LITERATURE REVIEW	9
2.1 Overview	9
2.2 Image Analysis	9
2.3 Tensor	11
2.4 Centring	14
2.4.1 The Important of Centring	15
2.4.2 The Mean Centring Issue	17
2.5 Relationships among the Face Variants	19
2.5.1 History of the Interaction Effects	20
2.5.2 The Issue with the Current Approach	22
2.5.3 Expression-Intensity Studies in the Psychology Domain	26
2.6 Implemented Techniques and Models	29
2.6.1 Singular Value Decomposition	29
2.6.2 Tucker Model	31
2.6.3 Combination of PCA and ANOVA Model	33
2.7 Discussion and Conclusions	35
CHAPTER 3: MEDIAN CENTRING AND ESTABLISHMENT OF INTERACTION EFFECTS	37
3.1 Overview	37
3.2 Research Workflow	37
3.3 Research Framework	37

3.4	Proposed Approach and Model	41
3.4.1	Median Projection Operator	41
3.4.2	Tensor-ANOVA Model	44
3.5	Experimental Setup and Dataset	46
3.6	Evaluation Approaches	48
3.6.1	Classification Accuracy	48
3.6.2	Leave-One-Out Cross-Validation	50
3.6.3	Statistical Test for Within-Subject Effects	50
3.7	Discussion and Conclusions	53
CHAPTER 4:	TESTING AND IMPLEMENTATION	54
4.1	Overview	54
4.2	Experiment 1 (Centring)	54
4.2.1	Important Statistical Properties	55
4.2.2	Experimental Works I	58
4.2.3	Results and Discussion I	59
4.2.4	Experimental Works II	63
4.2.5	Results and Discussion II	65
4.2.6	Summary	70
4.3	Experiment 2 (Interaction Effects)	70
4.3.1	Experimental Works	71

4.3.2	Results and Discussion	72
4.3.3	Conclusions	96
CHAPTER 5: CONCLUSION AND FUTURE WORKS		98
5.1	Conclusion	98
5.2	Limitations and Directions of Future Research	100
5.2.1	Applying a Two-Way Centring Technique	100
5.2.2	Deep Learning Approach	101
5.2.3	Considering a Non-Linear Case	101
REFERENCES		102
APPENDICES		120

LIST OF TABLES

	Page
Table 1.1 Research Summary	7
Table 2.1 A List of Studies that Use the Mean Centring Approach	18
Table 2.2 A List Of Studies With and Without Interaction Effects	24
Table 3.1 The Differences between Median and PCA Projection Operators	43
Table 3.2 Possible Outcomes for a Binary Classification	49
Table 4.1 Singular Values	60
Table 4.2 Comparison of Several Singular Values	61
Table 4.3 Total Variations	63
Table 4.4 Combination of Expressions	66
Table 4.5 Number of Correct Faces for Angry Query Set	74
Table 4.6 Number of Correct Faces for Disgust Query Set	74
Table 4.7 Number of Correct Faces for Fear Query Set	75
Table 4.8 Number of Correct Faces for Happy Query Set	75
Table 4.9 Number of Correct Faces for Sad Query Set	75
Table 4.10 Number of Correct Faces for Surprise Query Set	76
Table 4.11 The Mauchly's Test of Sphericity ^b (SPSS Output)	88
Table 4.12 List of Variables	90
Table 4.13 Several Tests for Within-Subject Effects (SPSS Output)	92
Table 4.14 Pairwise Comparison (Rewritten from SPSS Output)	94

LIST OF FIGURES

	Page
Figure 2.1 The Operational Steps in a Machine Vision System	10
Figure 2.2 A Third-Order Tensor, $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$	12
Figure 2.3 A Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{7 \times 5 \times 8}$	12
Figure 2.4 Fibers for a Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$	13
Figure 2.5 Slices for a Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$	13
Figure 2.6 Principal Component of a Covariance Matrix	17
Figure 2.7 Example of Facial Emotion Intensity for Anger	27
Figure 2.8 Neutral Face (0%) until Full-Blown Emotional Face (100%) Intensity	27
Figure 2.9 Facial Expressions in High and Low Intensity	29
Figure 2.10 Tucker Decomposition for a Third-Order Tensor	32
Figure 2.11 The HOSVD of a Third-Order Tensor	33
Figure 3.1 Research Workflow	38
Figure 3.2 Research Pipeline	39
Figure 3.3 Research Framework	40
Figure 3.4 Centred and Non-Centred Face Recognition System	41
Figure 3.5 The Pipeline of the Proposed Median Centring	42
Figure 3.6 The Workflow of the Proposed Tensor-ANOVA Model	45
Figure 3.7 Establishment of the Interaction Effects in the Tensor-ANOVA Model	45
Figure 3.8 Example of Expression with Four Levels of Intensity	47
Figure 3.9 Example of a Subject Showing Six Different Expressions	47

Figure 4.1	Comparison of Four Singular Values	60
Figure 4.2	Recognition Rates for PCA Model	67
Figure 4.3	Recognition Rates for Tensor Model	68
Figure 4.4	Comparison of Performance between PCA and Tensor Model	69
Figure 4.5	Intensity Level Three shows the Lowest Rate for the Kernel Approach	70
Figure 4.6	Basic Workflow for Two-Way ANOVA (Van Den Berg, 2015)	72
Figure 4.7	Intensity Level One Shows the Highest Correct Recognition	73
Figure 4.8	Figure (i) until (xxiv) Show the Box-Plots for Each Variable	80
Figure 4.9	Figure (i) until (xxiv) Show the Distribution of the Residuals	84
Figure 4.10	Figure (i) until (xxiv) Show the Normally Distributed Residuals	87
Figure 4.11	Flowchart for Mauchly's Test of Sphericity (Van Den Berg, 2015)	89
Figure 4.12	Profile Plot for the Estimated Marginal Means of Distance	91

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? (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.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

The background of image analysis, tensor, and centring techniques are presented in this chapter. In Section 2.2, a brief introduction to image analysis is given, followed by a discussion of tensor representation in Section 2.3. Section 2.4 and Section 2.5 then focus on the two primary challenges that arise when using the tensor technique, namely input pre-processing (centring approach) and output post-processing (relationships among the face variants). A list of current methodologies and models employed in this work can be found in Section 2.6. Section 2.7 summarises this chapter with a discussion and conclusion.

2.2 Image Analysis

In the last couple of decades, image computing has become one of the most rapidly growing fields of research. It provides useful results from image enhancement, image synthesis, and image analysis for quantitative measures used in decision-making for quality evaluation and improvement. This decision-making step is the final process, after classification, in a machine vision system or machine learning system. Figure 2.1 shows a general pipeline of the operational steps in a machine vision system (Mendoza & Lu, 2015).

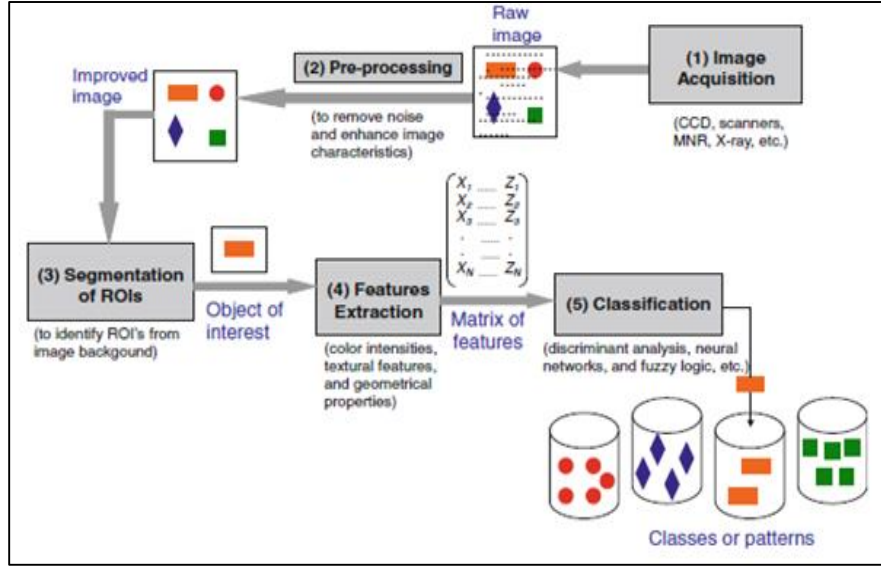


Figure 2.1: The Operational Steps in a Machine Vision System

As the proposed research focuses on image analysis techniques, the definition of the process needs to be fully understood. According to Prats-Montalbàn et al. (2011), image analysis usually returns numeric values or graphical information about the image characteristics or properties. The term *image analysis* is used when the output is a number or decision, not an image.

In general, image analysis can be described as the process of extracting useful information, mainly from digital images. The information obtained could be in the form of quantitative or qualitative characterisation of two-dimensional (2D) or three-dimensional (3D) digital images. For example, statistical details (such as variance and covariate relationships) can help researchers understand data at a deeper level, so that a correct conclusion or decision could be made.

Since a set of 3D face models with different causal factors (subject, expression, intensity) is studied, the usual two-way analysis (x, y mode) is no longer suitable. A multiway analysis is considered in order to avoid losing any information that is significant in

understanding the organisation of the data as a whole. In this case, a higher-order tensor (a mathematical object with more than two modes) is needed to represent the data. This provides a natural and compact representation for massive multidimensional data using low-rank approximations (Cichocki, 2014; Cichocki et al., 2016).

2.3 Tensor

A tensor can be defined as a multidimensional object. An N -way or N^{th} -order tensor is an object with N dimensions. In other words, it is an element of the tensor product of N vector spaces, and each vector has its coordinate system (Kolda & Bader, 2009). As said earlier, a vector which has one dimension can be defined as a first-order tensor. A matrix with two dimensions is a second-order tensor, and an object or array with three dimensions is a third-order tensor. Higher-order tensors are specifically referring to tensors of order three or higher. It is denoted by a Lucida calligraphy letter, such as $\mathcal{A}$. The next paragraphs will introduce a third-order tensor only, as it is the simplest representation in higher-order tensors. The notation used in this thesis is as proposed by Kiers (2000).

A third-order tensor is an object with three dimensions or has three indices (Figure 2.2 and Figure 2.3). Figure 2.2 shows an example of a third-order tensor $\mathcal{A}$ with dimensions I (mode-1 or number of rows, $i = 1, \dots, I$), J (mode-2 or number of columns, $j = 1, \dots, J$), and K (mode-3 or depth levels, $k = 1, \dots, K$) (Kolda & Bader, 2009). As an example, Figure 2.3 represents a third-order tensor $\mathcal{Y}$ with $I = 7$, $J = 5$, and $K = 8$, $\mathcal{Y} \in \mathbb{R}^{7 \times 5 \times 8}$ with element $y_{ijk} = y_{651}$ (6 = the sixth row, 5 = the fifth column, 1 = depth level one or the first layer) (Cichocki, 2014). Each element can be denoted by an alphabetical subscript, such as x_{ijk} . These three indices (i, j, k) can be fixed to define elements of a tensor.

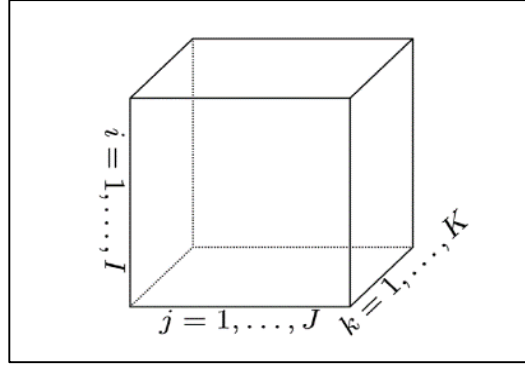


Figure 2.2: A Third-Order Tensor, $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$

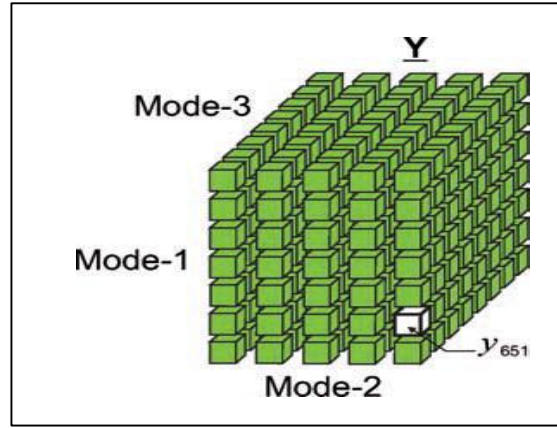


Figure 2.3: A Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{7 \times 5 \times 8}$

Fibres of a tensor form a vector, obtained by fixing two indices and leaving one index free, denoted by a semicolon, such as $(i, j, :)$. This results in three types of fibres, referred to as column fibres $(:, j, k)$, row fibres $(i, :, k)$, and tube fibres $(i, j, :)$, as shown in Figure 2.4 (Cichocki, 2014). A colon $(:)$ is used to indicate all elements of a mode, following the style of Matlab. As has been mentioned by Bader and Kolda (2006), fibres are usually depicted as column vectors.

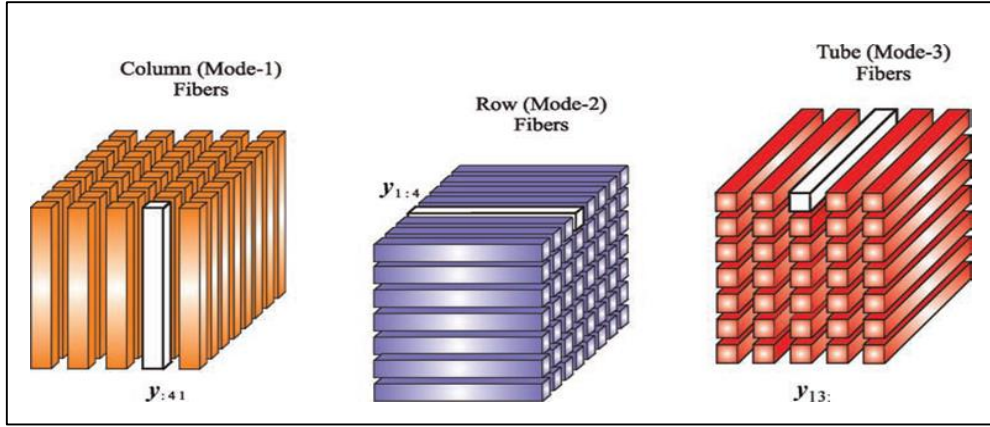


Figure 2.4: Fibers for a Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$

Other than fibres, another form of elements of a tensor is slices. This is obtained with only one index being fixed, and leaving the other two indices free. Horizontal slices (mode-1 slices) are obtained when the first index is fixed, lateral slices (mode-2 slices) when the second index is fixed, and frontal slices (mode-3 slices) when the third index is fixed, as shown in Figure 2.5 (Cichocki, 2014).

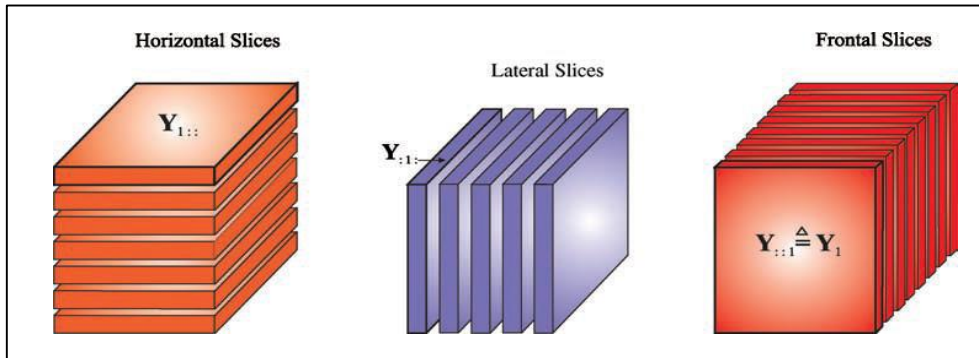


Figure 2.5: Slices for a Third-Order Tensor, $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$

An example is given in Figure 2.5, where $\mathbf{Y}[1, :, :]$ is the first slice of mode-1, $\mathbf{Y}[:, 1, :]$ is the first slice of mode-2, and $\mathbf{Y}[:, :, 1]$ is the first slice of mode-3. If $\mathbf{Y}[:, 6, :]$ is written, it represents the sixth slice along with the second mode of the tensor.

A more detailed discussion regarding the tensor model is presented in Section 2.6.

2.4 Centring

Multiway analysis should begin with a proper pre-processing stage. One of the most important stages in pre-processing is scaling. According to Kroonenberg (1981), the word scaling refers to both centring and standardisation. Rizzi and Vichi (1995) mentioned the important reasons for having a scaling procedure, which are: (i) to make data suitable for three-way analysis, (ii) to allow the implementation of the adopted three-way model, (iii) to convert different types of data, such as interval scale data to a ratio scale, and (iv) to connect the correlation effect within each slice of the three-way dataset.

Centring is the process of subtracting a constant term (normally the mean value) from each element in the dataset. It means the coordinate system is moved to a new reference point, usually the origin. In result, the new centred data will have a mean value of zero. Tucker (1966), Rowe (1979), Harshman (unpublished, quoted in Kruskal, 1981), and Kruskal (1981), were among the first to discuss centring for three-mode models. Most of the schemes for centring were proposed by at least one of them.

Rizzi and Vichi (1995) stated that the common centring procedures for multiway data are centring the variables over all subject-condition combinations (one-way centring), centring the variables over all subjects for each condition separately (two-way centring), or centring per condition over both variables and subjects (three-way centring). The decision on which centring is appropriate depends on the researcher's assessment on the variability of the data at hand (Kroonenberg, 1989).

Orlóci (1967) did not recommend applying double-centring since it may deviate from the original structure of the data defined by the original observations or measurements in an arbitrary manner. In other words, the new data after performing double-centring would no

longer contain information similar to the original situation. Cattell (1965) also discussed the other reasons for this inapplicability. However, both Kruskal (1981) and Harshman (unpublished, quoted in Kruskal, 1981) argue that certain centring is inappropriate for the three-mode model as the components after centring would bear no simple relation to the components before transformation.

On the other hand, standardisation is the process of dividing each element by a constant term, and the new data would normally have a variance of one (Kruskal, 1981). As for the decision on which one is to be used, should it be centring or standardisation, or both, depends on the measurement characteristics, research questions, and research design. For this study, only the centring procedure is carried out. This is because the standardisation procedure has a tendency to remove important or meaningful variations (Kroonenberg, 1989). Therefore, the following discussion focuses on the centring approach only.

2.4.1 The Important of Centring

From the definition of PCA, the components obtained represent the directions in the space that represent the largest variations. Therefore, if the centring procedure is not performed, the eigenvalues obtained may include the variance and mean of each variable. In other words, it is not maximising the variance anymore (Alexandris et al., 2017), and an incorrect interpretation of the relationship would be obtained if it is applied to raw data (Kroonenberg, 1989). Furthermore, as was written in (Sirovich & Kirby, 1987), it is reasonable to assume that an efficient procedure for recognising and storing pictures would focus on departures from the mean.

Apart from that, according to Kroonenberg (1989), an improper description may be obtained if a certain value, such as the mean, is not removed from the raw data (uncentred

data). This statement follows from the definition of the component analysis: if the centroid is located at a considerable distance from the origin of the space, then an important candidate for the direction of the first component will be the one from the origin through the centroid. Therefore, if the main purpose of an analysis is to investigate the covariations of the variables from the centroid, the means of the variables should be removed before the component analysis is carried out.

In (Orlóci, 1967), researchers studied the effect of centring on two types of matrices, the R-matrix and the Q-matrix. The R-matrix represents a condition where the number of samples exceeds the number of variables, whereas the Q-matrix is the other way around. They concluded that centring on variables should be applied to the Q-matrix, and centring on individuals should be applied to the R-matrix. These two matrices are commonly found in component analysis, and choosing the centring types incorrectly may lead to an incorrect value of the correlation or variance-covariance matrix.

The importance of centring is also discussed in (Cadima & Jolliffe, 2009; Jolliffe & Cadima, 2016). They believe that the SVD approach for obtaining principal components (axes) should be applied to the variable-centred data matrix. Although several studies tend to use the SVD approach only with uncentred data, Cadima and Jolliffe (2009) did not agree with this approach. They conducted a study to compare the results from standard centred data with uncentred data, and concluded that it is not advisable to immediately draw the conclusion that both variants of PCA should be similar unless the vector of means is near zero (in which case centred and uncentred moments are similar).

2.4.2 The Mean Centring Issue

Although mean centring has been used in many studies, there are debates on removing the mean value due to its sensitivity to outliers. The outliers may affect the calculation of the deviation vector, d_i (Figure 2.6), which may cause a projection error, and further analysis might not produce the intended results.

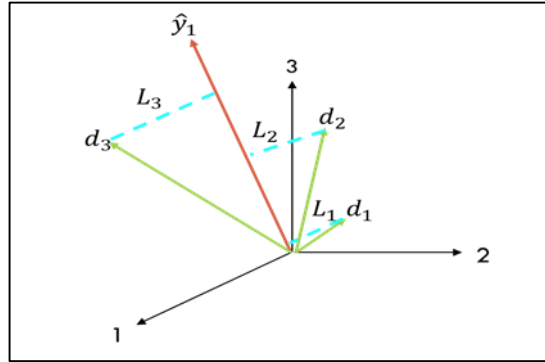


Figure 2.6: Principal Component of a Covariance Matrix

Figure 2.6 can help the readers understand how an error in projection could occur. Since the centring procedure involves the subtraction of a *constant term* (whether it is a mean or a median, say b) from each element in the dataset, it is written as $(x_i - b)$ and named as the deviation vector, d_i (with $d_i = x_i - b$). At this stage, the main objective is to minimise the sum of the squared lengths, $\sum L_i^2$, where L_i is the length from the deviation vectors to the chosen principal axis (Johnson & Wichern, 2008). An incorrect principal axis would be obtained if the chosen b maximises the sum of the squared lengths, instead of minimising it. Since the mean value is affected by outliers, there is a possibility that a longer deviation vector (d_i) would be obtained when the value of b is too large. This will increase the length (L_i), as well as the sum of the squared lengths ($\sum L_i^2$). As a result, incorrect principal axes would be obtained.

The disadvantage of using the mean value is also discussed in (Gnanadesikan & Kettenring, 1972). They mentioned that the presence of outliers might seriously bias the usual effects that are subtracted from an observation, such as row, column, and overall mean vectors in a two-way classification. They showed that if the outlier is extreme enough, it might bias the mean vector so badly. As a result, they suggested considering the use of other robust estimators of location, such as the vector of medians (Mood, 1941), the vector of Hodges-Lehmann (Bickel, 1964), or the vector of α -trimmed means, to replace the usual sample mean vector.

The following table shows examples of studies on face recognition (from proceedings and journal papers) performing mean centring (taken from the period of 2018 until 2020). It shows that researchers used the mean value instead of robust measures of centrality, such as median.

Table 2.1: A List of Studies that Use the Mean Centring Approach

Author(s)	Title	Centred by	
		Mean	Median
Avuglah et al. (2018)	Recognition of Face Images Under Varying Head-Poses Using Fast Fourier Transform-PCA/SVD Algorithm.	✓	
Jahan (2018)	On Dimension Reduction Using Supervised Distance Preserving Projection for Face Recognition.	✓	
Leo & Suchitra (2018)	SVM Based Expression-Invariant 3D Face Recognition System.	✓	
Lestriandoko et al. (2018)	The Behaviour of Principal Component Analysis and Linear Discriminant Analysis (PCA-LDA) for Face Recognition.	✓	

Table 2.1 continued

Li & Mi (2018)	Sparse Factorial Code Representation Using Independent Component Analysis for Face Recognition.	✓	
Meshkinfamfard et al. (2018)	Tackling Rare False-Positives in Face Recognition: A Case Study.	✓	
Abdulhamed & Othman (2019)	A Measure of Happiness and Sadness Based on Facial Expression Recognition.	✓	
Aminu & Ahmad (2019)	Locality Preserving Partial Least Squares Discriminant Analysis for Face Recognition.	✓	
Cevikalp & Dordinejad (2019)	Discriminatively Learned Convex Models for Set Based Face Recognition.	✓	
Cuculo et al. (2019)	Robust Single-Sample Face Recognition by Sparsity-Driven Sub-Dictionary Learning Using Deep Features.	✓	
Sendik & Keller (2019)	DeepAge: Deep Learning of Face-Based Age Estimation.	✓	
Yadav & Mehta (2019)	An Improved Illumination Normalisation and Robust Feature Extraction Technique for Face Recognition Under Varying Illuminations.	✓	
Ouanan et al. (2020)	Development of Deep Learning-Based Facial Recognition System.	✓	
Srivastava et al. (2020)	Real Time Attendance System Using Face Recognition Technique.	✓	

2.5 Relationships among the Face Variants

The tensor approach provides an opportunity to investigate the hidden structure of a multidimensional data. One way is through an investigation on the relationships among the factors, particularly the interaction effects. Wickremasinghe (2005) mentioned that the

primary objective of the multiway analysis should be to examine the interactions that may exist among the factors simultaneously. Then any results obtained should be interpreted based on the presence or absence of the interactions.

According to Marasinghe and Johnson (1981), an interaction exists when one of the row treatments reacts differently compared with the column treatments, or when compared with the other row treatments, or with any specific row treatment and column treatment. In a simpler interpretation, an interaction effect could exist if the effect of one factor on the subject under investigation is dependent on the second factor, or one factor reacts differently when another factor is considered. One important point to be noted is that the responses may react differently just because of the extreme level of a factor.

In the context of face recognition, the recognition performance is affected by both external and internal factors. External factors can be camera position and illumination, whereas internal factors are subject characteristics such as gender, age, and skin colour. These situations lead to a growing interest in analysing the effect of such factors (or covariates) on a given algorithm (Abdurrahim et al., 2018; Redfern & Benton, 2019).

2.5.1 History of the Interaction Effects

Interactions among the factors became an interest since the work of Tukey (1949). He proposed a test for testing the interaction effects or the non-additivity term in a two-way classification model. If an interaction exists, a transformation of the data should be applied in order to remove the non-additivity of the model (if it is not caused by outliers), and a standard linear additive model is applied. His test was proven to be a good tool for testing an interaction in a two-way model (Ghosh, 1963). Mandel (1959, 1961), Harter (1962), and Milliken (1970) continued to propose several generalised versions of Tukey's work.

Later, another approach was proposed by Johnson and Graybill (1972a) to handle the interaction effects. Their approach was to test for an interaction, and if it exists, to treat the interaction effect without removing it (as was proposed by Tukey), and to estimate the variance of the error term. Then, the estimate of the variance value is used to build the confidence intervals for the treatment effects.

All of the above approaches were developed to test for the existence of two-way interactions using non-additive models. In contrast, there is an alternative approach that can also be used, which is based on Singular Value Decomposition (SVD) through Principal Component Analysis (PCA). The two-way interactions using the combination of additive and multiplicative models are discussed in (Fisher & Makenzie, 1923; Gilbert, 1963; Gollob, 1968; Mandel, 1969, 1971). Further work on testing the interactions was also discussed in (Johnson & Graybill, 1972b).

Apart from that, Marasinghe and Johnson (1981) developed a procedure for testing interactions and for determining subareas of the data where the interactions might exist. Then the error variance is estimated based on the data that is free from any interaction effect. They claimed that the procedure is better than what had been developed by Johnson and Graybill (1972a). Next, Boik and Marasinghe (1989) extended the procedure to three-way data by combining the models of (Mandel, 1971) and (Johnson & Graybill, 1972b).

Another approach that is also based on SVD (or PCA) can be found in the Tucker model, which is solely based on a multiplicative model for modes higher than two. Although the interest in testing the interactions among the factors has been extended from two-way interactions to three-way interactions, the most appropriate general methods that look at the interactions within the three-mode simultaneously were only available after the 1980's.

Based on the Tucker model, which was the first model that discussed the three-way interactions (simultaneously) in the context of a tensor object, Kroonenberg and De Leeuw (1980) further developed a model where the parameters of the model can be estimated using an alternating least squares algorithm. The test derived for testing the three-factor interactions in the Tucker model is also discussed in (Marasinghe & Boik, 1993).

Wickremasinghe and Johnson (2002) also extended the work of Johnson and Graybill (1972a), and derived a test for the interaction effects in subareas of three-way data using Tucker's approach.

2.5.2 The Issue with the Current Approach

All of the studies mentioned earlier were performed on a full model. The advantage is that the interaction effects could be obtained easily from the core tensor values for a full tensor model. However, a problem would arise when one needs to assess the interaction effects for a reduced tensor model. Dimension reduction in each mode for a tensor model contributes to the different amount of variations in each mode, which leads to a problem in assessing the covariates relationships (Kroonenberg, 1989).

Several studies used the Receiver Operating Curve (ROC) to establish the relationships among covariates. Paone et al. (2011), for example, investigated three covariates, namely expression, gender, and illumination, using the Face Recognition Grand Challenge (FRGC) dataset (Phillips et al., 2005) with two types of expressions, smiling and neutral, and that illumination was set as under control and uncontrolled illumination. They observed that although a subject was difficult to be recognised under one set of external conditions, the subject could be easily recognised when the external conditions changed.

They concluded that smiling was easier to recognise compared to neutral faces, irrespective of what the conditions are for illumination and gender.

Smeets et al. (2011) conducted another research in which they compared seven 3D face databases - FRGC v1, FRGC v2, FRAV 3D, CASIA, BU-3DFE, SHREC, and Bosphorus (refer to Appendix B for references). They concluded that the two most important factors influencing performance were pose variations and expression variations. The three categories that fall under expression variations were expression types, intensity levels, and the total number of expressions. In order to determine the most influential factors, they measured the influences of the factors based on the performance of a baseline algorithm. They found that other than large rotations of the pose variations, an increasing number of expressions and expression intensity levels would decrease the recognition performance.

Recently, Lu et al. (2019) investigated the effect of both inter-subject and intra-subject variations on verification performance. They chose pose (yaw and roll), age, facial hair, gender, indoor (outdoor), occlusion (nose and mouth visibility, eyes visibility, and forehead visibility), and skin tone. They used three face datasets that have face verification covariate protocols, which are the Celebrities in Frontal-Profile in the Wild (CFPW) (Sengupta, 2016), the Intelligence Advanced Research Projects Activity (IARPA) Janus Benchmark B (IJB-B) 1:1 covariate (Whitelam, et al., 2017), and the IARPA Janus Benchmark C (IJB-C) 1:1 covariate (Maze et al., 2018). They concluded that pose variations and occlusions were the top confounding factors that could cause performance to drop by large margins. However, indoor images showed better performance compared to outdoor images. Apart from that, age, gender, and skin tone factors also have shown the impacts. Males were easier to verify than females, old subjects generally performed better than young

ones, and for skin tone, light pink achieved the best performance while medium-dark brown performed worst (the skin tone changes from light to dark, defined as: (i) light pink, (ii) light yellow, (iii) medium pink/brown, (iv) medium yellow/brown, (v) medium dark brown, and (vi) dark brown). Their conclusions were made based on ROC graphs performance.

Another study was done by Grm et al. (2018), who chose the Visual Geometry Group, VGG-Face dataset (Parkhi et al., 2015) to train the models, and the Labelled Faces in the Wild (LFW) dataset (Huang et al., 2007) to evaluate the performance. They found that the verification performance was affected by noise, image brightness, and missing data. The performance consistently degraded for low-resolution images but was least affected by colour changes. However, their conclusion was made based on a performance metric verification protocol (Huang & Miller, 2014) using mean and standard deviation values, and not the robust measures of centrality, such as median.

Table 2.2 lists several recent studies that discuss the effect of face variants on recognition performance. A number of the listed studies did not include the interaction effects, and for the studies that did, ROC was used to establish them.

Table 2.2: A List Of Studies With and Without Interaction Effects

Authors (Year)	Dataset	Covariates	Interaction effects
Jandová & Urbanová (2018)	FIDENTIS	Gender, expression	No interaction (Based on non-parametric MANOVA)
Bowyer & King (2019)	MORPH	Race, gender	X
Cook et al. (2019)	Eleven commercial face biometric systems on 363 subjects	Skin reflectance, gender, age, race, eyewear, and height, weight	X

Table 2.2 continued

Lu et al. (2019)	IARPA Janus (B, C), Celebrity Frontal-Profile (CFP)	Pose, age, gender, facial hair, occlusion, indoor (outdoor), skin tone	Gender*age Gender*skin tone Indoor(outdoor)*yaw angle (Based on ROC)
Nagpal et al. (2019)	CMU Multi-PIE, MORPH, Racial Faces in the Wild (RFW)	Race, age	X
Vera-Rodriguez et al. (2019)	VGG Face2	Gender	X
Wu & Wang (2019)	CAS-PEAL	Age, gender	Gender*age (Based on ROC)
Albeiro et al. (2020a, 2020c)	MORPH, Notre Dame, Asian Faces Dataset	Gender, expression, head pose, forehead occlusion, facial makeup	X
Albeiro et al. (2020b)	MORPH	Age	X
Bruveris et al. (2020)	Selfie-doc (selfie photos vs ID doc)	Continent, country, gender	Gender*country (Based on ROC)
Krishnapriya et al. (2020)	VGG Face2, MORPH	Race, skin tone	Race*gender (Based on ROC)
Serna et al. (2020)	Labelled Faced in the Wild (LFW)	Gender, ethnicity	Gender*ethnicity (Based on Equal Error Rate)

In apposition, Bowyer and King (2019) stipulated that it is quite difficult to measure exactly the relationships by simply observing them from graphs. This is because ROC graphs sometimes give unclear trends. As such, it is difficult to conclude which variable would be affected by the presence of another variable. The same opinion was cited in (Lu et al., 2019). It is not an appropriate way to assess relationships across demographic cohorts since different groups should have different decision thresholds. When comparisons were made, inconsistent results had been produced.

2.5.3 Expression-Intensity Studies in the Psychology Domain

The discussion given earlier alluded to the fact that the interaction effects between expression and intensity does not generate much interest among researchers in machine-based facial recognition. The following paragraphs provide a brief discussion on expression-intensity interaction effects in the human-based facial recognition community as additional information. Most of the studies were carried out by researchers in the psychology domain.

According to Ekman (1999), past research has shown that it was not just the expression itself, but also its dynamics that were important when attempting to infer its meaning. The dynamics of a facial expression can be defined as the intensity of the action units. Gagliardi et al. (2003) was the first study that enabled the investigation of facial expressions at different levels of intensity. However, they included only four levels of intensity. Montagne et al. (2007) overcame this shortcoming by using a more dynamic approach, where instead of static pictures, they used short video clips of faces, that started from a neutral expression (0%), which then changed to 20%, to 30%, to 40%, to 50%, to 60%, to 70%, to 80%, to 90%, and finally to 100% intensity (Figure 2.7). This was done for each of the six basic expressions (anger, disgust, fear, happiness, sadness, and surprise), with a total of 58 subjects involved. Their focus was to study the difference in recognition performance among the six basic expressions, and the effects of the covariates on the performance using repeated measures of ANOVA. The results showed that happiness was the easiest emotion to recognise, whereas fear and sadness were the most difficult. Furthermore, there was a significant interaction effect between intensity and expression on the performance. Their findings showed that as the levels of intensity increase, there was an increase and a decrease in the mean performance of correctly recognised emotions depending on the types of expression.

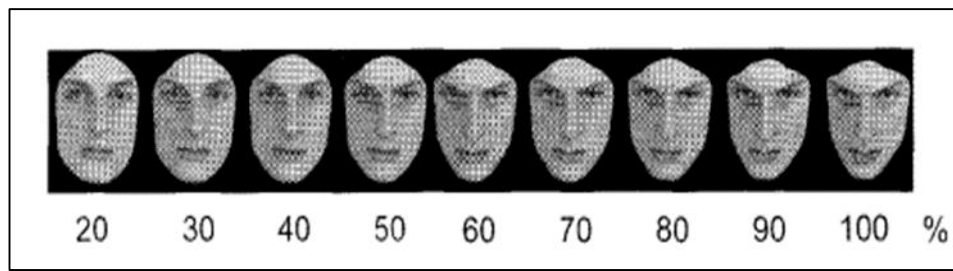


Figure 2.7: Example of Facial Emotion Intensity for Anger

Another study conducted by Hoffmann and colleagues (Hoffmann et al., 2010), and they reached the same conclusion. They used 42 images with an intensity of expression ranging from 40% (subtle cue) to 90% in steps of 10%, with the full-blown emotional picture (100%) for each image. This gave a total of 294 images used. A total of 186 students from the medical psychology course were the participants chosen to recognise those emotions. The analysis was carried out using a linear mixed model and the results showed that recognition accuracy was influenced by intensity. At the same time, there was an interaction between emotion and intensity levels (different intensities had different effects on emotion). Figure 2.8 shows an example of images with 0% intensity (neutral face) up to 100% intensity (full-blown emotional face). Images with 40% to 50% intensity are considered low intensity, 60% to 70% intensity are considered mid-intensity, and 80% to 100% are considered high intensity.

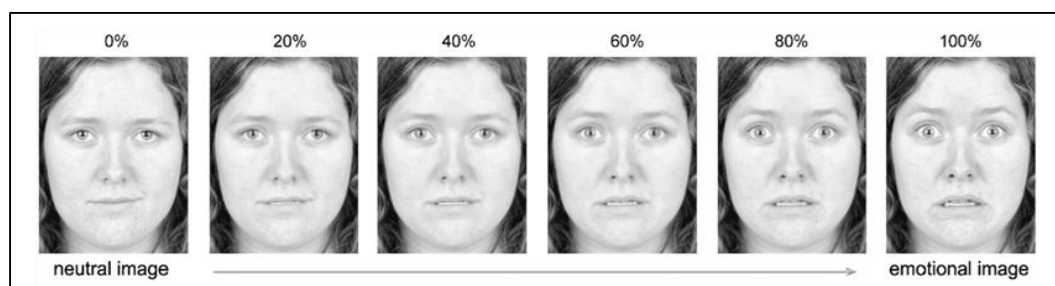


Figure 2.8: Neutral Face (0%) until Full-Blown Emotional Face (100%) Intensity

On a similar note, Wells et al. (2016) investigated the effects of expression, intensity, and gender on the speed, and accuracy of expression recognition. 39 undergraduate students from a university in the United Kingdom were asked to identify expressions from five female and five male faces displaying six basic expressions (anger, disgust, fear, happiness, sadness, surprise) and a neutral face, each with three levels of intensity: normal intensity (90%), moderate intensity (55%), and mild intensity (10%). A total of 180 face models were used. A series of ANOVAs and paired t-tests were carried out. They concluded that there was an interaction between expression and intensity, and that reaction time for expressions with moderate intensity showed the slowest response time. Generally, responses were rather slow in terms of speed, but accuracy was high for female faces.

The interaction of expression types and intensity levels also became an interesting subject in the fields of psychology and neuroscience. Calder et al. (1996), Adolphs and Tranel (2004), among others, have shown that varying the intensity of emotional expressions can make emotion recognition tasks more sensitive to subtle differences in the processing of different facial expressions. For example, varying the intensity of emotional expressions provides a more sensitive task for detecting possible impairments following amygdala (region of the brain primarily associated with emotional processes) damage.

Another research done by Lin et al. (2016) showed that the intensity of facial expressions controls amygdala activation. Event-related functional magnetic resonance imaging was used to assess brain activation while participants observed a static neutral expression with positive (happy) and negative (angry) expressions of either high or low intensity (Figure 2.9). They found that amygdala activation was highest for high intensity facial expressions as compared to low intensity facial expressions and neutral expression,

irrespective of valence (the pleasantness or unpleasantness of an emotional stimulus, from positive to negative or from happy to angry).

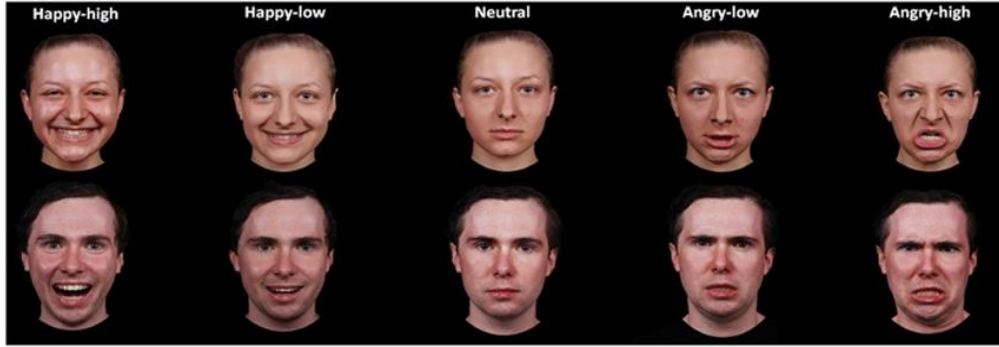


Figure 2.9: Facial Expressions in High and Low Intensity

2.6 Implemented Techniques and Models

The following sections provide a list of existing techniques and models that were used in this study.

2.6.1 Singular Value Decomposition

Singular value decomposition (SVD) is one of the most valuable theories in linear algebra. It can be viewed as a decomposition of a matrix $\mathbf{A}$ into a summation of rank-1 matrices that can be written as:

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \otimes (u_i, v_i), \quad \text{Equation 2.1}$$

where r is the rank of $\mathbf{A}$, u_i and v_i are the i^{th} columns of the orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$ respectively, and σ_i 's are the singular values (diagonal elements) of variance-covariance matrix of matrix $\mathbf{A}$. Another definition is given by the following theorem.

Theorem 2.1 (SVD) (Horn & Johnson, 2013): If $\mathbf{A}$ is an $(I_1 \times I_2)$ matrix with $\text{rank}(\mathbf{A}) = R$, its singular value decomposition can be written as:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{U} \text{ (basis matrix)} \mathbf{\Sigma} \mathbf{V}^T \text{ (coefficient matrix)}, \quad \text{Equation 2.2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices of dimensionality $(I_1 \times I_1)$ and $(I_2 \times I_2)$ respectively, $\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_n)$, $\mathbf{V} = (\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_n)$, and $\mathbf{\Sigma}$ is a $(I_1 \times I_2)$ diagonal matrix that contains the singular values $\sigma_1 \geq \sigma_2 \geq \dots \sigma_R > 0$. The parameters $\sigma_1, \dots, \sigma_R$ are the positive square roots of the decreasingly ordered nonzero eigenvalues of $\mathbf{A}\mathbf{A}^T$, which are the same as the decreasingly ordered nonzero eigenvalues of $\mathbf{A}^T\mathbf{A}$. The matrix $\mathbf{U}$ is also referred as the principal component directions of $\mathbf{A}$.

The theoretical efficacy of SVD allows many researchers to extend the two-way analysis into the multiway analysis. However, there are obviously many choices for tensor decompositions, which may lead to confusion about which model to choose for a particular application. According to Kolda and Bader (2009), there is no higher-order SVD that inherits all the properties (such as the unitary matrices of matrix $\mathbf{U}$ and $\mathbf{V}^T$ make SVD a suitable approach for discussing the geometry of n -space, and the diagonality of the variance-covariance matrix $\mathbf{\Sigma}$, makes it easier to determine the condition that could provide the best low-rank approximation to matrix $\mathbf{A}$) of the matrix SVD. However, two higher-order tensor decompositions that retain some of the nice properties of the matrix SVD are the CANDECOM/PARAFAC model (Carol & Chang, 1970; Harshman, 1970) and the Tucker model. Only the Tucker model is presented here since the CANDECOM/PARAFAC model is not related to this study.

2.6.2 Tucker Model

The Tucker decomposition was introduced by Ledyard R Tucker (Tucker, 1964, 1966). It generalises Equation 2.1 into a rank-1 expansion in the form of tensor $\mathcal{A}$:

$$= \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R \sigma_{pqr} \otimes (u_p, v_q, w_r) , \quad \text{Equation 2.3}$$

where $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$. The three factor matrices are $u_p \in \mathbb{R}^{I \times P}$, $v_q \in \mathbb{R}^{J \times Q}$, and $w_r \in \mathbb{R}^{K \times R}$.

The coefficients σ_{pqr} formed a core tensor, $\mathcal{D} \in \mathbb{R}^{P \times Q \times R}$. If P, Q, R are smaller than I, J, K , the core tensor $\mathcal{D}$ can be thought of as a compressed version of tensor $\mathcal{A}$. The elementwise of the core tensor (d_i) is obtained as:

$$d_i = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R \sigma_{pqr} u_{i_1 p} v_{i_2 q} w_{i_3 r} , \quad \text{Equation 2.4}$$

where $i = 1, 2, 3$ (tensor order three), and σ_{pqr} are the corresponding singular values on the diagonal of $\mathcal{D}$. The estimators for d_i are based on least square estimations if the non-zero eigenvalues are retained. However, if the truncation is based on the largest eigenvalues, the estimators are no longer based on the least squares (Kroonenberg, 1989).

The Tucker decomposition naturally generalises the orthonormal subspaces corresponding to the left or right matrices computed by the matrix SVD (Tucker, 1966; Kroonenberg, 1980). Compared to PARAFAC, the Tucker model is more general, and that the interpretation of the Tucker model is easier to handle compared to other models. It allows for a complete partitioning of the variability according to different modes. It is also an exploratory model that provides clues and trends that will help researchers come to a conclusion. Thus, this model is very useful for data reduction and data variability exploration

(Kroonenberg, 1980). On the other hand, the PARAFAC is more confirmatory. A confirmatory model is used to validate any ideas or theories that researchers have beforehand, using hypothesis testing. The special property contained in the Tucker model is that it allows for interactions among the modes or factors, and the strength of the interactions are represented by the core tensor values. Figure 2.10 shows an illustration of the Tucker decomposition for a third-order tensor (Cichocki et al., 2015).

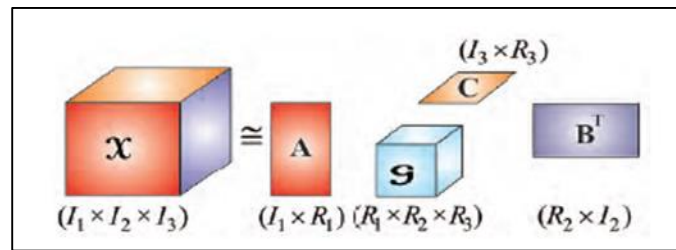


Figure 2.10: Tucker Decomposition for a Third-Order Tensor

Next, the Tucker decomposition can be written using the mode- n tensor-matrix multiplication as follows:

$$\mathcal{A} = \mathcal{D} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}. \quad \text{Equation 2.5}$$

The above decomposition is called the Higher-Order Singular Value Decomposition (HOSVD) as introduced by De Lathauwer et al. (2000a). The difference between the HOSVD and the original Tucker model is that, De Lathauwer et al. (2000a) included orthogonality constraints whereas the latter one did not put the assumption on the orthogonality. These would then result in the orthonormal matrices in Equation 2.5. Figure 2.11 shows an illustration of the HOSVD for a third-order tensor (Cichocki et al., 2016).

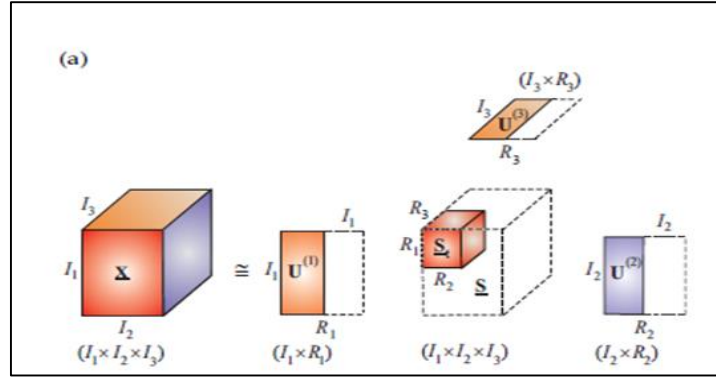


Figure 2.11: The HOSVD of a Third-Order Tensor

HOSVD is the most widely used formulation (Gnang et al., 2011) since this approach provides a natural way of dealing with images in the form of matrices, and the background factors can be easily separated. In computer vision, it has been used in Wang and Ahuja (2003) for analysis of face images. It has also been used in pattern recognition (Savas & Elden, 2007), data mining (Liu et al., 2005; Sun et al., 2005), signal processing (Muti & Bourennane, 2007; Haardt et al., 2008), and biomedical (Estienne et al., 2001; Omberg et al., 2007), among others.

2.6.3 Combination of PCA and ANOVA Model

This section is created to briefly explore the feasibility of the combination since this study uses multilinear PCA and an ANOVA model. It has been used and established in a variety of applications.

The idea of combining ANOVA with PCA is well described in (Läuter et al., 1996). They presented a general theory of how PCA, as a dimension reduction method, can promise a statistically correct analysis of ANOVA, since the test performs poorly for high dimensional data, especially for multivariate ANOVA. However, the idea of the combination was actually suggested back in 1958 by Calyampudi Radhakrishna Rao, an Indian-American

mathematician and statistician. He suggested the procedure in the context of analysing growth curves.

There are two procedures that can be implemented: PCA-ANOVA or ANOVA-PCA. The first one implements PCA at the first stage, followed by ANOVA, whereas the latter does the opposite. By applying them to sensory panel data, Luciano and Naes (2009) provided a good comparison between these two methods. They showed that both methods are suitable for assessing the relationships among factors, whether for the main effect factors or for interactions among factors. Both methods are easily computed using regular statistical software packages. Furthermore, neither needs any iteration or other complex numerical procedures.

In the context of face analysis, Givens et al. (2004) applied ANOVA in the first stage before logistic regression to examine what attributes make a subject easier or harder to recognise. They measured race, gender, age, glasses use, facial hair, bangs, mouth state, skin complexion, state of eyes, makeup use, and facial expression as the covariates. They made comparisons for three algorithms, which are PCA, Interpersonal Image Difference Classifier (IIDC), and Elastic Bunch Graph Matching (EBGM). It was found that some factors, such as changes to the eye status, would affect all algorithms in a similar manner, while race affected each algorithm differently.

Nowadays, the procedure has been implemented in various fields of studies (Volfovsky & Hoff, 2014; Fritzsche et al., 2018). Fritzsche et al. (2018), for example, applied ANOVA-PCA to a two-dimensional infrared spectroscopy (2D-IR) dataset. They demonstrated the ability to efficiently retrieve the base composition of a DNA sequence and discriminate ligand-DNA complexes from unbound sequences using this method.

Kiers and Van Mechelen (2001) provided a step-by-step guidance in analysing three-mode data and gave readers a clearer concept in applying ANOVA-multilinear PCA.

2.7 Discussion and Conclusions

This chapter covers the background studies on the centring approach as well as the relationships among the covariates. It demonstrates that, despite the fact that mean centring has been used in multiple studies, there is a drawback to using it. As a result, the research's first significant contribution is to investigate the advantages of using the median centring technique. ROC graphs have revealed in the past that the interaction effects among the face variants could not be reliably stated. The second contribution of this thesis will thus be to use the ANOVA model to investigate any interaction effects that may exist among the face variants.

The capability of the tensor model in isolating or separating the face variations makes this approach different from traditional PCA model. A lot of studies have shown the advantages of using a tensor approach in solving problems, whether in reducing the highly dimensional data, overcoming incomplete data, solving multimodal problems, as well as predicting the relative relationships between sources of variation. All of these advantages have increased the potential of tensors to improve their performance, not only in facial recognition systems, but also in other fields of study.

In addition, most recent studies show that researchers are more interested in multiway analysis than the one-way and two-way analysis (Cichocki et al., 2015; Cichocki et al., 2016; Lu et al., 2017; Ouamane et al., 2017; Ozdemir et al., 2018; Zare et al., 2018; Cyganek, 2019). Good overview articles, as an introduction for researchers in learning tensors, can be found in (Acar & Yener, 2009; Kolda & Bader, 2009; Zare et al., 2018), while (Sidiropoulos

et al., 2017), among others, provide more details for practitioners and programmers. This chapter also listed out the other existing models and techniques that are used in this study.

CHAPTER 3

MEDIAN CENTRING AND ESTABLISHMENT OF INTERACTION EFFECTS

3.1 Overview

The research methodology is presented in this chapter. In Section 3.2, the general research workflow is provided, and in Section 3.3, the research framework is presented. The median centring technique and the formation of interaction effects using the tensor-ANOVA model are described in Section 3.4 – these are the two key contributions of this research. The experimental setup, including the system specifications and dataset used, is described in Section 3.5. The evaluation presented in Section 3.6, as well as a discussion and conclusions in Section 3.7, bring this chapter to a close.

3.2 Research Workflow

The tensor model is a very useful approach for representing multidimensional data. It enables the separation of various variables or causal factors. Given these, many statistical methods can be applied for further analysis, resulting in a more accurate 3D image analysis. The workflow of this study is depicted in Figure 3.1.

3.3 Research Framework

The first section of this study is concerned with the pre-processing of the input data. It is one of the most important procedures in the early stages of multiway analysis. This procedure is carried out to ensure that the input data is suitable for the multilinear SVD or HOSVD model (or multilinear PCA). Two types of centring techniques (mean and median) are applied, as well as a non-centred approach for comparison purposes.

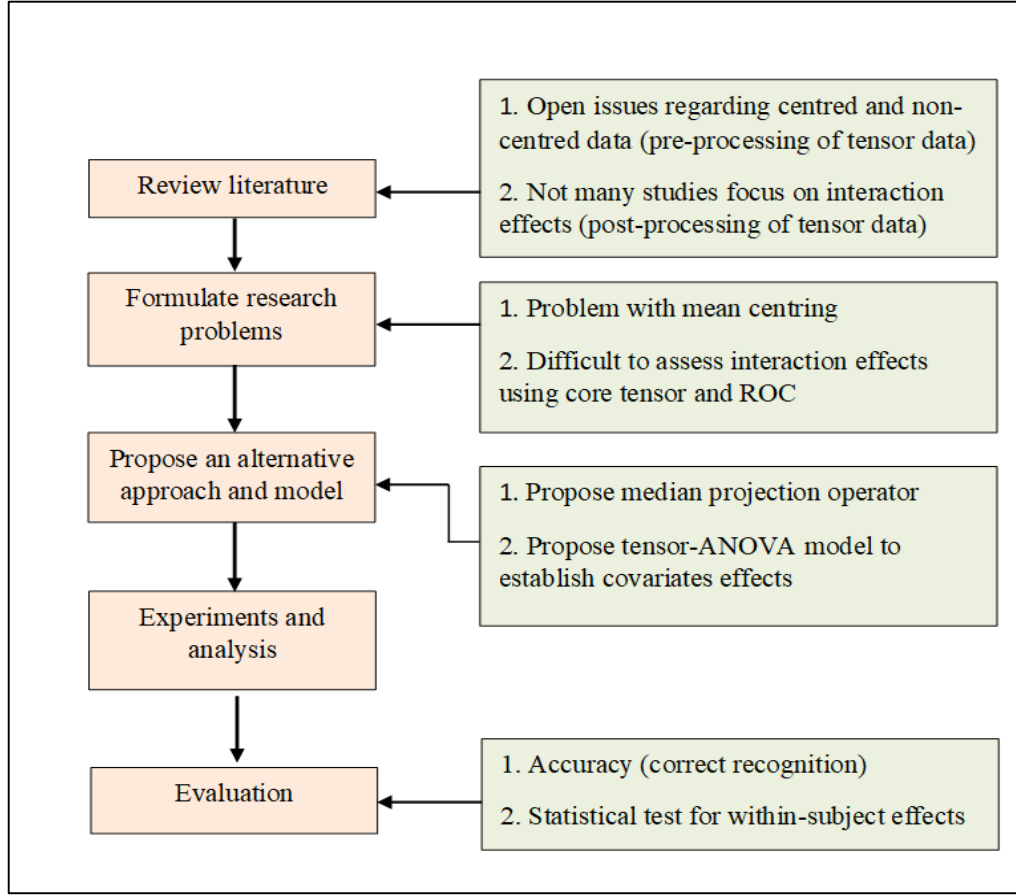


Figure 3.1: Research Workflow

The N^{th} -order tensor, as discussed in Chapter 2, can be decomposed into N modes of matrices $\mathbf{U}$ ($\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_N$). The matrices $\mathbf{U}$ represent the underlying factors, and each matrix $\mathbf{U}$ is an orthogonal matrix. In this study, the $\mathbf{U}$ matrices ($\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4$) represent subjects (*sub*), expression types (*exp*), expression strengths (*int*), and points (*pts*), as the four measurement variables. A third-order tensor, $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ for $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$, and a fourth-order tensor, $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ for $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$, and $\mathbf{U}_4$, are defined.

The multilinear decomposition of a third-order tensor $\mathcal{A}$ is given as:

$$\mathcal{A} = \mathcal{D} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 , \quad \text{Equation 3.1}$$

where $\mathbf{U}_1$ is the mode-1 matrix that spans the first variable space, $\mathbf{U}_2$ is the mode-2 matrix that spans the second variable space, $\mathbf{U}_3$ is the mode-3 matrix that spans the third variable space. All the three matrices are obtained by performing SVD on each flattening matrix $(I_1 \times I_2 I_3)$, $(I_2 \times I_3 I_1)$, and $(I_3 \times I_2 I_1)$ (De Lathauwer et al., 2000a). The same procedures are applied for the fourth-order tensor, depending on its respective modes.

Then, a face recognition system is built to evaluate the performance of mean-centred, median-centred, and non-centred approaches. A comparison is also made between the tensor model and the PCA model.

The second section of this study focuses on the effects of expression types and expression strengths on recognition performance. Specifically, the interaction effects. Figure 3.2 presents the research pipeline, highlighting the two main contributions of this study whereas Figure 3.3 depicts the overall research framework.

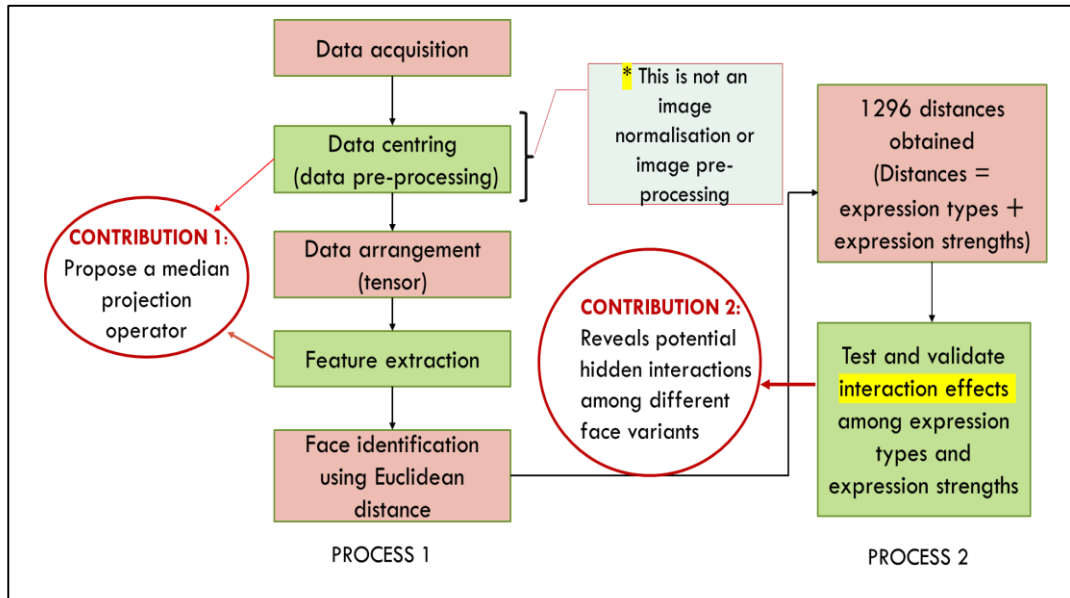


Figure 3.2: Research Pipeline

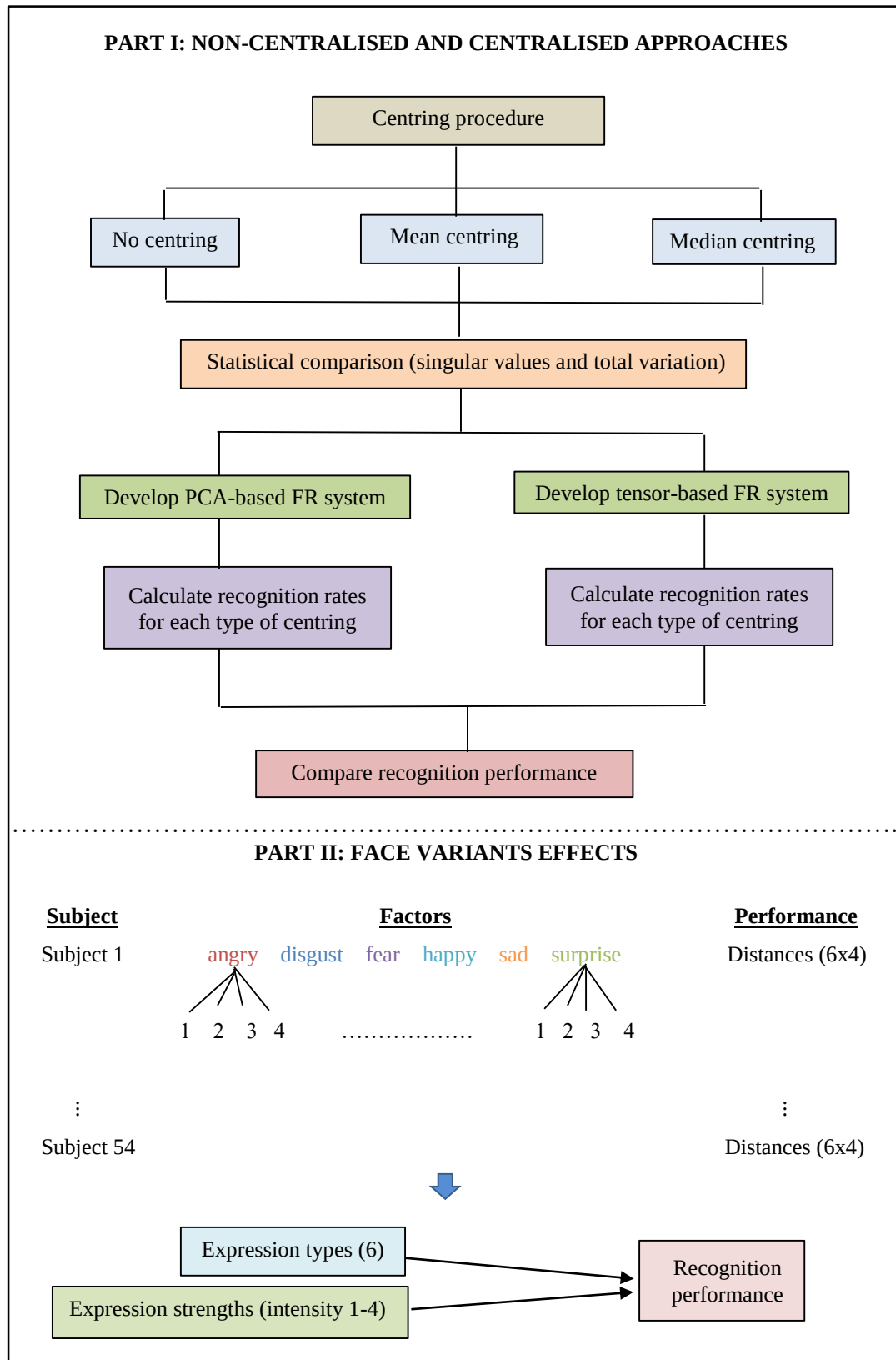


Figure 3.3: Research Framework

3.4 Proposed Approach and Model

The following sections explain the proposed median data centring approach and the proposed hybrid tensor-ANOVA model.

3.4.1 Median Projection Operator

Figure 3.4 shows the workflow of the centring procedure. The comparison is made based on two statistical values. They are singular values, σ_i (variance that each principal component carries) and total variations ($\sum \sigma_i$), carried by the final reduced model. The best approach is determined by the highest value of σ_i and the highest amount of $\sum \sigma_i$. Apart from that, the best approach is also determined by the total number of correctly recognised faces. A tensor-based face recognition system is built to compare the performance between centred tensor data and non-centred tensor data, as shown in Figure 3.4.

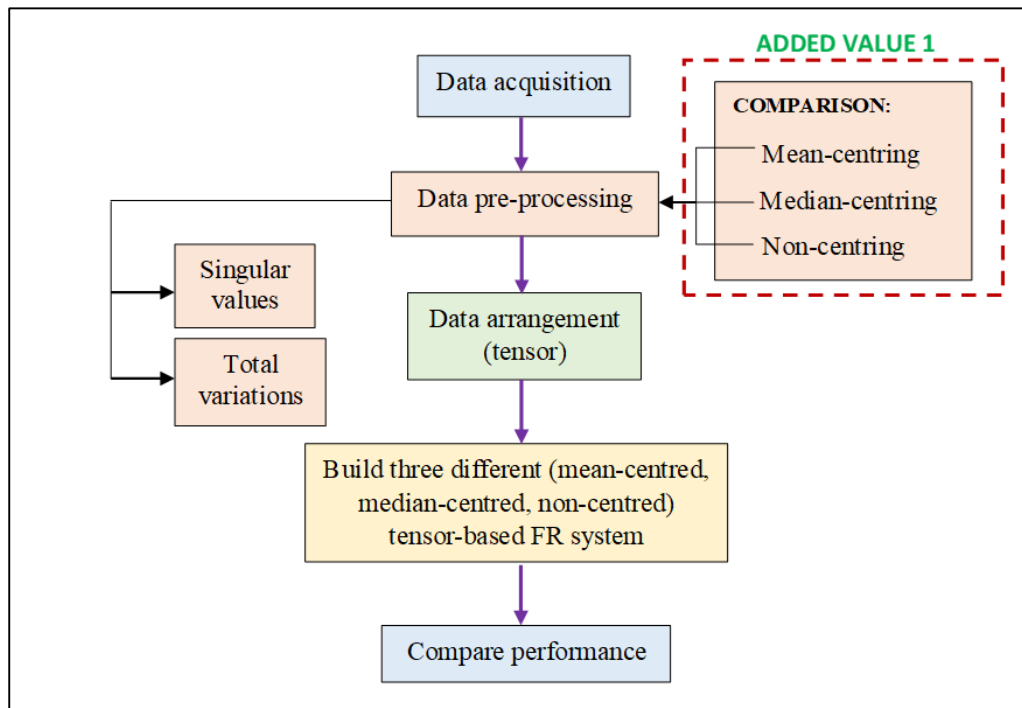


Figure 3.4: Centred and Non-Centred Face Recognition System

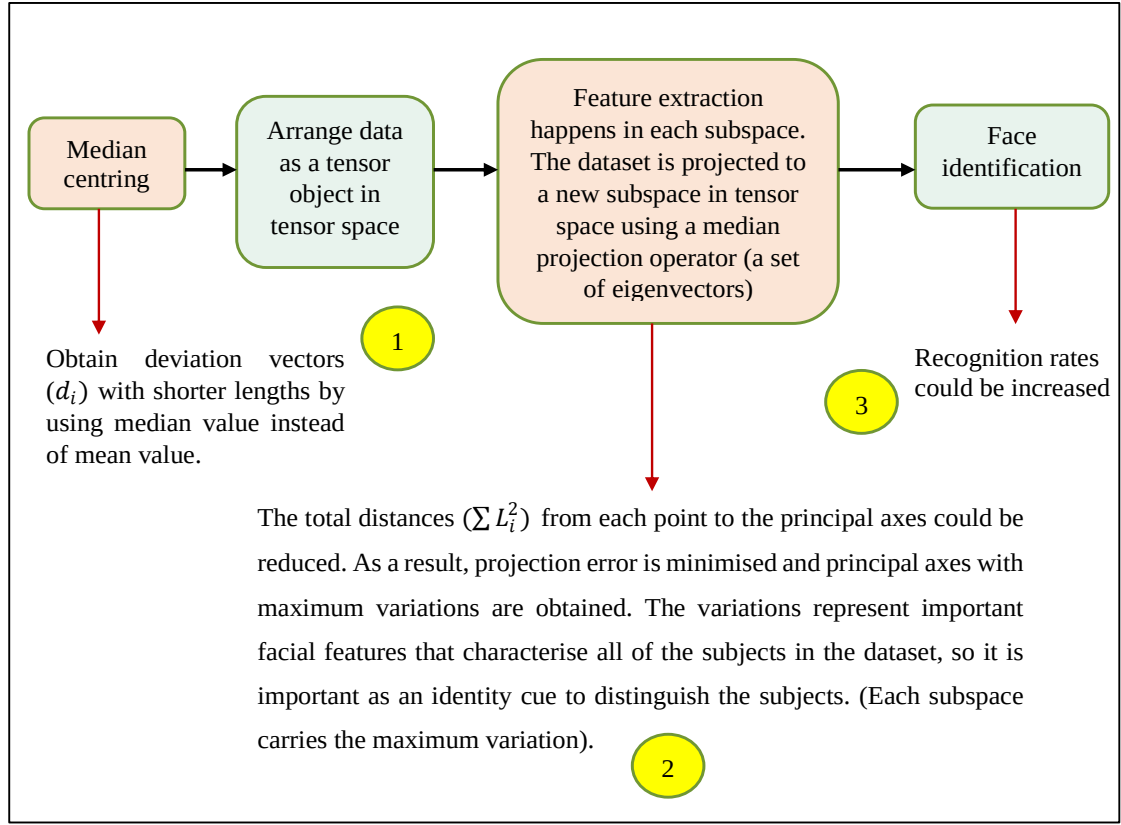


Figure 3.5: The Pipeline of the Proposed Median Centring

Figure 3.5 depicts the pipeline of the proposed median projection operator. It shows how the proposed median projection operator could increase the total variations of the tensor model, resulting in an increment in recognition rates. In the first stage, median centring yields a shorter deviation vector because it is unaffected by outliers as compared to mean centring. After the median-centred data is decomposed into a tensor model (tensorface), feature extraction happens in each subspace of the tensor model. At this stage (stage two), the median projection operator reduces the total distances from each point to the principal components, minimising the projection error, and principal components with maximum variations are obtained. The maximum variation in each subspace of the tensor model could help in increasing the total number of correct recognitions.

Table 3.1 describes the differences between the proposed median projection operator and the traditional PCA projection operator. The projection operator projects any new face onto a new reduced-dimensional space of image (face) coefficients.

Table 3.1: The Differences between Median and PCA Projection Operators

Projection operator for PCA model ($\mathbf{B}^{-1}$)	Median projection operator for tensor model ($\mathbf{B}_{e(s)}^{-1}$)
A data matrix $\mathbf{A}$ of training face models is decomposed into a set of eigenvectors (a reduced-dimensional basis matrix, $\mathbf{B}$) and a matrix $\mathbf{C}$ containing a vector of coefficients $\mathbf{c}_a$ associated with each vectorised image $\mathbf{a}_a$.	A basis tensor, $\mathcal{B}$ is constructed from a matrix $\mathbf{U}_{subject}$. The tensor $\mathcal{B}$ is then indexed for a particular expression (exp) to obtain a sub-tensor $\mathcal{B}_{exp}$ with a dimension of (54x1x15270). $\mathcal{B}_{exp}$ is then flattened along the subject (sub) mode to obtain a matrix $\mathbf{B}_{exp(sub)}$ with a dimension of (54x15270).
Any training face model can be written as $\mathbf{a}_a = \mathbf{B}\mathbf{c}_a$ which yields $\mathbf{c}_a = \mathbf{B}^{-1} \mathbf{a}_a$.	Any training face model $\mathbf{a}_a$ of subject 's' in expression 'e' can be written as $\mathbf{a}_{s,e} = \mathbf{B}_{e(s)}\mathbf{c}_s$ which yields $\mathbf{c}_s = \mathbf{B}_{e(s)}^{-1} \mathbf{a}_{s,e}$.
In this case, the projection operator ($\mathbf{B}^{-1}$) considers variations across all face models.	In this case, the projection operator ($\mathbf{B}_{e(s)}^{-1}$) considers variations in each mode of a tensor model since it is constructed based on different eigenvectors in each mode.

3.4.2 Tensor-ANOVA Model

Dimension reduction on each mode of a tensor gives a different amount of total energy (variation), which leads to a problem in assessing the covariate relationships. This study uses within-subject variations criterion in repeated measures of ANOVA to establish the interaction effects among the face covariates. As a result, the effect of subject characteristics (expression variability) on algorithm performance (distance) is quantifiable.

ANOVA is a statistical hypothesis test that can be used to investigate the interactions that may exist among different types of variations in the data. This can be achieved by testing the hypothesis of different means among three or more groups or levels. Groups or levels are different groups within the same independent variable. The term variance is used since the calculation of the test statistic is based on the variations that exist among and within groups of study. In other words, the difference in means between groups is tested by evaluating the variability in the data.

Figure 3.6 shows the workflow of the proposed tensor-ANOVA model, to establish the interaction effects. Initially, the median centring technique is applied. Then, the median-centred input data is arranged as a fourth-order tensor because another variable, *expression strengths*, is added. Correct identification is determined by the minimum Euclidean distance. As a result, a total of 1296 distance values (*54 subjects x six types of expressions x four levels of intensity*) are obtained. The *distance* variable is treated as a dependent variable in the ANOVA model, as shown in Figure 3.7.

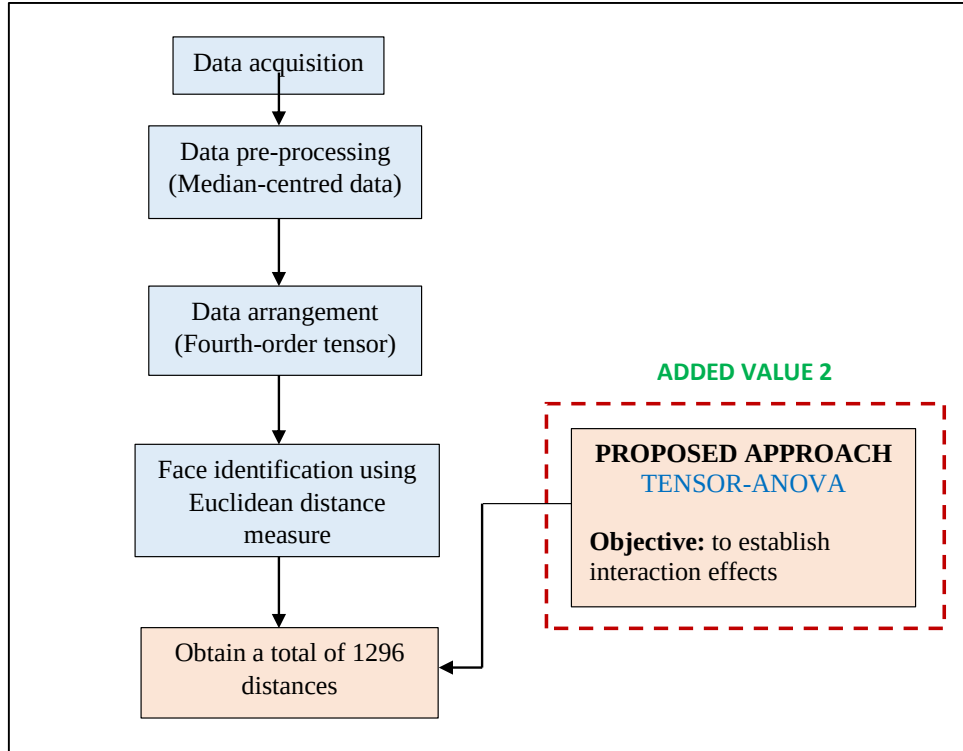


Figure 3.6: The Workflow of the Proposed Tensor-ANOVA Model

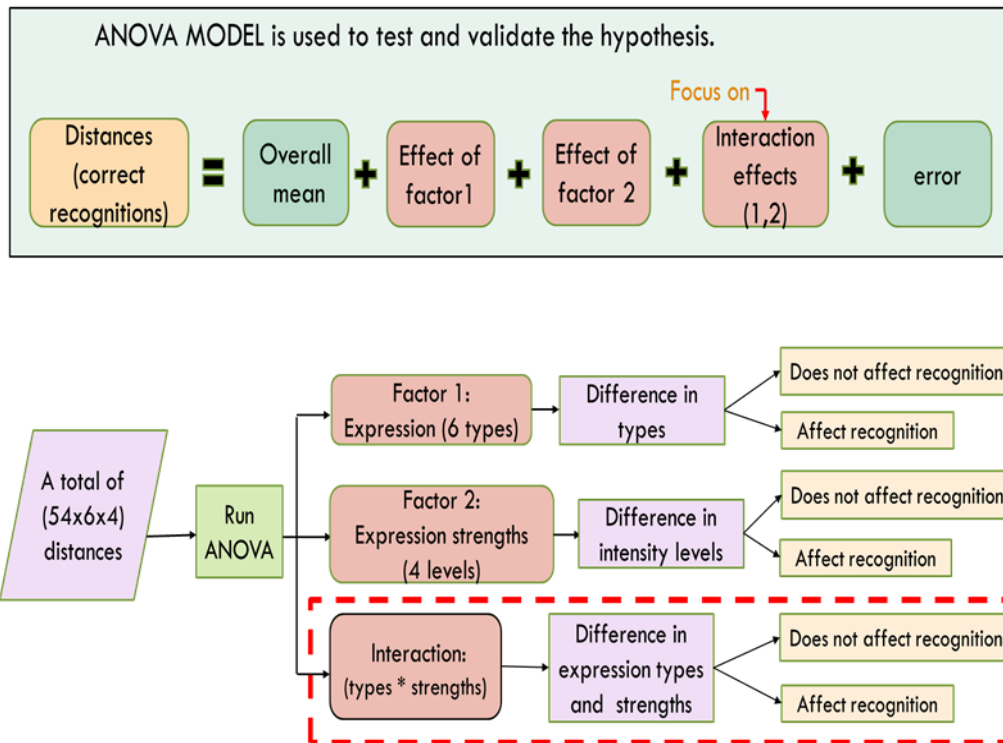


Figure 3.7: Establishment of the Interaction Effects in the Tensor-ANOVA Model

In Figure 3.7, the minimum distance is treated as a dependent variable. The independent variables are *expression types* (factor one) with six different levels, and *expression strengths* (factor two) with four different levels. Two types of covariate effects are obtained, namely the main effects (effects of factors one and two) and the interaction effects. Since this research focuses on the tensor approach and the establishment of the interaction effects, a statistical test is used to test the hypothesis that the interaction effects do exist and that they do affect recognition performance. In other words, the covariate effects can be established.

3.5 Experimental Setup and Dataset

All experiments were carried out using Matlab version 2016a, a multi-paradigm numerical computing environment developed by Mathworks and Tensor Toolbox version 2.6 (Bader & Kolda, 2015), developed by Sandia National Laboratories. The Tensor Toolbox provides several classes, which use Matlab's object-oriented features (Bader & Kolda, 2006). These classes are used for manipulating dense, sparse and structured tensors. Apart from that, statistical analysis for the proposed tensor-ANOVA model was carried out using SPSS (Statistical Product and Service Solutions alias Statistical Package for the Social Sciences).

This experiment used the Binghamton Faces dataset BU-3DFE (Yin et al., 2006), which consists of 54 three-dimensional face surfaces. As mentioned in (Smeets et al., 2012), the BU-3DFE database is the most appropriate database to evaluate algorithms for robustness against expression variations.

There are six different types of facial expressions in the dataset, namely angry, disgust, fear, happiness, sadness, and surprise. Each expression has four levels of intensity (or strength of expression) named level one (low), level two (middle), level three (high), and level four (highest). Therefore, a total of 1296 3D facial expression models are used. All of the faces have gone through the image pre-processing stage. They are cleaned face surfaces with 5090 mesh points (x, y, z) for each face surface.

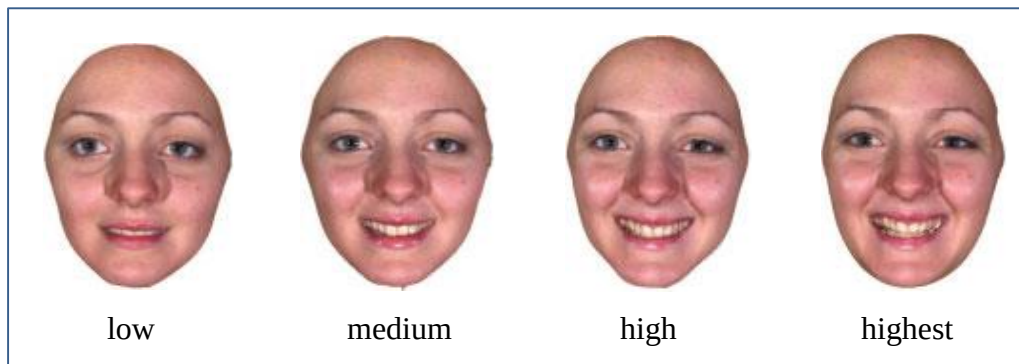


Figure 3.8: Example of Expression with Four Levels of Intensity

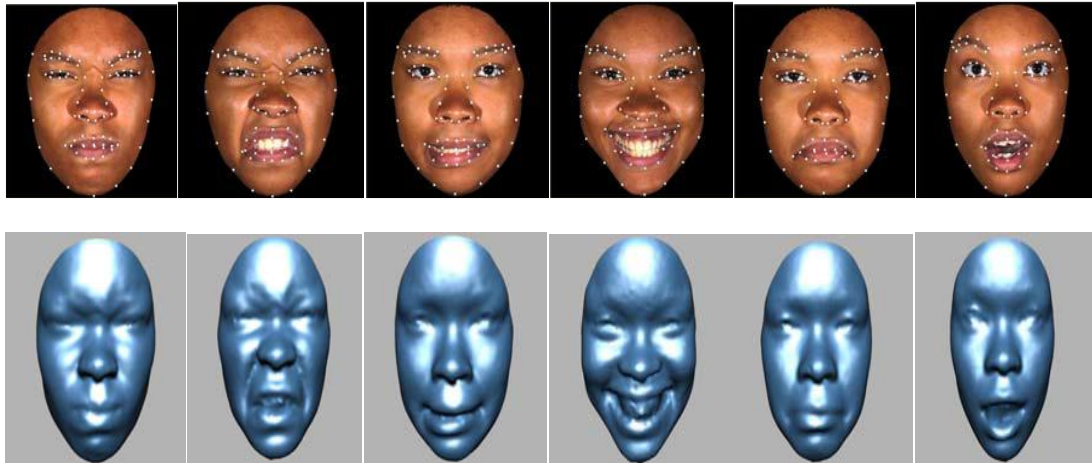


Figure 3.9: Example of a Subject Showing Six Different Expressions

Figure 3.8 and Figure 3.9 show examples of images (and 3D face models) from the BU-3DFE database. Figure 3.8 shows an example of 3D textured face scans of a sample subject showing happiness expressions at four levels of intensity, from the lowest to the highest. Then, the first row in Figure 3.9 shows an example of a subject showing six different expressions (angry, disgust, fear, happiness, sadness, and surprise) with chosen feature points, while the second row represents the cropped shape models for each expression (Yin et al., 2006).

3.6 Evaluation Approaches

The methodologies to be used to confirm the efficacy of the suggested median centring approach, as well as the establishment of interaction effects, are presented in the following sub-sections.

3.6.1 Classification Accuracy

This section begins with a basic introduction to binary classification. Each instance is formally assigned to one of two class labels: positive or negative. The mapping mechanism from instances to expected classes is known as a classification process. Several classification algorithms generate a continuous output (the probability value of an instance's class membership), which can be used to predict class membership using various thresholds. Other models produce a discrete class label or value that merely indicates the instance's expected class (Fawcett, 2006).

Table 3.2 shows the confusion matrix obtained. There are four possible outcomes that could be obtained. If an instance is positive and it is classified as positive, it is counted as a true positive (correctly classified). The same goes for negative condition. If an instance is negative and it is classified as negative, it is counted as a true negative (correctly

classified). However, if the instance is positive but it is classified as negative, it is counted as a false negative. If the instance is negative and it is classified as positive, it is counted as a false positive. Both conditions are considered as wrongly classified cases.

Table 3.2: Possible Outcomes for a Binary Classification

PREDICTED CLASS			
		Yes	No
ACTUAL CLASS	Yes	True Positive (TP)	False Positive (FP)
	No	False Negative (FN)	True Negative (TN)

There are several common performance metrics that can be obtained from the confusion matrix, which are as listed below:

- i. True positive rate = positives correctly classified / total positives = TP/P
- ii. False positive rate = negatives incorrectly classified / total negatives = TN/N
- iii. Accuracy = $TP + TN / P+N$
- iv. Precision = $TP / TP+FP$ (positive predictive value)
- v. Recall = TP / P (sensitivity)
- vi. Specificity = $TN / FP + TN$

3.6.2 Leave-One-Out Cross-Validation

Other than the traditional tenfold cross-validation, leave-one-out cross-validation is one of the estimate approaches. It is just n -fold cross-validation, with n denoting the number of cases (or a subset) in the data set (Witten & Frank, 2005). The remaining samples comprise the training set, and each subgroup must be left out once. The testing set is used to determine its correctness.

The advantage of this method is that it assures the highest possible amount of data to be represented in the training set. This leads to a more accurate classification result. Furthermore, there is no randomness (random sampling) involved in this procedure since each subject will be chosen as a member of the subset. This will prevent an unbiased representation of a group.

3.6.3 Statistical Test for Within-Subject Effects

In a standard ANOVA, the null hypothesis states that there is no effect of different factors or independent variables on the dependent variable. It can be written as follows:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k \quad \text{Equation 3.2}$$

$$H_1: \text{Not all } \mu_i \text{ are equal ,}$$

where μ_i is the mean value and k is the number of independent comparison groups. Therefore, rejecting the null hypothesis will lead to an interesting observation on which factors influence the dependent variable (the dependent variable is a numerical value, whereas the independent variable is commonly categorical).

There is one independent variable in a one-way ANOVA. On the contrary, there are two independent variables in the two-way ANOVA. It is an extension of one-way ANOVA.

Therefore, the objective of the analysis is not only to examine the effect of two categorical independent variables on the dependent variable, but also to assess the interaction that may exist between those two independent variables.

A repeated measure ANOVA is another statistical technique in the ANOVA class. It is used for repeated measures design of experiments. Repeated measures design means the measurements are taken on the same subjects for all different conditions in the experiment. At least two levels must exist in each factor, and these levels are dependent. In this analysis, the repeated-measurements of the independent variables fall under within-subject factors.

Subjects or respondents are not considered as one of the factors to be investigated in repeated measures design. Subjects are considered as experimental blocks since the model assesses how a subject responds to all the treatments. In other words, the subject itself has its own control in the study. This gives an advantage because only the variability within subjects is included in the error term, causing a smaller error term and hence a more powerful analysis.

The test statistic for testing the null hypothesis is the **F**-test, given as:

$$F = \frac{\sum_i n_i (\bar{Y}_i - \bar{Y})^2 / (k - 1)}{\sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2 / (N - k)} , \quad \text{Equation 3.3}$$

where n_i is the sample size of the i^{th} group (if there are four levels, $i = 1, 2, 3, 4$), $\bar{Y}_i$ is the sample mean in the i^{th} group, $\bar{Y}$ is the overall mean, k represents the number of comparison levels ($i = k$), and N represents the total number of observations in the analysis.

This **F** statistic is computed by taking the ratio of between-group variability to the within-group (residual) variability. The numerator captures between treatment variability (differences among the sample means) and the denominator contains an estimate of the variability in the outcome.

The ANOVA hypothesis testing can be concluded using two approaches. These two approaches will determine whether to accept or reject the null hypothesis. However, both approaches will give the same result. The first approach is based on the comparison between **F** calculated (Equation 3.3) and the **F** value from the statistical table (known as **F** critical). The null hypothesis is rejected if **F** calculated is greater than **F** critical ($F_{\text{calculated}} > F_{\text{critical}}$).

The second method is a computer-based method that calculates the probability (*p-value*) of a value of **F** greater than or equal to the observed value. The null hypothesis is rejected if this probability value is less than or equal to the chosen significance level of alpha (α).

ANOVA has assumptions that must be met to have a valid model. This is important to make sure that any conclusion made can be trusted. The following assumptions are made when using repeated-measures ANOVA.

- i. The response variable is continuous.
- ii. The residual is normally distributed.
- iii. The subjects are independent.
- iv. The within-subject covariance matrices are equal for all between-subject groups (omitted since between-subject groups are not considered).

- v. All within-subject variance-covariance matrices are circular. Circularity means the variances of differences between any two measurements within a subject are constant. It can be tested using Mauchly's test statistic W , namely Sphericity testing and epsilon value, ϵ . W and ϵ range from 1 to 0. The smaller the values, the greater the departure from sphericity and circularity.

3.7 Discussion and Conclusions

The research workflow was described to give a general summary of the study. After that, the tensorface framework was introduced. It was necessary to describe the framework's unique features in order to emphasise the advantages of utilising a tensor model to represent the faces dataset, or any other multidimensional dataset.

The proposed approach and model were also thoroughly detailed in order to provide a clearer picture of the study's two primary contributions: median centring and the formation of interaction effects using the tensor-ANOVA model.

In addition, the experimental setup was shown in order to provide details on the platform employed. A description of the dataset was also supplied to illustrate why the Binghamton Face dataset is acceptable for the tensor model. Finally, the evaluation procedure is critical and must be completed in order to confirm the efficacy of the given approach and model.

CHAPTER 4

TESTING AND IMPLEMENTATION

4.1 Overview

This chapter discusses the research work's experiments and implementation. Different centring strategies (mean, median, and non-centring) are used in Section 4.2, and the results are reviewed. A face recognition system is also constructed to examine the effects of centring on recognition performance between the commonly used PCA-based system and a tensor-based system.

A tensor model organises a face dataset into multiple categories of variation, with each dimension reflecting one variation type. It allows researchers to handle inter-subject and intra-subject differences across multiple dimensions at the same time. This advantage leads to the second level of the investigation. The effects of each type of variation on recognition performance are discussed in Section 4.3, with a focus on the interaction effects that may arise among face variants.

4.2 Experiment 1 (Centring)

One-way centring can be categorised as centred by row or centred by column (Orlóci, 1967). If the data matrix is centred by row, the total for the elements in each row is nullified ($\sum_{j=1}^p a_{ij} = 0$). On the other hand, if the data matrix is centred by column, the total for the elements in each column is nullified ($\sum_{i=1}^t b_{ij} = 0$). The choice is determined by how the dataset is organised. The centring by row (centring by variables) technique is used in this study. The following important statistical properties are provided as proof or reference for analysing the results obtained, with a focus on singular values and total variation.

4.2.1 Important Statistical Properties

PCA is an exploratory data analysis method that is used to capture variations in a dataset with p variables and n individuals (more details can be found in Johnson and Wichern (2008), and Jolliffe (2002)). This dataset represents n vectors of $(x_1, x_2, \dots, x_p)$, or an $n \times p$ matrix $\mathbf{X}$. The purpose of PCA is to seek a linear combination of the p variables given by:

$$\sum_{i=1}^p \mathbf{a}_i \mathbf{X}_i = \mathbf{a}' \mathbf{X} , \quad \text{Equation 4.1}$$

where $\mathbf{a}$ is a vector of constants $a_1, a_2, \dots, a_p$. These linear combinations represent a new coordinate system, where each linear combination of $(\mathbf{a}' \mathbf{X})$ is named as the principal component (PC). The first PC carries the most variability in the data, followed by the remaining components. The variance of any PC is given by:

$$\text{var}(\mathbf{a}' \mathbf{X}) = \mathbf{a}' \mathbf{S} \mathbf{a} , \quad \text{Equation 4.2}$$

where $\mathbf{S}$ is the value of the sample covariance matrix. A more specific explanation is given below:

Let $\mathbf{S}$ be the sample covariance matrix with eigenvalue-eigenvector pairs $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$, the i^{th} sample principal component is given by:

$$y_i = \mathbf{e}_i \mathbf{x} = e_{i1}x_1 + e_{i2}x_2 + \dots + e_{ip}x_p, \quad i = 1, 2, \dots, p, \quad \text{Equation 4.3}$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ and $\mathbf{x}$ is any observation on the variables $X_1, X_2, \dots, X_p$.

Also,

$$\text{sample variance}(y_i) = \lambda_i, \quad i = 1, 2, \dots, p, \quad \text{Equation 4.4}$$

$$\text{total sample variance} = \sum_{i=1}^p s_{ii} = \lambda_1 + \lambda_2 + \dots + \lambda_p. \quad \text{Equation 4.5}$$

From the given explanation, it shows that $\mathbf{a}$ in Equation 4.1 must be eigenvectors ($\mathbf{e}_i$), as given by Equation 4.3. In addition, an eigenvalue (λ) is an important value because it represents the variance of the PC, as given by Equation 4.4. The total variance of the p original variables can be represented by the sum of the eigenvalues of all PCs, as given by Equation 4.5. Then, the standard calculation to measure the proportion of the total variation that can be accounted for by the first q retained PCs (where $q < p$) is given by:

$$\text{Proportion of total variation} = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_q}{\lambda_1 + \lambda_2 + \dots + \lambda_q + \dots + \lambda_p}, \quad q < p. \quad \text{Equation 4.6}$$

The following statistical theorem, corollary, and proposition are important in analysing the results obtained from uncentred and centred procedures (Takane & Shibayama, 1991; Cadima & Jolliffe, 2009).

Theorem 4.1 (Yanai and Takeuchi Separation Theorem): *Let $\mathbf{Y}$ be an $n \times p$ matrix. Let $\mathbf{P}_n$ and $\mathbf{P}_p$ be, respectively, $n \times n$ and $p \times p$ orthogonal projection matrices of rank r and s . Then,*

$$\sigma_{j+t}(\mathbf{Y}) \leq \sigma_j(\mathbf{P}_n \mathbf{Y} \mathbf{P}_p) \leq \sigma_j(\mathbf{Y}), \quad (4.1)$$

where $\sigma_{j(\cdot)}$ indicates the j^{th} largest singular value of the matrix and $t = n + p - (r+s)$.

Corollary (4.1): Let $\mathbf{S}$ be the covariance matrix for a given data set, and $\mathbf{T}$ is its corresponding matrix of non-central second moments. Let $\lambda_j(\cdot)$ be the j^{th} largest eigenvalue of either matrix, and $\pi_j(\cdot) = \frac{\lambda_j(\cdot)}{\sum_{i=1}^p \lambda_i(\cdot)}$ the proportion of total variation accounted for by the corresponding PC. Then $\lambda_{j+1}(\mathbf{T}) \leq \lambda_j(\mathbf{S}) \leq \lambda_j(\mathbf{T})$. Equivalently, $\pi_{j+1}(\mathbf{T}) \cdot \alpha \leq \pi_j(\mathbf{S}) \leq \pi_j(\mathbf{T}) \cdot \alpha$ where $\alpha = \frac{\text{tr}(\mathbf{T})}{\text{tr}(\mathbf{S})} = \frac{1}{1 - \|\mathbf{c}_m\|^2 / \text{tr}(\mathbf{T})}$, with $\text{tr}(\cdot)$ is the matrix trace, $\mathbf{c}_m = \frac{1}{n} \mathbf{X}_{uc}^t \mathbf{1}_n$ and

$\|\mathbf{c}_m\|^2 = \mathbf{c}_m^t \mathbf{c}_m$. The p -dimensional vector $\mathbf{c}_m$ is the vector of column means of the uncentred data matrix $\mathbf{X}_{uc}$ and $\|\mathbf{c}_m\|^2$ is the sum of squares of those column means.

The implication of **Corollary (4.1)** is that an eigenvalue obtained from the mean-centred procedure cannot be larger than the eigenvalue obtained from the uncentred procedure, with the same rank, and it also cannot be smaller than the uncentred's next ranking eigenvalue (given by: $\lambda_{j+1}(\mathbf{T}) \leq \lambda_j(\mathbf{S}) \leq \lambda_j(\mathbf{T})$). However, in terms of total variation that can be accounted for by j^{th} PCs, the mean-centred procedure can give either a larger or smaller value than the uncentred procedure, of equal rank (given by: $\pi_{j+1}(\mathbf{T}) \cdot \alpha \leq \pi_j(\mathbf{S}) \leq \pi_j(\mathbf{T}) \cdot \alpha$). This is because α represents the trace value of both covariance matrices of uncentred and mean-centred, which is always greater than or equal to 1.

Proposition (4.1): Let $\mathbf{X}_{uc}$ be an $n \times p$ data matrix, whose vector of column means is $\mathbf{c}_m = \frac{1}{n} \mathbf{X}_{uc}^t \mathbf{1}_n$ and whose matrix of non-central second moments is $\mathbf{T} = \frac{1}{n} \mathbf{X}_{uc}^t \mathbf{X}_{uc}$. Let $\mathbf{S}$ be the covariance matrix of the data, and $\mathbf{X}_{cc}$ be the column-centred data matrix so that $\mathbf{S} = \frac{1}{n} \mathbf{X}_{cc}^t \mathbf{X}_{cc}$. Let $\mathbf{S}$ and $\mathbf{T}$ have eigenvalues $\{\lambda_i^{[c]}\}_{i=1}^p$ and $\{\lambda_j^{[u]}\}_{j=1}^p$, and eigenvectors $\{\mathbf{w}_i^{[c]}\}_{i=1}^p$ and $\{\mathbf{w}_j^{[u]}\}_{j=1}^p$, respectively. Let $\boldsymbol{\eta}_i^{[c]}$ denote the i^{th} column-centred PC and $\boldsymbol{\eta}_j^{[u]}$ is its j^{th} uncentred counterpart. Then,

i. The mean score on the j^{th} uncentred principal component is $\bar{s}_j =$

$\|\mathbf{c}_m\| \cdot \cos(\mathbf{w}_j^{[u]}, \mathbf{c}_m)$. Equivalently,

$$\cos(\mathbf{w}_j^{[u]}, \mathbf{c}_m) = \bar{s}_j / \|\mathbf{c}_m\|. \quad \text{Equation 4.7}$$

- ii. The variance of the scores on the j^{th} uncentred principal component is:

$$var(\boldsymbol{\eta}_j^{[u]}) = \lambda_j^{[u]} - \|\mathbf{c}_m\|^2 \cdot \cos^2(\mathbf{c}_m, \mathbf{w}_j^{[u]}). \quad \text{Equation 4.8}$$

The covariance of the scores on the i^{th} and j^{th} uncentred principal component ($i \neq j$) is:

$$cov(\boldsymbol{\eta}_i^{[u]}, \boldsymbol{\eta}_j^{[u]}) = -\|\mathbf{c}_m\|^2 \cdot \cos(\mathbf{c}_m, \mathbf{w}_i^{[u]}) \cos(\mathbf{c}_m, \mathbf{w}_j^{[u]}). \quad \text{Equation 4.9}$$

The implication of **Proposition (4.1)** is that there is an important difference between uncentred and centred PCs. This is shown in Equation 4.8 and Equation 4.9. Equation 4.9 shows that PCs for an uncentred data matrix are not uncorrelated among themselves (since the covariance value is not equal to zero). Furthermore, from Equation 4.8, the variance for any uncentred PC is not represented by the eigenvalue alone, but also includes a mean score.

4.2.2 Experimental Works I

Experiments were carried out to compare the results of the non-centred and centred (mean and median) techniques. Singular values and total variances of the retained principal components are used to do the comparison.

The experiments were carried out using the Binghamton Faces dataset BU-3DFE (Yin et al., 2006). It includes a collection of three-dimensional face surfaces of 54 subjects, 25 females and 29 males. There are six different types of facial expressions: *angry*, *disgust*, *fear*, *happy*, *sad*, and *surprise*. Each of these expressions has four levels of intensity, labelled as level one (low), level two (middle), level three (high), and level four (highest), with these being based on the intensity of the expression, from lowest to highest. This provides each person a total of 24 different facial surfaces. A total of 1296 face surfaces were used (54 subjects $\times$ 6 expressions $\times$ 4 levels of intensity). Each face model has 5090 mesh points and

is a cleaned face surface (x, y, z) . This dataset includes information on both shape and texture. However, only the shape information is used in this research.

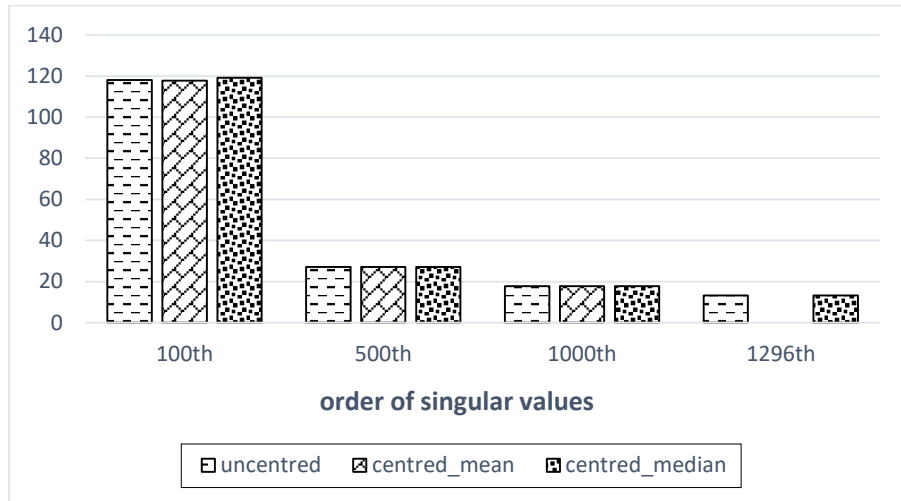
The face models were represented by a two-dimensional data matrix or a two-way array with dimension $(I_1 \times I_2)$, where each column is a measurement vector $(x, y, z \text{ points})$ for one subject. Since there is a total of 54 subjects, and six different expressions with four levels of intensity, thus a matrix with dimension $(I_{(5090 \times 3)} \times I_{(54 \times 6 \times 4)})$ was obtained.

4.2.3 Results and Discussion I

On the face dataset, mean centring and median centring, as well as non-centring techniques, were used. The single values obtained are shown in Table 4.1. There are a total of 1296 singular values, and only five values are selected to be shown here: the first value, the 100th value, the 500th value, the 1000th value, and the 1296th value. It can be seen that the values decrease as the order increases. To put it another way, early components have higher unpredictability than later components. This can happen with any of the three types of centring. The highest values are obtained from uncentred data, particularly for the first component. This is not surprising as the value does not represent the true variance, but may contain other values such as the average score (refer to Equation 4.8). This condition has been mentioned in (Kroonenberg, 1989) as well. Median centring gives slightly higher values compared to the mean. As the order increases, the differences among those three approaches become smaller, as shown by the 100th values onwards. This result is in line with the discussion in (Cadima & Jolliffe, 2009). However, mean-centred data gives the smallest value (3.4135×10^{-9}) , which is approximately equal to zero, for the last component. A graph illustration is given in Figure 4.1 (the first singular value is not included since the value is too large).

Table 4.1: Singular Values

Order of singular values	Singular values (σ)		
	Uncentred	Centred mean	Centred median
1 st	3.9114×10^6	1.1590×10^5	1.1595×10^5
100 th	118.0341	117.7366	119.1206
500 th	27.1109	27.0815	27.1058
1000 th	17.7279	17.7239	17.7281
1296 th	13.2296	3.4135×10^{-9}	13.2304

**Figure 4.1:** Comparison of Four Singular Values

The following table, Table 4.2, shows an intriguing result. As stated in Corollary (4.1), an eigenvalue obtained (in this case, singular value) from centred-mean cannot be larger than an eigenvalue (singular value) obtained from the uncentred procedure, with equal rank. It also cannot be smaller than the next ranking eigenvalue (singular value) obtained from the uncentred procedure. As an example, the first singular value from the uncentred procedure is 3.911×10^6 , which is larger than the value for centred-mean, 1.159×10^5 , and this

value is greater than the second uncentred singular value (9.086×10^3). This pattern could be seen for all singular values. However, only certain values are shown in Table 4.2.

Table 4.2: Comparison of Several Singular Values

Order	Uncentred	Centred mean	Centred median
1 st	3.911×10^6	1.159×10^5	1.160×10^5
2 nd	9.086×10^3	8.613×10^3	8.695×10^3
3 rd	6.386×10^3	6.297×10^3	6.315×10^3
4 th	4.787×10^3	4.205×10^3	4.215×10^3
5 th	3.732×10^3	3.657×10^3	3.656×10^3
6 th	3.150×10^3	2.676×10^3	2.676×10^3
7 th	2.667×10^3	2.378×10^3	2.387×10^3
8 th	2.109×10^3	2.089×10^3	2.092×10^3
9 th	1.989×10^3	1.757×10^3	1.762×10^3
10 th	1.676×10^3	1.588×10^3	1.612×10^3
101 st	117.532	116.777	117.411
102 nd	116.3	115.536	116.772
103 rd	114.219	114.178	115.174
104 th	113.451	112.498	113.691
105 th	112.392	112.300	112.341
106 th	111.386	110.705	112.184
107 th	109.138	108.647	110.212
108 th	108.214	107.880	108.602
109 th	107.346	107.043	107.854
110 th	106.387	105.197	107.035
501 st	27.079	27.054	27.064
502 nd	27.049	27.030	27.053
503 rd	26.999	26.999	27.006
504 th	26.988	26.988	26.997
505 th	26.96	26.930	26.976
506 th	26.926	26.920	26.927

Table 4.2 continued

507 th	26.908	26.877	26.919
508 th	26.875	26.867	26.873
509 th	26.859	26.818	26.852
510 th	26.8	26.757	26.816
1001 st	17.723	17.706	17.719
1002 nd	17.706	17.692	17.706
1003 rd	17.692	17.675	17.690
1004 th	17.674	17.670	17.674
1005 th	17.666	17.664	17.668
1006 th	17.663	17.653	17.663
1007 th	17.653	17.639	17.653
1008 th	17.638	17.624	17.639
1009 th	17.624	17.607	17.622
1010 th	17.605	17.605	17.607

Since the singular values obtained using uncentred data do not represent the true variances of the PCs, comparisons are made between mean-centred singular values and median-centred singular values only. Table 4.1 and Table 4.2 show that the singular values obtained from the median approach are slightly higher compared to the mean approach. This implies that each PC in the median approach carries a slightly higher variation compared to the mean approach. As a result, the total variations accounted by all the PCs (as given in Equation 4.5) have increased, and are thus higher than mean centring, as shown in Table 4.3.

Table 4.3: Total Variations

Number of Principal Component (PCs)	Mean	Median
1 PC = s_1	1.1590×10^5	1.1595×10^5
100 PCs = $(s_1 + s_2 + \dots + s_{100})$	1.8178×10^5	1.823×10^5
500 PCs = $(s_1 + s_2 + \dots + s_{500})$	2.0072×10^5	2.0131×10^5
1000 PCs = $(s_1 + s_2 + \dots + s_{1000})$	2.1159×10^5	2.1219×10^5
1296 PCs = $(s_1 + s_2 + \dots + s_{1296})$	2.1622×10^5	2.1683×10^5

The expansion of a tensor-based face recognition system using uncentred, mean-centred, and median-centred input data is discussed next.

4.2.4 Experimental Works II

Experiments in this section were carried out to investigate the advantage of using a tensor approach for a face recognition system, as well as the importance of applying centring on this tensor object.

A tensor framework was explicitly adopted, accounting for the face surface dataset containing subjects, expression types, and points (x, y, z) as the three variables. The actual measurements are the points (x, y, z) . Others are called the causal factors since they are not directly measured (changes in points depend on these causal factors). A third-order tensor $\mathcal{A} \in \mathbb{R}^{sub \times exp \times pts}$ was defined by arranging the dataset into a tensor $\mathcal{A}$ with three modes ($I_{sub} \times I_{exp} \times I_{pts}$). I_{sub} is the number of subjects, I_{exp} is the number of expression types, and I_{pts} is the number of landmark points.

Applying the N -Mode SVD algorithm (Vasilescu & Terzopoulos, 2002b), the decomposition of a tensor $\mathcal{A}$ was obtained as follows:

$$\mathcal{A} = \mathcal{D} \times_1 \mathbf{U}_{sub} \times_2 \mathbf{U}_{exp} \times_3 \mathbf{U}_{pts}. \quad \text{Equation 4.10}$$

$\mathcal{D}$ is a core tensor that manages the interactions between the indices of the three matrices. The $\mathbf{U}_{sub}$ is the mode-1 matrix spanning the subject vector space, $\mathbf{U}_{exp}$ is the mode-2 matrix spanning the expression vector space, and $\mathbf{U}_{pts}$ is the mode-3 matrix spanning the landmark points space. All these three matrices were obtained by performing SVD on each flattening matrix. This computation allows the multiple modes to vary, and at the same time, promises the orthogonality of the matrices.

A significant amount of memory and high computation power were needed since the core tensor $\mathcal{D}$ has the same dimension as tensor $\mathcal{A}$. In order to resolve the problem, a sub-tensor (subject) within the tensor model was used (Vasilescu & Terzopoulos, 2002a; Minoi & Gillies, 2008). This sub-tensor was constructed based on the variation subspace approach according to:

$$\mathcal{B} = \mathcal{D} \times_2 \mathbf{U}_{exp} \times_3 \mathbf{U}_{pts} \quad \text{Equation 4.11}$$

The dimension of sub-tensor $\mathcal{B}$ is $[I_{sub} \times (I_{exp_n} \times I_{pts})]$. This sub-tensor was easier to manipulate and requires fewer computation and memory requirements since the dimension of $[I_{sub} \times (I_{exp_n} \times I_{pts})]$ is $[54 \times (1 \times 15270)]$, being smaller than the dimension of $[I_{sub} \times (I_{exp} \times I_{pts})]$, which is $[54 \times (20 \times 15270)]$. A sub-tensor $\mathcal{B}$ with a dimension of $(I_{sub} \times 1 \times I_{pts})$ was obtained by indexing into Equation 4.11 for a particular expression. Next, it was flattened along the subject mode to obtain a matrix $\mathbf{B}_{exp(sub)}$ with a dimension of

$(I_{sub} \times I_{pts})$. This is done to get a vector of coefficients for each subject in the training set so that it can be compared with a vector of coefficients as a testing set (each testing subject). As a result, any training face model ($\mathbf{a}_a$) of subject 's' in expression 'e' can be written as:

$$\mathbf{a}_{s,e} = \mathbf{B}_{e(s)} \mathbf{c}_{train} , \quad \text{Equation 4.12}$$

where $\mathbf{c}_{train}$ is the person specific coefficient vector. Solving Equation 4.12 will yield:

$$\mathbf{c}_{train} = \mathbf{B}_{e(s)}^{-1} \mathbf{a}_{s,e} . \quad \text{Equation 4.13}$$

Given any testing face model ($\mathbf{a}_b$), the projection operator $\mathbf{B}_{e(s)}^{-1}$ was used to project the data into a candidate coefficient vector $\mathbf{c}_{test} = \mathbf{B}_{e(s)}^{-1} \mathbf{a}_b$. The recognition was carried out by finding the Euclidean distance between $\mathbf{c}_{test}$ and $\mathbf{c}_{train}$. A correct identification was determined by the minimum value of $\|\mathbf{c}_{test} - \mathbf{c}_{train}\|$.

4.2.5 Results and Discussion II

The face dataset was divided into two groups in these experiments: the query set and the training set. A leave-one-out approach was used to create a sample size of 216 faces by selecting one facial expression as a query set (54 subjects $\times$ four levels of intensity). The remaining five expressions served as a training set, with a sample size of 1080 faces (54 subjects $\times$ five expressions $\times$ four levels of intensity). This process was repeated until each expression was converted into a query set. This approach allows each expression in the query and training set to be a representative sample. If the distance between the query face and any of the faces in the training set is the least, the query face is classed as a particular subject. If the minimal distance is reported by the same subject, the identification is correct. The comparison of recognition rates between the PCA and the tensor model is shown in this section.

Table 4.4: Combination of Expressions

	Training set	Query set
1	Disgust, Fear, Happy, Sad, Surprise	Angry
2	Angry, Fear, Happy, Sad, Surprise	Disgust
3	Angry, Disgust, Happy, Sad, Surprise	Fear
4	Angry, Disgust, Fear, Sad, Surprise	Happy
5	Angry, Disgust, Fear, Happy, Surprise	Sad
6	Angry, Disgust, Fear, Happy, Sad	Surprise

Figure 4.2 shows the performance of PCA in recognising faces using the three types of centring. It shows that the PCA model gives the same results although different techniques of centring were applied. The projection operator for the PCA model, $(\mathbf{B}_{PCA}^{-1})$ has a dimension of (1080×15270) . As given in Table 4.1, when the total number of PCs (dimensions) is greater than 500, the variability that the PCs carry tends to be similar for all three types of centring approaches. In other words, they tend to give similar correct subjects or are unable to distinguish the subjects. In terms of correct recognition, subjects with fear faces show the highest percentage, which is 90.3 %, followed by subjects with sad faces, almost 88%, subjects with angry faces with 83.8%, subjects with disgusted faces, with 82.9 %, subjects with happy faces, with 62.04 %, and lastly, subjects with surprised faces, with only 37.04%.

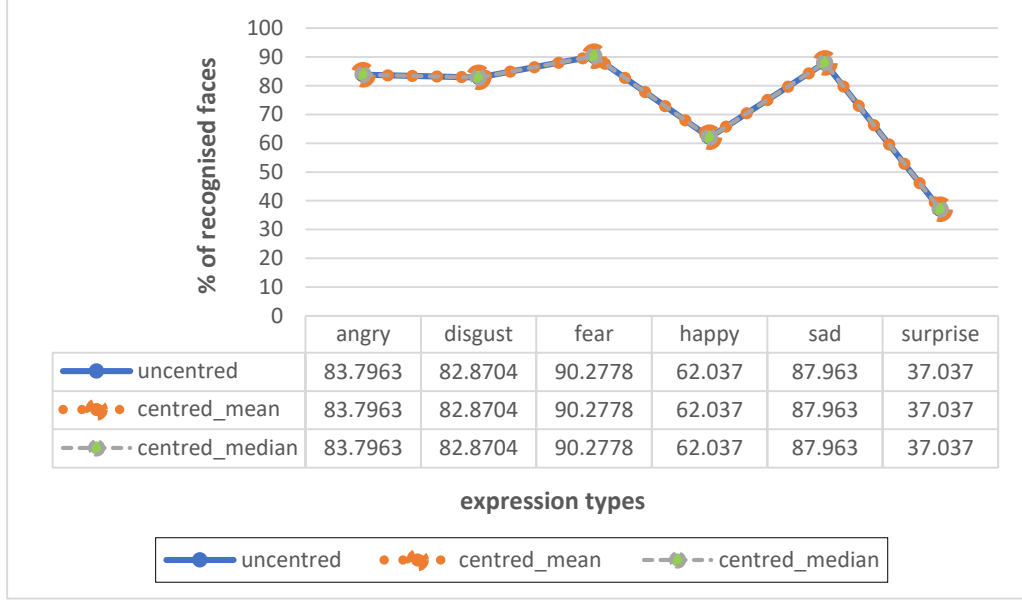


Figure 4.2: Recognition Rates for PCA Model

The following experiment will use a tensor model to recognise faces and compare the results to the prior one. Face recognition using a tensor model performs better and more consistently (Figure 4.3). Furthermore, it shows that the tensor model could capture the difference in the types of centring by referring to the different recognition rates obtained for mean, median, and uncentred approaches. This is due to the fact that the size of the tensor projection operator, $(\mathbf{B}_{e(s)}^{-1})$ is smaller, with a dimension of (54×15270) . Each mode carries a different variability. As a result, the tensor model recognition performance outperformed the PCA model.

The median approach shows the best results in recognising subjects with angry faces (95.8%), sad faces (94.4%), and surprised faces (65.7%), compared to the other two approaches (uncentred and mean-centred). This could be due to the formation of the expression itself. These three expressions hold a wide range of expressions. Different subjects will express their feelings according to their own interpretation. Since the median approach is not affected by outliers, it shows a better performance. On the other hand, for

subjects with disgusted and happy faces, the uncentred approach seems better than the centred approach, with 93.1% for subjects with disgusted faces, and 95.4% for subjects with happy faces. However, in this case, the median approach still shows better performance compared to the mean approach, whereas for disgusted faces, the median approach shows a rate of 92.6% compared to 92.1% using the mean approach, and for happy faces, the median approach shows 94.9%, and the mean shows a rate of 93.5%. Although the recognition rates for subjects with fear faces show no difference for the three types of centring (98.6%), the distances obtained differ. As an example, median centring gives the correct identified subject with a disgusted face, but mean centring gives the same person with a sad expression. Overall, median centring gives a better performance than mean centring for all types of expression.

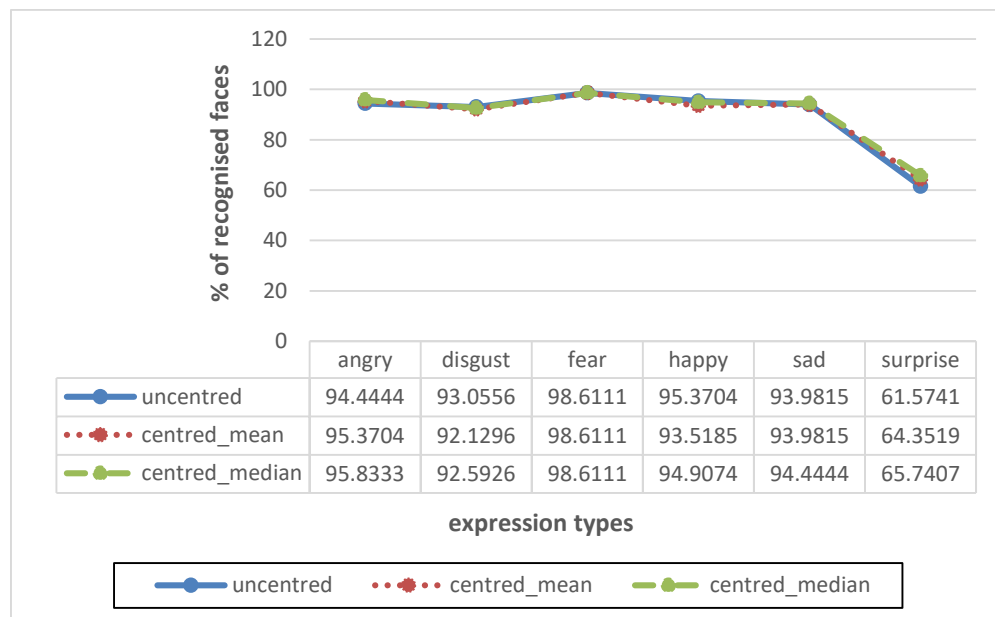


Figure 4.3: Recognition Rates for Tensor Model

Since the median approach shows a better performance compared to the mean approach, Figure 4.4 shows the overall performance in recognising faces, between the PCA and tensor model using the median approach. It shows that the tensor model gives a better

and more consistent performance compared to PCA. The percentage of correctly classified faces using the tensor model is more than 90% for all expressions, except for subjects with surprised faces. However, for PCA, only fear expression shows more than a 90% rate. Besides that, when the PCA model only shows 62% for happy faces, the tensor model achieves nearly 95%. Furthermore, although both methods cannot recognise surprised faces well, the tensor model is still far better than the PCA model, with 65.7% for the tensor model, compared to 37% for the PCA model.

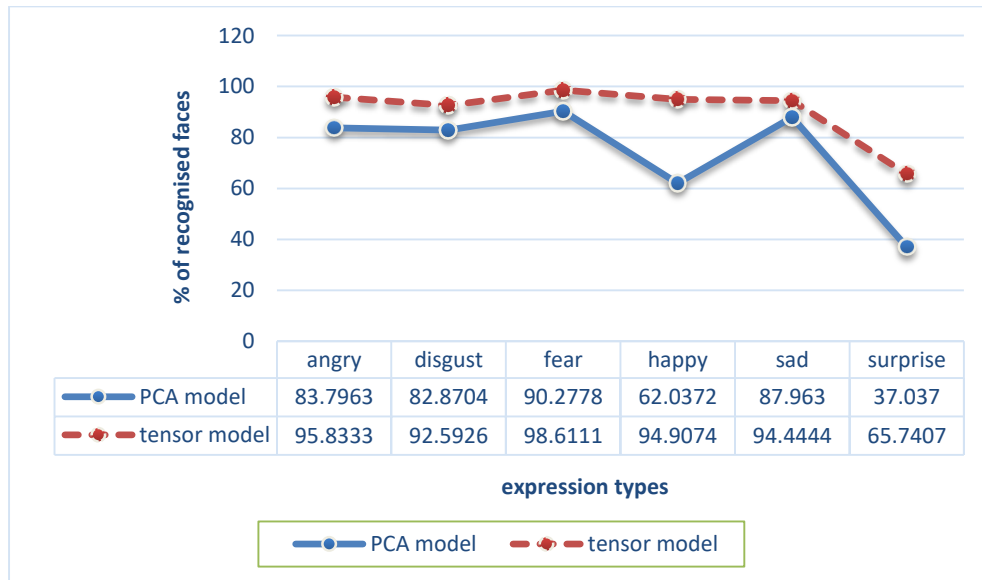


Figure 4.4: Comparison of Performance between PCA and Tensor Model

It is possible that the low recognition rate for surprised faces is due to the way the expression is formed. Typically, someone with a surprised face shows an excited expression with a wider opening of the mouth and eyes, a higher rising of the eyebrows, and a lower dropping of the jaw. Aside from that, there is an error in the surprised faces with a level three intensity dataset (which is not completely clean), which has an impact on recognition performance. The result is shown in (Peter, 2020). The researcher made a comparison between PCA and Kernel PCA models. The recognition rates for intensity level one and two

are more than 80%, and more than 50% for level four, using the kernel approach. However, the recognition rate for intensity level three is less than 10%, and lower than PCA, as shown in Figure 4.5.

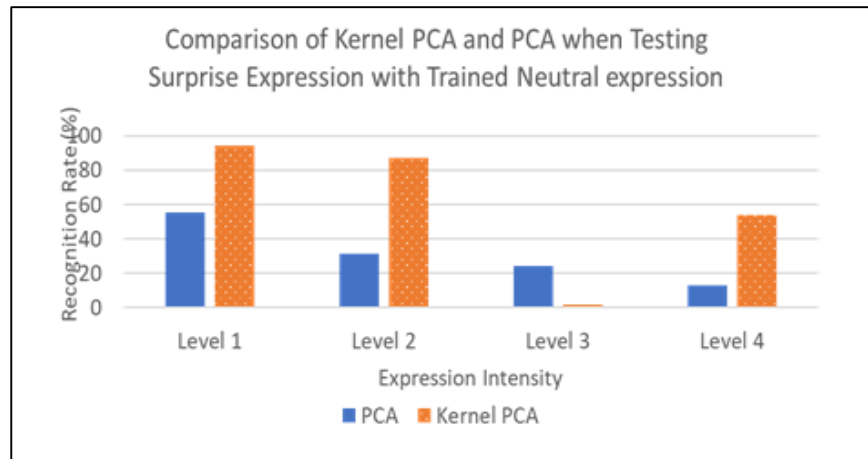


Figure 4.5: Intensity Level Three shows the Lowest Rate for the Kernel Approach

4.2.6 Summary

In summary, median centring increases the amount of variation carried by each principal as well as the overall variation carried by all components. This variation is significant because it represents facial characteristics that can be used to distinguish subjects. Apart from that, when compared to mean or non-centred processes, a tensor-based face recognition system using median centring delivers the best results, and the tensor model is better at distinguishing diverse ways of the centring approach utilised than the PCA model.

4.3 Experiment 2 (Interaction Effects)

The aim of this section is to examine the influences of different expression types and strengths (intensity levels) on recognition performance, as well as the interactions between them. This statistical testing for within-subject effects is performed in tandem with the aim of examining biometric performance. The purpose is to detect, compare, and explain how

attributes of humans, the environment, and images are connected with the performance of recognition algorithms (Givens et al., 2013).

4.3.1 Experimental Works

Since the objective of this experiment is to investigate the effects of causal factors on recognition, another factor was added, which is intensity level. For this, a new 4D tensor model was formulated as follows:

$$\mathcal{A} = \mathcal{D} \times_1 \mathbf{U}_{sub} \times_2 \mathbf{U}_{exp} \times_3 \mathbf{U}_{int} \times_4 \mathbf{U}_{pts}. \quad \text{Equation 4.14}$$

A leave-one-out procedure as in the previous experiment was carried out. Each subject would undergo a recognition process, and the minimum distance between the testing subject and the subject in the training sample was recorded. This gives a total of 1296 distances. The *distance* variable is treated as a dependent variable in the repeated measure ANOVA model. The categorical independent variables (or within-subject variations) are *expression types* (Angry = 1, Disgust = 2, Fear = 3, Happy = 4, Sad = 5, Surprise = 6), and *expression strengths* (level one = 1, level two = 2, level three = 3, level four = 4). The subject is not considered as one of the factors, because in repeated measure design, the subject is considered as an experimental block (as has been discussed in Section 3.6.3).

The repeated two-way ANOVA model is given as follows:

$$X_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijm} \quad \text{Equation 4.15}$$

where μ is the overall mean, the α 's are the expression types, the β 's are the expression strengths, the $\alpha\beta$'s are the interactions between those two factors, and ε is the random error term corresponding to the variation. These error terms are assumed to be uncorrelated and normally distributed with equal variance.

A two-way repeated measure ANOVA was applied to compare the means of the two factors: expression types and expression strengths, which are both measured on the same subjects. The following steps were taken:

- i. Data inspection to detect any outliers, missing value or weird pattern. This is done using box-plot.
- ii. Assumption fulfilment checking (Section 3.6.3).
- iii. Running repeated measure analysis.
- iv. Report findings.

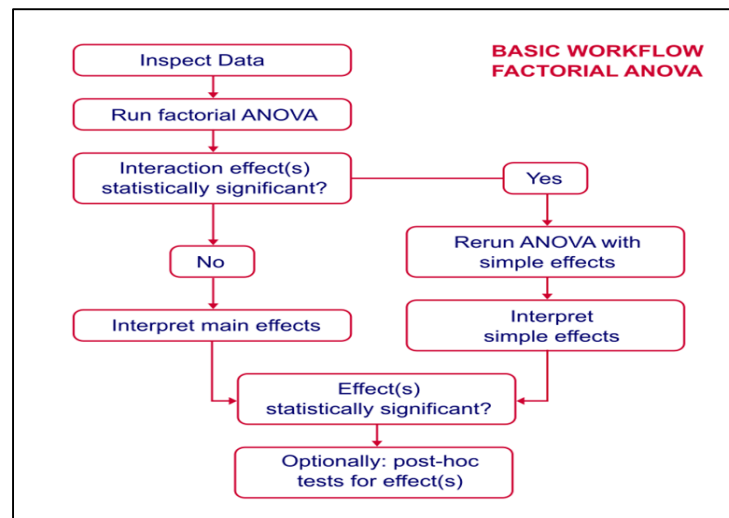


Figure 4.6: Basic Workflow for Two-Way ANOVA (Van Den Berg, 2015)

4.3.2 Results and Discussion

The results from this experiment reveal that when different levels of intensity are considered, the intensity level one gives the highest correct recognition for all the query sets (Figure 4.7). The result is consistent with what has been reported in Smeets et al. (2012).

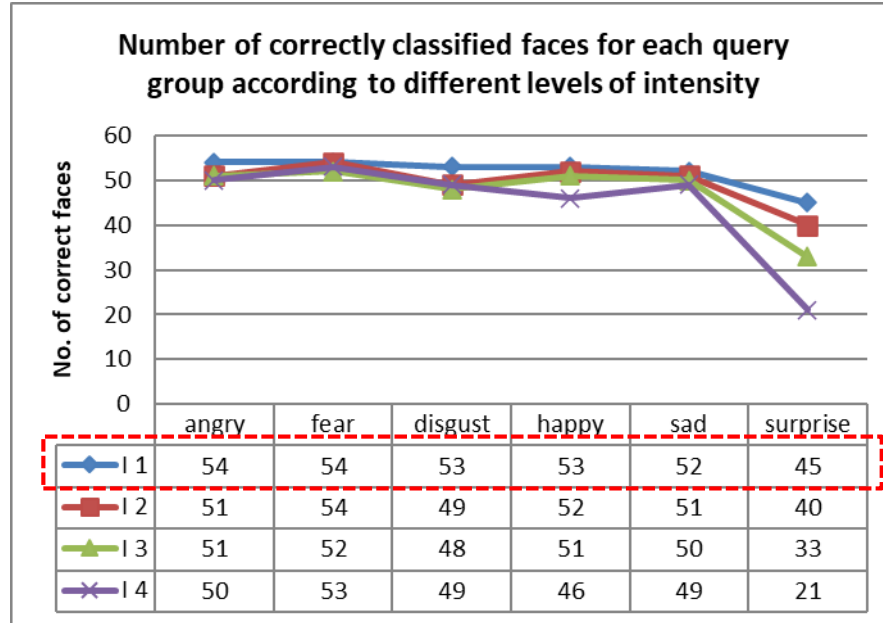


Figure 4.7: Intensity Level One Shows the Highest Correct Recognition

However, an interesting finding may be obtained if another investigation were to be carried out on each of the query sets. The result would reveal that the number of right faces from the training set varies for each query set, not only according to the intensity level, but also depending on the types of expression. Certain expressions would be more effective in recognition if they are more intense. The results are as shown in Table 4.5 until Table 4.10. The tables present the number of correct faces for each expression in the training set, according to the levels of intensity for each of the query sets.

Table 4.5 shows that for the angry query set, happy and disgusted faces, both at intensity level three, one can recognise unknown faces better compared to intensity level one. Then, for the disgusted query set, fear faces at level two intensity, and happy faces at level three intensity, performance would be better (Table 4.6). This is followed by the fear query set, happy and surprised faces at level two intensity, angry faces at level three intensity, and sad faces at level four intensity, would also perform better than level one intensity (Table

4.7). The same pattern is seen by the happy query set, where sad faces at level two intensity, angry faces at level three intensity, and disgusted faces at level four intensity, one can recognise faces from the query set better than intensity level one (Table 4.8). Next, for the sad query set, fear faces at level two intensity, surprised faces, and disgusted faces at level three intensity, would also perform better. Lastly, for the surprise query set, angry at level two intensity would also perform better compared to level one intensity (Table 4.9 and Table 4.10).

Table 4.5: Number of Correct Faces for Angry Query Set

	Fear	Happy	Surprise	Sad	Disgust	Total
I 1	9	3	11	20	11	54
I 2	9	3	4	19	16	51
I 3	7	5	6	11	22	51
I 4	9	4	3	13	21	50
Total	34	15	24	63	70	206

Table 4.6: Number of Correct Faces for Disgust Query Set

	Angry	Happy	Sad	Surprise	Fear	Total
I 1	19	4	10	7	13	53
I 2	11	10	6	5	17	49
I 3	11	12	5	6	14	48
I 4	13	10	9	4	13	49
Total	54	36	30	22	57	199

Table 4.7: Number of Correct Faces for Fear Query Set

	Angry	Happy	Surprise	Sad	Disgust	Total
I 1	5	24	5	7	13	54
I 2	3	27	7	9	8	54
I 3	7	23	6	9	7	52
I 4	2	23	5	10	13	53
Total	17	97	23	35	41	213

Table 4.8: Number of Correct Faces for Happy Query Set

	Angry	Fear	Surprise	Sad	Disgust	Total
I 1	5	30	5	3	10	53
I 2	5	30	4	6	7	52
I 3	6	29	3	3	10	51
I 4	2	27	3	2	12	46
Total	18	116	15	14	39	202

Table 4.9: Number of Correct Faces for Sad Query Set

	Angry	Fear	Happy	Surprise	Disgust	Total
I 1	22	7	6	10	7	52
I 2	18	10	4	10	9	51
I 3	18	8	2	11	11	50
I 4	22	9	3	9	7	49
Total	80	34	15	40	34	203

Table 4.10: Number of Correct Faces for Surprise Query Set

	Angry	Fear	Happy	Sad	Disgust	Total
I 1	5	12	3	6	19	45
I 2	8	10	2	5	15	40
I 3	2	6	2	6	17	33
I 4	1	6	2	2	10	21
Total	16	34	9	19	61	139

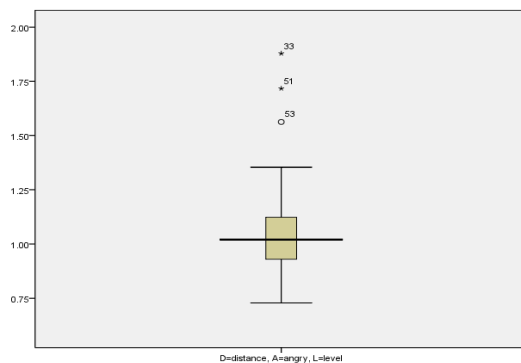
This leads to a further examination to discover the reason behind this occurrence, or to find an answer to questions such as: ‘*why is it that certain expressions can recognise an unknown face better, at a higher level of intensity?*’ or ‘*why is it that certain expressions perform better in recognising unknown faces compared to other expressions?*’ (for example, most of the correct faces for the happy faces query set come from fear faces, with 116 faces, the rest are less than 40).

From the literature review that was conducted, it can be said that there is a possible interaction effect that exists at the deeper level of the *Tensorface* model or among the face variants. Interactions may occur when the effect of one independent variable on the dependent variable changes depending on the level of another independent variable. In other words, the effect of different expressions on the recognition of unknown faces changes depending on the levels of intensity in the training sets. However, a problem arises as the interaction effects cannot be directly extracted from the core tensor values since the total variation in each mode of the tensor model is different.

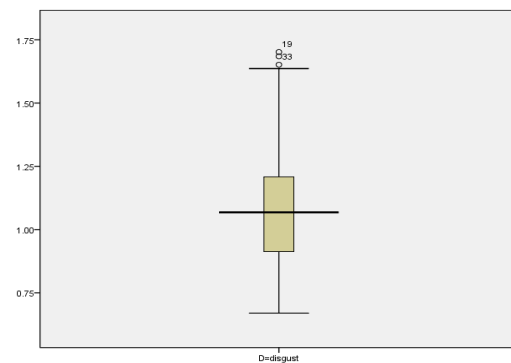
For this, an ANOVA was performed to confirm the above claim. This study involves a six (expressions: angry, disgust, fear, happy, sad, surprise) by four (intensity levels: low, middle, high, highest) factorial design. Before that, the data was inspected by checking for outliers, missing values, or unnatural patterns, to avoid any negative impact on the recognition performance. This was carried out using a box plot.

By referring to the box plots in Figure 4.8, from (i) to (xxiv), each of them shows only a few dots (representing outliers) outside the box plot. Outliers occur when a person wears or appears with extra clothing such as hats, scarfs or even sunglasses. Apart from that, it could also be due to the change in the view of the 3D scanner, which could cause the visibility of other parts, such as ears or chest (Smeets et al., 2011). However, these few outliers will not significantly affect the analysis.

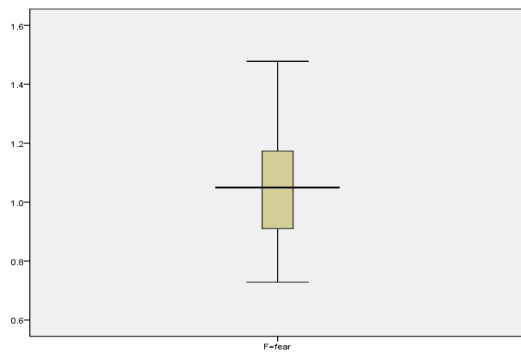
As had been mentioned in (Smeets et al., 2011), the BU-3DFE database has the least outliers, despite the strong expression variations. As such, the analysis was carried out with all the available data values.



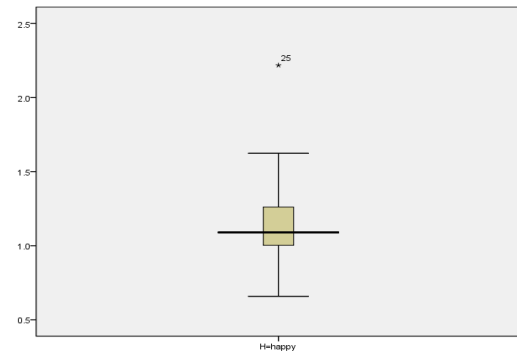
(i) angry, level 1



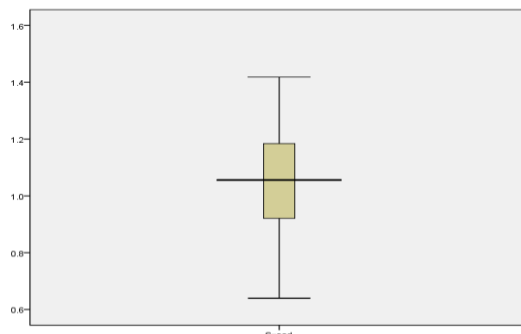
(ii) disgust, level 1



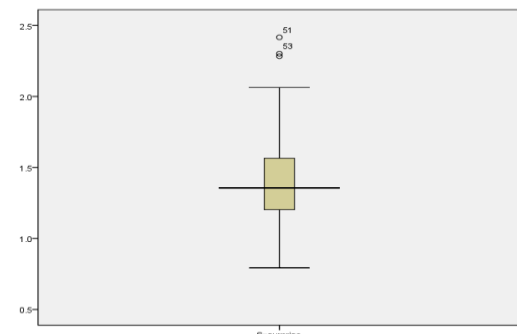
(iii) fear, level 1



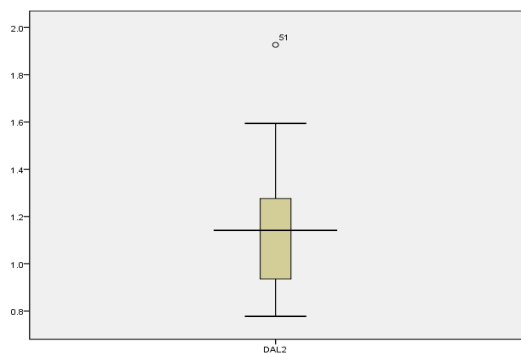
(iv) happy, level 1



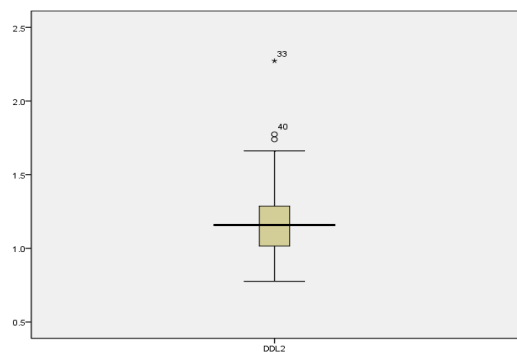
(v) sad, level 1



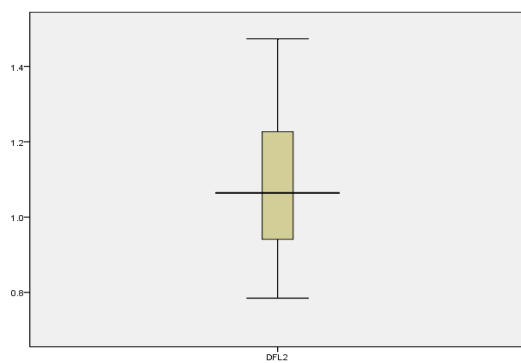
(vi) surprise, level 1



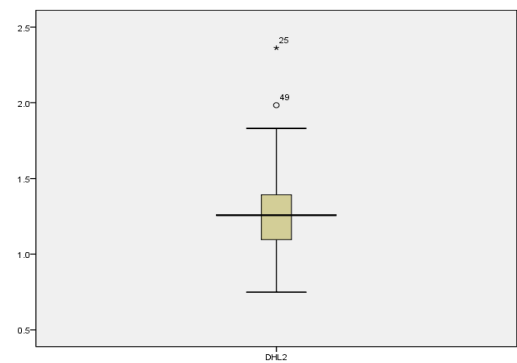
(vii) angry, level 2



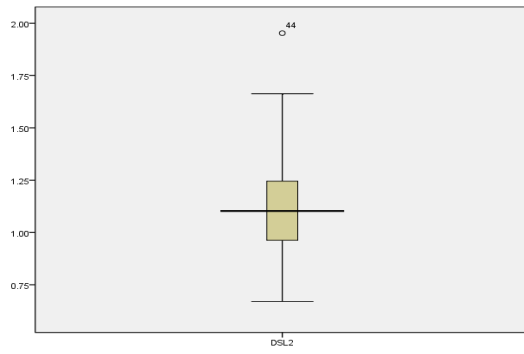
(viii) disgust, level 2



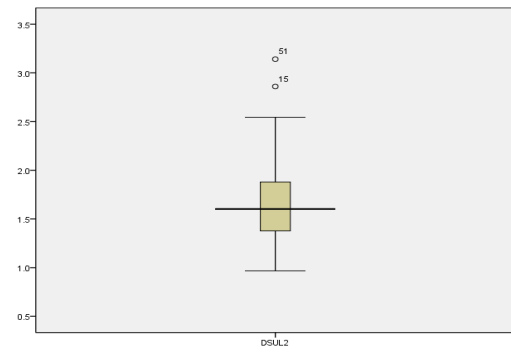
(ix) fear, level 2



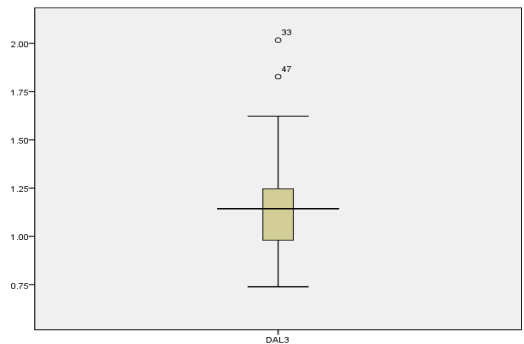
(x) happy, level 2



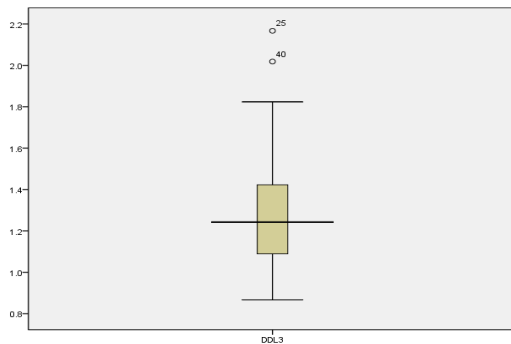
(xi) sad, level 2



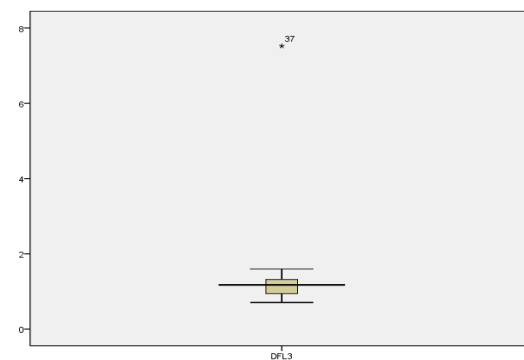
(xii) surprise, level 2



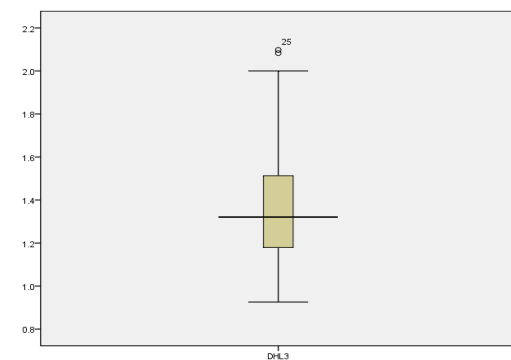
(xiii) angry, level 3



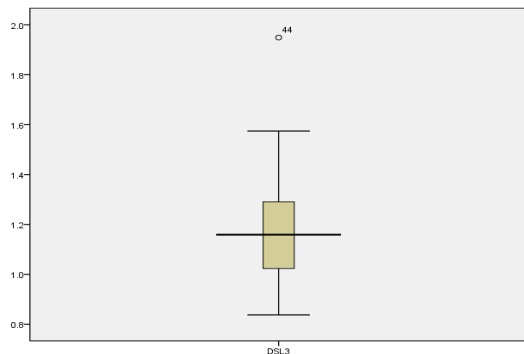
(xiv) disgust, level 3



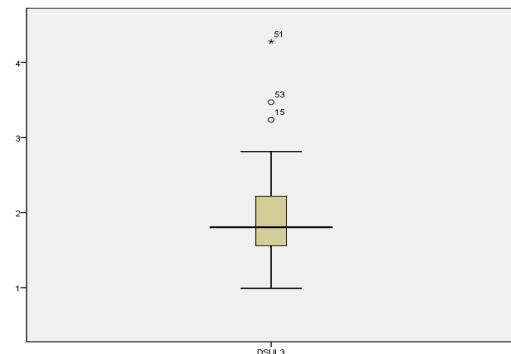
(xv) fear, level 3



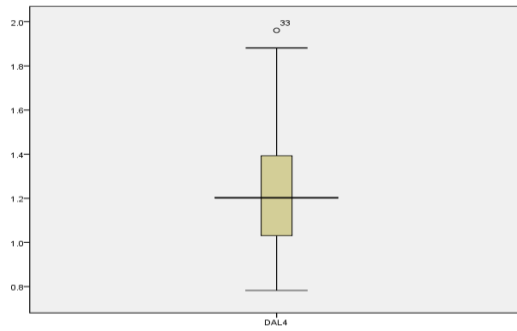
(xvi) happy, level 3



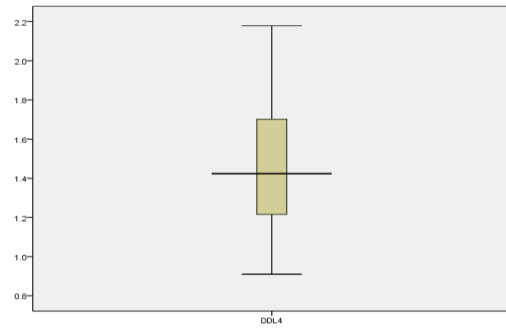
(xvii) sad, level 3



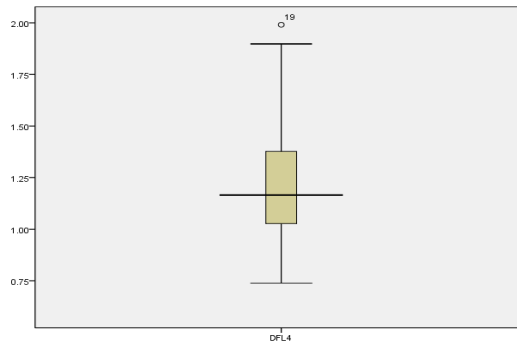
(xviii) surprise, level 3



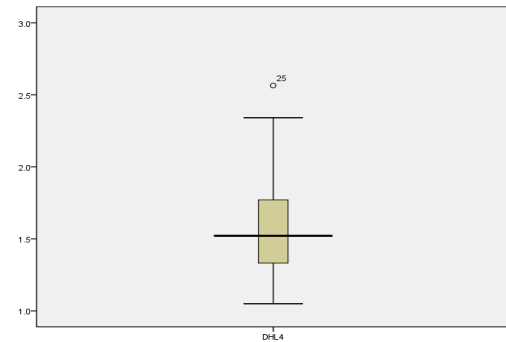
(xix) angry, level 4



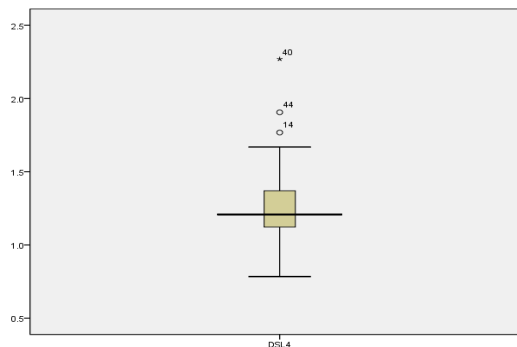
(xx) disgust, level 4



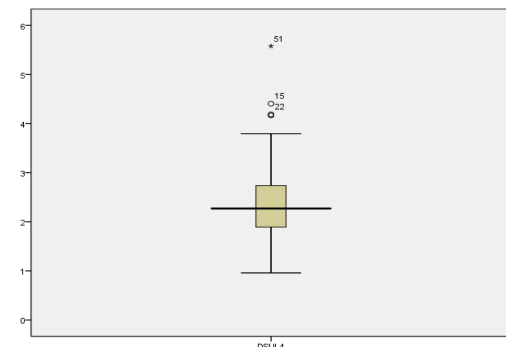
(xxi) fear, level 4



(xxii) happy, level 4



(xxiii) sad, level 4



(xxiv) surprise, level 4

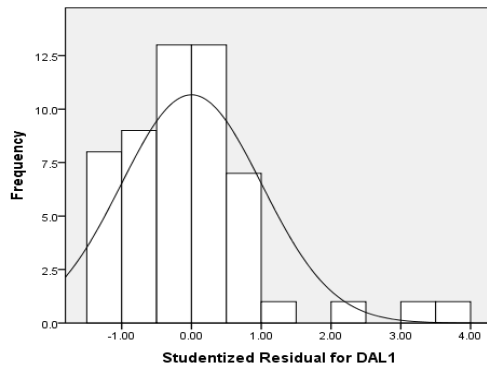
Figure 4.8: Figure (i) until (xxiv) Show the Box-Plots for Each Variable

There are four assumptions that must be met, and they are as follows:

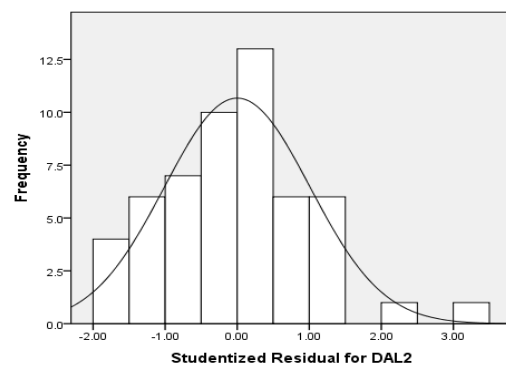
- i. The response variable is continuous.
- ii. The residual is normally distributed.
- iii. The subjects are independent.

- iv. All within-subject variance-covariance matrices are circular. Circularity means the variance of differences between any two measurements within a subject are constant. It can be tested using Mauchly's test statistic W , called Sphericity testing and epsilon value, ϵ . W and ϵ ranges from 1 to 0. The smaller the values, the greater the departure from sphericity and circularity.

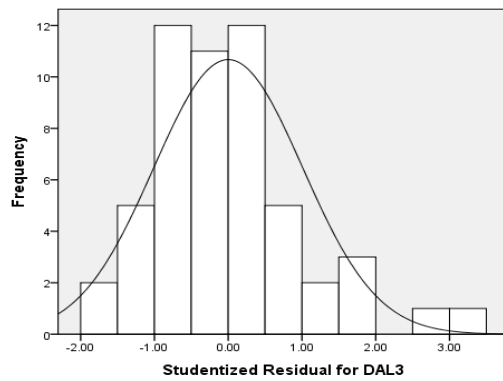
The first condition is fulfilled since the response variable in this study is distance, which is a continuous (not a discrete) measurement. Then, condition number three is also acceptable because all the data are measured on different subjects that are independent of one another. In terms of residuals, two are not normally distributed (refer to Figure 4.9(i) and Figure 4.9(xi)), as they show a skewed distribution (not a bell-curve distribution).



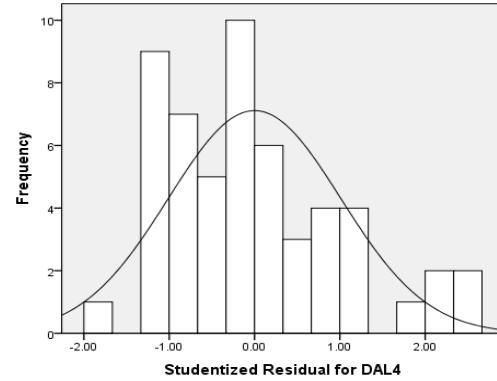
(i)



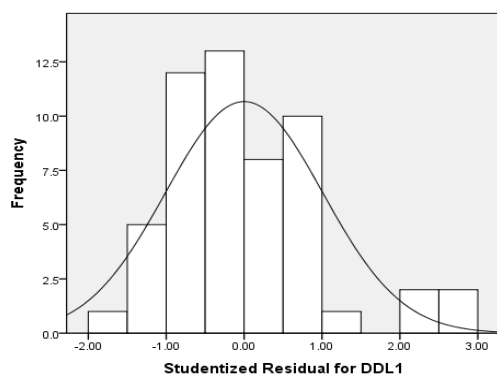
(ii)



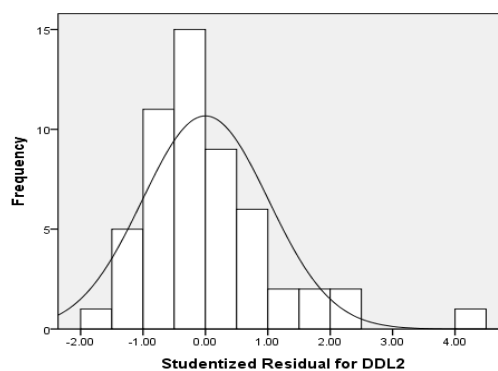
(iii)



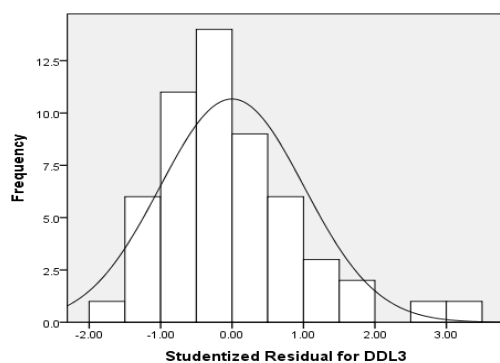
(iv)



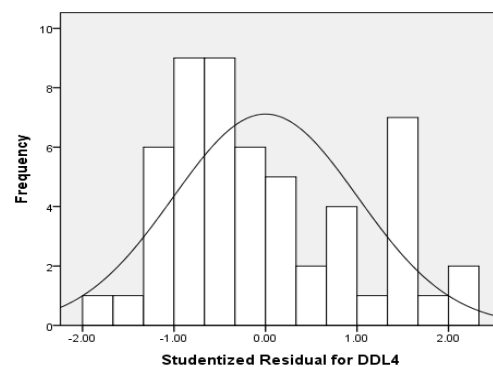
(v)



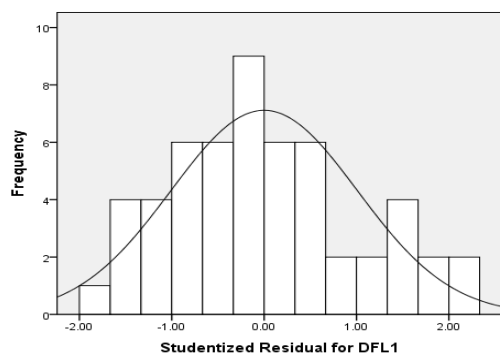
(vi)



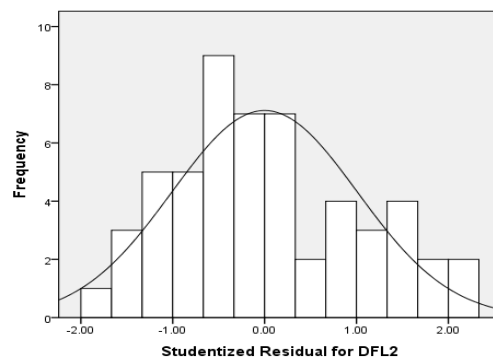
(vii)



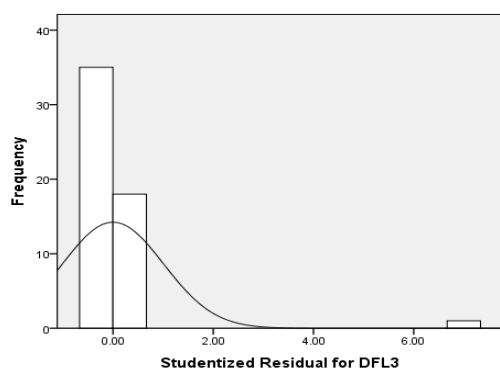
(viii)



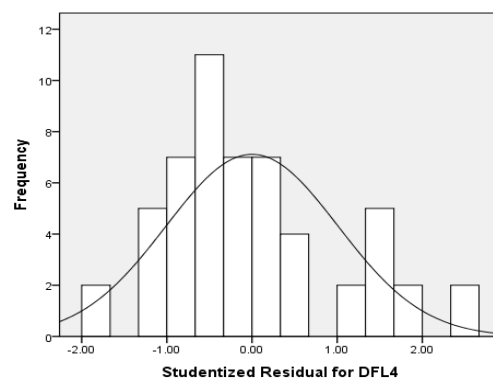
(ix)



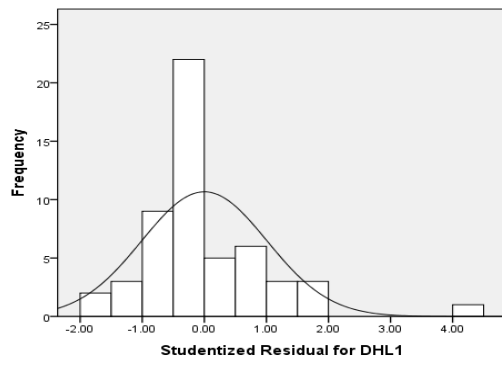
(x)



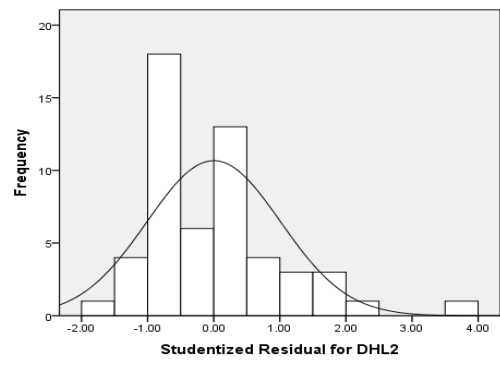
(xi)



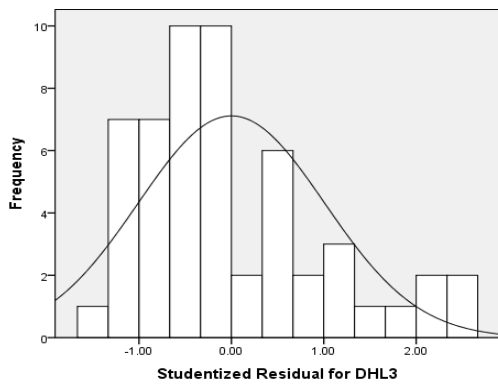
(xii)



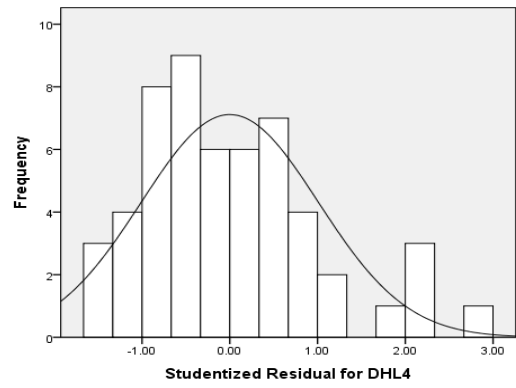
(xiii)



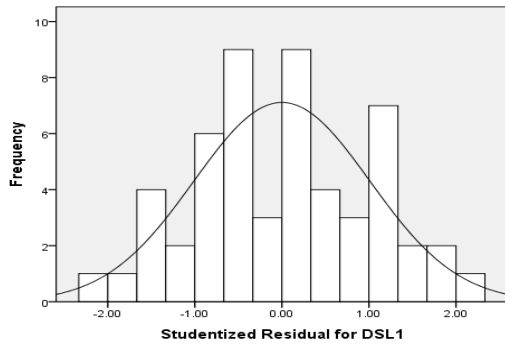
(xiv)



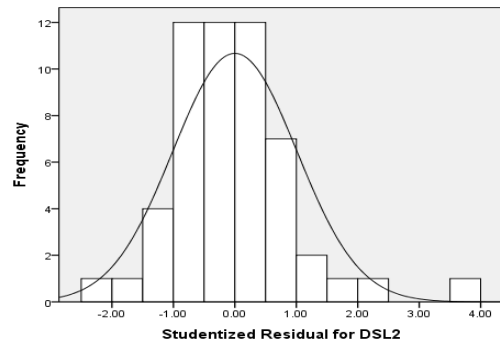
(xv)



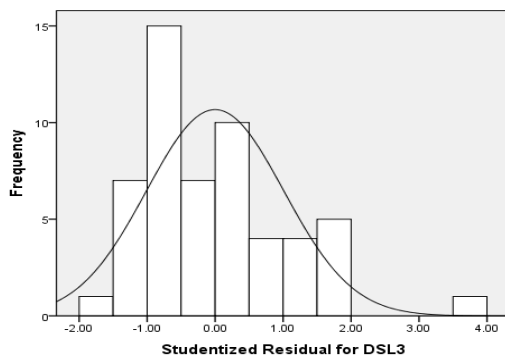
(xvi)



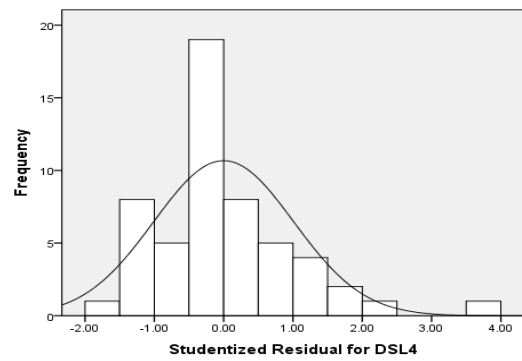
(xvii)



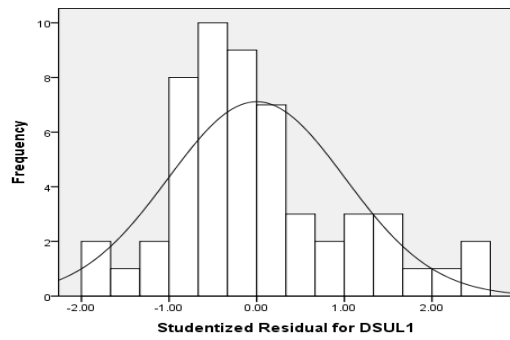
(xviii)



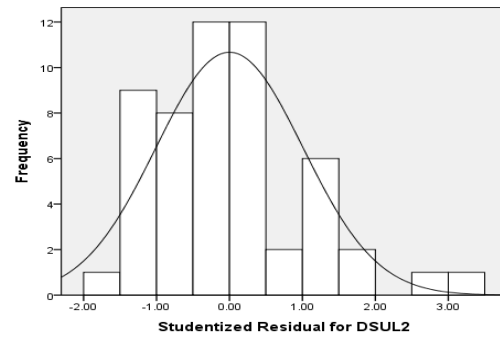
(xix)



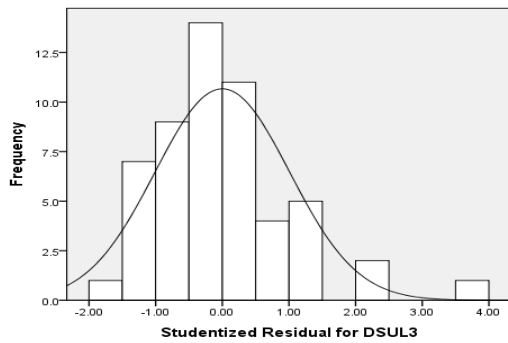
(xx)



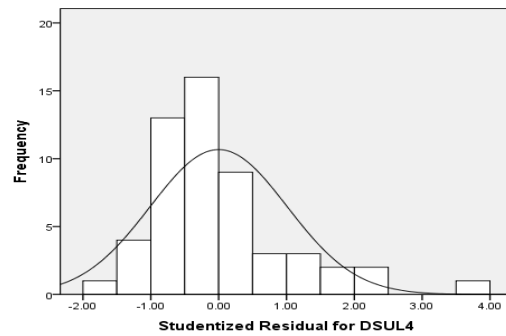
(xxi)



(xxii)



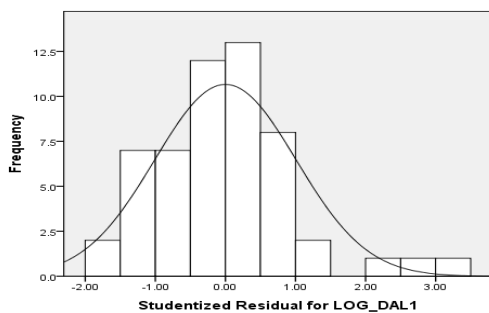
(xxiii)



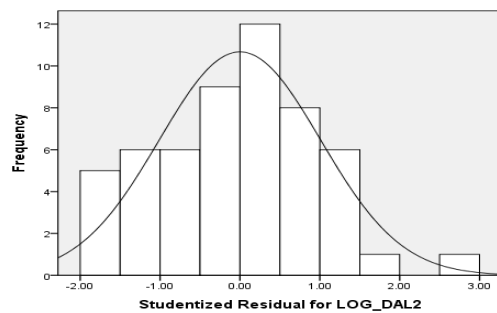
(xxiv)

Figure 4.9: Figure (i) until (xxiv) Show the Distribution of the Residuals

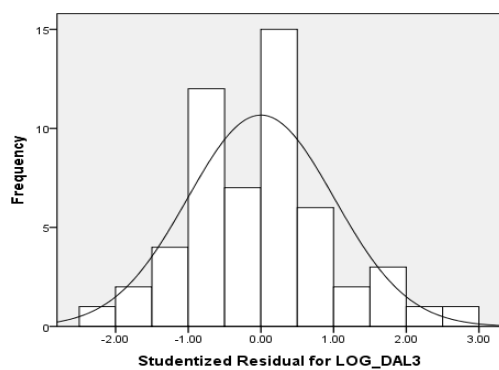
Several variable transformations are available in order to transform this non-normality into a normal condition, such as the log (log10) transform, square root transform, and reciprocal transform. However, a logarithmic transformation was chosen since the non-normal residuals show a positively-skewed distribution. With this, a more bell-curved shape was obtained after the logarithmic transformation for all the residuals, as shown in Figure 4.10(i) until Figure 4.10(xxiv).



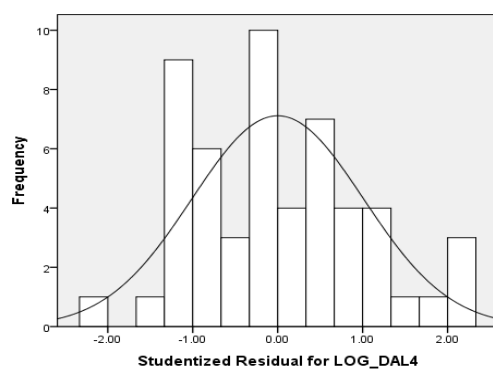
(i)



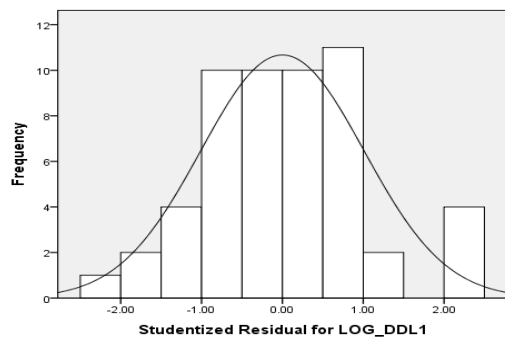
(ii)



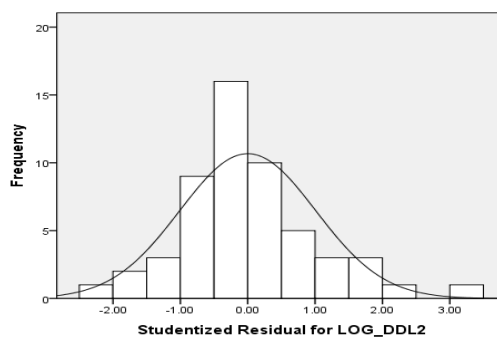
(iii)



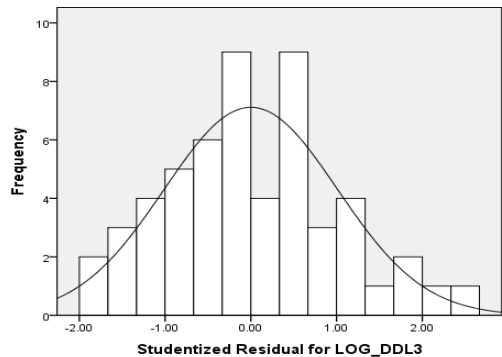
(iv)



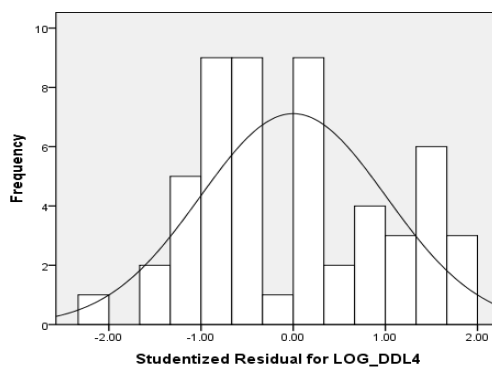
(v)



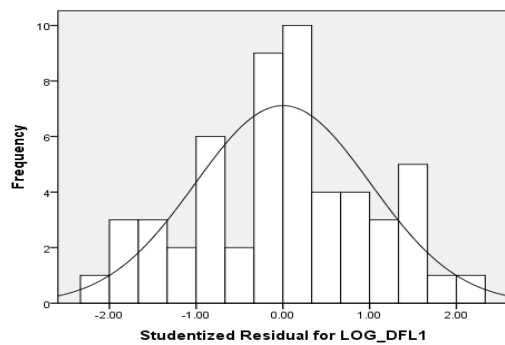
(vi)



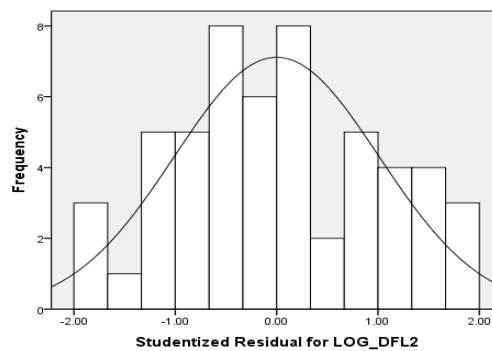
(vii)



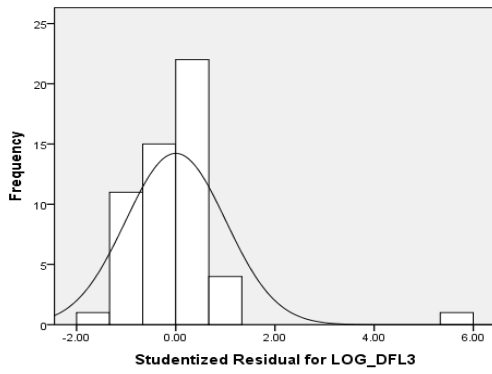
(viii)



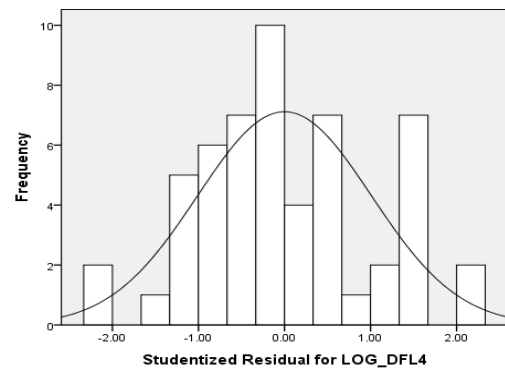
(ix)



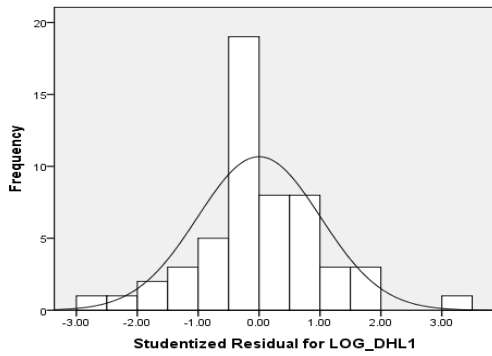
(x)



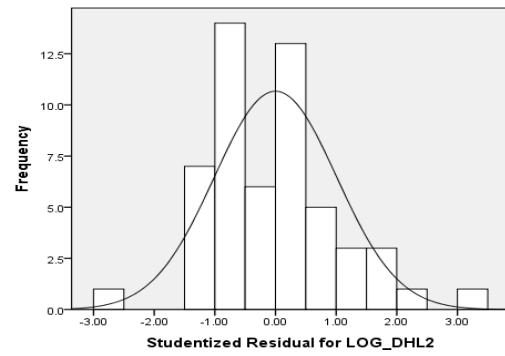
(xi)



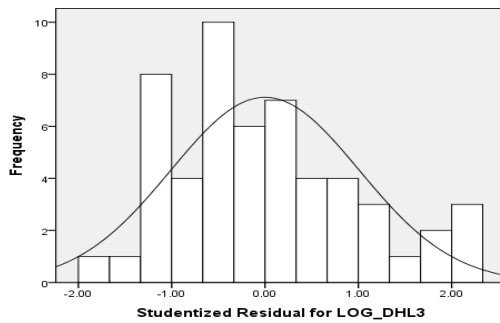
(xii)



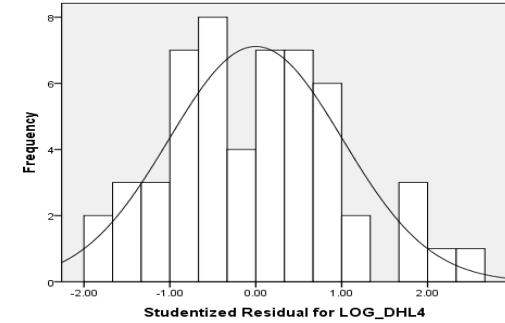
(xiii)



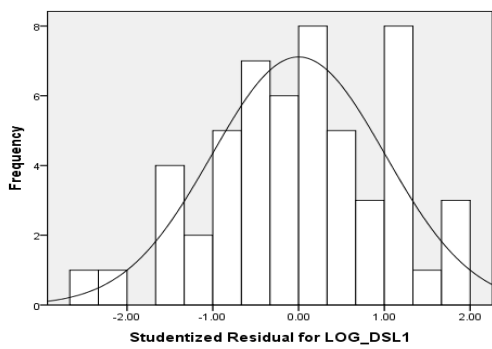
(xiv)



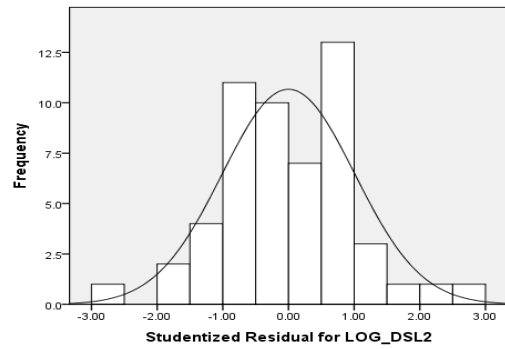
(xv)



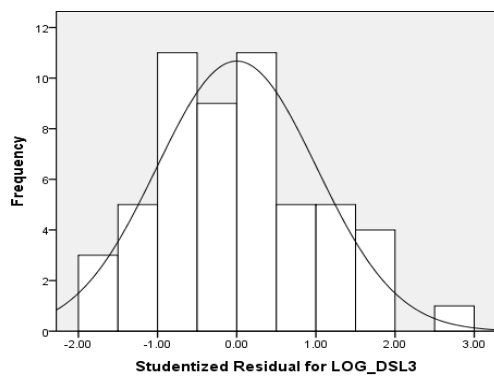
(xvi)



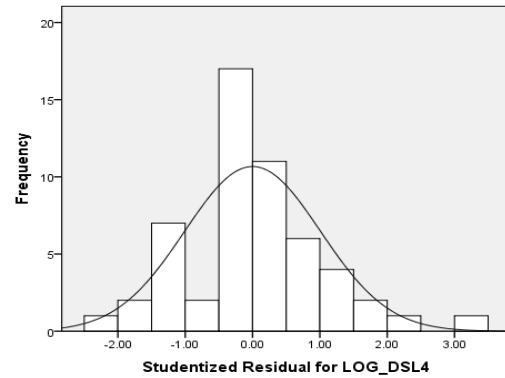
(xvii)



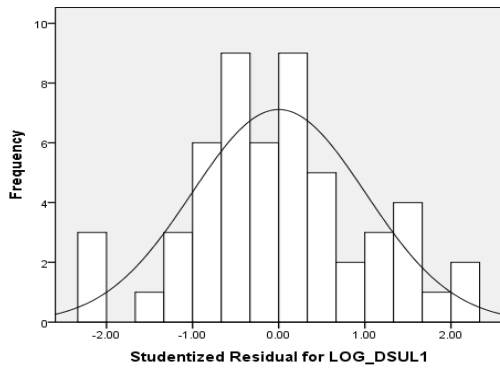
(xviii)



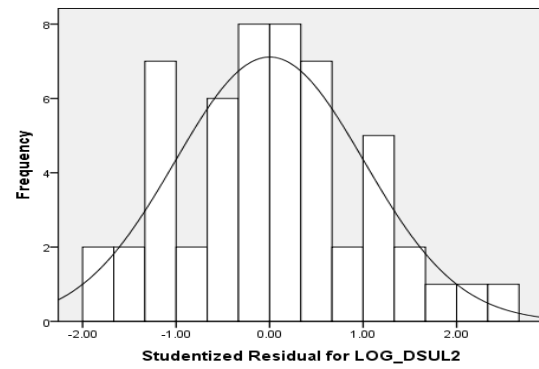
(xix)



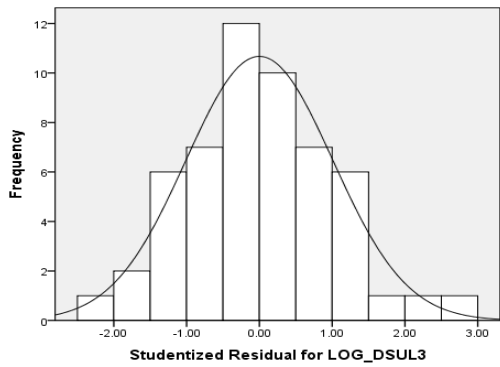
(xx)



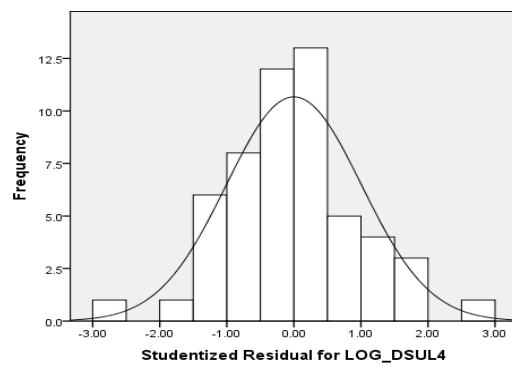
(xxi)



(xxii)



(xxiii)



(xxiv)

Figure 4.10: Figure (i) until (xxiv) Show the Normally Distributed Residuals

Table 4.11 (generated automatically using SPSS) shows the results of Mauchly's test of sphericity (for checking the fourth assumption). Mauchly's test was used to test the null hypothesis that the variances of the difference between conditions are equal or whether the variance-covariance matrix departs from circularity. There are three within-subject effects:

the expression effect, the intensity effect, and the interaction effects among expressions and intensity levels (expression*intensity), as listed in the table. All the significance values (under the Sig. column) are less than $\alpha = 0.05$ (0.000, 0.002, and 0.001) for expression, intensity, and expression*intensity (p -value < 0.05). This means that the null hypothesis is rejected. In other words, the variances of the difference between levels are significantly different. Therefore, the assumption of sphericity has been violated. This is supported by all the three values under Mauchly's W column (0.392, 0.697, and 0.027), where the values are far from one. The smaller the values, the greater the departure from sphericity and circularity.

Table 4.11: The Mauchly's Test of Sphericity^b (SPSS Output)

Measure: DISTANCE

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
expression	.392	47.891	14	.000	.748	.812	.200
intensity	.697	18.681	5	.002	.800	.840	.333
expression * intensity	.027	173.294	119	.001	.693	.875	.067

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: expression + intensity + expression * intensity

Since the Mauchly's test shows a violation in the sphericity assumption, this will increase the probability of getting a Type II error (accepting the null hypothesis when it is actually false) and a test statistic F that simply cannot be compared to the tabulated value of the F -distribution. Therefore, a measurement correction (degrees of freedom for the F

statistic) was applied (refer to Figure 4.11) in order to continue with further statistical analysis. There are three corrected tests available: Greenhouse–Geisser, Huynh-Feldt, and Lower-bound (under Epsilon column, Table 4.11). Among these three, Huynh-Feldt was chosen since all the epsilon values (ϵ) are greater than 0.75 (0.812, 0.840, and 0.875).

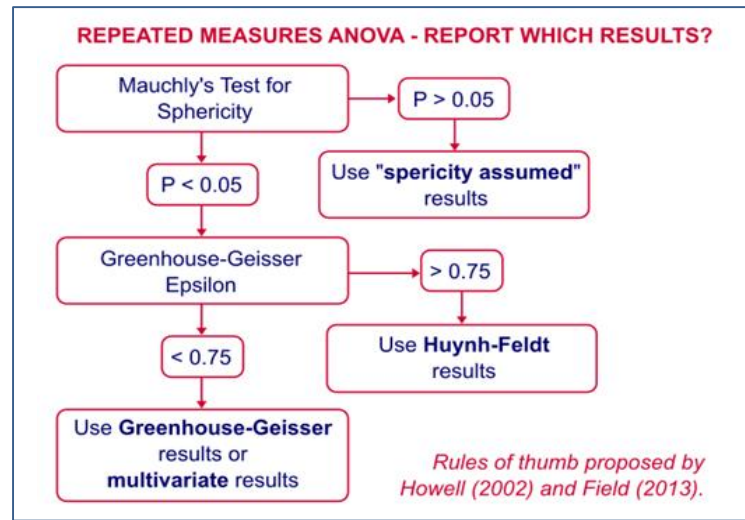


Figure 4.11: Flowchart for Mauchly’s Test of Sphericity (Van Den Berg, 2015)

Further statistical analysis was carried out after checking all the assumptions. Table 4.12 lists all the variables (intra-subject covariates) involved. The first independent variable is expressions, where 1 represents an angry query set, 2 represents a disgust query set, 3 represents a fear query set, 4 represents a happy query set, 5 represents a sad query set, and 6 represents a surprise query set (1 = angry, 2 = disgust, 3 = fear, 4 = happy, 5 = sad, 6 = surprise). This is followed by the second independent variable, intensity levels, from level one to level four (1 to 4), and dependent variable DISTANCE (DAL = Distance for Angry face with Level 1 to 4, DDL = Distance for Disgust face with Level 1 to 4, DFL = Distance for Fear face with Level 1 to 4, DHL = Distance for Happy face with Level 1 to 4, DSL = Distance for Sad face with Level 1 to 4, DSUL = Distance for Surprise face with Level 1 to 4), which has been transformed using the Log function.

Table 4.12: List of Variables
(SPSS Output)

expression	intensity	Dependent Variable
1	1	LOG_DAL1
	2	LOG_DAL2
	3	LOG_DAL3
	4	LOG_DAL4
2	1	LOG_DDL1
	2	LOG_DDL2
	3	LOG_DDL3
	4	LOG_DDL4
3	1	LOG_DFL1
	2	LOG_DFL2
	3	LOG_DFL3
	4	LOG_DFL4
4	1	LOG_DHL1
	2	LOG_DHL2
	3	LOG_DHL3
	4	LOG_DHL4
5	1	LOG_DSL1
	2	LOG_DSL2
	3	LOG_DSL3
	4	LOG_DSL4
6	1	LOG_DSUL1
	2	LOG_DSUL2
	3	LOG_DSUL3
	4	LOG_DSUL4

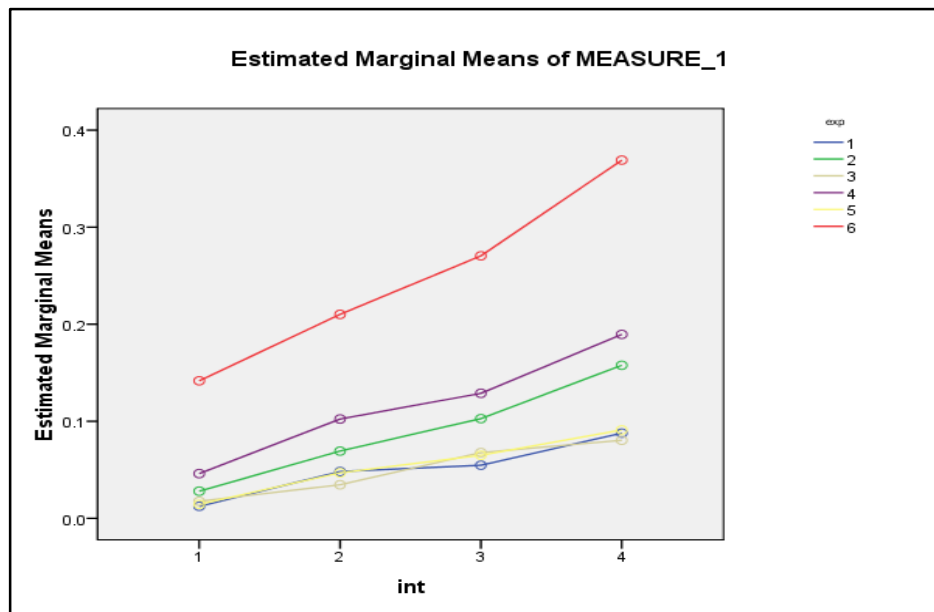


Figure 4.12: Profile Plot for the Estimated Marginal Means of Distance

Figure 4.12 is the SPSS output obtained (MEASURE_1 = distance, int = intensity levels, exp = expression types). The profile plot shows a rough idea of what is happening within the dataset. The plot shows that all the six expressions have the lowest mean distances for level 1 intensity (highly correct recognition), and as the level of intensity increases, most of the expressions show an increase in the mean distances, which leads to a more difficult recognition process. Among the six different expressions, the surprise query set (red colour) shows the highest value of mean distances (meaning the most difficult to recognise), followed by happy (purple colour), and disgust (green colour) query sets.

More interestingly, the above figure also demonstrates the existence of interaction effects among expression types and intensity levels. Such an interaction means that the effect of intensity levels is different for different types of expression, or that the effect of expression types on recognition performance changes according to changes in intensity levels. The interaction effects are shown by a set of non-parallel lines, with the intersection between line

1 (angry), line 3 (fear), and line 5 (sad). The next paragraphs will go over a more in-depth quantitative study of this issue.

Table 4.13: Several Tests for Within-Subject Effects (SPSS Output)

Measure:distance

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
expression	Sphericity Assumed	6.317	5	1.263	84.090	.000	.613
	Greenhouse- Geisser	6.317	3.740	1.689	84.090	.000	.613
	Huynh-Feldt	6.317	4.059	1.556	84.090	.000	.613
	Lower-bound	6.317	1.000	6.317	84.090	.000	.613
Error(expression)	Sphericity Assumed	3.982	265	.015			
	Greenhouse- Geisser	3.982	198.230	.020			
	Huynh-Feldt	3.982	215.153	.019			
	Lower-bound	3.982	53.000	.075			
intensity	Sphericity Assumed	2.449	3	.816	127.900	.000	.707
	Greenhouse- Geisser	2.449	2.399	1.021	127.900	.000	.707
	Huynh-Feldt	2.449	2.521	.972	127.900	.000	.707
	Lower-bound	2.449	1.000	2.449	127.900	.000	.707
Error(intensity)	Sphericity Assumed	1.015	159	.006			
	Greenhouse- Geisser	1.015	127.158	.008			
	Huynh-Feldt	1.015	133.606	.008			
	Lower-bound	1.015	53.000	.019			

Table 4.13 continued

expression * intensity	Sphericity Assumed	.579	15	.039	9.425	.000	.151
	Greenhouse-Geisser	.579	10.393	.056	9.425	.000	.151
	Huynh-Feldt	.579	13.125	.044	9.425	.000	.151
	Lower-bound	.579	1.000	.579	9.425	.003	.151
Error(expression*intensity)	Sphericity Assumed	3.256	795	.004			
	Greenhouse-Geisser	3.256	550.823	.006			
	Huynh-Feldt	3.256	695.622	.005			
	Lower-bound	3.256	53.000	.061			

Table 4.13 provides the statistical analysis for the main effects (expression, intensity), together with the interaction effects (expression*intensity). As had been mentioned earlier, the Huynh-Feldt correction measure is chosen. For expression types effect, the significance value is 0.000 (under Sig. column), which is less than $\alpha = 0.05$, with the partial eta squared value, η_p^2 being 0.613 ($F = 84.090$, $p\text{-value} = 0.000$, $\eta_p^2 = 0.613$). Eta squared is used to describe the proportion of the total variability attributable to that factor. This implies that the null hypothesis, which assumes that all the mean values for the expression types are the same, will be rejected. There is a significant difference in the effects of expression types towards recognition performance. The same conclusion can be made for intensity levels. Since the significance value is 0.000, which is less than $\alpha = 0.05$, then the null hypothesis, which assumes that there is no difference in the means for intensity at all levels, would be rejected ($F = 127.9$, $p\text{-value} = 0.000$, $\eta_p^2 = 0.707$). There is a significant difference in the effect of intensity levels towards recognition performance. The expression types effect and the intensity levels effect, are considered the main effects.

The interaction effects are the emphasis of this section. The significance value for the (expression*intensity) effect is 0.000, which is less than $\alpha = 0.05$ ($F = 9.425$, p -value = 0.000, $\eta_p^2 = 0.15$). Therefore, the null hypothesis, which assumes that there is no interaction among expression types and intensity levels, is then rejected. In other words, when there are changes in intensity levels, the effect of expression types on recognition performance will also change (as predicted from Table 4.5 until Table 4.10, and the discussion on the matter). Therefore, this study has proven what had been predicted earlier with evidence from the statistical analysis.

Since this study could conclude that different intensity levels effect recognition accuracy for the six expressions in different ways, determining the distinct groups or levels would be of interest. For the next experiment, a different statistical test was used: the post-hoc test or pairwise comparison. The results are shown in Table 4.14.

Table 4.14: Pairwise Comparison (Rewritten from SPSS Output)

Expression	Intensity levels	Sig.	Expression	Intensity levels	Sig.
angry	(1,2)	0.04	disgust	(1,2)	0.002
	(1,3)	0.07		(1,3)	0.000
	(1,4)	0.000		(1,4)	0.000
	(2,3)	0.629		(2,3)	0.001
	(2,4)	0.004		(2,4)	0.000
	(3,4)	0.003		(3,4)	0.000
fear	(1,2)	0.114	happy	(1,2)	0.000
	(1,3)	0.006		(1,3)	0.000
	(1,4)	0.000		(1,4)	0.000
	(2,3)	0.066		(2,3)	0.014
	(2,4)	0.000		(2,4)	0.000
	(3,4)	0.429		(3,4)	0.000

Table 4.14 continued

sad	(1,2)	0.002	surprise	(1,2)	0.000
	(1,3)	0.000		(1,3)	0.000
	(1,4)	0.000		(1,4)	0.000
	(2,3)	0.85		(2,3)	0.000
	(2,4)	0.001		(2,4)	0.000
	(3,4)	0.009		(3,4)	0.000

The pairwise comparisons of the expression types and intensity levels are shown in Table 4.14, with the Least Significant Difference (LSD) adjustment. The effects of intensity on disgusted, joyful, and surprised faces are statistically significant at all levels, according to the table, with all significance values less than $\alpha = 0.05$. To put it another way, variations in intensity levels for these three types of emotion will have an impact on face recognition performance.

Certain levels of intensity, on the other hand, have a similar influence on the recognition of angry, fearful, and sad faces. Angry faces show that intensity level (1, 3) and level (2, 3) have similar effects on recognition performance, since the significance values for level (1, 3) and level (2, 3) are both greater than $\alpha = 0.05$, which are 0.07 and 0.629, respectively. Apart from that, for fear faces, intensity levels (1, 2), (2, 3), and (3, 4) all have similar effects. The significance values for these combinations are 0.114, 0.066, and 0.429 respectively. Lastly, sad faces show that intensity level (2, 3) has no difference since the significance value is 0.85, greater than $\alpha = 0.05$.

The results obtained are partially in line with those in Hoffmann et al. (2010). In their study, they examined recognition performance at 50% and 100%, and found that changes in intensity have a significant effect on angry, disgust, happy, and sad. However, when they

changed the participants and intensity levels (ranging from 40% to 90% in steps of 10%), they found that changing expression intensity had particular effects on angry, fear, and sad. An explanation for the differences in findings may be due to the use of a human observer for recognition as compared to machine-based for this study. It is possible that there are small changes in facial expressions that an automatic face recognition system would miss. Manfredonia et al. (2018), for example, said that distinctions such as gaze hate can be recognised by human coders but not by the automated system or the selected AUs employed. Furthermore, as stated by Wells et al. (2016), intensity effects on expression types do not appear to be consistent and may vary depending on the emotional content of the expressions.

4.3.3 Conclusions

In conclusion, the objective of this chapter is to investigate the impact of expression types and expression strengths on recognition performance. A number of studies have indicated that when facial expression strengths increase, recognition performance decreases. Intensity level one contributes the highest number of correct faces for each query set compared to other levels of intensity. This result is in line with the previous studies conducted by Smeets et al. (2011) and Smeets et al. (2012) using the same dataset, BU-3DFE. They stated that the decrease in performance could be due to facial muscle contractions of the soft tissue of the face deforming during extreme expression variations, which can affect the recognition performance.

However, looking at each type of expression in the training set (for each query set) gives a more complex picture. Although most of the correct faces are from intensity level one, several expressions in the training set show poor recognition at intensity level one but higher in other levels.

Encouraged by the intriguing outcome, a statistical analysis was conducted to determine the cause of the problem. There are interaction effects between expression types and expression strengths, according to the findings. Since the core tensor values cannot be used to establish the interaction effects, the ANOVA model has been shown to be able to overcome the said limitation, and confirms that the effect of intensity levels on recognition accuracy varies across expression types. Changes in intensity levels have particular effects on disgust, happy, and surprise expressions, but they have no particular effect on angry, fear, and sad expressions. These interaction effects have also been reported in the psychology domain (Montagne et al., 2007; Hoffmann et al., 2010; Wells et al., 2016). However, the researchers used human-based recognition, not machine-based recognition.

A study carried out by Smeets et al. (2012) had also investigated the factors that were considered to be influencing the recognition performance of the same dataset used in this study, BU-3DFE. The factors that they considered were expression types, intensity levels, and the number of expressions. They concluded that recognition performance increased as levels of intensity decreased. However, they did not perform a further investigation on the interaction effects among the covariates. This study can therefore be considered as an extension of their findings.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1 Conclusion

The goal of this research is to show how the tensor technique may be used to model and analyse face datasets. The face variants can be arranged in different parsimonious spaces according to the multilinear representation of a tensor model. This method has the advantage of being able to properly suit the statistical analysis, resulting in a more accurate face analysis.

The first contribution is in the method of centring. Since median is a reliable measure of location, the results demonstrate that median centring might be used instead of the existing mean centring method to reduce projection errors and increase the variability carried by each principal component. The singular values for median centring are 1.1595×10^5 , 119.1206, 27.1058, 17.7281, and 13.2304 for the first, the 100th, the 500th, the 1000th, and the 1296th dimensions, respectively. On the other hand, for mean centring, the values are 1.1590×10^5 , 117.7366, 27.0815, 17.7239, and 3.4135×10^{-9} . It shows that singular values obtained from the median centring technique are slightly higher (the difference is small since there are not many outliers in the BU-3DFE face dataset) than mean centring, meaning that the variation of each component is maximised.

The following contribution proposes a median projection operator that could increase the overall variations carried by the components. The variety carried by each component decreases as the dimensions rise. However, the multidimensional space of a tensor object allows the proposed median projection operator to consider variations in each mode,

resulting in components with fewer dimensions and maximising the total variations. On the other hand, the PCA projection operator considers variations along the subject mode only, resulting in components with high dimensions and fewer variations. As a result, the PCA-based FR system cannot distinguish between different centring techniques applied (it gives similar recognition rates for all the three techniques applied: mean, median, and uncentred), and gives lower recognition rates compared to the median tensor-based FR system. The recognition rates for the median-centred tensor-based system are 95.83%, 92.59%, 98.61%, 94.91%, 94.44%, and 65.74% for angry, disgust, fear, happiness, sadness, and surprise, respectively, whereas for the PCA-based system, the values are 83.80%, 82.87%, 90.28%, 62.04%, 87.96%, and 37.04%.

One of the keys to having robust automatic face recognition is understanding the sources of variability that impact face recognition performance. This source of variability, whether it is intra-subject or inter-subject, can be exploited in designing a good face recognition algorithm. As a result, the next second contribution in this research is the use of the hybrid framework of the tensor-ANOVA model as an alternative to the core tensor or ROC graph to establish the interaction effects that exist among intra-subject variations (expression types and expression strengths). Although Smeets et al. (2011) and Smeets et al. (2012) had mentioned that these two face variants give significant effects on recognition performance (using the same dataset), the researchers did not discuss the interaction effects, and did not represent the data using a tensor model. As a result, this research adds to the existing body of knowledge.

In general, as expressive strength increases, recognition performance declines. When compared to other degrees of intensity, intensity level one contributes the most right faces

for each query set. However, interaction effects between expression types and expression strengths have been discovered. The impact of intensity levels on recognition accuracy varies depending on the type of expression. Changes in intensity have a distinct effect on disgust, happiness, and surprise emotions, but had no effect on anger, fear, or sadness expressions.

Another important point to be noted is that, since this study applied a tensor-ANOVA based approach, a special combination of multilinear global approach with intrinsic facial properties was successfully introduced. It allows, on the one hand, a global approach to the data structure and, on the other hand, detailed insight into the variability of each descriptor. The median tensor-based algorithm, together with an ANOVA model to establish the relationship among the face variants, was developed specifically to be used in the face recognition paradigm. It can, however, be used to a variety of other research topics.

5.2 Limitations and Directions of Future Research

A tensor technique was applied to analyse 3D face models. The recommended model and approach were created specifically for the purpose of the study, in order to provide a more accurate discussion and conclusion. The tests that were carried out can be extended in the following ways:

5.2.1 Applying a Two-Way Centring Technique

The centring strategy proposes a median statistic as an alternative to the mean statistic in order to reduce projection errors. Because the median projection operator would enhance the variability conveyed by each principal component, the focus of this research is on one-way centring as a starting point and the benefits of utilising the median. The focus of future research should be on two-way centring.

5.2.2 Deep Learning Approach

Nowadays, deep learning is a promising field of study. It would be interesting to combine the proposed median tensor-ANOVA model with a neural network technique to extend it. Despite the fact that Cichocki (2014) discussed the core concept of a tensor network, the researcher did not investigate the interaction effects. A neural network diagram is appropriate in portraying the cause and effect of a problem being investigated. The constraints of a core tensor and ROC in establishing interaction effects would then be resolved in a future investigation. Future research on the 'how' and 'why' interaction effects exist should be considered.

5.2.3 Considering a Non-Linear Case

A multilinear object is a tensor. A tensorface is a multilinear (multiple linear) representation of multiple causative factors in a set of faces. Multilinear PCA is a well-known method for modelling face datasets. The face dataset is considered in a linear form in this work, which uses the fundamental tensor technique. A non-linear instance could be the focus of future research. In other words, instead of using a typical PCA on each mode, a non-linear PCA might be used. This method can be used to determine the non-linear system's underlying dynamics or patterns (Friston et al., 2000).

REFERENCES

- Abdulhamed, M.A., & Othman, Z.I. (2019). A Measure of Happiness and Sadness Based on Facial Expression Recognition. In *Proceeding of the 10th International Scientific Conference on Geophysical, Social, Human and Natural Challenges in a Changing Environment*, (pp. 1-11).
- Abdurrahim, S.H., Samad, S.A., & Huddin, A.B. (2018). Review on the Effects of Age, Gender and Race Demographics on Automatic Face Recognition. *The Visual Computer*, 34(11), 1617–1630.
- Acar, E., & Yener, B. (2009). Unsupervised Multiway Data Analysis: A Literature Survey. *IEEE Transactions on Knowledge and Data Engineering*, 21(1), 6–20.
- Adini, Y., Moses, S., & Ullman, S. (1997). Face Recognition: The Problem of Compensating for Changes in Illumination Direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 721 – 732.
- Adolphs R., & Tranel, D. (2004). Impaired Judgements of Sadness but not Happiness Following Bilateral Amygdala Damage. *Journal of Cognitive Neuroscience*, 16(3), 453–462.
- Albiero, V., Krishnapriya, K.S., Vangara, K., Zhang, K., King, M.C., & Bowyer, K.W. (2020a). Analysis of Gender Inequality in Face Recognition Accuracy. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, (pp. 81-89).
- Albiero, V., Bowyer, K.W., Vangara, K., & King, M.C. (2020b). Does Face Recognition Accuracy Get Better with Age? Deep Face Matchers Say No. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, (pp. 250-258).

- Albiero, V., Zhang, K., & Bowyer, K.W. Cornell University (2020c). *How Does Gender Balance in Training Data Affect Face Recognition Accuracy?* [online] Available at: <https://arxiv.org/abs/2002.02934>. [Accessed on 8 April 2020].
- Alexandris, N., Gupta, S., & Koutsias, N. (2017). Remote Sensing of Burned Area Via PCA, Part 1: Centering, Scaling, and EVD vs SVD. *Open Geospatial Data, Software and Standards*, 2(17), 1-11.
- Aminu, M., & Ahmad, N.A. (2019). Locality Preserving Partial Least Squares Discriminant Analysis for Face Recognition. *Journal of King Saud University - Computer and Information Sciences*, 1-12.
- Avuglah, R.K., Asiedu, L., Nortey, E.N.N., & Yirenkyi, F.N. (2018). Recognition of Face Images Under Varying Head-Poses Using FFT-PCA/SVD Algorithm. *Far East Journal of Mathematical Sciences*, 103(11), 1769–1788.
- AIM. (2004). *Tensor Decompositions Workshop Discussion Notes*. American Institute of Mathematics, Palo Alto.
- Bader, B.W., & Kolda, T.G. (2006). Algorithm 862: MATLAB Tensor Classes for Fast Algorithm Prototyping. *ACM Transactions on Mathematical Software*, 32(4), 635-653.
- Bader, B.W., Kolda T.G., & others (2015). *MATLAB Tensor Toolbox Version 2.6*. [online] Available at: <http://www.sandia.gov/~tgkolda/TensorToolbox/>. [Accessed on 5 January 2016].
- Bickel, P.J. (1964). On Some Alternative Estimates for Shift in the P -Variate One Sample Problem. *The Annals of Mathematical Statistics*, 35(3), 1079-1090.

- Boik, R.J., & Marasinghe, M.G. (1989). Analysis of Nonadditive Multiway Classifications. *Journal of the American Statistical Association*, 84(408), 1059-1064.
- Bowyer, K., & King, M. (2019). Why Face Recognition Accuracy Varies Due to Race. *Biometric Technology Today*, 2019(8), 8-11.
- Bruveris, M., Gietema, J., Mortazavian, P., & Mahadevan, M. (2020). Reducing Geographic Performance Differentials for Face Recognition. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, (pp. 98-106).
- Cadima, J., & Jolliffe, I.T. (2009). On Relationships Between Uncentred and Column-Centred Principal Component Analysis. *Pakistan Journal of Statistics*, 25(4), 473-503.
- Calder, A.J., Young, A.W., Perrett, D.I., Etcoff, N.L., & Rowland, D. (1996). Categorical Perception of Morphed Facial Expressions. *Visual Cognition*, 3(2), 81–118.
- Carroll, J.D., & Chang, J. (1970). Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of ‘Eckart-Young’ Decomposition. *Psychometrika*, 35(3), 283–319.
- Cattell, R.B. (1965). Factor Analysis: An Introduction to Essentials, II. The Role of Factor Analysis in Research. *Biometrics*, 21(2), 405-435.
- Cevikalp, H., & Dordinejad, G.G. (2019). Discriminatively Learned Convex Models for Set Based Face Recognition. In *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 10122–10131).
- Cichocki, A. (2014). *Era of Big Data Processing: A New Approach Via Tensor Networks and Tensor Decompositions*. [online] Available at: <http://arxiv.org/abs/1403.2048>. 1–30. [Accessed on 25 November 2019].

- Cichocki, A., Mandic, D., Caiafa, C., Zhou, G., Zhao, Q., & Lathauwer, L. De. (2015). Tensor Decompositions for Signal Processing Applications: From Two-Way to Multiway Component Analysis. *IEEE Signal Processing Magazine*, 32(2), 145–163.
- Cichocki, A., Lee, N., Oseledets, I., Phan, A.-H., Zhao, Q., & Mandic, D.P. (2016). Tensor Networks for Dimensionality Reduction and Large-Scale Optimization. Part 1 Low-Rank Tensor Decompositions. *Foundations and Trends in Machine Learning*, 9(4–5), 249–429.
- Conde, C., Serrano, A., & Cabello, E. (2006). Multimodal 2D, 2.5D, and 3D Face Verification. In *Proceedings of the IEEE International Conference on Image Processing*, (pp. 2061–2064).
- Cook, C.M., Howard, J.J., Sirotin, Y.B., Tipton, J.L., & Vemury, A.R. (2019). Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems. In *Proceedings of the IEEE Transactions on Biometrics, Behaviour, and Identity Science*, (pp. 32–41).
- Cuculo, V., D’amelio, A., Grossi, G., Lanzarotti, R., & Lin, J. (2019). Robust Single-Sample Face Recognition by Sparsity-Driven Sub-Dictionary Learning Using Deep Features. *Sensors*, 19(1), 1-19.
- Cyganek, B. (2019). Thumbnail Tensor - A Method for Multidimensional Data Streams Clustering with an Efficient Tensor Subspace Model in the Scale-Space. *Sensors*, 19(19), 1-19.
- De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000a). A Multilinear Singular Value Decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4), 1253–1278.

- Ekman, P. (1999). Basic Emotions. In Dalglish, T., & Power, M. (Eds.), *Handbook of Cognition and Emotions*, (pp. 45-60). John Wiley & Sons.
- Estienne, F., Matthijs, N., Massart, D.L., Ricoux, P., & Leibovici, D. (2001). Multi-Way Modelling of High-Dimensionality Electroencephalographic Data. *Chemometrics and Intelligent Laboratory Systems*, 58(1), 59-72.
- Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- Fisher, R.A., & Mackenzie, W.A. (1923). Studies in Crop Variation: II. The Manurial Response of Different Potato Varieties. *The Journal of Agricultural Science*, 13(3), 311–320.
- Friston, K., Phillips, J., Chawla, D., & Büchel, C. (2000). Nonlinear PCA: Characterizing Interactions Between Modes of Brain Activity. *Philosophical Transactions of the Royal Society B*, 355(1393), 135-146.
- Fritzsche, R., Donaldson, P.M., Greetham, G.M., Towrie, M., Parker, A.W., Baker, M.J., & Hunt, N.T. (2018). Rapid Screening of DNA-Ligand Complexes Via 2D-IR Spectroscopy and ANOVA-PCA. *Analytical Chemistry*, 90(4), 2732-2740.
- Gagliardi, C., Frigerio, E., Burt, M., Cazzaniga, I., Perrett, D. I., & Borgatti, R. (2003). Facial Expression Recognition in Williams Syndrome. *Neuropsychologia*, 41(6), 733-758.
- Ghosh, M.N., & Sharma, D. (1963). Power of Tukey's Test for Non-Additivity. *Journal of the Royal Statistical Society, Series B*, 25(1), 213-219.
- Gilbert, N. (1963). Non-Additive Combining Abilities. *Genetics Research*, 4(1), 65–73.

- Givens, G.H., Beveridge, J.R., Draper, B., Grother, P., & Phillips, P.J. (2004). How Features of the Human Face Affect Recognition: A Statistical Comparison of Three Face Recognition Algorithms. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 381-388).
- Givens, G.H., Beveridge, J.R., Phillips, P.J., Draper, B., Lui, Y.M., & Bolme, D. (2013). Introduction to Face Recognition and Evaluation of Algorithm Performance. *Computational Statistics and Data Analysis*, 27, 236-247.
- Gnanadesikan, R., & Kettenring, J.R. (1972). Robust Estimates, Residuals, and Outlier Detection with Multi-Response Data. *Biometrics, Special Multivariate Issue*, 28(1), 81-124.
- Gnang, E.K., Elgammal, A., & Retakh, V. (2011). A Spectral Theory for Tensors. In *Annals de la Faculte des Sciences de Toulouse*, (pp. 801-841).
- Gollob, H.F. (1968). A Statistical Model Which Combines Features of Factor Analytic and Analysis of Variance techniques. *Psychometrika*, 33(1), 73-115.
- Grm, K., Štruc, V., Artiges, A., Caron, M., & Ekenel, H.K. (2018). Strengths and Weaknesses of Deep Learning Models for Face Recognition Against Image Degradations. *IET Biometrics*, 7(1), 81-89.
- Haardt, M., Roemer, F., & Del Galdo, G. (2008). Higher-Order SVD-Based Subspace Estimation to Improve the Parameter Estimation Accuracy in Multidimensional Harmonic Retrieval Problems. *IEEE Transactions on Signal Processing*, 56(7), 3198-3213.

- Harshman, R.A. (1970). Foundations of the PARAFAC Procedure: Model and Conditions for an 'Explanatory' Multi-Mode Factor Analysis. *UCLA Working Papers in Phonetics*, 16, 1–84.
- Harter, H.L., & Lum, M.D. (1962). An Interpretation and Extension of Tukey's One Degree of Freedom for Non-Additivity. *Aeronautical Research Laboratory Technical Report No. 62-313*, (pp. 1-42).
- Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H.C. (2010). Expression Intensity, Gender and Facial Emotion Recognition: Women Recognize Only Subtle Facial Emotions Better Than Men. *Acta Psychologica*, 135(3), 278–283.
- Honeine, P. Cornell University. (2014). *An Eigen Analysis of Data Centring in Machine Learning*. [online] Available at: <https://arxiv.org/abs/1407.2904>. [Accessed on 2 June 2018].
- Horn, R.A., & Johnson, C.R. (2013). *Matrix Analysis*, 2nd ed., Cambridge University Press.
- Huang, G.B., Ramesh, M., & Berg, T. (2007). Labelled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. In *Technical Report, University of Massachusetts*, (pp. 7-49).
- Huang, G.B., & Miller, E.L. (2014). Labelled Faces in the Wild: Updates and New Reporting Procedures. In *Technical Report, UM-CS-2014-003, University of Massachusetts*, (pp. 1-5).
- Jahan, S. (2018). On Dimension Reduction Using Supervised Distance Preserving Projection for Face Recognition. *Universal Journal of Applied Mathematics*, 6(3), 94–105.
- Jandová, M., & Urbanová, P. (2018). Sexual Dimorphism in Human Facial Expressions by 3D Surface Processing. *Journal of Comparative Human Biology*, 69(3), 98–109.

- Johnson, D.E., & Graybill, F.A. (1972a). Estimation of σ^2 in a Two-Way Classification Model with Interaction. *Journal of the American Statistical Association*, 67(338), 388–394.
- Johnson, D.E., & Graybill, F.A. (1972b). An Analysis of a Two-way Model with Interaction and no Replication. *Journal of the American Statistical Association*, 67(340), 862-868.
- Johnson, R.A., & Wichern, D.W. (2008). *Applied Multivariate Statistical Analysis*, 6th ed., Pearson International.
- Jolliffe, I.T. (2002). *Principal Component Analysis*. 2nd ed., Springer Series in Statistics.
- Jolliffe, I.T., & Cadima, J. (2016). Principal Component Analysis: A Review and Recent Developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 374(2065), 1-16.
- Kiers, H.A.L. (2000). Towards a Standardized Notation and Terminology in Multiway Analysis. *Journal of Chemometrics*, 14(3), 105–122.
- Kiers, H.A.L., & Van Mechelen, I. (2001). Three-Way Component Analysis: Principles and Illustrative Application. *Psychological Methods*, 6(1), 84-110.
- Kolda, T., & Bader, B. (2009). Tensor Decompositions and Applications. *SIAM Review*, 51(3), 455–500.
- Krishnapriya, K.S., Albiero, V., Vangara, K., King, M.C., & Bowyer, K.W. (2020). Issues Related to Face Recognition Accuracy Varying Based on Race and Skin Tone. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, (pp. 8-20).

- Kroonenberg, P.M. (1981). User's Guide to TUCKALS3. A Program for Three-Mode Principal Component Analysis. In *WEP-reeks, Vakgroep W.E.P., University of Leiden*.
- Kroonenberg, P.M. (1989). *Three-Mode Principal Component Analysis. Theory and Applications*, DSWO Press.
- Kroonenberg, P.M., & De Leeuw, J. (1980). Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithms. *Psychometrika*, 45(1), 69–97.
- Kruskal, J.B. (1981). Multilinear Models for Data Analysis. *Behaviormetrika*, 8(10), 1-20.
- Läuter, J., Glimm, E., & Kropf, S. (1996). New Multivariate Tests for Data with an Inherent Structure. *Biometrical Journal*, 38(1), 1–23.
- Leo, M.J., & Suchitra, S. (2018). SVM Based Expression-Invariant 3D Face Recognition System. *Procedia Computer Science*, 143, 619–625.
- Lestriandoko, N.H., Spreeuwes, L., & Veldhuis, R. (2018). The Behaviour of Principal Component Analysis and Linear Discriminant Analysis (PCA-LDA) for Face Recognition. In *Proceedings of the 2018 Symposium on Information Theory and Signal Processing in the Benelux*, (pp. 133-148).
- Li, C., & Mi, J.X. (2018). Sparse Factorial Code Representation Using Independent Component Analysis for Face Recognition. *Multimedia Tools and Applications*, 77(16), 21371–21392.
- Lin, H., Mueller-Bardorff, M., Mothes-Lasch M., Buff, C., Brinkmann, L., Miltner, W.H.R., & Straube, T. (2016). Effects of Intensity of Facial Expressions on Amygdala Activation Independently of Valence. *Frontiers in Human Neuroscience*, 10(646), 1-12.

- Liu, N., Zhang, B., Yan, J., Chen, Z., Liu, W., Bai, F., & Chien, L. (2005). Text Representation: From Vector to Tensor. In *Proceedings of the IEEE International Conference on Data Mining*, (pp. 725-728).
- Lu, B., Chen, J.-C., Castillo, C.D., & Chellappa, R. (2019). An Experimental Evaluation of Covariates Effects on Unconstrained Face Verification. *IEEE Transactions on Biometrics, Behaviour and Identity Science*, 1(1), 42-55.
- Lu, H., Zhang, L., Cao, Z., Wei, W., Xian, K., Shen, C., & van den Hengel, A. (2017). When Unsupervised Domain Adaptation Meets Tensor Representations. In *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 599-608).
- Luciano, G., & Naes, T. (2009). Interpreting Sensory Data by Combining Principal Component Analysis and Analysis of Variance. *Food Quality and Preference*, 20(3), 167-175.
- Mandel, J. (1959). The Analysis of Latin Squares with a Certain Type of Row-Column Interaction. *Technometrics*, 379-387.
- Mandel, J. (1961). Non-Additivity in Two-Way Analysis of Variance. *Journal of the American Statistical Association*, 56(296), 878–888.
- Mandel, J. (1969). The Partitioning of Interaction in Analysis of Variance. *Journal of Research of the National Bureau of Standards-Section B Mathematical Sciences*, 73(4), 309–328.
- Mandel, J. (1971). A New Analysis of Variance Model for Non-Additive Data. *Technometrics*, 13(1), 1–18.
- Manfredonia, J., Bangerter, A., Manyakov, N.V., Ness, S., Lewin, D., Skalkin, A., Boice, M., Goodwin, M.S., Dawson, G., Hendren, R., Leventhal, B., Shic, F., & Pandina, G.

- (2018). Automatic Recognition of Posed Facial Expression of Emotion in Individuals with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(1), 279–293.
- Marasinghe, M.G., & Boik, R.J. (1993). A Three-Degree of Freedom Test of Additivity in Three-Way Classifications. *Computational Statistics and Data Analysis*, 16(1), 47–61.
- Marasinghe, M.G., & Johnson, D.E. (1981). Testing Sub-Hypotheses in the Multiplicative Interaction Model. *Technometrics*, 23(4), 385–393.
- Maze, B., Adams, J., Duncan, J.A., Kalka, N., Miller, T., Otto, C., Jain, A.K., Niggel, W.T., Anderson, J., & Cheney, J. (2018). IARPA Janus Benchmark–C: Face Dataset and Protocol. In *Proceedings of the IEEE International Conference on Biometrics*, (pp. 158-165).
- Mendoza, F., & Lu R. (2015). Basics of Image Analysis. In Park B., & Lu R. (Eds.), *Hyperspectral Imaging Technology in Food and Agriculture. Food Engineering Series*, (pp. 9-56). Springer.
- Meshkinfamfard, S., Gorban, A., & Tyukin, I. (2018). Tackling Rare False-Positives in Face Recognition: A Case Study. In *Proceedings of the IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems*, (pp. 1592-1598).
- Milliken, G.A., & Graybill, F.A. (1970). Extensions of the General Linear Hypothesis Model. *Journal of the American Statistical Association*, 65, 797-807.

- Minoi, J.L., & Gillies, D.F. (2008). Sub-Tensor Decomposition for Expression Variant 3D Faces Recognition. In *Proceedings of the 3rd IEEE International Conference on Geometric Modelling and Imaging*, (pp. 108-133).
- Montagne, B., Kessels, R.P., De Haan, E.H., & Perrett, D.I. (2007). The Emotion Recognition Task: A Paradigm to Measure the Perception of Facial Emotional Expressions at Different Intensities. *Perceptual and Motor Skills*, 104(2), 589–598.
- Mood, A.M. (1941). On the Joint Distribution of the Medians in Samples from a Multivariate Population. *The Annals of Mathematical Statistics*, 12(3), 268-278.
- Muti, D., & Bourennane, S. (2007). Survey on Tensor Signal Algebraic Filtering. *Signal Processing*, 87(2), 237-249.
- Nagpal, S., Singh, M., Singh, R., & Vatsa, M. (2019). *Deep Learning for Face Recognition: Pride or Prejudiced?* [online] Available at: <http://arxiv.org/abs/1904.01219>. [Accessed on 11 January 2021].
- Omberg, L., Golub, G.H., & Alter, O. (2007). A Tensor Higher-Order Singular Value Decomposition for Integrative Analysis of DNA Microarray Data from Different Studies. *Proceedings of the National Academy of Sciences*, 104(47), 18371-18376.
- Orlóci, L. (1967). Data Centering: A Review and Evaluation with Reference to Component Analysis. *Systematic Biology*, 16(3), 208-212.
- Ouamane, A., Chouchane, A., Boutellaa, E., Belahcene, M., Bourennane, S., & Hadid, A. (2017). Efficient Tensor-Based 2D+3D Face Verification. *IEEE Transactions on Information Forensics and Security*, 12(11), 2751-2762.

- Ouanan, H., Gaga, A., Diouri, O., Ouanan, M., & Aksasse, B. (2020). Development of Deep Learning-Based Facial Recognition System. In Ezziyyani, M. (Ed.), *Advances in Intelligent Systems and Computing*, (pp. 45–52). Springer.
- Paone, J., Biswas, S., Aggarwal, G., & Flynn, P. (2011). Difficult Imaging Covariates or Difficult Subjects? - An Empirical Investigation. In *Proceedings of the International Joint Conference on Biometrics*, (pp. 1-8).
- Parkhi, O.M., Vedaldi, A., & Zisserman, A. (2015). Deep Face Recognition. In Xie, X., Jones, M.W., & Tam, G.K.L. (Eds.), *Proceedings of the British Machine Vision Conference*, (pp. 1-12). BMVA Press.
- Peter, M. (2020). *Facial Expression Synthesis Using Kernel Approach* (Unpublished master's thesis). Universiti Malaysia Sarawak.
- Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., & Worek, W. (2005). Overview of the Face Recognition Grand Challenge. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 947-954).
- Phillips, P.J., Wechsler, H., & Rauss, P. (1998). The FERET Database and Evaluation Procedure for Face Recognition Algorithms. *Image and Vision Computing*, 16(5), 295–306.
- Prats-Montalbàn, J.M., de Juan, A., & Ferrer, A. (2011). Multivariate Image Analysis: A Review with Applications. *Chemometrics and Intelligence Laboratory System*, 107(1), 1–23.
- Redfern, A.S., & Benton, C.P. (2019). Representation of Facial Identity Includes Expression Variability. *Vision Research*, 157, 123-131.

- Rizzi, A., & Vichi, M. (1995). Representation, Synthesis, Variability and Data Processing of a Three-Way Data Set. *Computational statistics & Data Analysis*, 19(2), 203-222.
- Rowe, H.A.H. (1979). Three-Mode Factor Analysis: Problems of Interpretation and Possible Solutions. *Australian Psychologist*, 14, 222-223.
- Savas, B., & Elden, L. (2007). Handwritten Digit Classification Using Higher-Order Singular Value Decomposition. *Pattern Recognition*, 40(3), 993-1003.
- Savran, A., Alyüz, N., Dibeklioglu, H., Çeliktutan, O., Gökberk, B., Sankur, B., & Akarun, L. (2008). Bosphorus Database for 3D Face Analysis. In Schouten, B., Juul, N.C., Drygajlo, A., & Tistarelli, M. (Eds.), *Biometrics and Identity Management. Lecture Notes in Computer Science*, (Vol. 5372, pp. 47-56). Springer.
- Sendik, O., & Keller, Y. (2019). DeepAge: Deep Learning of Face-Based Age Estimation. *Signal Processing: Image Communication*, 78, 368–375.
- Sengupta, S., Chen, J.-C., Castillo, C., Patel, V.M., Chellappa, R., & Jacobs, D.W. (2016). Frontal to Profile Face Verification in The Wild. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, (pp. 1-9).
- Serna, I., Morales, A., Fierrez, J., Cebrian, M., Obradovich, N., & Rahwan, I. (2020). Algorithmic Discrimination: Formulation and Exploration in Deep Learning-Based Face Biometrics. In *Proceedings of the Central Europe Workshop*, (pp. 146–152).
- Sidiropoulos, N.D., De Lathauwer, L., Fu, X., Huang, K., Palexakis, E.E., & Faloutsos, C. (2017). Tensor Decomposition for Signal Processing and Machine Learning. *IEEE Transactions on Signal Processing*, 65(13), 3551–3582.
- Sirovich, L., & Kirby, M. (1987). Low-Dimensional Procedure for the Characterization of Human Faces. *Journal of the Optical Society of America*, 4(3), 519–524.

- Smeets, D., Claes, P., Hermans, J., Vandermeulen, D., & Suetens, P. (2012). A Comparative Study of 3D Face Recognition Under Expression Variations. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 42(5), 710–727.
- Smeets, D., Hermans, J., Vandermeulen, D., & Suetens, P. (2011). A Quantitative Comparison of 3D Face Databases for 3D Face Recognition. In *Proceedings of the Society of Photo-optical Instrumentation Engineers (SPIE) Defence, Security, and Sensing 2011 Conference - Biometric Technology for Human Identification VIII*, (Vol. 8029, pp. 8029-8038).
- Srivastava, M., Kumar, A., Dixit, A., & Kumar, A. (2020). Real Time Attendance System Using Face Recognition Technique. In *Proceedings of the 2020th International Conference on Power Electronics and IoT Applications in Renewable Energy and Its Control*, (pp. 370–373).
- Sun, J.-T., Zeng, H.-J., Liu, H., Lu, Y., & Chen, Z. (2005). CubeSVD: A Novel Approach to Personalized Web Search. In *Proceedings of the 14th International Conference on World Wide Web*, (pp. 382-390).
- Takane, Y., & Shibayama, T. (1991). Principal Component Analysis with External Information on Both Subjects and Variables. *Psychometrika*, 56(1), 97–120.
- Tucker, L.R. (1964). The Extension of Factor Analysis to Three-Dimensional Matrices. In Gulliksen, H., & Frederiksen, N. (Eds.), *Contributions to Mathematical Psychology* (pp. 110-127). Holt, Rinehart & Winston.
- Tucker, L.R. (1966). Some Mathematical Notes on Three Mode Factor Analysis. *Psychometrika*, 31(3), 279–311.

- Tukey, J.W. (1949). One Degree of Freedom for Non-Additivity. *Biometrics*, 5(3), 232–242.
- Turk, M.A., & Pentland, A.P. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.
- Van Den Berg, R.G. (2015). *SPSS Repeated Measures ANOVA II*. [online] Available at: <https://www.spss-tutorials.com/spss-repeated-measures-anova-example-2/>. [Accessed on 12 August 2019].
- Vasilescu, M.A.O., & Terzopoulos, D. (2002a). Multilinear Image Analysis for Facial Recognition. In *Proceedings of the 16th IEEE International Conference on Pattern Recognition*, (pp. 511-514).
- Vasilescu, M.A.O., & Terzopoulos, D. (2002b). Multilinear Analysis of Image Ensembles: Tensorfaces. In *Proceedings of the European Conference on Computer Vision*, (pp. 447–460).
- Veltkamp, R.C., & ter Haar, F.B. (2007). SHREC 2007: 3D Shape Retrieval Contest. In *Technical Report, UU-CS-2007-015, Utrecht University*, (pp. 1-72).
- Vera-Rodriguez, R., Blazquez, M., Morales, A., Gonzalez-Sosa, E., Neves, J.C., & Proenca, H. (2019). FaceGenderID: Exploiting Gender Information in DCNNs Face Recognition Systems. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 2254–2260).
- Volfovsky, A., & Hoff, P.D. (2014). Hierarchical Array Priors for ANOVA Decompositions of Cross-Classified Data. *Annals of Applied Statistics*, 8(1), 19-47.
- Wang, H., & Ahuja, N. (2003). Facial Expression Decomposition. In *Proceedings of the 9th IEEE International Conference on Computer Vision*, (pp. 958-965).

- Wells, L.J., Gillespie, S.M., & Rotshtein, P. (2016). Identification of Emotional Facial Expressions: Effects of Expression, Intensity and Sex on Eye Gaze. *PLoS One*, 11(12), 1-20.
- Whitelam, C., Taborsky, E., Blanton, A., Maze, B., Adams, J., Miller, T., Kalka, N., Jain, A.K., Duncan, J.A., Allen, K., Cheney, J., & Grother, P. (2017). IARPA Janus Benchmark-B Face Dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 592-600).
- Wickremasinghe, W.N. (2005). A General Approach to Analysis of Three-Mode Data. *Journal of the National Science Foundation of Sri Lanka*, 33(1), 9–16.
- Wickremasinghe, W.N., & Johnson, D.E. (2002). Testing Sub-Hypotheses and Estimating σ^2 in the Nonreplicated Three-Way Multiplicative Interaction Model. *Communications in Statistics - Simulation and Computation*, 31(4), 605-618.
- Witten, I.H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed., Morgan Kaufmann Publishers.
- Wu, S., & Wang, D. (2019). Effect of Subject's Age and Gender on Face Recognition Results. *Journal of Visual Communication and Image Representation*, 60, 116–122.
- Xu, C., Tan, T., Li, S., Wang, Y., & Zhong, C. (2006). Learning Effective Intrinsic Features to Boost 3D-Based Face Recognition. In Leonardis, A., Bischof, H., & Pinz, A. (Eds.), *Computer Vision. Lecture notes in Computer Science*, (pp. 416-427). Springer.
- Yadav, J., Rajpal, N., & Mehta, R. (2019). An Improved Illumination Normalization and Robust Feature Extraction Technique for Face Recognition Under Varying Illuminations. *Arabian Journal for Science and Engineering*, 44(11), 9067–9086.

- Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M.J. (2006). A 3D Facial Expression Database for Facial Behaviour Research. In *Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition*, (pp. 211–216).
- Zare, A., Ozdemir, A., Iwen, M.A., & Aviyente, S. (2018). Extension of PCA to Higher Order Data Structures: An Introduction to Tensors, Tensor Decompositions, and Tensor PCA. *Proceedings of the IEEE*, 106(8), 1341–1358.

APPENDICES

Article Publications (Appendix A)

1. Ab Rahman, S., Minoi, J-L., & Ujir, H. (2018). Face perception using tensor approach. In: *Proceedings of the 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Sarawak, Malaysia, 3 - 6 December 2018. IEEE Publisher.
2. Ab Rahman, S., Peter, M., Minoi, J.L., & Ujir, H. (2019). Multiway analysis for face recognition. In: *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy (ICCSP '19)*, Kuala Lumpur, Malaysia, 19 – 21 January 2019. ACM Publisher.

Appendix B

Name	Data type	Texture	Total images	Authorization
Face Recognition Grand Challenge (FRGC) Version 1 (Phillps et al., 2005)	range image	yes	943	University of Notre Dame
Face Recognition Grand Challenge (FRGC) Version 2 (Phillps et al., 2005)	range image	yes	4007	University of Notre Dame
Race Recognition and Artificial Vision (FRAV) 3D (Conde et al., 2006)	mesh	yes	1696	King Juan Carlos University
The Institute of Automation, Chinese Academy of Sciences (CASIA) (Xu et al., 2006)	mesh	no	4059	Chinese Academy of Sciences
BU-3DFE (Yin et al., 2006)	mesh	yes	2500	Binghamton University
SHREC2007: 3D Shape Retrieval Contest (Veltkamp & ter Haar, 2007)	mesh	yes	1000	Utrecht University
Bosphorus (Savran et al., 2008)	point cloud	yes	4652	Boğaziçi University