

Hyperspectral Data Acquisition and Its Application for Face Recognition

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Woon Cho
December 2015

Copyright © 2015 by Woon Cho
All rights reserved.

Acknowledgements

First and foremost, I thank God for blessing all my efforts in completing this dissertation. I wish to thank my beloved wife, Jinyoung Yang, who has been with me all the way, has always helped me in a time of need, and has stood by me through all my travails even at the most difficult period of my Ph.D. program. Her support, tolerance, encouragement, quiet patience, and unwavering love were undeniably the bedrock of my life. Along with her, I would like to thank my children, Diana Eunhae and Joshua Yoon for their devoted love. My sincere gratitude is due to my parents, Yongnam Cho and Heeyoung Jung for their love, support, patience, and unwavering belief in me throughout my entire life.

I am indebted to Dr. Koschan and Dr. Mongi A. Abidi for their guidance, invaluable suggestions, insights regarding the ideas presented in this dissertation. Their mentorship was paramount in solving various problems throughout my research. My sincere gratitude is due to Dr. Koschan for his willingness to assist anytime when needed and for numerous discussions and useful suggestions. I also wish to thank Dr. Qi, and Dr. Alexiades for their support and accepting the responsibility to serve on my dissertation committee. I sincerely appreciate all of their time and assistance as I navigated this process.

Many thanks are due to the entire faculty of the Department of Electrical Engineering and Computer Science at the University of Tennessee for their various help and support throughout my graduate studies. I am particularly grateful to Dr. Djouadi for his invaluable help and his suggestions to develop the approach for a reduced-order spectral data modeling based on local proper orthogonal decomposition during his class.

Finally, I wish to express my sincere appreciation and thanks to all of my collaborators at the Imaging, Robotics, and Intelligent Systems Laboratory (IRIS Lab.). I share all of my achievements with them. I especially thank Jacob D'Avi for spending a great deal of time to help me throughout my graduate program and to revise my publications and became a close friend. This acknowledgements would not be completed without acknowledging my best friends, Lee Leadbetter, Yonggil Choi, Sookyung Park, and Hunju Yi, who prayed, encouraged, and supported me morally throughout this program.

Abstract

Current face recognition systems are rife with serious challenges in uncontrolled conditions: e.g., unrestrained lighting, pose variations, accessories, etc. Hyperspectral imaging (HI) is typically employed to counter many of those challenges, by incorporating the spectral information within different bands. Although numerous methods based on hyperspectral imaging have been developed for face recognition with promising results, three fundamental challenges remain: 1) low signal to noise ratios and low intensity values in the bands of the hyperspectral image specifically near blue bands; 2) high dimensionality of hyperspectral data; and 3) inter-band misalignment (IBM) correlated with subject motion during data acquisition.

This dissertation concentrates mainly on addressing the aforementioned challenges in HI. First, to address low quality of the bands of the hyperspectral image, we utilize a custom light source that has more radiant power at shorter wavelengths and properly adjust camera exposure time corresponding to lower transmittance of the filter and lower radiant power of our light source.

Second, the high dimensionality of spectral data imposes limitations on numerical analysis. As such, there is an emerging demand for robust data compression techniques with loss of less relevant information to manage real spectral data. To cope with these challenging problems, we describe a reduced-order data modeling technique based on local proper orthogonal decomposition in order to compute low-dimensional models by projecting high-dimensional clusters onto subspaces spanned by local reduced-order bases.

Third, we investigate 11 leading alignment approaches to address IBM correlated with subject motion during data acquisition. To overcome the limitations of the considered alignment approaches, we propose an accurate alignment approach (A_3) by incorporating the strengths of point correspondence and a low-rank model. In addition, we develop two qualitative prediction models to assess the alignment quality of hyperspectral images in determining improved alignment among the conducted alignment approaches. Finally, we show that the proposed alignment approach leads to promising improvement on face recognition performance of a probabilistic linear discriminant analysis approach.

Table of Contents

1	Introduction.....	1
1.1	Motivation.....	1
1.1.1	Problem description	3
1.1.2	State of the art techniques	5
1.1.3	Advanced technique via hyperspectral imaging	9
1.2	Research contributions.....	9
1.3	Organization of the dissertation.....	14
2	Fundamentals of Hyperspectral Imaging	16
2.1	Hyperspectral imaging (HI)	16
2.1.1	Hyperspectral imaging technologies.....	18
2.1.2	Hyperspectral image formation.....	20
2.2	Summary	22
3	Review of Existing Techniques for Major Challenges.....	23
3.1	Existing face databases	23
3.1.1	CMU	24
3.1.2	IRIS-M	26
3.1.3	PolyU-HSFD.....	28
3.1.4	Stanford.....	29
3.1.5	UWA-HSFD	31
3.2	High dimensionality.....	32
3.3	Inter-band misalignments.....	33
3.4	Summary	35
4	Development of New Face Database	36
4.1	Imaging modules.....	37

4.2 Data acquisition and calibration.....	41
4.3 New face database (IRIS-HFD-2014).....	48
4.3.1 Version 1 (IRIS-HFD-2014-V1).....	49
4.3.2 Version 2 (IRIS-HFD-2014-V2).....	50
4.4 Summary	52
5 Proposed Spectral Data Compression Method	54
5.1 Background on proper orthogonal decomposition.....	54
5.2 Theory of proper orthogonal decomposition	55
5.3 Local proper orthogonal decomposition	57
5.3.1 Clustering a set of the sample data	57
5.3.2 Construction of local reduced-order bases.....	59
5.3.3 Choice of parameter k	60
5.4 Experimental Results	61
5.4.1 Test databases	61
5.4.2 Error metrics	63
5.4.3 Test on hyperspectral signal databases	65
5.4.4 Test on hyperspectral image databases	67
5.5 Summary	68
6 Design of New Framework for Face Alignments	69
6.1 Two conventional alignment approaches (FBB and EC)	69
6.2 Iterative convex optimization approaches (RASL and ORIA).....	73
6.3 Landmark-based alignment approaches (LAs)	76
6.4 Image alignment approaches (LK and SIFTFlow)	80
6.5 Summary	83
7 Proposed Qualitative Assessments of Alignment Methods	84
7.1 Curvature-based alignment quality assessment	85
7.2 Hue-based alignment quality assessment.....	88
7.3 Evaluation of alignment quality assessment	92
7.3.1 Choice of σ for curvature model similarity.....	94
7.3.2 Results of CMS on rigid object sets.....	95

7.3.3 Results of CMS on IRIS-HFD-2014-V1.....	97
7.3.4 Results of HUQA on rigid object sets.....	100
7.3.5 Results of HUQA on IRIS-HFD-2014-V1	102
7.4 Summary	105
8 Proposed Accurate Alignment Approach (A₃) Based on a Mixture Model.....	107
8.1 Accurate alignment approach using a mixture model.....	107
8.2 Experimental results.....	112
8.2.1 Results on rigid object sets	112
8.2.2 Results on UWA-HSFD.....	118
8.2.3 Results on IRIS-HFD-2014.....	120
8.3 Summary	127
9 Evaluation of Face Recognition	128
9.1 Probabilistic linear discriminant analysis (PLDA)	128
9.2 Experimental Results	132
9.2.1 Results on UWA-HSFD.....	132
9.2.2 Results on IRIS-HFD-2014.....	135
9.3 Summary	136
10 Conclusion	137
10.1 Directions for future research	139
Bibliography	141
Vita	156

List of Tables

Table 3.1: Overview of hyperspectral face databases (HFDs) in this dissertation.	23
Table 5.1: Summary of the performance results MSE-GIA, MINIMAX, GMM, GPOD, and LPOD on the MUNSELL database.	64
Table 5.2: Summary of the performance results MSE-GIA, MINIMAX, GMM, GPOD, and LPOD on the NCS database.	65
Table 5.3: Summary of the performance results ICA, NMF, RSNMU, RSPNMU, GPOD, and LPOD on the CAVE database.	66
Table 5.4: Summary of the performance results ICA, NMF, RSNMU, RSPNMU, GPOD, and LPOD on the in-house database.	67
Table 7.1: A comparison of the reported scores for the sample sets in Fig. 7.2.	88
Table 7.2: Results of two FR-IQA approaches and the proposed CMS on two rigid object sets.	97
Table 7.3: A comparison of prediction accuracies of two FR-IQA approaches and the proposed CMS on IRIS-HFD-2014-V1.	98
Table 7.4: Parameters of the estimated regression lines for FBB, EC, RASL, and ORIA via the proposed CMS.	98
Table 7.5: Results of five NR-IQA approaches and the proposed HUQA on two rigid object sets.	101
Table 7.6 A comparison of prediction accuracies of five NR-IQA approaches and the proposed HUQA on IRIS-HFD-2014-V1.	103
Table 7.7: Parameters of the estimated regression lines for FBB, EC, RASL, and ORIA via the proposed HUQA.	105
Table 8.1: Results of averaging alignment errors of the conducted alignment approaches on translated rigid sets.	114
Table 8.2: Results of averaging alignment errors of the conducted alignment approaches on rotated rigid sets.	114
Table 8.3: Results of averaging alignment errors of the conducted alignment approaches on the rigid sets artificially disturbed by Δ and ϕ	116

Table 8.4: Results of averaging alignment errors of the conducted alignment approaches on UWA-HSFD.	119
Table 8.5: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V1 in session 1.	121
Table 8.6: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V1 in sessions 2 and 3.	121
Table 8.7: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V2.	127
Table 9.1 Comparison of the first rank identification rate based on 48 factors of PLDA on UWA-HSFD.	134
Table 9.2 Comparison of the first rank identification rate based on 40 factors of PLDA on UWA-HSFD once resizing the face cubes to $30 \times 30 \times 24$ ranging from 490 to 720 nm.	134
Table 9.3 Comparison of the first rank identification rate based on 64 factors of PLDA on IRIS-HFD-2014.	136
Table 10.1 Comparison of the identification rate of PLDA based on 20 factors using misaligned and aligned sets on IRIS-HFD-2014-V1.	140

List of Figures

Figure 1.1: Examples for authenticating individuals and permitting them to access physical or virtual domain: (a) ID or smart cards, (b) passwords and PINs, (c) a key, and (d) a magnetic card.	1
Figure 1.2: This figure shows examples of other biometric-based recognition techniques: fingerprint in (a) and (b), and iris in (c) and (d).	2
Figure 1.3: Examples of practical applications of face recognition: (a) identity of Boston Marathon bombing suspects and analysis of customer's behavior	3
Figure 1.4: Examples of the challenging problems in current face recognition systems. These images are taken from blurring, accessory, pose, aging, expression, illumination. .	4
Figure 1.5: Overview of the main contributions discussed in this dissertation.	11
Figure 1.6: (a) Spectral transmittances of the LCTF from 400 nm to 700 nm in 10 nm intervals and (b) spectral power distributions (SPDs) of synthetic lights: a halogen (H), a LED 40W (L), a projector with a blue polarizer filter (P), and mixtures of the studied lights (H+P, L+P, and H+L+P).	11
Figure 2.1: A comparison between a traditional broad-band image captured by a RGB color camera and hyperspectral face images (HFIs) captured by a hyperspectral imaging system with liquid crystal tunable filters (LCTFs) in the visible spectrum.	17
Figure 2.2: A hyperspectral face cube.	17
Figure 2.3: Examples of reflectance spectra for a single pixel in a hyperspectral face cube.	18
Figure 2.4: Hyperspectral imaging system using the LCTF in our experiments: (a) target, (b) light source, (c) LCTF, (d) Lens, (e) detector, and (f) system controller.	20
Figure 2.5: A sample sequence of ID: F009_02 in IRIS-HFD-2014-V1 using LCTFs ranging from 420 nm to 680 nm in 10 nm steps where only 27 bands are shown in this figure. Note that the total number of bands in our database is 29 bands from 420 nm to 700 nm in 10 nm steps.	21
Figure 3.1: Hyperspectral imaging system of the CMU database [Denes02]: (a) target, (b)-(d) light sources placed at -45° , 0° , and $+45^\circ$ according (a), and (e) hyperspectral imaging sensor.	24

Figure 3.2: CMU database: example of four different datasets gathered under four different illumination directions. The images taken from [Denes02] are sampled in the range of 500 nm, 700 nm, 900 nm, and 1090 nm.	25
Figure 3.3: Hyperspectral imaging system used to obtain the IRIS-M database [Chang08]. The left figure shows (a) target, (b) and (c) light sources, and (d) hyperspectral imaging sensor. The right figure shows (a) target, (b) natural illuminant (sun), and (c) hyperspectral imaging sensor.	26
Figure 3.4: The normalized SPDs in terms of four different light sources utilized in the IRIS-M database. The figures are taken from [Koschan11].	27
Figure 3.5: Example samples of multispectral face images (MFIs) in IRIS-M database [Chang08]: (a) indoor MFIs sampled for every 40 nm steps in the range from 580 nm to 700 nm under halogen lamps; and (b) outdoor MFIs sampled in the range from 580 nm to 700 nm in 40 nm steps under daylight	27
Figure 3.6: Hyperspectral imaging system employed to obtain the PolyU-HSFD [Di10]: (a) target in three different poses, (b) light source, and (c) hyperspectral imaging sensor. ...	28
Figure 3.7: HFI sequences of PolyU-HSFD in the front (F), right (R), and left (L) views where HFIs are sampled for every 50 nm step in the range from 500 nm to 650 nm. These sample images are taken from [Di10].	28
Figure 3.8: Example of an artifact in Stanford database resulting from eye blinking during data acquisition. This sample image is taken from [Skauli13].	29
Figure 3.9: Sample sRGB images in the Stanford database taken from [Skauli13].	30
Figure 3.10: (a) CIE 1931 XYZ-CMF and (b) spectral power distribution (SPD) of CIE illuminants.	30
Figure 3.11: Examples of inter-band misalignments in UWA-HSFD [Uzair13], [Uzair15]. Each of the estimated warps from a fixed bounding box-based alignment approach (FBB) is depicted by a rectangle in each band. As shown in the bands at 680 nm and 720 nm, there are IBMs in this database.	31
Figure 3.12: Examples of the presence of IBM artifacts in the sRGB color space (best viewed in color). As illustrated in this figure, IBMs result in the problem of motion blurring combined with distorted colors that are caused by spectral distortion. The sRGB image is generated from ID: F009 02 in IRIS-HFD-2014-V1.	34
Figure 4.1: A comparison of three facial databases collected by the LCTF sensors where in each row from top to bottom, sample HFIs covering the visible range from 420 to 690 nm in 30 nm intervals are taken from (a) PolyU-HSFD, (b) IRIS-M, and (c) IRIS-HFD-2014, respectively. Note that two HFIs at 420 nm and 450 nm in the IRIS-M database are added as a visual aid.	36
Figure 4.2: An example of our data acquisition system: (a) a lateral view and (b) a rear view.	38

Figure 4.3: Hyperspectral imaging modules in IRIS-HFD-2014. (a) VariSpec VIS mounted in front of XIMEA xiQ USB3.0 camera, (b) Lumia 5.1 Reef, (c) a controller for our light source, and (d) EasyView30 light meter and Ocean Optics USB2000.....	38
Figure 4.4: SPD of channel 1 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.....	39
Figure 4.5: SPD of channel 2 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.....	39
Figure 4.6: SPD of channel 3 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.....	40
Figure 4.7: SPD of channel 4 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.....	40
Figure 4.8: SPD of channel 5 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.....	41
Figure 4.9: (a) Result of the measured light falloff using X-Rite White Balance and Grayscale cards located on the left and right sides, respectively; and (b) the pseudo-colored result of (a) captured at 420 nm.	41
Figure 4.10: (a) Spectral power distribution (SPD) of Lumia 5.1 Reef and (b) tuned exposure times at each wavelength from 420 nm to 700 nm used in our data acquisition.	42
Figure 4.11: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the frontal view.	43
Figure 4.12: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the left profile.	44
Figure 4.13: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the right profile.	45
Figure 4.14: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the structural feature, i.e., glasses.	46
Figure 4.15: An example of color images taken from ID: 019_01 of IRIS-HFD-2014-V1.	47
Figure 4.16: Overview of selecting the subject areas within target images where I_i indicates the i th sub-band image of an input HFI set, B represents a binary mask, and p stands for pixel coordinates of I_i	48
Figure 4.17: A summary of gender and ethnic diversities in IRIS-HFD-2014-V1.	49
Figure 4.18: The aligned datasets (ID: 001 to 050) of IRIS-HFD-2014-V1 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.	50

Figure 4.19: The aligned datasets (ID: 051 to 130) of IRIS-HFD-2014-V1 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.....	51
Figure 4.20: A summary of gender and ethnic diversities in IRIS-HFD-2014-V2.	51
Figure 4.21: Sample HFI sequences for each pose in IRIS-HFD-2014-V2, after data calibration.	52
Figure 4.22: Sample color images representing each pose in IRIS-HFD-2014-V2.	53
Figure 4.23: Entire aligned datasets of IRIS-HFD-2014-V2 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.....	53
Figure 5.1: Comparison for overviews of (a) global POD and (b) local POD proposed in this dissertation. As the difference between global and local POD illustrated in (a) and (b), a clustering method in local POD splits A into k groups. For each group, the POD approximation is performed to construct local reduced-order bases (LROB). Then, we find the approximated A	58
Figure 5.2: Results for the cluster validity indices in Calinski-Harabasz (CH) criterion on the Munsell and Natural color system (NCS) databases. In both databases, the optimal number of clusters k is two.	61
Figure 6.1: Overview of the fixed bounding box (FBB) approach where $\kappa = 1$ in our experiments.	70
Figure 6.2: Overview of the eye coordinates (EC) approach.....	71
Figure 6.3: Results of FBB approach on ID: F009_02 in IRIS-HFD-2014-V1 where we select the canonical frame to be 140×160 pixels and $\xi = 80$ pixels for the distance between two eye coordinates throughout this dissertation.	72
Figure 6.4: Results of EC approach on ID: F009_02 in IRIS-HFD-2014-V1.....	72
Figure 6.5: Overview of RASL approach.	73
Figure 6.6: Results of RASL approach on ID: F009_02 in IRIS-HFD-2014-V1.....	74
Figure 6.7: Overview of ORIA approach.	74
Figure 6.8: Results of ORIA approach on ID: F009_02 in IRIS-HFD-2014-V1.	75
Figure 6.9: Example of a failure of ORIA on ID: F129_01 in IRIS-HFD-2014-V1 at 600 nm, 630 nm, 660 nm, and 690 nm.	75
Figure 6.10: A framework of landmark-based alignment approach.	76
Figure 6.11: Sample results of TSPM [Zhu12] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.	77

Figure 6.12: Sample results of SDM [Xiong12] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.	77
Figure 6.13: Sample results of CDM [Yu13] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.	78
Figure 6.14: Sample results of DRMF [Asthana13] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.	78
Figure 6.15: Sample results of IPCM [Asthana14] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.	78
Figure 6.16: Examples of inaccurate localizations of landmark points: (a) the results of TSPM [Zhu12] and (b) the results of IPCM [Asthana14] at 600 nm and 700 nm. As illustrated in (a) and (b).	79
Figure 6.17: Overview of LK approach.	80
Figure 6.18: Overview of SIFTFlow approach.	81
Figure 6.19: An example of a failure of LK approach on ID: F048_01 in IRIS-HFD-2014-V1.	81
Figure 6.20: An example of a failure of SIFTFlow approaches on ID: F048_01 in IRIS-HFD-2014-V1.	82
Figure 6.21: Examples of the aligned HFIs at 420 nm, 500 nm, 600 nm, and 700 nm according to (a) FBB, (b) EC, (c) RASL, and (d) ORIA. The color images in the fifth column are displayed using sRGB rendered with full sub-bands under CIE D65 where the regions of interest (ROIs) marked with rectangles are magnified and displayed in the last column.	83
Figure 7.1: Overview of the proposed framework (best viewed in color): (a) input, (b) alignment, and (c) assessing improved alignment via two proposed qualitative prediction models.	84
Figure 7.2: A comparison of the results of (b) gradient similarity model (GSM) [Liu12], (c) gradient magnitude similarity (GMS) [Xue14], and (d) the proposed method based on curvature model similarity (called CMS) at 420 nm, 550 nm, and 700 nm on the results of RASL [Peng12] applied to ID: F067_01 in IRIS-HFD-2014-V1 as shown in (a).	86
Figure 7.3: A comparison of the results of four alignment approaches with the corresponding sRGB images rendered under CIE D65. The HFIs are taken from ID: F067_01 in IRIS-HFD-2014-V1.	87
Figure 7.4: HSV color space.	89
Figure 7.5: Analyses of the effects of IBMs in the HSV color space. For the misaligned sRGB image in (a), we can observe that the distribution of the colors is more extensively spread over the HSV color space as shown in (b) and (c) in different views where the vertical axis is the V value, the horizontal distance from the axis is the S value, and the	

angle is the H value. However, the color distribution of the aligned sRGB image in (d) is concentrated near the red color of the hue component as illustrated in (e) and (f). Note that the dark colors with low V values in (d) are induced by the subject's clothes where the colors do not appear in (a).	90
Figure 7.6: Converting the color space from sRGB to HSV.	91
Figure 7.7: An illustration of H on a color wheel divided into six sectors associated with three primary and three pairwise mixed colors, ϕ , including red, yellow, green, cyan, blue, and magenta.	92
Figure 7.8: A comparison of the probability of the six representative colors (red, yellow, green, cyan, blue, and magenta) between (a) the misaligned and (b) the aligned sRGB images.	93
Figure 7.9: Examples of input HFI sets for rigid objects in (b). Note that these sRGB images are rendered with full sub-bands under CIE D65.	94
Figure 7.10: (a) shows the number of errors on IRIS-HFD-2014-V1 as increasing σ values from 1.0 to 2.0. The σ denotes the Gaussian scale and is a parameter for the proposed CMS metric. (b) indicates the subject IDs associated with the errors where a value of '1' represents the error. In (a), we can observe that CMS appears not to enhance the accuracy of the prediction to determine the most improved alignment by which the σ value is greater than 1.4. Note that whereas CMS obtain four errors in this experiment where we compare the scores of FBB with EC, RASL, and ORIA, there is no error if we associate the errors with only RASL and ORIA.	95
Figure 7.11: Effects of varying σ values at 420 nm, 550 nm, and 700 nm on the results of RASL.	96
Figure 7.12: Results of two FR-IQAs and the proposed CMS on rigid object sets.	97
Figure 7.13: Errors of two FA-IQAs and the proposed CMS on IRIS-HFD-2014-V1. Note that the errors are computed by counting the number of the subject IDs where the scores of FBB are higher than either the scores of RASL or ORIA. As can be shown in (c), there is no error in the CMS metric compared to GSM and GMS where the number of errors of (a) GSM and (b) GMS is 9 and 15, respectively.	99
Figure 7.14: Scatter plots with linear regression lines corresponding to results of the proposed CMS on IRIS-HFD-2014-V1.	100
Figure 7.15: Results of five state-of-the-art NR-IQAs and the proposed HUQA on rigid object sets. Note that the scores of SSEQ and CPBD are rescaled by adding 0.3 and 0.5, respectively.	102
Figure 7.16: One error HFI set found in HUQA. The HFI set is taken from ID: F_009_01 of IRIS-HFD-2014-V1.	103
Figure 7.17: Errors of five NR-IQA approaches and the proposed HUQA.	104

Figure 7.18: Scatter plots with the regression lines corresponding to the results of the proposed HUQA on non-rigid subject sets in IRIS-HFD-2014-V1.....	105
Figure 8.1: The shape of the LoG filter as $\sigma = 2.0$	108
Figure 8.2: Response of a 1-D LoG filter to a step edge function.	109
Figure 8.3: Haar wavelet filters.	110
Figure 8.4: Example of SUFT descriptor with the oriented 4×4 square sub-regions. For each sub-region, we compute the wavelet responses from 5×5 regularly spaced sample points.	110
Figure 8.5: Example of feature matching across two images such as the template and input images.	111
Figure 8.6: Basic scenario of estimating a similarity transformation model, τ , from the matched features.	111
Figure 8.7: Process pipeline of the proposed alignment approach.	112
Figure 8.8: Sample HFI sequences of rigid objects taken from 420nm, 500nm, 600nm, and 700nm.	113
Figure 8.9: Sample artificial HFI images under controlled conditions in each row from top to bottom: translations, rotations, and translations and rotations, respectively.	113
Figure 8.10: Comparison of the alignment errors of LoG, ORIA, and the proposed A3 on the translated rigid sets where each rigid set is translated along with both x - and y -directions in the interval Δ	115
Figure 8.11: Comparison of the alignment errors of LoG, ORIA, and the proposed A3 on the rotated rigid sets where each rigid set is rotated as $\phi \in 1, 2$	116
Figure 8.12: Comparison of the alignment errors of LoG, ORIA, and the proposed A3 on the rigid sets disturbed by artificially disturbed by Δ and ϕ in the front (F), left profile (L), and right profile (R) views for $\Delta = 1$ and $\phi = 2$	117
Figure 8.13: The aligned datasets of UWA-HSFD where full sub-bands (400 to 720 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.	118
Figure 8.14: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the CMS metric on UWA-HSFD.	119
Figure 8.15: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the HUQA metric on UWA-HSFD.	120
Figure 8.16: Examples of IBMs in IRIS-HFD-2014-V2. Each of the estimated warps from the FBB approach is depicted by a rectangle in each band.....	121

Figure 8.17: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the CMS metric on IRIS-HFD-2014-V1 in session 1.	122
Figure 8.18: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the HUQA metric on IRIS-HFD-2014-V1 in session 1.	123
Figure 8.19: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 the CMS metric on IRIS-HFD-2014-V1.....	124
Figure 8.20: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the HUQA metric on IRIS-HFD-2014-V1.....	124
Figure 8.21: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the CMS metric on IRIS-HFD-2014-V2.....	125
Figure 8.22: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A3 via the HUQA metric on IRIS-HFD-2014-V2.....	126
Figure 9.1: Graphical model of the PLDA: the j th image of the i th individual, x_{ij} , is modeled using a linear latent model with identities h_i , noise variables $w_{i,j}$, and parameter θ	129
Figure 9.2: PLDA model of IRIS-HFD-2014: (a) mean face and (b) noise covariance sRGB images.	131
Figure 9.3: Examples of the reconstructed sRGB images by the PLDA model trained from IRIS-HFD-2014.	132
Figure 9.4: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on UWA-HSFD.....	133
Figure 9.5: HFI sequences of UWA-HSFD ranging from 400 to 720 nm with 10 nm steps.	133
Figure 9.6: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on UWA-HSFD once resizing the face cubes to $30 \times 30 \times 24$ ranging from 490 to 720 nm.	134
Figure 9.9: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on IRIS-HFD-2014.....	135

1 Introduction

1.1 Motivation

In recent years, biometric-based techniques have received tremendous attention as the most reliable methods for authenticating individuals and permitting them to access physical and virtual domains based on physiological characteristics (e.g., fingerprints [Yoon12], face [Hua11], palm [Yang07], and iris [Sun14]) or behavioral traits [Liu06] (e.g., gait and signature). Three problems specific to non-biometric-based security techniques: 1) passwords and/or personal identification number (PINs), such as those required to access fiscal accounts utilizing an automated teller machine (ATM) and/or to access secure electronic domains, are intended to be a complex combinations of letters, numbers, special characters and/or symbols and, as such, are often challenging to commit to memory, can be stolen, and are able to be predicted by external parties; 2) IDs, smart cards, tokens, and keys can be misdirected, forgotten, lost, stolen or duplicated; and 3) magnetic cards can become corrupted and unreadable. Each of these issues, individually or any combination thereof, renders any real or intended measure of security null and void. The examples are shown in Fig. 1.1.

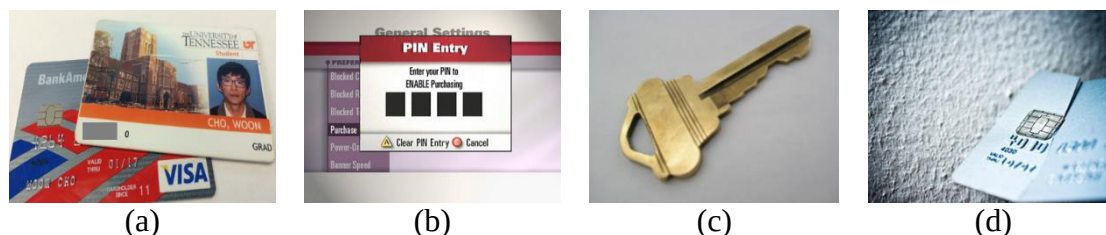


Figure 1.1: Examples for authenticating individuals and permitting them to access physical or virtual domain: (a) ID or smart cards, (b) passwords and PINs¹, (c) a key², and (d) a magnetic card³.

¹ <http://www.twcondemand.com/images/parentalcontrols/screen-purchase-pin-entry.jpg>.

² http://upload.wikimedia.org/wikipedia/commons/3/3c/House_key.jpg.

³ <http://cdn.sheknows.com/articles/broken-credit-card.jpg>.

As one of the most promising biometric modalities in Fig. 1.2, face recognition [Hua11] is, in brief, the process of authenticating the identity of an individual through facial imaging. The demand for robust face recognition techniques that successfully perform in realistic conditions and/or uncontrolled environments is continually growing. Due to the merits of its high accuracy and low intrusiveness in multiple and varied applications including, but not limited to, multimedia data management, digital entertainment, access control, financial transaction, and video surveillance, significant and extended attention has been given and continues to be directed towards face recognition technology.

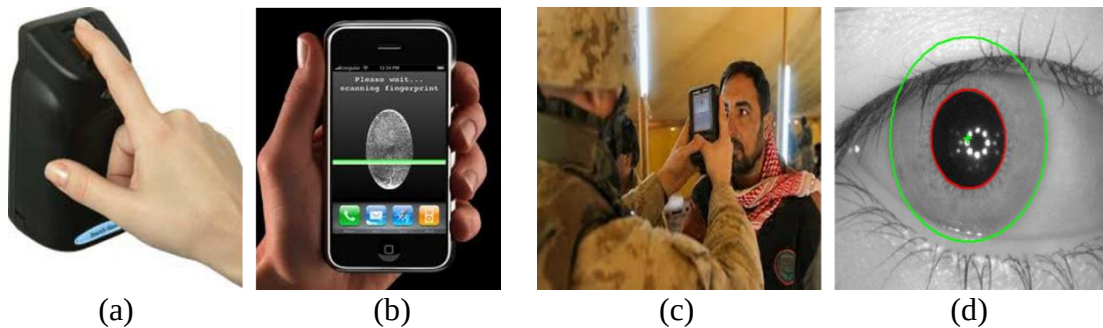


Figure 1.2: This figure shows examples of other biometric-based recognition techniques: fingerprint in (a)⁴ and (b)⁵, and iris in (c)⁶ and (d)⁷.

Compared to other biometric modalities depending on the level of participant cooperation, such as fingerprint and retina or iris recognition, face recognition has the distinct advantage of being non-collaborative or unaware. In other words, the personal identification system based on face recognition can effectively identify a particular individual without the particular individual's cooperation or knowledge. Figure 1.3 shows practical applications of face recognition in order to automatically identify the suspects or analyze customer's behavior. In spite of such innately distinguishing characteristics of face recognition systems, there still exist serious challenges under uncontrolled conditions, e.g., unrestrained lighting, actively changing facial expressions, pose variations, accessories, etc. These challenges have been the motivation of face recognition research but have yet to reach a mature stage of contending with them under highly unpredictable and uncertain circumstances [Liao13].

⁴ http://www.tiresias.org/research/guidelines/images/biometrics_fingerprint_03.jpg.

⁵ <http://swapnakm786.files.wordpress.com/2013/04/fingerprint-recognition-in-iphone.jpg>.

⁶ <http://www.visasouthafrica.org/wp-content/uploads/2011/11/VISA-SOUTH-AFRICA-BLOG32.jpg>.

⁷ http://www.neurotechnology.com/res/verieye_not_circles.jpg.

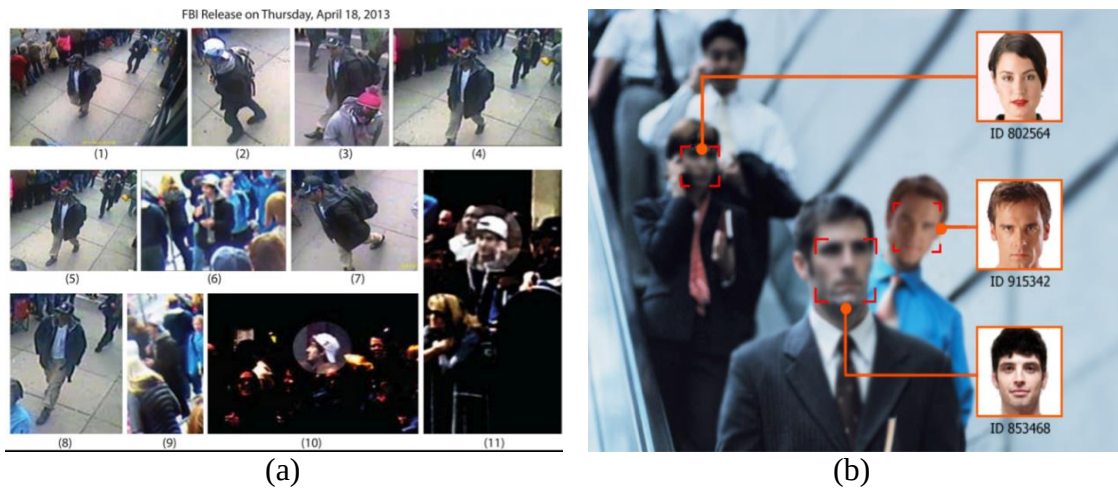


Figure 1.3: Examples of practical applications of face recognition: (a) identity of Boston Marathon bombing suspects⁸ and analysis of customer's behavior⁹

1.1.1 Problem description

The problem statement of face recognition can be formulated as follows: Given a new probe image and a set of labeled gallery faces of known individuals, how can it be matched against the gallery faces to be verified or determined as a known face or rejected? The way to solve the problem involves a well-designed system which can be configured as three key steps: 1) face detection; 2) feature detection; and, 3) feature matching.

The goal of face detection, the first fundamental step in face recognition systems, is to determine face localization in the image containing faces. It is crucial for many potential applications in face processing including face recognition, tracking, pose estimation, expression, etc., since they rely heavily on the outcomes of face detection. Next, feature detection is desired to detect and define the key components or patterns of the face appearance in the face region of interest consisting of a 2D intensity array. It is comprehensively affected to distinguish the face recognition techniques. The last step as feature matching is also referred to as face identification. Face identification plays a role in determining the identity of a person by comparing a set of features from an input image (probe) against a whole set of features for known individuals stored in a database (gallery).

The challenges intrinsically related to face recognition performances [Yang02], [Hua11] can be attributed to the following key factors as shown in Fig. 1.4:

⁸ <http://cdn.arstechnica.net/wp-content/uploads/2013/05/fbiimagesapril18-640x499.png>.

⁹ <http://www.netmechanic.co.za/assets/blog/face%20recognition%20technology.png>.

- Illumination variations: including, but not limited to, lighting variations and internal camera characteristics affect facial appearance when an image is captured.
- Pose variations: this factor introduces partial or entire occlusion of facial components including mouth, nose, or eyes.
- Facial expression: the shape of facial features is directly determined by facial actions or facial movements.
- Structural features: the variations in facial features including beards, mustaches, hair styles, and glasses influence the shape, color, and size of facial components.
- Time delay: this factor essentially lies in the constantly changing face appearance with advancing time increments.



Figure 1.4: Examples of the challenging problems in current face recognition systems. These images are taken from blurring¹⁰, accessory¹¹, pose¹², aging¹³, expression¹⁴, illumination¹⁵.

To construct robust techniques for face recognition, numerous approaches to address the challenging factors specific to appearance variations have been proposed. Hence, the

¹⁰ <http://www.bagnewsnotes.com/files/2013/08/Ilkka-Uimonen-Obama-photo-Wellford-interview.jpg>.

¹¹ <http://nimg.sulekha.com/others/original700/barack-obama-2010-1-22-21-14-15.jpg>.

¹² <http://wallpaperwonder.com/wp-content/uploads/2014/02/Obama-In-White-House-Wallpaper.jpg>.

¹³ http://www.oxy.edu/sites/default/files/assets/communications/Obama_Occidental.jpg.

¹⁴ http://cdn.breitbart.com/mediaserver/Breitbart/Big-Government/2013/Barack%20and%20Michelle/obama_OMG_AP.jpg

¹⁵ <http://muslimwriters.org/wp-content/uploads/2013/04/Barack-Obama-010.jpg>

preceding subsections deal with comprehensive and up-to-date reviews on existing techniques and critical discussions of the challenges.

1.1.2 State of the art techniques

With an aim to contribute a comprehensive and critical survey of current face recognition methods, we provide a detailed review of techniques to identify individuals according to two key steps: feature detection and feature matching. For the initial step in terms of the face detection, we refer the reader to the recent surveys [Yang02] [Hua11] [Samal92] because it is beyond the scope of this work to verify the identity of the individuals on specific facial images where the faces on the probe images and gallery images in database are already detected and aligned.

Existing techniques to detect features from a facial image can be classified into three categories [Samal92] [Chellappa95] [Tan06] [Abate07] [Jafri09]: 1) holistic approaches, 2) local methods, and 3) hybrid methods. Holistic approaches use the entire face region as the input to a face recognition system. One of the most widely used representations of the face region is eigen pictures [Kirby90], which are based on principal component analysis (PCA) which is a fast, simple, and practical approach. In contrast with holistic approaches, local methods can be classified into two categories: local feature-based and local appearance-based. Local feature-based approaches first process the input image to extract local distinctive features such as lines or fiducial points, or facial features (eyes, mouth, nose, etc.), and then compute the geometric relationships among a set of those features. Finally, to identify the faces, the vectors of geometric features are matched by using the techniques of standard statistical pattern recognition. Local appearance approaches mainly involve four steps: 1) local region partition, 2) feature extraction, 3) feature selection, and 4) classification. Compared with local feature-based approach, local appearance-based methods first divide a facial image into several subregions, and then, detect local features in the subregions. Hybrid approaches take advantage of both metrics of holistic and feature-based methods as human visual system.

Once the introduction of the holistic approach based on PCA [Kirby90] to model linear variation in high-dimensional data, various extended holistic methods have been proposed. Turk and Pentland [Turk91] proposed eigenfaces based on information theory, which indicate a small set of facial features spanning the significant variations among the known facial images. Fisherfaces based on Fisher's linear discriminant analysis (LDA) [Belhum.97] was introduced to find the most discriminative projection directions in eigenspace. Support vector machine (SVM) method classifying two different patterns in eigenspace was proposed by Phillips [Phillips99]. Bartlett et al. proposed independent component analysis (ICA)-based approach for face representation [Bartlett98] that was utilized by higher order statistics and generalized in PCA while enhancing the performance of face recognition by employing cosine similarity measure. Yang et al. [Yang04]

introduced two-dimensional PCA (2DPCA) based on 2D image matrices rather than 1D vectors. Similar to 2DPCA, Zhang and Zhou [Zhang05] introduced the bidirectional 2DPCA (2D2PCA). Xiaofei et al. [Xiaofei05] introduced Laplacianfaces based on locality preserving projections (LPP) which exploited a facial subspace to extract the essential face manifold structure while preserving local information. Kim et al. [Kim07] proposed linear discriminant function to maximize the canonical correlations of within-class sets and minimize the canonical correlations of between-class sets. The nonlinear approaches based on neural networks [Lin97] [Fleming90] in order to overcome misclassifications among the neighborhood classes are proposed. Moghaddam et al. [Mogh.97] proposed probabilistic-based method to estimate the probability of the difference between intra-personal facial variation and extra-personal facial variation. Naseem et al. [Naseem10] described a linear model representation to be classified by linear regression algorithm. Jiang et al. [Jiang10] proposed global harmonic subspace analysis (GHSA) based on Laplacian eigenmap. Recently, many researchers [Wright09] [Hu12] [Meng13A] [Meng13B] [Lu13] [Zhang13] introduced sparse representation to handle various changes in illumination, expression, occlusion, pose, etc. For further studies on comprehensive performance evaluation of holistic approaches, we strongly refer the interested reader to [Ashok10] and the references therein for more details where the reference provides comprehensive performance evaluation of about twenty five different subspace approaches under several real test conditions. The major advantages of the holistic approaches are two folds [Tan06]: 1) it completely maintains all the detailed shape information of the facial appearance that is useful for distinctive faces; and 2) it can obtain more global characteristics of faces than local feature-based approaches. However, as mentioned in [Zhang09], most of the holistic approaches are severely sensitive to changes in poses and scales.

In feature-based approaches, a set of fiducial points or facial features is only or mainly considered to be extracted from a limited region in the facial images. With reducing the 2D intensity matrices or images to a set of geometric vectors represented in templates, the feature-based approaches may be more suitable for saving storage cost and computational cost than holistic approaches. In addition, feature-based methods can have additional flexibility to recognize a part of facial features and incorporate the global information. The origin of the automated face recognition based on local features can be traced back to the early period in 1973 when Kanade [Kanade73] proposed simple algorithms to extract a vector of 16 facial parameters which were ratios of distances, areas, and angles where the extracted vectors stored in the databases were matched by using Euclidean distance to identify the probe face against gallery faces. Brunelli and Poggio [Brunelli93] illustrated 35 geometrical features for face representation including nose width and length, mouth position and chin shape to be used for a template matching with a Bayes classifier. However, those geometric feature-based methods have two drawbacks: 1) it is hard to exactly detect the geometric features on real facial images with some complicated cases; and 2) the only considerations for geometrical features without other information such as gray-level values of the images should be impertinent to fully represent distinctive faces. To overcome the

weakness of the purely geometrical features approaches, Manjunath et al. [Manj.92] proposed more powerful local feature representation based on Gabor wavelet decomposition, consisting of location and feature information where feature information containing the spatial and angular distance from neighboring feature points were connected to the edge with minimal distance, and then, constructed as a fixed topological graph. However, the fixed topological graph is criticized because it does not allow to update or modify the feature information once it is built. Thus, it is hard to be suitable for different variations in scale. To provide more flexibility to variations in illumination, translation, distortion, rotation, and scaling, Wiskott et al. [Wiskott97] proposed elastic bunch graph matching (EBGM) which was one of the most successful local feature-based approaches and was based on the dynamic link architecture (DLA) [Lades93] as an extension of classical artificial neural networks. EBGM is also called a deformable topology graph matching method. Whereas it is robust to changes in appearance, it requires higher computational cost and it does not consider the partial occlusions of key positions such as eyes, nose, mouse, etc.

In contrast with topology graph-based approach, Gao et al. [Gao02] proposed a compact face feature based on line edge map (LEM) with generic line segment Hausdorff distance measure which was used to cope with the difficulty of changes in lighting conditions. However, it has need of high computational expense for face image retrieval in the database. Gao and Qi [Gao05] proposed a directional corner points (DCP) including directional information which indicated the connectivity to its neighbors by exploiting point correspondence. The DCP approach is low storage cost and less sensitive to variation in illumination. However, its performance depends heavily on the precise localization of the detected feature points. Thus, it may be hard to handle major changes in shape or appearance. Recently, Meng et al. [Meng12] proposed monogenic binary coding for efficiently detecting local features to be decomposed into three complementary components such as amplitude, orientation, and phase. For face recognition, they compute histogram-based similarity measure. Liao et al. [Liao13] proposed an alignment-free face representation method based on multi-keypoint descriptors (MKD): Gabor ternary pattern (GTP) and scale invariant feature transform (SIFT) [Lowe04]. They showed that MKD approach achieved high accuracies in partial occlusion and pose changes under synthesized partial faces from FRGCv2.0 database, occluded holistic faces from AR database, and occluded or non-frontal faces collected in unconstrained scenarios from LFW and PubFig database.

For local appearance-based methods, we first provide the details of four steps, including local region partition, feature extraction, feature selection, and classification. In the first step of the local region partition, the facial images were divided into mostly rectangular windows, according to the size and shape of the local regions predefined. After the step of local region partition, local features were detected in the limited subregions. The feature detection and representation are most crucial to influence the performance of a face recognition system. For local features, gray-value features [Martinez02] [Chen04] [Tan05],

Gabor wavelet [Manj.92] [Wiskott97] [Lades93], Harr wavelet [Le04], etc., were commonly used for local feature descriptors. Whereas gray-value-based descriptor were simple yet no loss of texture information, Gabor wavelet-based descriptors is suitable for changes in illumination and geometric translation. For feature selection, it dealt with outliers or selected empirical information from a set of local features. PCA [Sirovich87] [Kirby90], LDA [Chen04], local statistics-based methods [Kim95] were employed to select coherent structures. The final step of classification applied to each component was for face recognition.

One of the most successful approaches based on local appearance was proposed by Martinez [Martinez02]. He described a local probabilistic approach to deal with partial occlusion and expression changes. The Mahalanobis distance defined by the Gaussian distributions as a solution to the localization problem was used to yield the best match. However, it needs high storage and computational cost under large databases. Tan et al. [Tan05] introduced an extension of the local probabilistic approach based on Self-Organizing Maps (SOM) instead of the mixture of Gaussians. In LDA-based approach, Fisher linear discriminant analysis (FLDA) often fails to partial occlusions since nonexistence of the intra-class scatter. To solve the problem, Chen et al. [Chen04] extended FLDA to local appearance approach. More specifically, they divided face images into a set of sub-images with the same dimensionality and trained them. Finally, FLDA was applied to each of the sub-patterns for the classification. Samaria and Young [Samaria94] introduced the hidden markov model (HMM)-based approach as a way of automatically segmenting face images and detecting useful features for face recognition. However, HMM-based approach has one drawback according to estimating the model types and model parameter in training samples. Ahonen et al. [Ahonen06] extended local binary pattern (LBP) texture features [Ojala06] by incorporating a spatially enhanced feature histogram or feature vector for global information representation. Recently, Chen et al. [Chen10] proposed Weber local descriptor (WLD) inspired by analyzing human perception of a pattern. WLD involves two components: 1) differential excitation and 2) orientation. Differential excitation was computed as the ratio between the intensity of the current pixel and the relative intensity differences of a current pixel against its neighbors. The orientation component was computed as the gradient orientation of the current pixel. They reported that WLD outperformed Gabor and SIFT on the Brodatz and KTH-TIPS2-a texture databases. Notwithstanding the successes of local-based approaches have been shown to be a possible way to deal with various changes in pose, expression, illumination, occlusion, etc. in the last few decades, there remain a variety of challenges in uncontrolled situations.

For hybrid-based approaches, they considered how to incorporate the advantages of holistic and local-based methods and prevent the disadvantages of them simultaneously. Kittler et al. [Kittler98] studied the problem of joining classifiers representing different patterns. While they struggled to provide various classified combination schemes with different assumptions and approximations, those problems are too difficult to derive the consistent

solutions due to local features and holistic features contain different characteristics to a certain degree.

1.1.3 Advanced technique via hyperspectral imaging

Despite such innately distinguishing characteristics of face recognition systems, there still exist serious challenges in uncontrolled conditions; e.g., unrestrained lighting, a range of facial expressions, pose variations, accessories, etc. Several of the major challenges are caused by variations in illumination conditions specific to current face recognition systems [Hua11]. Since the radiance utilized by face recognition system is proportional to the product of surface albedo and incident illumination, it is heavily dependent on illumination conditions. However, illumination conditions are not static and, as such, illumination variations are frequently occurring. Such illumination variations continually change the appearance of facial images and, accordingly, challenge the performance abilities of existing face recognition systems to produce accurate results. One viable way to resolve this complexity of face recognition is incorporating spectral information associated with hyperspectral imaging modality [Chang08], [Di10], [Pan03], [Koschan11].

Hyperspectral imaging is typically employed in response to these challenges by incorporating the spectral information within different sub-bands. Primarily, two factors contribute to this enhanced sub-band spectral information: 1) the invariance of illumination conditions that results from the recovery of objects' spectral properties, and 2) the ability to detect distinct patterns contained in human faces where such discriminative patterns cannot be captured by trichromatic (RGB) color or monochromatic (gray-scale) cameras. Indeed, hyperspectral imaging spanning the visible range as the portion of the electromagnetic spectrum is directly involved in human visual systems that can recover the spectral properties of objects in a scene under varying light conditions [Wandell95], [Robles.13]. As opposed to RGB or monochromatic (gray-scale) imagery, HI consists of more than three spectral measurements at different wavelengths that can convey more information about the objects presented in a scene where the same spatial region of interest is captured multiple times [Koschan11]. Hence, the information included in HI as a cube can be used for latent applications, such as face/object recognition, scene analysis, food security, object tracking, etc.

1.2 Research contributions

In this dissertation, our work concentrates mainly on three challenging problems in HI that must be resolved to improve face recognition performance: 1) low signal to noise ratios and low intensity values in the bands of the hyperspectral image specifically near blue bands, 2) high dimensionality of hyperspectral data [Gillis12], [Gillis13], [Gillis14],

[Nascimen.05], and 3) inter-band misalignment (IBM) when HI is applied to non-rigid objects [Uzair13]. The main contributions of this dissertation are shown in Fig. 1.5.

First, besides new opportunities of HI for practical use of face recognition, HI produces new challenges specific to low quality of the bands of hyperspectral images especially at shorter wavelengths [Uzair15]. This is due primarily to a low transmittance characteristic of the LCTF and low radiant power of synthetic lights at specific wavelength regions (particularly at blue bands) as presented in Fig. 1.6(a) and Fig. 1.6(b), respectively. In Fig 1.6(a), we can observe that the spectral transmittances of the LCTF decrease from long to short wavelength, implying that longer exposure time is a prerequisite for the short wavelength regions of the spectrum. In addition, it is necessary for the light source to have more radiant power near blue bands depicted in Fig. 1.6(b). As a result, our new database (IRIS-HFD-2014) was developed properly setting the camera exposure time according to low transmittances of the LCTF and low radiant power of a light source at the corresponding wavelength regions. IRIS-HFD-2014 was recently developed over multiple sessions in the IRIS laboratory at the University of Tennessee. IRIS-HFD-2014 contains a total of 19,346 facial images obtained from a participant subject base that includes 115 males (64%) and 64 females (36%), all of diverse ethnic backgrounds and diverse physical appearance. Similar to the PolyU-HSFD and IRIS-M databases, we employed an LCTF to acquire HFIs that cover the visible spectral range from 420 nm to 700 nm in 10 nm steps (29 narrow-bands). However, IRIS-HFD-2014 can provide more spectral information of diverse faces in the visible spectrum compared to the PolyU-HSFD and IRIS-M databases. This is because the HFIs in IRIS-HFD-2014 were collected by appropriately tuning the camera exposure time and using a custom light source that has more radiant power near blue bands.

Second, the high dimensionality of the spectral data causes limitations on physical experiments and detailed numerical analysis since spectral data include multiple sub-bands captured at each wavelength. For this reason, there is a pressing need for robust data compression techniques; specifically post-processing techniques, by means of extracting relevant basis functions from large quantities of high-dimensional spectral data. These techniques are crucial to analyze and represent a large set of spectral data and, finally, to model the processes. As mentioned in [Mittelman12], neighboring spectral samples are highly correlated, as the spatial resolution of the spectral data increases. In short, the neighboring spectral samples are more likely to involve corresponding spectral characteristics. These fundamental features of the spectral data should facilitate the potential of ROM techniques in order to address the challenges of compressing high-dimensional data. Accordingly, ROM has attracted increasing interest in recent decades due to its possibilities for a wide variety of applications including machine learning [Duda12], data mining [Kim11], and hyperspectral imaging [Robles.13]. For example, singular value decomposition (SVD) [Harde.01] and principle component analysis (PCA) [Cohen64], [Maloney86] were proposed to reduce the dimensionality of spectral data.

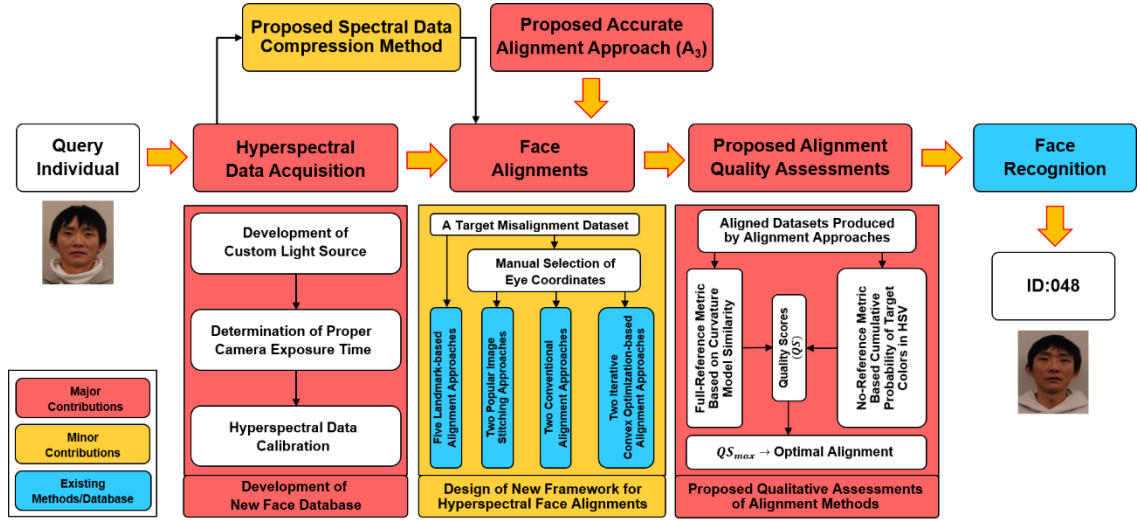


Figure 1.5: Overview of the main contributions discussed in this dissertation.

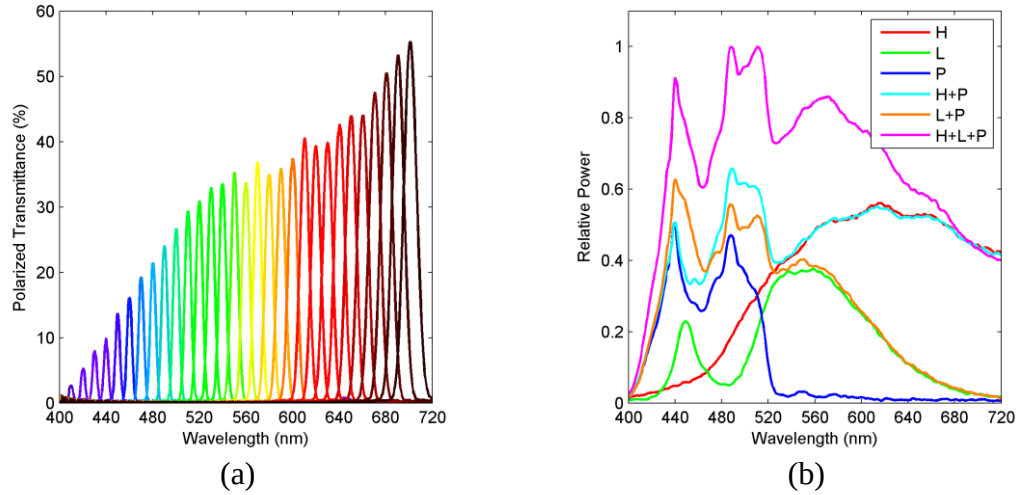


Figure 1.6: (a) Spectral transmittances of the LCTF from 400 nm to 700 nm in 10 nm intervals and (b) spectral power distributions (SPDs) of synthetic lights: a halogen (H), a LED 40W (L), a projector with a blue polarizer filter (P), and mixtures of the studied lights (H+P, L+P, and H+L+P).

SVD-based ROM is closely related to the PCA-based ROM as both are used to find the optimal subspace spanned by the principal directions by globally solving an optimization

problem. An alternative approach to solving an optimization is to conclude that the basis functions are fundamental for the approximation of the given function over some domain of interest. It will be shown that this specific approach-based ROM accurately corresponds to the proper orthogonal decomposition (POD) [Holmes98]. According to [Altmann13], the spectral reflectance could result from a non-linear function of the abundant vectors with respect to the spectral components. However, global solutions such as SVD, PCA, and POD could fail to capture the non-linear degrees of freedom, since they assume that the given data set belongs to a linear space [Gu08]. For this reason, we develop a local POD approach utilizing a standard clustering technique which can efficiently approximate the optimal solution of the non-linear function from local linear Euclidean distances by means of a few basis functions. Furthermore, the main advantage of the proposed POD lies in the fact that it only requires standard matrix computations in spite of its application to non-linear problems.

Third, HI with a sequential recoding scheme requires longer exposure times not only to acquire detailed facial tissue absorbance values over multiple narrow-bands in the visible spectrum and beyond but also to ensure high signal-to-noise ratio in low light conditions. The longer data acquisition times result in increasing the chances of subject motion that would otherwise nullify the benefit of hyperspectral imaging. Practically, an extended period of time during data acquisition produces a challenge for subject participants to remain motionless while staring at the camera under a bright light. In our data acquisition, for example, the illuminance on the target surface provided by the light source is approximately 1,140 lux, which is roughly equivalent to typical TV studio lighting, and the distance from the target to the light source is 60 cm (about 23.6 inches).

In cases where HFIs are acquired with constant exposure time there is less concern about IBMs as subject motion is insignificant with shorter data acquisition time. However, it is shown without adapting the camera exposure time at each wavelength that the IRIS-M and PolyU-HSFD databases have an essential limitation in obtaining spectral properties of facial tissue at multiple narrow-bands of the visible spectrum. Nonetheless, longer data acquisition time predictably creates additional IBMs that must be preferentially resolved to improve face recognition rate. In general, the better the inter-band alignment in the HFIs is, the higher the accuracy of face identification will be.

Towards addressing IBMs in HFIs we individually employ four alignment techniques: 1) conventional alignment approaches (AAs) based on selecting regions of interest, such as fixed bounding box-based AA (FBB) [Szeliski10] and eye coordinate-based AA (EC) [Denes02], [Di10], [Uzair13], 2) iterative convex optimization (ICO) processes for face alignment, such as RASL [Peng12] and ORIA [Wu12], 3) landmark-based AAs, such as DRMF [Asthana13], IPCM [Asthana14], SDM [Xiong12], CDM [Yu13], and TSPM [Zhu12], and 4) two popular image AAs such as Lucas-Kanade (LK) [Baker04] and SIFTFlow [Liu11]. The experimental determination of the promising AAs will be demonstrated in Chapter 5.

Additionally, we develop a novel framework within which to automatically predict the improved alignment among four selected AAs thusly addressing IBMs in our database. To determine the better-quality alignment among the selected AAs, we propose two different metrics for alignment quality assessment (AQA). We introduce a full-reference AQA based on a principal curvature map [Deng07], [Steger98] obtained by computing the maximum or minimum eigenvalues of a 2×2 Hessian matrix. By using average pooling scheme, the principal curvature map is used to evaluate the similarity index between reference and target images. To achieve high accuracy of AQA, we only utilize the maximum eigenvalues of the Hessian matrix as we heuristically found that the curvature map built from the minimum eigenvalues tends to be unpredictable where inconsistent curvature lines or edges result in reductions in the accuracy of AQA.

A no-reference (NR) AQA based on the cumulative probability of target colors in hue, saturation, and value (HSV) color space [Jayaraman10] is proposed for assessing the alignment quality (AQ) of a single sRGB image rendered with the entire sub-bands under CIE D65 [Moan14]. Inspired by the analysis of the color distribution of the misaligned sRGB image, we observed that the color distribution of the misaligned sRGB image is more widely spread over the HSV color space compared to the aligned sRGB image, which is typically concentrated on the red color of the hue component. This is due to spectral distortion associated with IBMs. In other words, since the colors of an sRGB image at the given pixels are generally estimated by the linear combinations of all of the measured reflectance spectra at each wavelength, the colors of the misaligned sRGB images can be altered to distorted colors when a few images out of the stacked HFIs are shifted by the subject's movement during data acquisition. Therefore, the distorted colors in sRGB images can be designed as a criterion to determine the alignment index for HFIs.

According to the experimental results in this dissertation, we find that the existing alignment approaches are essentially limited to address IBMs on the hyperspectral image sets including significant subject motion in our database. To overcome the drawbacks of the considered alignment approaches, we tackle the challenging problem by combining the strengths of a Laplacian of Gaussian (LoG)-based point correspondence and a low-rank model. The efficacy of the proposed alignment approach is verified with extensive experiments on two large-scale hyperspectral face databases (UWA-HSFD and IRIS-HFD-2014) that include a wide range of realistic IBMs. Furthermore, we show that the proposed alignment approach leads to better accuracies of face recognition performance by using a probabilistic linear discriminant analysis (PLDA) approach.

1.3 Organization of the dissertation

The remainder of this document is organized as follows:

- Chapter 2 introduces fundamentals of hyperspectral imaging relevant to this dissertation, including hyperspectral imaging technologies and hyperspectral image formations.
- Chapter 3 reviews five publically available hyperspectral face databases (HFDs): CMU, PolyU-HSFD, IRIS-M, Stanford, UWA-HSFD databases towards providing information on the key points of each of the considered databases. We also introduce substantial challenges in terms of high dimensionality of hyperspectral data and inter-band misalignment for more detail.
- Chapter 4 introduces our new database, called IRIS-HFD-2014, in order to serve as a benchmark for comprehensively and statistically evaluating the performance of current and future algorithms for hyperspectral face alignment and recognition.
- Chapter 5 describes the proposed spectral data compression method based on local proper orthogonal decomposition in order to account for the high dimensionality of hyperspectral data.
- Chapter 6 presents a new framework for face alignment to address inter-band misalignments in hyperspectral face images resulting from subject motion during data acquisition. Furthermore, for experimental determination of the promising alignment approaches, we individually employ four alignment techniques: 1) conventional alignment approaches based on selecting regions of interest, such as fixed bounding box-based alignment approach (FBB) and eye coordinate-based alignment approach (EC), 2) iterative convex optimization processes for face alignment, such as RASL and ORIA, 3) landmark-based alignment approaches, such as DRMF, IPCM, SDM, CDM, and TSPM, and 4) two popular image alignment approaches such as Lucas-Kanade (LK) and SIFTFlow.
- Chapter 7 describes the proposed qualitative assessments of alignment methods for predicting the most improved alignment among the selected alignment approaches: curvature-based alignment quality assessment (CMS) and hue-based alignment quality assessment (HUQA).
- Chapter 8 introduces the proposed robust alignment approach to tackle the challenging problem specific to inter-band misalignments in HI by combining the

strengths of a Laplacian of Gaussian (LoG)-based point correspondence and a low-rank model.

- Chapter 9 verifies that the improved alignment leads to better accuracies of face recognition performance. To evaluate face recognition performance, we employ a technique known as probabilistic linear discriminant analysis (PLDA), which models intraclass and interclass variance as multidimensional Gaussian to seek maximum facial discriminability.
- Chapter 10 provides a summary of accomplished work and future work.

2 Fundamentals of Hyperspectral Imaging

Hyperspectral imaging (HI) is a key tool in many applications of computer vision for art conservation, cultural heritage, remote sensing, scene understanding, etc. It can play a vital role in advanced studies on face recognition by extracting physical properties from facial tissue over multiple narrow bands in the visible spectrum and beyond. In recent years, numerous methods based on hyperspectral imaging have been developed with a significant improvement of the identification performance of individuals.

This Chapter presents an overview of the theoretical and practical issues to exploit hyperspectral face images. Before getting into the hyperspectral imaging technology, we distinguish hyperspectral imaging from the term multispectral imaging with a small number of channels. Indeed, the terms “multispectral” and “hyperspectral” imaging can be discriminated by the number of spectral bands and how narrow/wide the bands are [Robles.13]. For example, multispectral sensors typically provide a few wide bands such as three channels (red, green, and blue) and infrared bands, while hyperspectral sensors measure energy in narrower and more numerous bands than multispectral sensors where hyperspectral images can contain as many as 200 (or more) contiguous spectral bands.

2.1 Hyperspectral imaging (HI)

HI is an emerging technique that can improve face recognition performance by detecting and uncovering spectral signature, or reflectance spectrum, reflected by face tissue. It takes advantage of a rich variety of spectral information of facial tissue that cannot be resolved by a conventional imaging sensor. The composite images acquired in different bands for each spatial location of interest carry spectral reflectance information that is of particular relevance to illumination invariants [Westland04]. The fundamental idea of HI to recover the spectral properties of objects in a scene was inspired by the fact that human visual system would be able to recover those of objects from cone excitations [Wandell95]. Similar to human visual system, HI techniques incorporate conventional imaging and spectroscopy techniques in order to attain both spatial and spectral information [Hardeberg01]. For this reason, when HI is applied to face recognition systems, objects can be identified by the characteristics of their absorption or reflectance spectra.

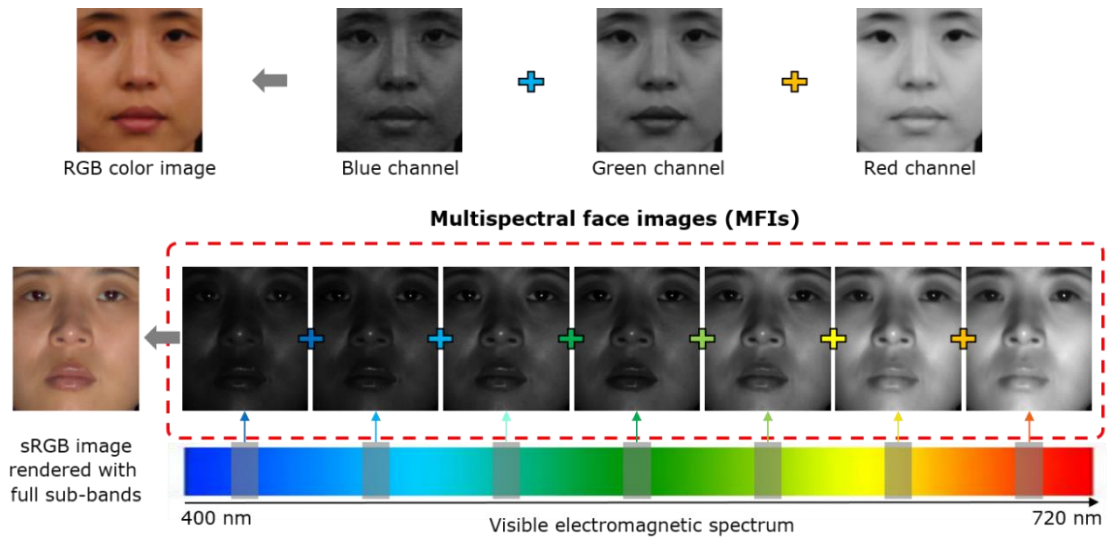


Figure 2.1: A comparison between a traditional broad-band image captured by a RGB color camera and hyperspectral face images (HFIs) captured by a hyperspectral imaging system with liquid crystal tunable filters (LCTFs) in the visible spectrum.

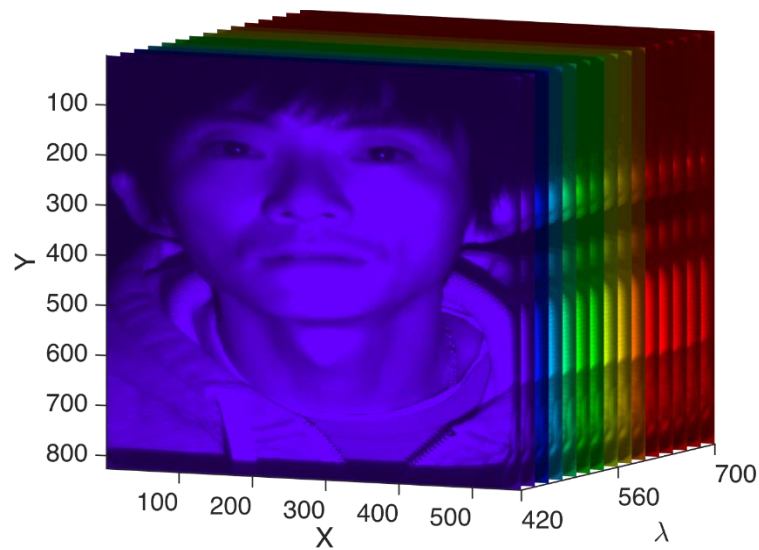


Figure 2.2: A hyperspectral face cube.

Compared with traditional broad-band images captured by trichromatic (RGB) color or monochromatic (grayscale) cameras, HI can be made up of a large number of wavelength indexed channels or bands [Koschan11] in Fig. 2.1. Each spectral image is referred to as 2-dimensional intensity data obtained over each of the different spectral bands. Hence, if all the spectral images are stacked directionally vertical or horizontal, spectral images model a three dimensional cube as shown in Fig. 2.2: two spatial dimensions corresponding to the coordinates of pixel on the image lattice and one spectral dimension corresponding to the wavelength [Robles.13]. When we plot spectral values of a spatial location in hyperspectral data cube as a function of wavelength, the spectra for the selected pixel has a specified spectral shape which uniquely specify target materials as shown in Fig. 2.3.

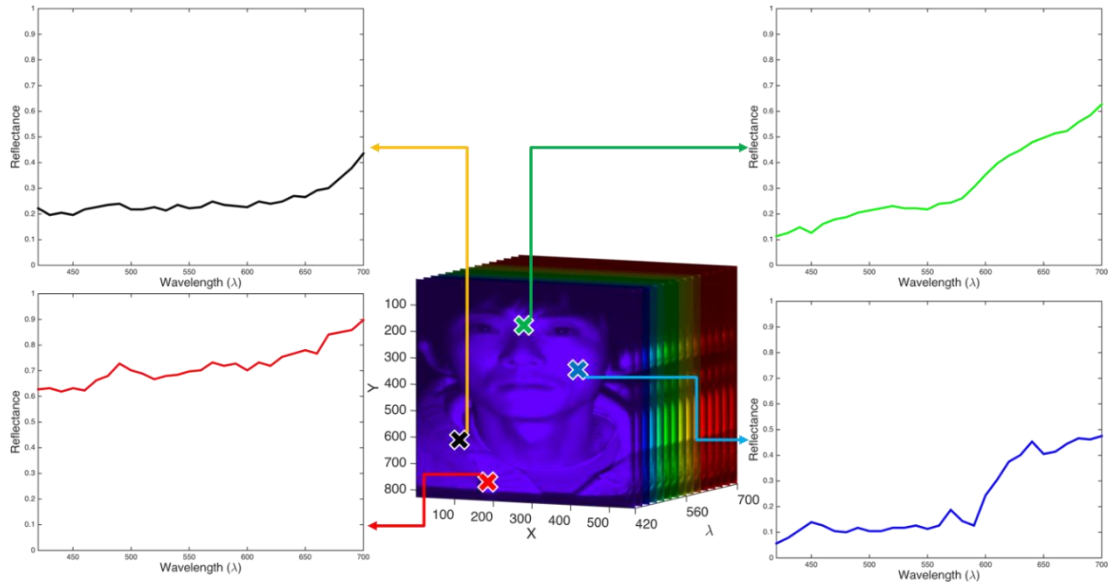


Figure 2.3: Examples of reflectance spectra for a single pixel in a hyperspectral face cube.

2.1.1 Hyperspectral imaging technologies

As mentioned above, the aims of HI are to achieve spatial and spectral information at the same time which are directly used for real-time applications, including face recognition, food security, photography, etc. Accordingly, the process of HI acquisition is much faster and simpler than hyperspectral imaging systems. HI technologies [Koschan11], [Robles.13] involve ground-based commercial systems based on a rotating wheel with various types of narrow or broad band filters and electronically tunable filters: liquid crystal tunable filters (LCTFs), acousto-optic tunable filters (AOTFs), Fabry-Perot imagers, etc. Recently, many researcher proposed alternative systems for HI based on multiplexed illumination [Park07],

optimized wide band illumination [Chi10], a prism-mask [Cao11], and hand-held devices [Khan13]. Detailed explanations for typical HI techniques are set forth below.

For HI using a filter wheel, which is a fixed-filter system where narrow band-pass glass filters are placed in front of the camera lens, hyperspectral images are taken over rotating or swapping the spectral glass filters. However, as pointed out by Koschan *et al.* [Koschan11], there are intrinsic restrictions in fixed-filter systems: 1) limitations of selecting color filters and the number of filters; 2) limitation to build narrow-band filters; 3) time delay for selecting the filters due to the fundamental problem of a mechanical system; and 4) occurrence of the vibrations in the imaging system.

In contrast, the electronically tunable filters (ETFs) in staring array devices¹⁶ are primarily based on acousto-optic or liquid crystal technologies. Staring arrays can collect the full set of spectral band-sequential images by consecutively capturing the entire spatial resolution at one wavelength indexed channel or band at a time. The electronically tunable filter systems may be faster and more flexible for hyperspectral image acquisition due to the following observations: 1) the capacity to select specific bands or sequentially sweeping bands by an electrical controller; 2) high image quality due to very low distortions; and 3) light weight to allow airborne use and/or remote sensor platforms.

AOTFs are electro-optical devices including a tellurium dioxide (TeO_2) or quartz crystal bonded to a transducer which generates a high-frequency acoustic wave propagating into the crystal. As the incoming light reaches the crystal, concurrently a radio-frequency acoustic wave propagates into the crystal. During this process, the acoustic wave affects a variation in the refractive index, thusly performing as a transmission diffraction. The selection of the specific wavelength can be controlled by adjusting the frequency of the acoustic wave [Koschan11].

LCTF is comprised of a set of liquid crystal wave plates to tune a specific wavelength. LCTF offers a linear optical path by polarizing a stack of wave plates and provides the ability to select any wavelength in visible range or near-infrared (NIR) range. While LCTF is sensitive for polarization and has instinct restrictions with regard to relaxation time of polarizing a stack of wave plates for tuning a wavelength about 5 to 50 ms, LCTF is the most commonly used ETF due to the following reasons: 1) light transmission is rapidly and readily controlled by electrical applications with a USB interface; 2) there is no vibration as tuning a specific wavelength; 3) it has high flexibility to be employed in potential applications such as face recognition, agriculture, biomedical and chemical imaging, artwork, etc., with electronic CCD cameras; and 4) it provides large aperture and high image quality due to very low distortions. For information specific to remaining ETF techniques, see [Koschan11], [Robles.13] herein referenced.

¹⁶ The staring array devices are also called the staring focal plane arrays (FPAs) which is regarded as an image sensing device with a rectangular array of light-sensing pixels at the focal plane of a lens.

2.1.2 Hyperspectral image formation

We shortly introduce the basic principle of image formation. Note that we will focus on wavelengths spanning the visible spectrum defined by the range between approximately 400 nm and 720 nm. From Sir Isaac Newton's experiments [Wandell95], a white light can be dispersed by a prism into rays of different wavelength as a foundation of understanding of light and color. The decomposing rays of wavelength from light are defined by Newton as a spectrum. When we look at the distinct spectral components, each of them has a different color appearance: red color corresponding to relatively long wavelengths and blue color corresponding to relatively short wavelengths. The spectral properties of light with energy emitted at each wavelength are physically characterized by its spectral power distribution (SPD). The SPD is the distribution of power as a function of wavelength.

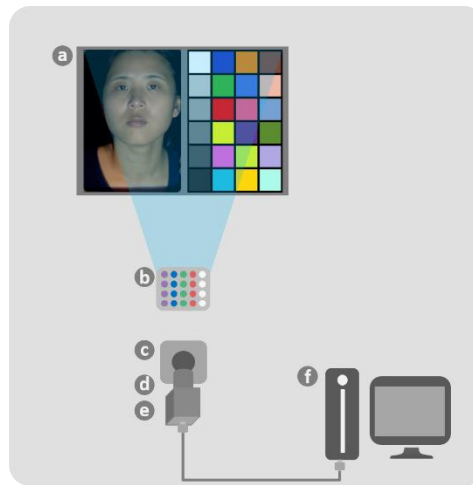


Figure 2.4: Hyperspectral imaging system using the LCTF in our experiments: (a) target, (b) light source, (c) LCTF, (d) Lens, (e) detector, and (f) system controller.

The image formation processes can be regarded as energy transfers from light sources to surface areas and from surface areas to image sensors and observers. The incident light from the source is either absorbed or reflected by the surface. The portion of the light reflected by the surface for different wavelengths defines the spectral reflectance. The spectral reflectance is dependent upon viewing geometries, such as the angles of incident light and of observation and on multiple and varied physical processes acting immediate and concurrent, as the incident light arrives at the surface [Wandell95]. The light incident at a detector or an observer is denoted as spectral radiance of light that is reflected from the surface. Namely, the spectral radiance is proportional to the product of surface albedo

and incident light. Figure 2.4 illustrates a hyperspectral imaging system using LCTF in our experiments.

The main components in terms of hyperspectral image formation as shown at the spatial coordinate (x, y) and the wavelength λ particularly in the visible spectrum can be formulated as:

$$E(x, y, \lambda) = I(x, y, \lambda)R(x, y, \lambda)S(x, y, \lambda) + O(x, y, \lambda) \quad (2.1)$$

where E is the spectral radiance, I is the spectral power distribution of the illumination, R is the spectral reflectance of the object surface, S is the spectral sensitivity of the monochromatic CCD array, and O is the offset including dark current and stray light [Pan03].

The hyperspectral image acquisition systems decomposing the incident light into the full set of spectral band-sequential images in the visible range mostly utilize a CCD sensor to capture the raw image due to its high quantum efficiency. However, radiance spectra are intrinsically affected by low spectral power from the illuminant source or the transmission function of the spectral filters. Accordingly, radiance spectra require calibration to account for noise and bias.



Figure 2.5: A sample sequence of ID: F009_02 in IRIS-HFD-2014-V1 using LCTFs ranging from 420 nm to 680 nm in 10 nm steps where only 27 bands are shown in this figure. Note that the total number of bands in our database is 29 bands from 420 nm to 700 nm in 10 nm steps.

To calibrate the radiance spectra acquired by a hyperspectral imager, we first determine the appropriate exposure time with white patch or white spectralon for each wavelength of

interest in image acquisition process. Once we determine the proper exposure time at each wavelength, we capture dark current images with the same exposure time at each wavelength by placing a cap on the front of the spectral filter. In general, hyperspectral imaging systems are commonly equipped with a spectral filter, an optics (or a lens), and a detector in this order, or an optics, a spectral filter, and a detector. In our experiments, we choose the former order as [Chang08], [Hardberg01], [Koschan11]. To eliminate dark current and bias of a sensor, the dark current images are subtracted from the raw images according to each wavelength. Then, reflectance images are computed as:

$$R(x, y, \lambda) = E(x, y, \lambda) / I(x, y, \lambda) . \quad (2.2)$$

Figure 2.5 shows a sample sequence of our database after calibration process. In Chapter 4, we will demonstrate hyperspectral data acquisition process for more detail.

2.2 Summary

In this Chapter, we introduced fundamentals of hyperspectral imaging (HI) including HI technologies and image formation to exploit hyperspectral face data which can be used to resolve substantial challenges of face recognition systems particularly produced by variations in lighting. Hyperspectral imaging sensors measure the intensity of the energy reflected by a target object in different parts of the visible spectrum or beyond. The measured spectral information forms a hyperspectral data cube. Each pixel in the hyperspectral data cube has a specified spectral shape which uniquely specifies target materials. The derived apparent reflectance data can be used to authenticate the individuals in the following Chapters.

3 Review of Existing Techniques for Major Challenges

In this Chapter, we review the limitations of five publically available hyperspectral face databases (HFDs): CMU, PolyU-HSFD, IRIS-M, Stanford, UWA-HSFD databases. We also introduce substantial challenges in terms of high dimensionality of hyperspectral data and inter-band misalignment to deal with up-to-date reviews on reduced-order data modeling techniques and analysis of IBMs, respectively.

3.1 Existing face databases

Table 3.1: Overview of hyperspectral face databases (HFDs) in this dissertation.

Database	# of Subjects	Conditions	Spectral Range
CMU	54	Illumination direction and time delay	450 - 1100 nm
IRIS-M	82	Illumination conditions and time delay	480 - 720 nm
PolyU-HSFD	25	Pose and time delay	400 - 720 nm
Stanford	45	Viewing distance	415 - 950 nm
UWA-HSFD	79	Time delay	400 - 720 nm

The development of hyperspectral face databases (HFDs) has, to date, received minimal attention due to 1) the high cost of hyperspectral sensors compared to a trichromatic or monochromatic camera; and 2) the considerable time and effort required for building HFD. Based on the foregoing reasons, there are few publicly available HFDs [Zheng11] that comparatively evaluate face recognition algorithms. We individually analyze five publically available HFDs: the CMU, IRIS-M, PolyU-HSFD, Stanford, and UWA-HFD

databases towards providing information on the key points of each of the considered databases. Table 3.1 shows an overview of HFDs considered in this Chapter. In summary, the robustness of the developed algorithms for hyperspectral face recognition based on the studied databases can be verified through variations of a large number of factors: 1) face pose (PolyU-HSFD); 2) time delay (CMU, IRIS-M, PolyU-HSFD, and UWA-HSFD); 3) illumination direction (CMU), 4) illumination condition (IRIS-M), 5) viewing distance (Stanford), and 6) accessory (IRIS-M). For the extraction of spectral properties of facial tissue in both the visible and the NIR ranges, CMU and Stanford databases can be used. In the case constrained to the visible range, spectral measurements of facial tissue can be achieved on IRIS-M, PolyU-HSFD, and UWA-HSFD.

3.1.1 CMU

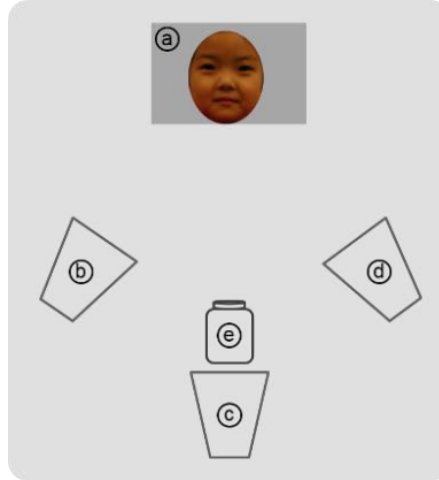


Figure 3.1: Hyperspectral imaging system of the CMU database [Denes02]: (a) target, (b)-(d) light sources placed at -45° , 0° , and $+45^\circ$ according (a), and (e) hyperspectral imaging sensor.

First of all, The CMU database¹⁷ [Denes02], collected at the Carnegie Mellon University, is comprised of hyperspectral images of 54 diverse faces covering the visible and near-infrared (NIR) ranges from 450 nm to 1100 nm in 10 nm steps (65 spectral bands). The hyperspectral imaging system is configured as shown in Fig. 3.1. Three light sources are placed at -45° (Fig. 3.1(b)), 0° (Fig. 3.1(c)), and $+45^\circ$ (Fig. 3.1(d)) according to the target (Fig. 3.1(a)). Each light source, individually and in tandem, can be configured to

¹⁷ <http://www.consortium.ricmu.edu/hsagree/index.cgi>.

on/off status. The light status determines the illumination direction and, accordingly, results in differences specific to facial appearance. The hyperspectral face database (640×480 pixels) in frontal view was captured by acousto-optic tunable filters (AOTFs) under 600W halogen lamps in a studio.

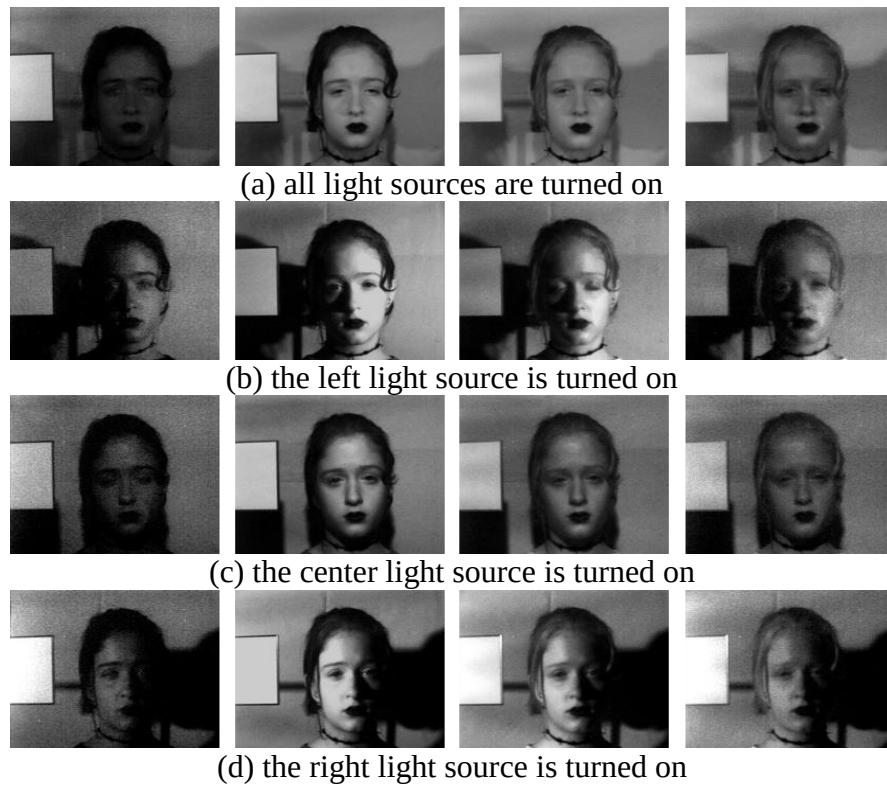


Figure 3.2: CMU database: example of four different datasets gathered under four different illumination directions. The images taken from [Denes02] are sampled in the range of 500 nm, 700 nm, 900 nm, and 1090 nm.

The CMU database, utilizing a hyperspectral sensor, considered the effects of varying illumination directions on facial appearance. In addition, the facial data in the CMU database were taken during multiple sessions over a period of several weeks (approximately two months). As shown in Fig. 3.2, this database provides four different hyperspectral face datasets per each data subject; datasets were gained under varying illumination directions. The noise level in this database is relatively higher (see Fig. 3.2) due to low radiant power of their light sources. Each of the hyperspectral face images can be aligned by using a 2D similarity transform (rotation, translation, and scale) (see Chapter 6) based on the eye coordinates distributed with the CMU database.

3.1.2 IRIS-M

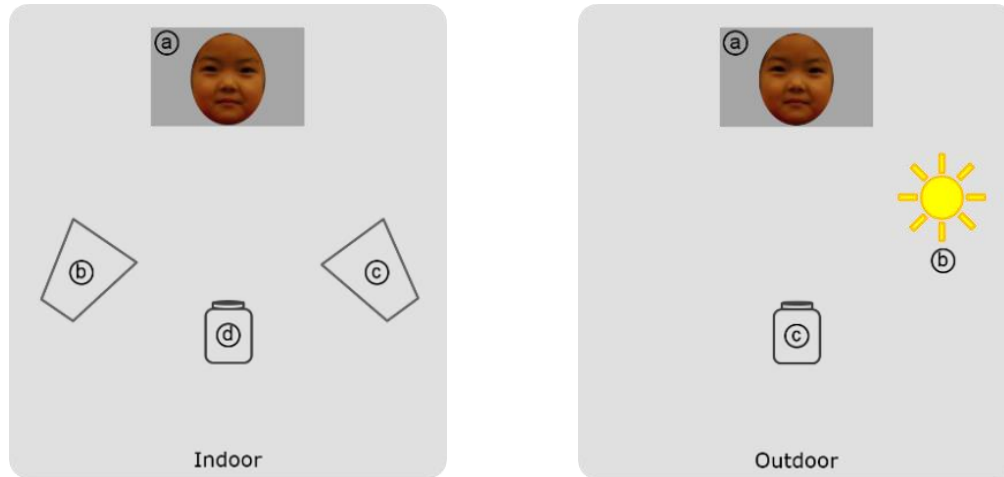


Figure 3.3: Hyperspectral imaging system used to obtain the IRIS-M database [Chang08]. The left figure shows (a) target, (b) and (c) light sources, and (d) hyperspectral imaging sensor. The right figure shows (a) target, (b) natural illuminant (sun), and (c) hyperspectral imaging sensor.

Chang *et al.* [Chang06], [Chang08] created the IRIS-M database at the University of Tennessee which consists of 82 data subjects reflecting different ethnicities, ages, facial hair characteristics, and genders; hyperspectral facial data were gathered over 10 sessions. The IRIS-M database was developed in two different environments (see Fig. 3.3): 1) indoor environment under either two halogen or two fluorescent lamps; and 2) outdoor environment under daylight. The spectral power distributions (SPDs) of four different illuminants utilized in the IRIS-M database are shown in Fig 3.4. The IRIS-M database (640×480 pixels) in frontal view was collected by a VariSpec VIS liquid crystal tunable filter (LCTF) in the visible spectral range from 480 nm to 720 nm in steps of 10 nm (25 bands) and by a Raytheon Palm-IR Pro camera¹⁸ for thermal infrared images. The RGB images (2272×1704 pixels) in the IRIS-M database were captured by a Sony XC-75 camera¹⁹.

¹⁸ <http://www.palmir250.com/ir250pro.htm>.

¹⁹ <http://www.subtechnique.com/sony/PDFs/xc-7573e.pdf>.

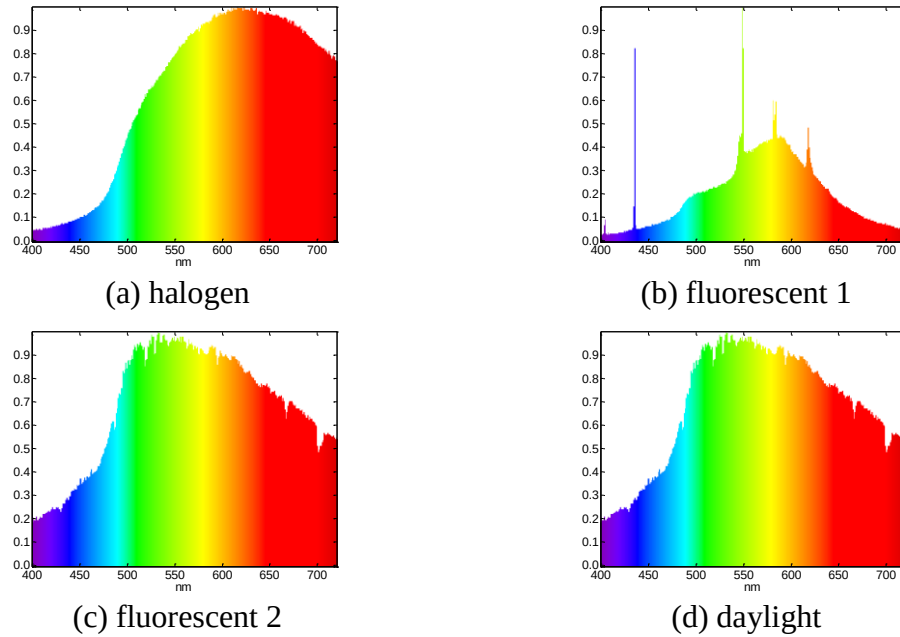


Figure 3.4: The normalized SPDs in terms of four different light sources utilized in the IRIS-M database. The figures are taken from [Koschan11].

In the development of the IRIS-M database, the effects of variations in illuminant and time delay on the facial skin from hyperspectral and thermal imaging were studied. Examples are shown in Fig. 3.5. This database contains low intensity values in the bands of the hyperspectral images at lower wavelengths (see Fig. 3.5).

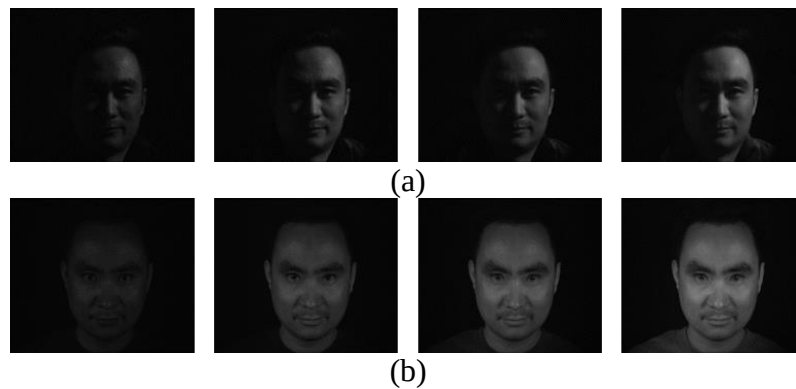


Figure 3.5: Example samples of multispectral face images (MFIs) in IRIS-M database [Chang08]: (a) indoor MFIs sampled for every 40 nm steps in the range from 580 nm to 700 nm under halogen lamps; and (b) outdoor MFIs sampled in the range from 580 nm to 700 nm in 40 nm steps under daylight

3.1.3 PolyU-HSFD

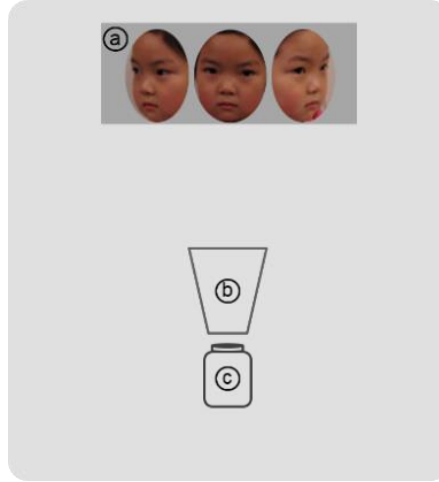


Figure 3.6: Hyperspectral imaging system employed to obtain the PolyU-HSFD [Di10]: (a) target in three different poses, (b) light source, and (c) hyperspectral imaging sensor.

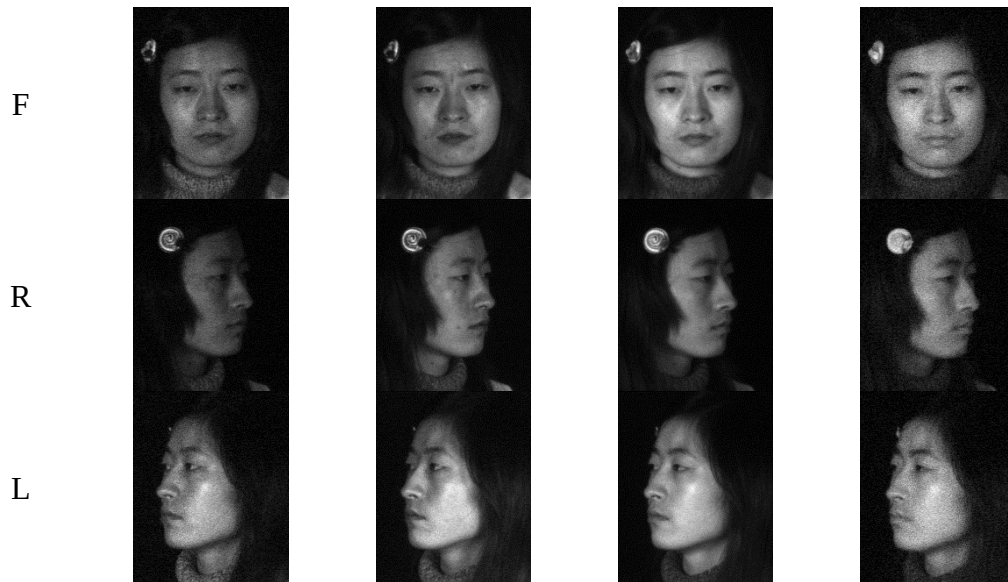


Figure 3.7: HFI sequences of PolyU-HSFD in the front (F), right (R), and left (L) views where HFIs are sampled for every 50 nm step in the range from 500 nm to 650 nm. These sample images are taken from [Di10].

The Hong Kong Polytechnic University Hyperspectral Face Database (PolyU-HSFD)²⁰ [Di10] consists of 25 data subjects of Asian descent, ranging in age of years (21 to 38) and multiple genders (8 females and 17 males). It primarily considers the effects of varying poses on the facial appearance: frontal, right, and left view of a subject as illustrated in Fig. 3.6. The angles of right and left views are approximately ± 45 degrees with respect to the frontal subject, respectively. Sample sequences of the indoor PolyU-HSFD in three different poses are shown in Figs. 3.7.

Each facial set (300 hyperspectral image cubes, $180 \times 220 \times 33$ voxels) obtained by a CRI's VariSpec LCTF under a halogen light contains a 33-channel hyperspectral image in 10 nm steps from 400 nm and 720 nm. According to the data collection dates, the PolyU-HSFD provides four different sets obtained at roughly one month intervals. Note that the first six bands (400 nm to 450 nm) and the last three bands (690 nm to 720 nm) in this database are rejected due to very low signal-to-noise ratios ($\text{SNR} < 6\text{db}$) as mentioned in [Di10], [Uzair13], [Uzair15].

3.1.4 Stanford

The indoor Stanford database of 45 subjects, established by Skauli and Farrell [Skauli13] was acquired by a HySpex line-scan imaging spectrometers²¹ under studio tungsten light. The HySpex camera is a pushbroom sensor developed at NEO. The primary advantage of the pushbroom sensor [Robles.13] is that it is able to collect all of the spectra relevant to each individual line, employing a line-by-line imaging collection approach. Nevertheless, the pushbroom sensor suffers from spectral distortion and is also heavily sensitive to a subject's movement, as shown in Fig. 3.8, as one line of the scene of interest is scanned at a time.

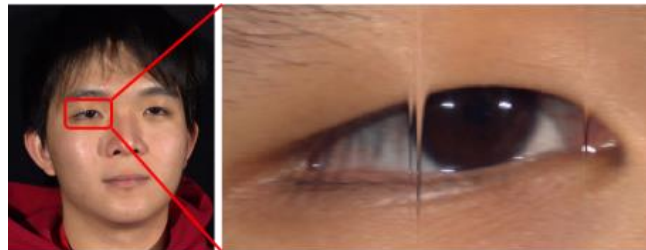


Figure 3.8: Example of an artifact in Stanford database resulting from eye blinking during data acquisition. This sample image is taken from [Skauli13].

²⁰ <http://www4.comp.polyu.edu.hk/~biometrics/>.

²¹ <http://www.neo.no/hyspex/>.

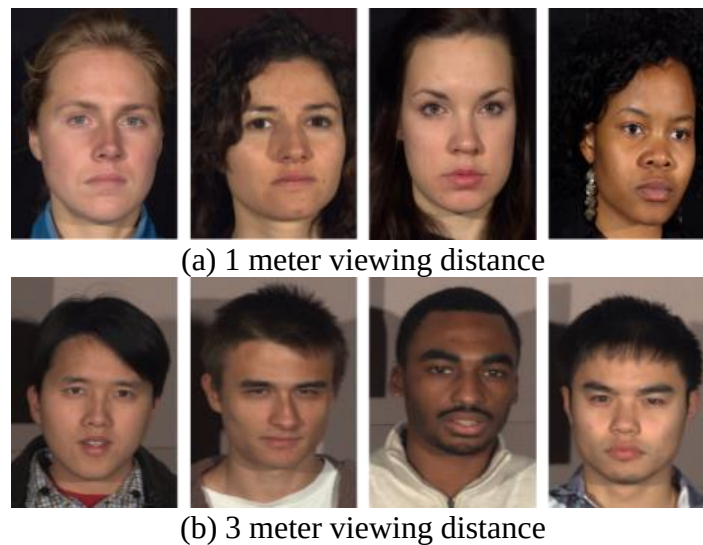


Figure 3.9: Sample sRGB images in the Stanford database taken from [Skauli13].

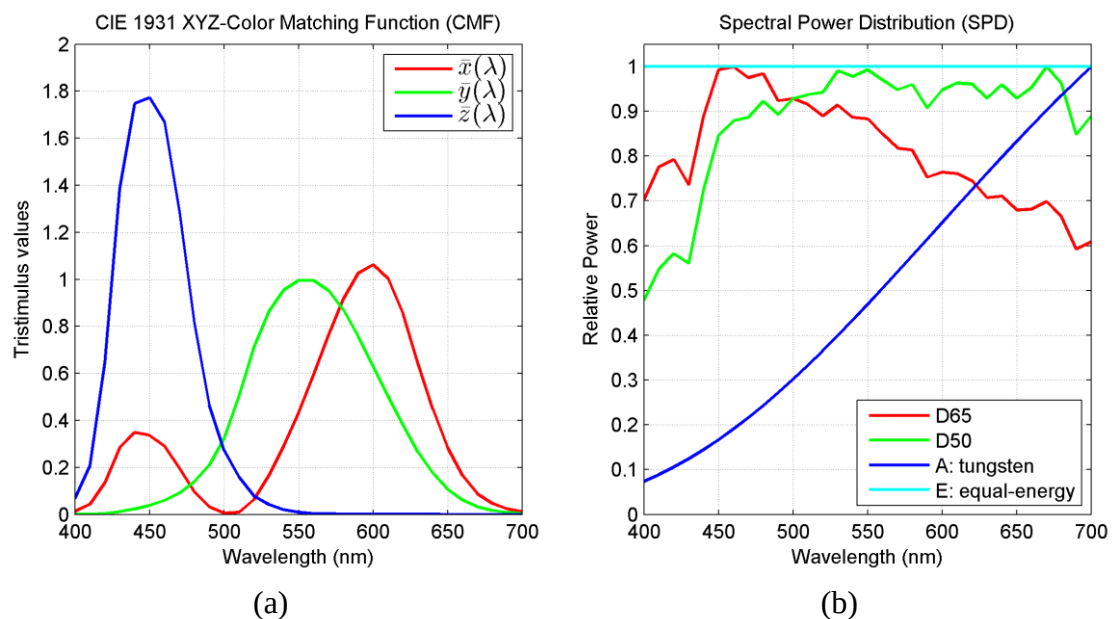


Figure 3.10: (a) CIE 1931 XYZ-CMF and (b) spectral power distribution (SPD) of CIE illuminants.

The Stanford database in frontal view contains 148 bands spanning the visible and NIR range from 415 nm to 950 nm in steps of 4 nm. The Stanford database considered the effects of the variations in scale based on the viewing distance from a face to a detector (1

to 3 meters) specific to hyperspectral face images. In Fig. 3.9, the samples of Stanford database in different viewing conditions are shown where they are displayed using sRGB values [Stokes98] rendered under the CIE illuminant D65. Note that those sRGB values can be computed by employing CIE 1931 XYZ-color matching function (CMF) and CIE illuminants in Fig. 3.10. The hyperspectral face data are denoted in an $n \times m \times w$ matrix: n corresponds to the number of rows in an image, m corresponds to the number of columns in an image, and w corresponds to the number of spectral bands (about 148 bands). Varying the number of rows and columns of an image for each subject is dependent on the scanning time.

3.1.5 UWA-HSFD

UWA-HSFD consisting of 79 data subjects in the frontal view taken over 4 sessions [Uzair13], [Uzair15] was developed by the University of Western Australia²². Each hyperspectral image was captured by the VariSpec LCTF integrated with a photon focus camera. Each dataset of HFIs contains 33 bands covering the visible spectral range from 400 nm to 720 nm with 10 nm steps. Since UWA-HSFD considered the adaptation of the camera exposure time according to lower transmittances of the filter and lower illumination intensities in each band, this database contains relatively lower noise level than the PolyU-HSFD. However, UWA-HSFD suffers from IBMs that result from slight head movements during data acquisition as shown in Fig. 3.11.

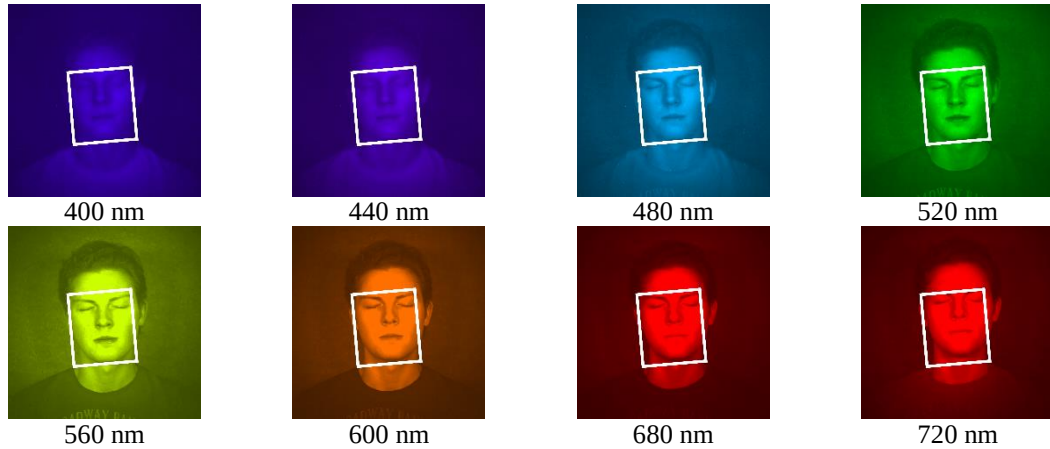


Figure 3.11: Examples of inter-band misalignments in UWA-HSFD [Uzair13], [Uzair15]. Each of the estimated warps from a fixed bounding box-based alignment approach (FBB) is depicted by a rectangle in each band. As shown in the bands at 680 nm and 720 nm, there are IBMs in this database.

²² <http://www.csse.uwa.edu.au/~ajmal/>.

3.2 High dimensionality

The high dimensionality of the spectral data causes limitations on physical experiments and detailed numerical analysis since spectral data consist of not just three visible bands but multiple sub-bands captured at each wavelength. For this reason, there is a pressing need for robust data compression techniques; specifically post-processing techniques, by means of extracting relevant basis functions from large quantities of high-dimensional spectral data. These techniques are crucial to analyze and represent a large set of spectral data and, finally, to model the processes. As mentioned in [Mittelman12], neighboring spectral samples are highly correlated, as the spatial resolution of the spectral data increases. In other words, the neighboring spectral samples are more likely to involve corresponding spectral characteristics. These fundamental features of the spectral data should facilitate the potential of reduced-order spectral data modeling (ROM) techniques to address the challenges of compressing high-dimensional data. Accordingly, ROM has attracted increasing interest in recent decades due to its possibilities for a wide variety of applications, including machine learning [Duda12], data mining [Kim11], and hyperspectral imaging [Robles.13]. A number of different ROM approaches have been introduced for spectral data compression in a higher dimension of spectral space [Gillis12], [Gillis13], [Gillis14], [Nascimento05].

Principle component analysis (PCA)-based ROM initially proposed by Cohen [Cohen64] and later Maloney [Maloney86] yields a small number of linear basis functions of reflectance spectra. PCA has become a powerful tool for the analysis of high-dimensional spectral data. The basis functions to describe a particular set of reflectance spectra can be obtained by computing the eigenvectors. Fairman and Brill [Fairman04] estimated the reflectance spectra from colorimetric tristimulus values based on applying a PCA-based model for dimensionality reduction. In addition, local-based PCA and adaptive PCA models for improving the reconstruction accuracy were proposed by Zhang and Xu [Zhang08] and Mansouri *et al.* [Mansouri08], respectively. Recently, Barakzehi *et al.* [Barakzehi13] introduced nonlinear PCA (NCPA) [Farajikhah11] based on artificial neural networks [Scholz08] for extracting the nonlinear patterns by training an auto-associative neural network.

Non-negative matrix factorization (NMF)-based ROM that takes into account the nonnegativity of spectral data and improves the interpretability of the decomposition was proposed by Amirshahi and Amirshahi [Amirshahi10]. Kim *et al.* [Kim12] adapt the fast NMF (FNMF) originally proposed by Kim and Park [Kim11] to the process of dimensionality reduction. FNMF, which is based on the block principal pivoting method with nonnegative constraints, iteratively explores a local minimum of the objective function by extracting hidden patterns from the data and eliminating the insignificant directions. Recently, Gillis *et al.* proposed a hierarchical clustering approach using convex geometry and rank-two nonnegative matrix factorization for high-dimensional

hyperspectral imaging in [Gillis15]. However, as mentioned in [Gillis13], NMF-based ROM cannot characteristically separate all basic functions correctly owing to the nonuniqueness of the solutions. A technique to overcome the constraints of NMF is to integrate prior information into the NMF-based ROM to account for the characteristics of the solutions and make the problem well-posed. Similar to PCA-based ROM, recursive sparse nonnegative matrix underapproximation (RSNMU)-based ROM was introduced by Gills and Plemmons [Gillis12]. RSNMU includes additional underapproximation constraints. Later, Gills *et al.* [Gillis14] proposed to further improve RSNMU by using spatial information from neighboring pixels (RSPNMU).

Several different approaches based on independent component analysis (ICA) [Bianco10], [Hyvärinen00] and Wiener estimation [Peyvandi13], [Shimano06], such as singular value decomposition (SVD) suggested by Hardeberg [Harde.01], were proposed to reduce the dimensionality of spectral data. SVD-based ROM is closely related to the PCA-based ROM as both are used to find the optimal subspace spanned by the principal directions thusly globally solving an optimization problem. An alternative approach towards optimization solution is to conclude that the basis functions are fundamental for the approximation of the given function over some domain of interest. It will be shown that this specific approach-based ROM accurately corresponds to the proper orthogonal decomposition (POD) [Holmes98]. POD is a popular model reduction technique used to lessen the computational expense required for high-dimensional systems. It has been extensively investigated in various applications due to its capability for reduction of order which is simple and efficient [Djouadi08]. According to [Altmann13], the spectral reflectance could result from a non-linear function of the abundant vectors with respect to the spectral components. However, global solutions such as SVD, PCA, and POD could fail to capture the non-linear degrees of freedom, since they assume that the given data set belongs to a linear space where the Euclidean distance is adopted as the metric to minimize [Gu08]. To overcome the limitations of the global solutions, we develop a reduced-order spectral data modeling (ROM) technique based on local proper orthogonal decomposition (LPOD) as will be mentioned in Chapter 5.

3.3 Inter-band misalignments

Although face recognition approaches based on hyperspectral imaging have shown very promising results, there are still several challenges. One challenge, inherent in the image acquisition process itself, is that the individual bands of the hyperspectral image must be acquired quickly and sequentially to avoid inter-band misalignment due to subject motion during data acquisition. However, the shorter the acquisition time the lower the quality of the bands of the hyperspectral image. For this reason, hyperspectral images inherently request longer exposure times, not only to acquire the detailed spectral properties of facial

tissue over the visible spectrum or beyond, but also to ensure high signal-to-noise ratio in low light conditions.

Most of the research efforts on hyperspectral face recognition have, to date, not been put into devising improved alignment to reduce IBM artifacts in HFIs. Two existing, publicly available HFDs, such as PolyU-HSFD and IRIS-M databases were developed using liquid crystal tunable filters (LCTFs) with constant exposure time in order to shorten the data acquisition time. An extended period of time during data acquisition produces a challenge for subject participants to remain motionless while staring at the camera under a bright light. In our data acquisition, for example, the illuminance on the target surface provided by the light source is approximately 1,140 lux, which is roughly equivalent to typical TV studio lighting, and the distance from the target to the light source is 60 cm (about 23.6 inches).



Figure 3.12: Examples of the presence of IBM artifacts in the sRGB color space (best viewed in color). As illustrated in this figure, IBMs result in the problem of motion blurring combined with distorted colors that are caused by spectral distortion. The sRGB image is generated from ID: F009 02 in IRIS-HFD-2014-V1.

In cases where HFIs are acquired with constant exposure time, there is less concern about IBMs because the subject's movement is insignificant with shorter data acquisition time. However, it is shown without adapting the camera exposure time at each wavelength that the existing, publicly available HFDs introduced in [Chang08], [Di10] have an essential limitation in obtaining spectral properties of facial tissue at multiple narrow-bands of the visible spectrum. Nonetheless, longer data acquisition time predictably creates additional IBMs that are one of the most important challenges for practical face recognition systems based on hyperspectral imaging. Figure 3.12 shows an example of the noticeable artifacts associated with the IBMs when the misaligned HFIs are mapped to the sRGB image. As presented in Fig. 3.12, IBMs lead to motion blurring combined with distorted colors that

are caused by spectral distortion. This is due to the spectral distortion associated with IBMs. In other words, since the linear combination of all of the measured reflectance spectra fundamentally estimate the colors of the sRGB image at the given pixels, the colors of the misaligned sRGB image are altered to distorted colors when subject motion during data acquisition causes a few images out of the stacked HFIs to shift.

To address IBM, a fixed bounding box-based alignment approach using one manual input set of eye coordinates is typically employed to align hyperspectral images. However, in the alignment scenario of FBB, there is no guarantee that IBM artifacts are removed in hyperspectral images due to the presence of subject motion during data acquisition, even though the aligned hyperspectral images are resized to lower pixel resolutions. For this reason, we investigate 11 state-of-the-art alignment approaches (AAs) to overcome the IBM problem in Chapter 6.

3.4 Summary

We provide a review of five publically available hyperspectral face databases: CMU, PolyU-HSFD, IRIS-M, Stanford, UWA-HSFD databases. We show that existing databases suffer from low signal to noise ratios and low intensity values in the bands of the hyperspectral image (CMU, PolyU-HSFD, and IRIS-M databases), and inter-band misalignment (Stanford and UWA-HSFD databases).

The high dimensionality of the spectral data causes limitations on physical experiments and detailed numerical analysis since spectral data include multiple bands captured at each wavelength. For this reason, there is a need for robust data compression techniques; specifically post-processing techniques, by means of extracting relevant basis functions from large quantities of high-dimensional spectral data. These techniques are crucial to analyze and represent a large set of spectral data and, finally, to model the processes.

Inter-band misalignments must be preferentially resolved to aid the subsequent processes, e.g., band selection, band fusion, and extraction of spectral features for face recognition within the compute vision community. To address inter-band misalignments in HI, conventional alignment approaches based on eye coordinates have typically been employed. However, it is difficult to consistently select the eye coordinates at the same positions by hand over a large hyperspectral face image set such as IRIS-HFD-2014. In the particular cases for the different profiles and structural feature (e.g., glasses), the conventional alignment approach should be limited due to the partial occlusion of one eye in the profile views and the problem of reflection on the glasses. Accordingly, it is necessary to develop automatic alignment approaches to deal with inter-band misalignments in HFIs.

4 Development of New Face Database

In this Chapter, we introduce a new large hyperspectral face database called IRIS-HFD-2014 recently developed over multiple sessions in the IRIS laboratory at the University of Tennessee. Similar to PolyU-HSFD [Di10] and IRIS-M [Chang08] databases, we employed the LCTF to take the hyperspectral face images (HFI) covering the visible range from 420 to 700 nm in 10 nm steps (29 narrow-bands).



Figure 4.1: A comparison of three facial databases collected by the LCTF sensors where in each row from top to bottom, sample HFIs covering the visible range from 420 to 690 nm in 30 nm intervals are taken from (a) PolyU-HSFD, (b) IRIS-M, and (c) IRIS-HFD-2014, respectively. Note that two HFIs at 420 nm and 450 nm in the IRIS-M database are added as a visual aid.

However, as shown in Fig. 4.1, PolyU-HSFD and IRIS-M databases contain either low signal-to-noise or low intensity values in lower bands of hyperspectral images specifically near blue bands. Compared to PolyU-HSFD and IRIS-M databases, the noise level in IRIS-HFD-2014 is relatively lower because we adjust appropriate exposure time in each band and use a custom light source that has more radiant power near blue bands.

IRIS-HFD-2014 is designed to address several challenging problems in face recognition research, including variations in time, pose (both frontal and profile views), and structural features (i.e., glasses). In addition, the database contains RGB color images of 142 participants captured by a traditional color camera under varying illuminant conditions and blurring in uncontrolled settings. IRIS-HFD-2014 consists of a total of 19,346 facial images of 115 males and 64 females of diverse ethnic backgrounds and appearance.

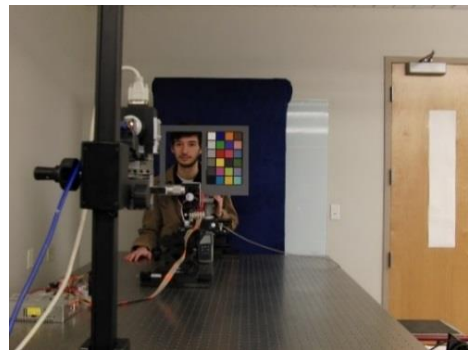
In short, IRIS-HFD-2014 includes 1) hyperspectral images of 179 individuals for three different neutral poses without glasses (frontal view, 45 degree left profile, and 45 degree right profile), 2) hyperspectral images of 51 individuals wearing glasses, and 3) color images (RGB) corresponding to the scenarios mentioned in 1) and 2).

4.1 Imaging modules

The configuration of our data acquisition systems is established as shown in Fig. 4.2 and each of our imaging modules are shown in Fig. 4.3. For stable alignment of the imaging modules over time, our imaging system was mounted on a Newport Optical Table. In the imaged target (Fig. 2.4(a)), an X-rite ColorChecker classic was placed to the side of the individuals so as to allow the calibration and analysis of facial color. For a light source (Fig. 2.4(b)), we used a Lumia 5.1 Reef version with 5 channel LEDs: 1) neutral white, 2) royal blue, 3) hyper violet, 4) deep red and turquoise, and 5) true violet and cool blue as shown in Fig. 4.4 to Fig. 4.8, respectively. The VariSpec VIS such as the LCTF (Fig. 2.4(c)) was mounted in front of a detector. As mentioned in [Chang08], the LCTF provides narrow-band filters with a full width at half maximum (FWHM) of 7 nm. The aperture and field of view of the LCTF are 35 mm and $\pm 7^\circ$, respectively. By continuously sweeping the entire visible spectrum supported by the LCTF, we can acquire a maximum of 321 narrow-band images ranging from 400 to 720 nm in 1 nm step. Between LCTFs and the detector, we equipped a 25mm fixed focal length lens (Fig. 2.4(d)), supporting a wide aperture of f/0.95. For the detector (Fig. 2.4(e)), we utilized an 1.3 megapixel monochrome 12 bit XIMEA xiQ USB3.0 camera supporting a resolution of 1280×1024 pixels. A controller of our light source in Fig. 4.3(c) can manually configure on/off status of each channel and adjust the current of each channel from 0 to 700 mA. In Fig. 4.3(d) EasyView30 light meter and Ocean Optics USB2000 are used to measure the illuminance and spectral power distributions (SPDs) of the studied illuminations, respectively.



(a)

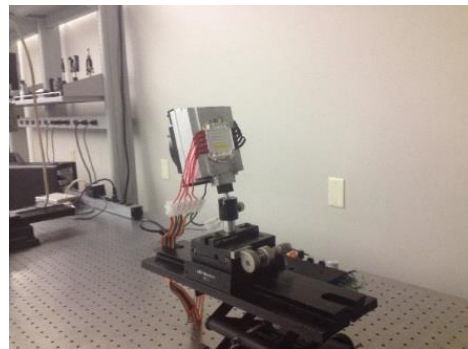


(b)

Figure 4.2: An example of our data acquisition system: (a) a lateral view and (b) a rear view.



(a)



(b)



(c)



(d)

Figure 4.3: Hyperspectral imaging modules in IRIS-HFD-2014. (a) VariSpec VIS mounted in front of XIMEA xiQ USB3.0 camera, (b) Lumia 5.1 Reef, (c) a controller for our light source, and (d) EasyView30 light meter and Ocean Optics USB2000.

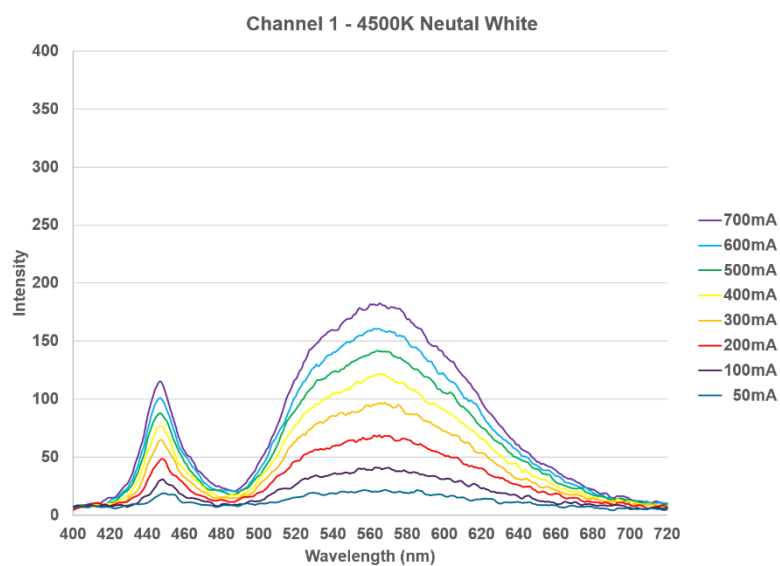


Figure 4.4: SPD of channel 1 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.

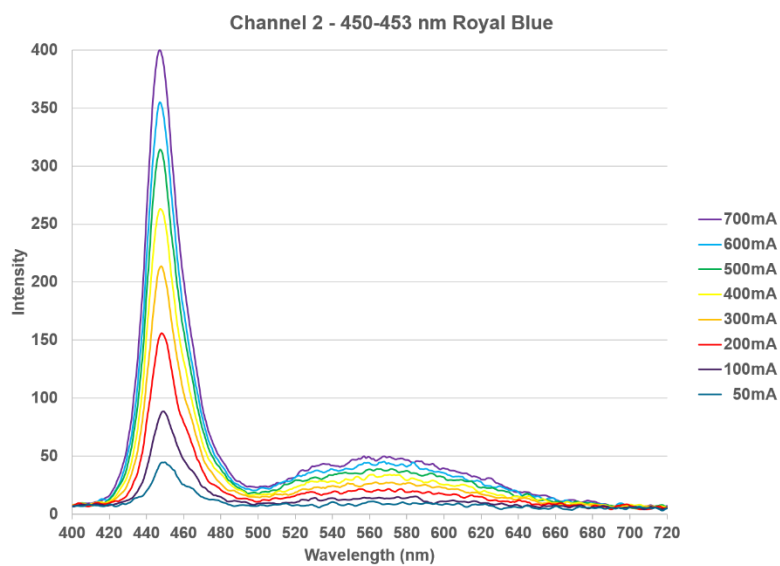


Figure 4.5: SPD of channel 2 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.

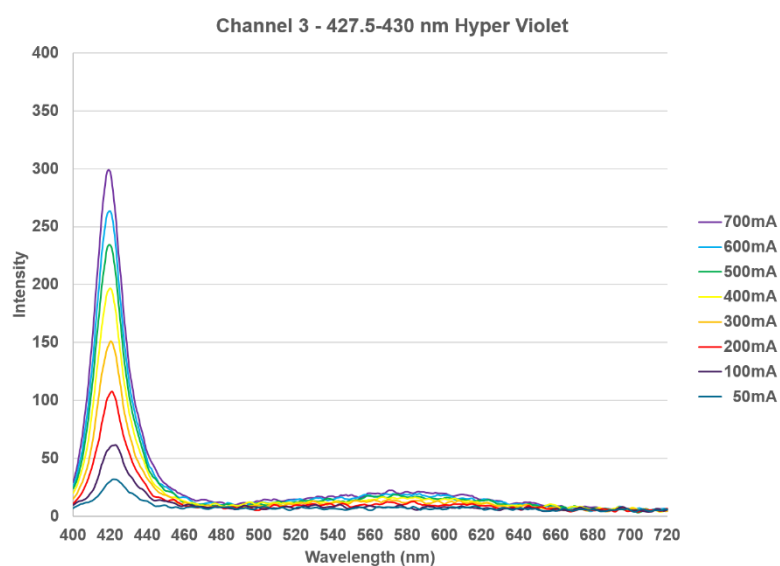


Figure 4.6: SPD of channel 3 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.

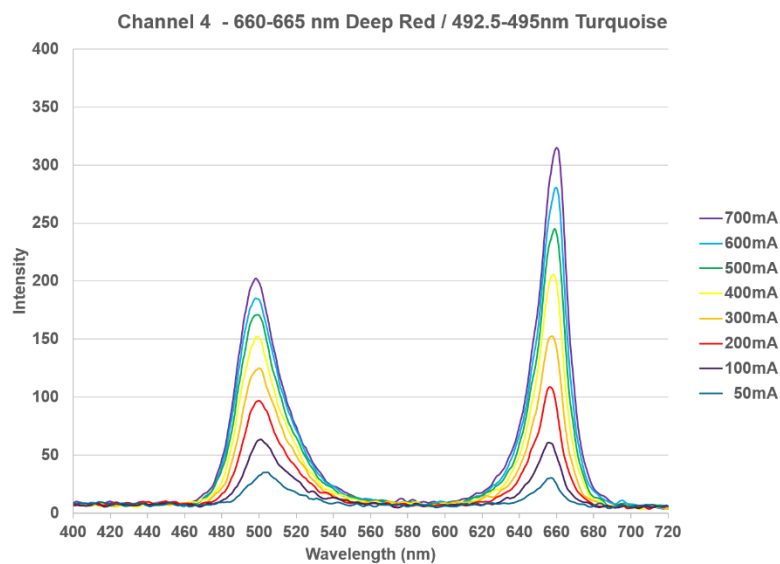


Figure 4.7: SPD of channel 4 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.

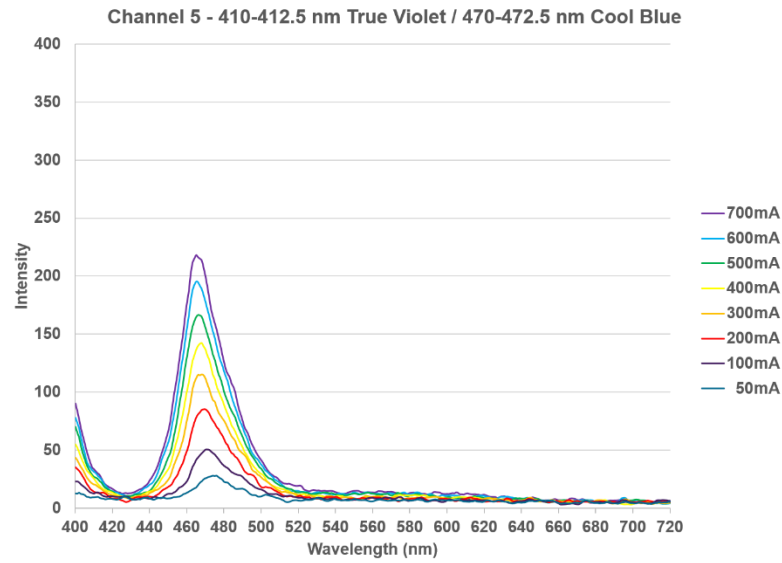


Figure 4.8: SPD of channel 5 of Lumia 5.1 Reef as increasing the current from 50 to 700 mA.

4.2 Data acquisition and calibration

Before invoking the adaptation of the camera exposure time, we observed the uniformity of light by replacing the targets with X-Rite White Balance and Grayscale cards, located on the left and right sides, respectively, as shown in Fig. 4.9(a).

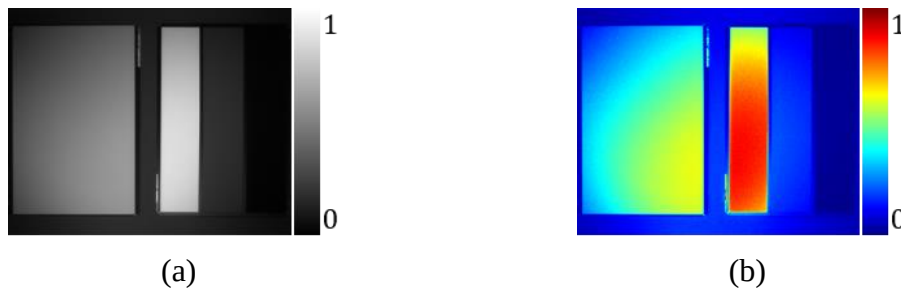


Figure 4.9: (a) Result of the measured light falloff using X-Rite White Balance and Grayscale cards located on the left and right sides, respectively; and (b) the pseudo-colored result of (a) captured at 420 nm.

The light source did not emit uniformly in all directions, as seen in Fig. 4.9(b) where the measured light falloff at 420 nm is displayed with the scaled colors mapping from zero (blue) to one (red). To address this problem is beyond the scope of this work, because the aim of our work is to conceal IBM artifacts in HFIs. However, in brief, this problem can be solved by utilizing a local binary pattern (LBP) filter [Ahonen06] to normalize the illumination variations when our database is used to evaluate algorithms for hyperspectral face recognition.

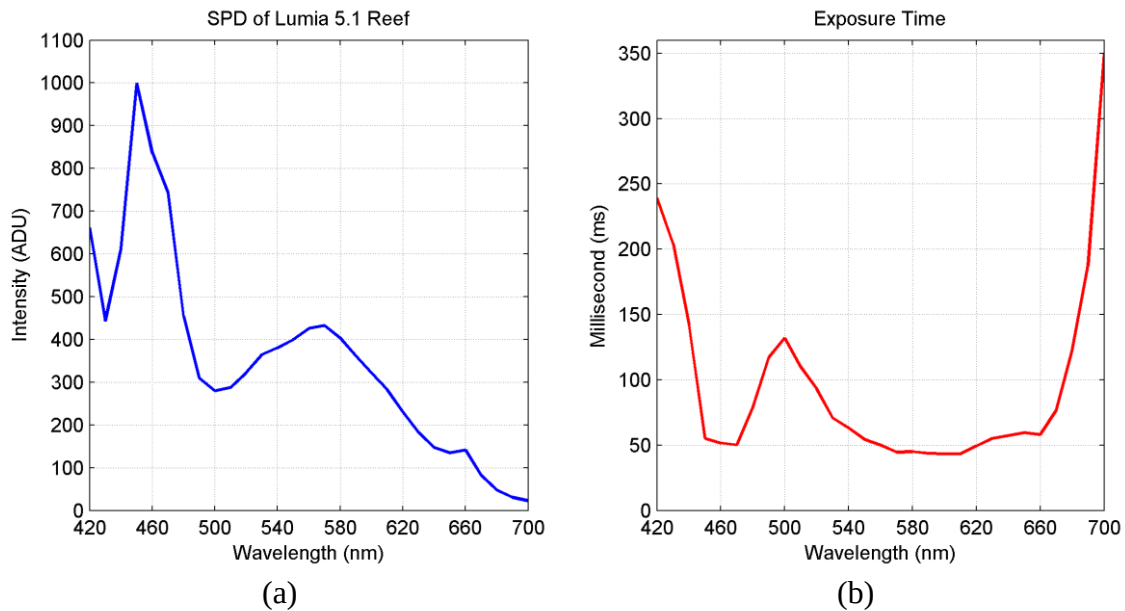


Figure 4.10: (a) Spectral power distribution (SPD) of Lumia 5.1 Reef and (b) tuned exposure times at each wavelength from 420 nm to 700 nm used in our data acquisition.

As illustrated in Fig. 1. 6(a), the spectral transmittances of the LCTF decrease from long to short wavelengths. As a result, it is necessary to properly adapt the camera exposure time according to the spectral transmittances at each wavelength. For example, at the shorter wavelengths, we set longer exposure time in order to accumulate more radiant energy in the detector. In addition to considering low illumination intensities in our data acquisition, we disabled the fourth channel of the light source, deep red (660 to 665 nm) and turquoise (492.5 to 495 nm), to obtain more radiant power in the short wavelength regions of the spectrum. It is also worth noting that since our light source inherently has low radiant power relative to the spectral transmittances of the LCTF from 470 nm to 520 nm and from 670 nm to 700 nm as shown in Fig. 4.10(a), we increased the camera exposure time at these wavelengths (see Fig. 4.10(b)). Figure 4.10(a) elucidates the SPD of the

Lumia 5.1 Reef when we turned off the fourth channel. Figure 4.10(b) illustrates the exposure time utilized in our data acquisition.

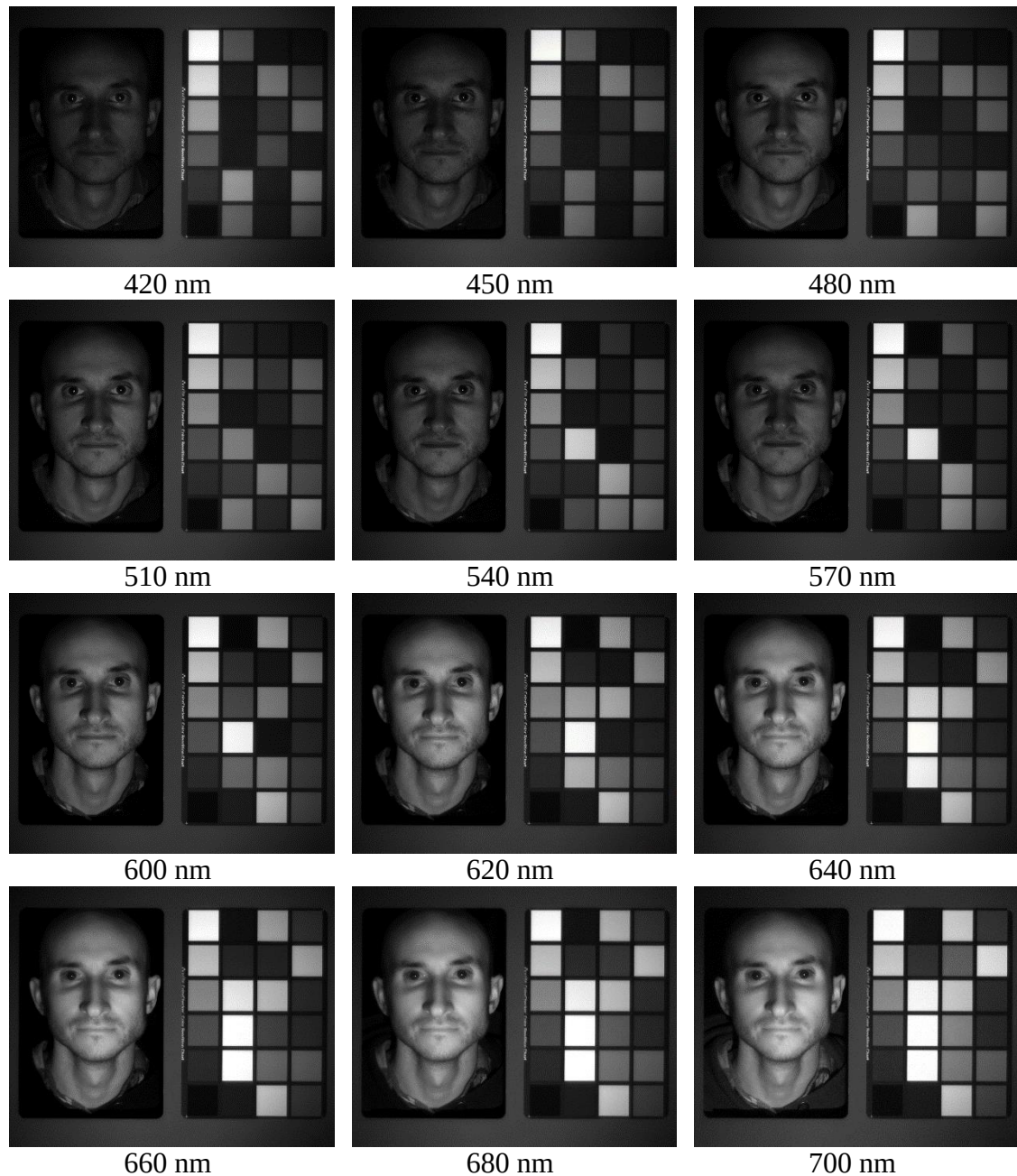


Figure 4.11: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the frontal view.

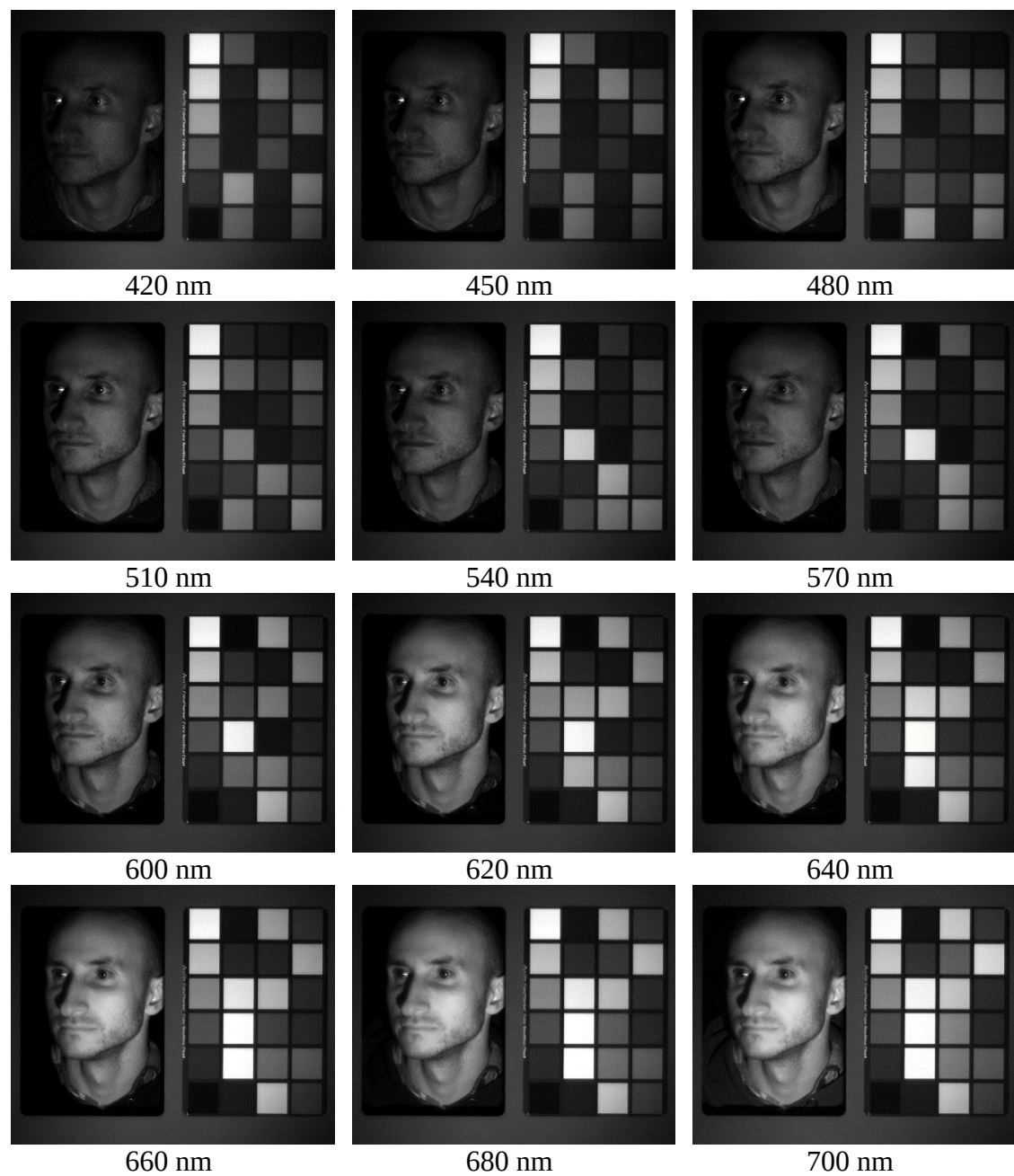


Figure 4.12: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the left profile.

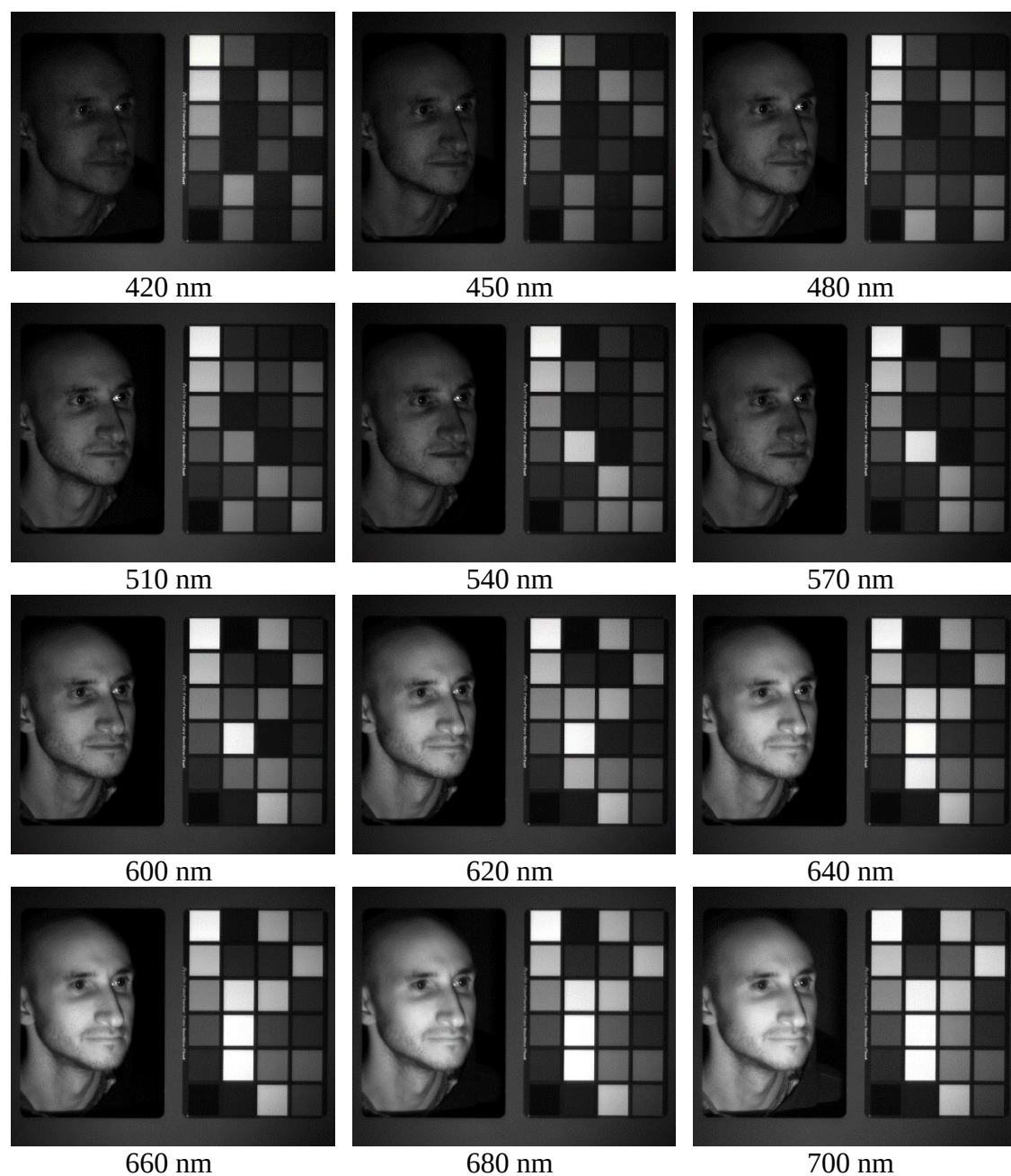


Figure 4.13: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the right profile.

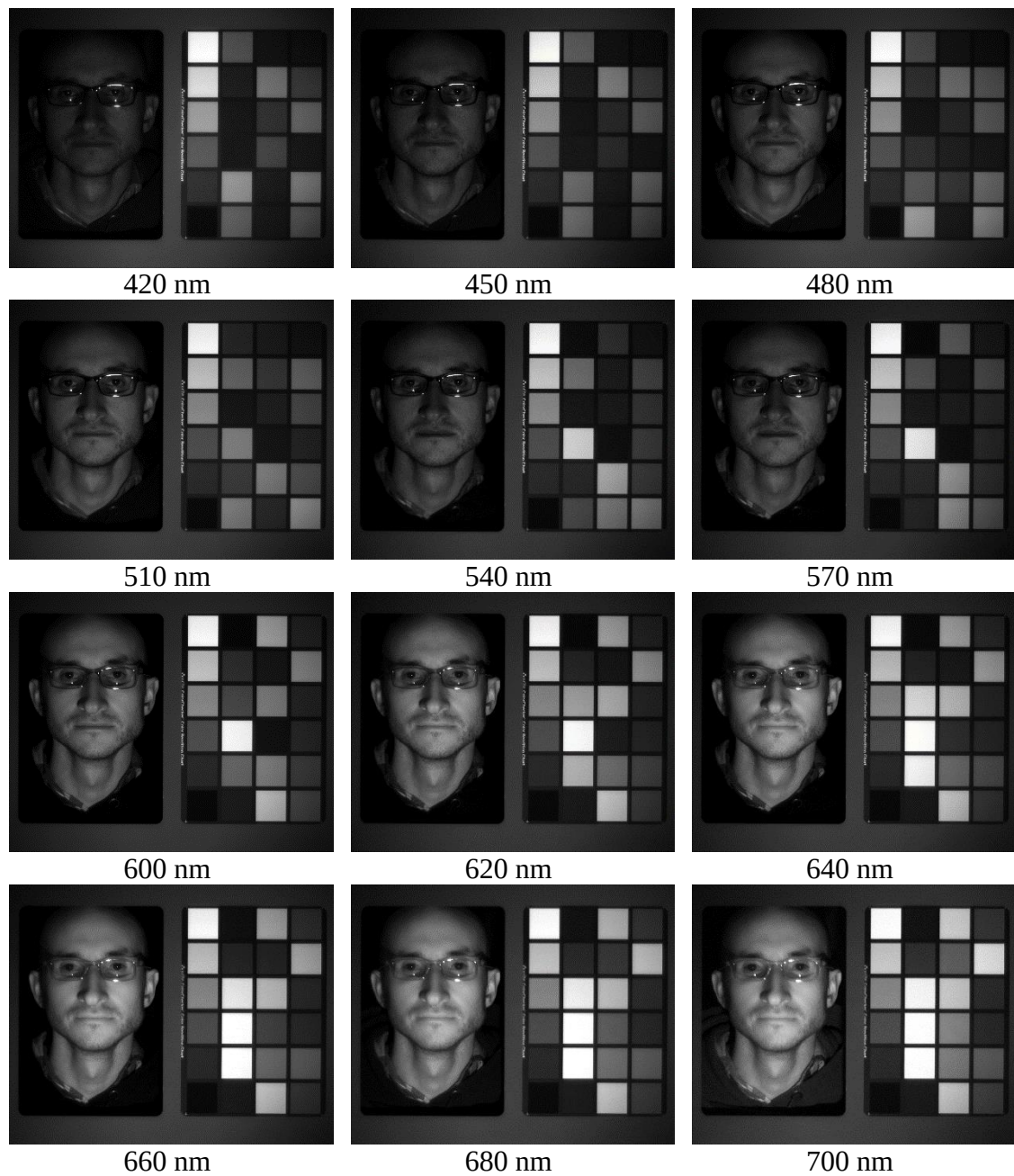


Figure 4.14: An example of hyperspectral face images (HFIs) with ColorChecker taken from ID: F019_01 of IRIS-HFD-2014-V1 in the structural feature, i.e., glasses.

To determine the exposure time for each wavelength of interest, we selected the brightest region as a ROI within the reference white of the Grayscale card (Fig. 4.9(a)). We set the size of the ROI to 100×100 pixels throughout our experiments. Next, we increased the exposure time until the average of the intensity values within the ROI reached about 86 percent of the camera saturation value as in [Foster06]. Once we found the appropriate exposure time at each wavelength, we captured dark current images with the same exposure time corresponding to the wavelengths by placing a cap on the front of the LCTF. To eliminate the constant noises exhibited by CCD imagers depending on exposure duration, we subtracted dark current images from the radiance images. Next, we estimated the maximum intensity values of the radiance images at each wavelength within the ROI located at the same position. The reflectance images were recovered by dividing the radiance images by the estimated maximum values at each wavelength.

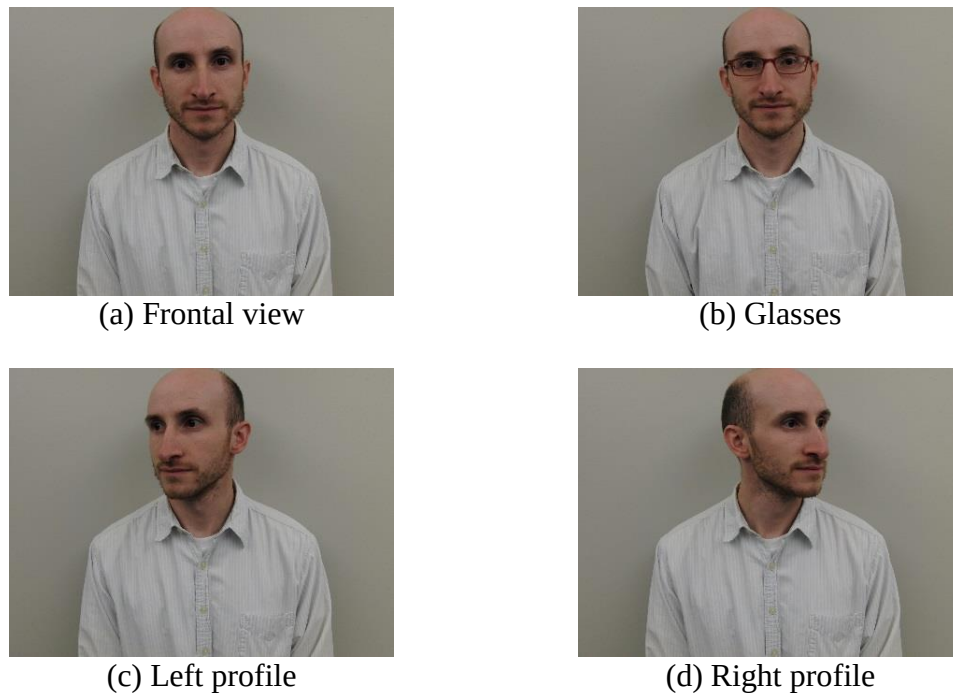


Figure 4.15: An example of color images taken from ID: 019_01 of IRIS-HFD-2014-V1.

Figure 8 shows an example of HFIs with the ColorChecker at 420 nm, 500 nm, 600 nm, and 700 nm in IRIS-HFD-2014 after recovering reflectance images. Sample HFIs with the ColorChecker, after the process of recovering the reflectance images from the radiance images in each pose, are shown in Fig. 4.11 to Fig. 4.14. The images in the profile views were collected by asking the subjects to rotate their head ranging from -45° to $+45^\circ$.

Samples of aligned HFIs in three different poses are shown in Fig. 4.12 to Fig. 4.14. Compared to the frontal view in Fig. 4.11, HFIs in both profile views shown in Figs. 4.12 and 4.13 cover a set of pose variations ranging from approximately -45° to $+45^\circ$ where pose variations introduce partial or entire occlusion of facial components including mouth, nose, or eyes. Figure 4.15 shows sample color images corresponding to different poses in the database.

For all other experiments in this dissertation, we deal only with the subject area in the target by specifying the ROI with a binary mask as presented in Fig. 4.16. This is because the main consideration of this task is to address the inter-band misalignments (IBMs) in IRIS-HFD-2014.

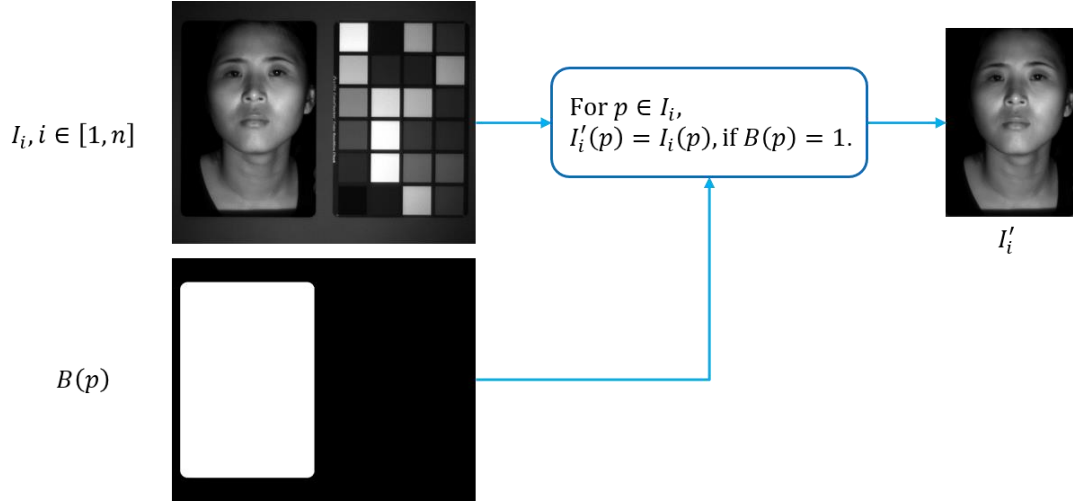


Figure 4.16: Overview of selecting the subject areas within target images where I_i indicates the i th sub-band image of an input HFI set, B represents a binary mask, and p stands for pixel coordinates of I_i .

4.3 New face database (IRIS-HFD-2014)

Practical use of hyperspectral face recognition has, to date, been limited due to database restrictions in the public domain due to 1) the high cost of hyperspectral sensors compared to a trichromatic or monochromatic camera; and 2) the considerable time and effort required for building HFD. Based on the foregoing reasons, there are few publicly available HFDs as mentioned in Chapter 3.1 that comparatively evaluate face recognition algorithms.

Therefore, the purpose of the new database presented in this paper is to meet the emerging demands for a new HFD that can serve as a benchmark for comprehensively and statistically evaluating the performance of current and future algorithms for hyperspectral face recognition.

In this section, we introduce IRIS-HFD-2014, a new hyperspectral face database recently developed, which incorporates adjusted exposure time at each wavelength. Without adjusting appropriate exposure time at each wavelength, spectral information of hyperspectral imaging systems cannot always be sufficiently captured due to lower transmittances of the hyperspectral imaging sensors and lower intensities of synthetic and natural lights at specific wavelengths.

From December 2013 to February 2014, IRIS-HFD-2014 was obtained from 115 males (64%) and 64 females (36%), a total of 179 data subject participants representing diverse ethnic backgrounds and diverse physical appearance. IRIS-HFD-2014 consists of 644 hyperspectral cubes and contains a total of 19,346 facial images: 18,676 hyperspectral face images (644 cubes $\times$ 29 bands) and 670 color images. The version 1 and version 2 contain a total of 14,832 face images and 4514 face images, respectively.

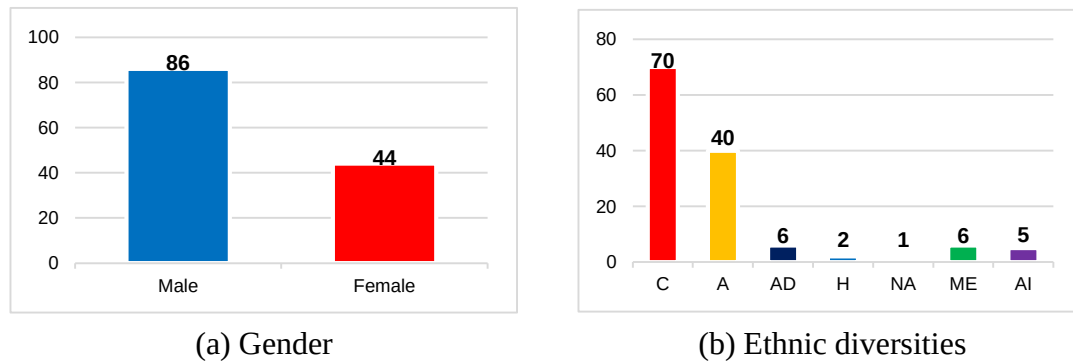


Figure 4.17: A summary of gender and ethnic diversities in IRIS-HFD-2014-V1.

4.3.1 Version 1 (IRIS-HFD-2014-V1)

IRIS-HFD-2014-V1 was collected from 130 data subjects over multiple sessions including 86 males (66%) and 44 females (34%), all of diverse ethnic backgrounds and diverse physical appearance, who represent several ethnic groups: 70 Caucasians (C), 40 Asians (A), 6 of African Decent (AD), 6 of Middle Eastern Decent (ME), 5 Asian Indians (AI), 2 Hispanics (H), and 1 Native American (NA). Figure 4.16 shows the summary of gender and ethnic diversities in IRIS-HFD-2014-V1. The participants vary in age (18 to 74). Most data subjects in the version 1 perform less head movements and eye blinking during data

acquisition, compared to IRIS-HFD-2014-V2. Figure 4.18 and 4.19 show all the datasets in the frontal view of IRIS-HFD-2014-V1, which are aligned by the proposed alignment approach in Chapter 8. It is important to mention that for the sRGB color mappings in this dissertation, once the conversion from reflectance data for the visible range to CIEXYZ tristimuli is performed with regard to a CIE 1931 2° standard observer and a CIE illuminant D65 reference white, CIEXYZ tristimuli are transformed into the sRGB color space.

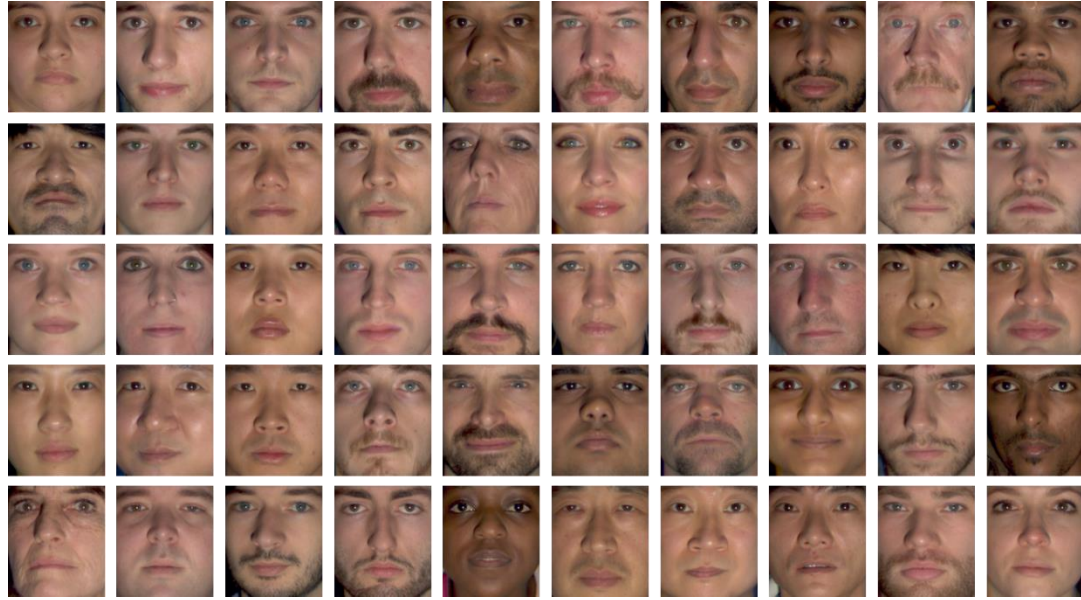


Figure 4.18: The aligned datasets (ID: 001 to 050) of IRIS-HFD-2014-V1 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.

4.3.2 Version 2 (IRIS-HFD-2014-V2)

IRIS-HFD-2014-V2 was collected from 49 data subjects of 29 males (59%) and 20 females (41%), representing several ethnic groups: 33 Caucasians (C), 11 Asians (A), 3 of African Decent (AD), 2 Asian Indians (AI). Figure 4.20 shows the summary of gender and ethnic diversities in IRIS-HFD-2014-V2. The participants vary in age (4 to 46). Most data subjects in the version 2 perform significant head movements and eye blinking during data acquisition, compared to IRIS-HFD-2014-V1. Sample HFIs with the ColorChecker, after the process of recovering the reflectance images from the radiance images in each pose, are shown in Fig. 4.21. Figure 4.22 shows sample color images corresponding to different poses in IRIS-HFD-2014-V2. The entire aligned datasets in this version performed as Fig. 4.19 are presented in Fig. 4.23.

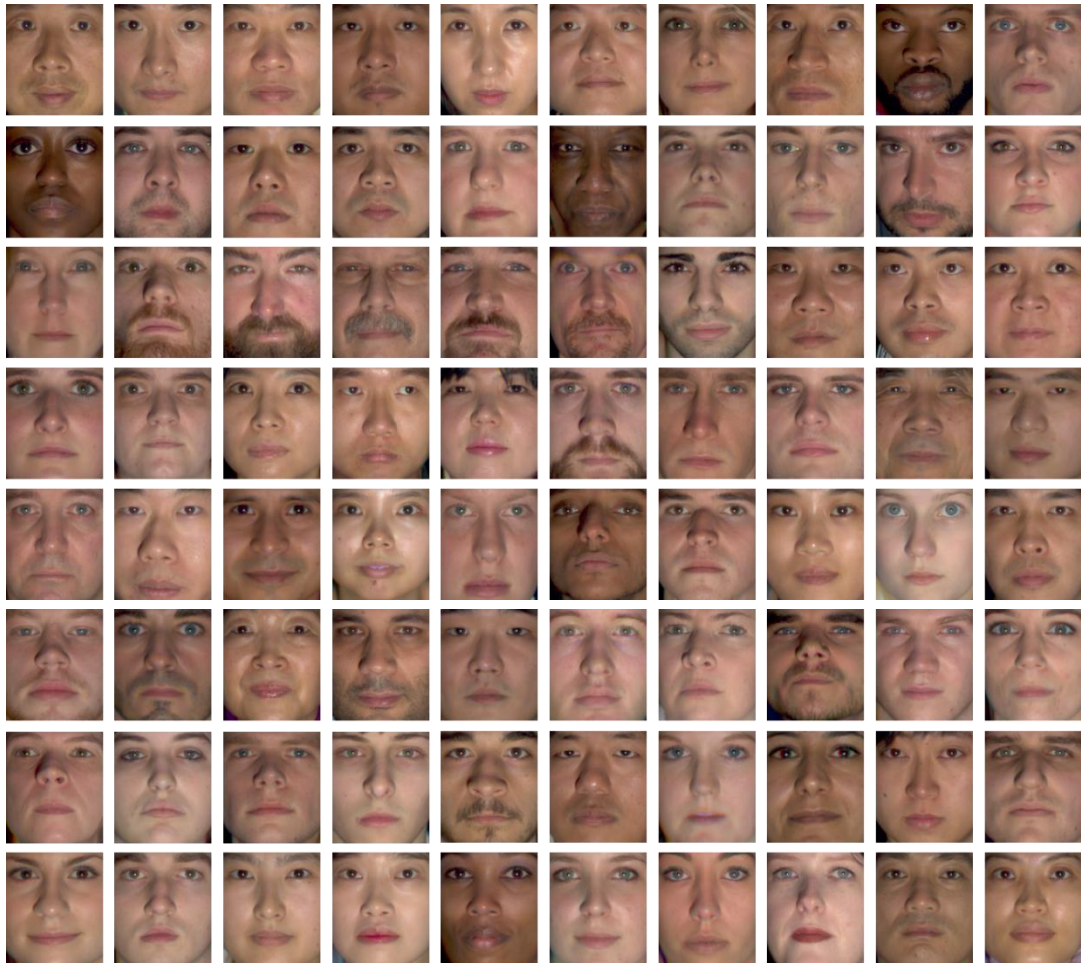
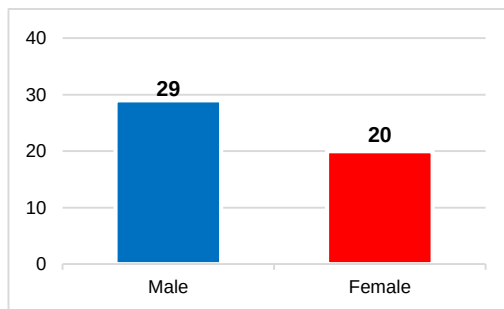
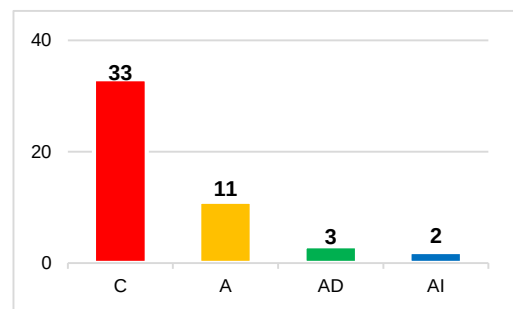


Figure 4.19: The aligned datasets (ID: 051 to 130) of IRIS-HFD-2014-V1 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.



(a) Gender



(b) Ethnic diversities

Figure 4.20: A summary of gender and ethnic diversities in IRIS-HFD-2014-V2.

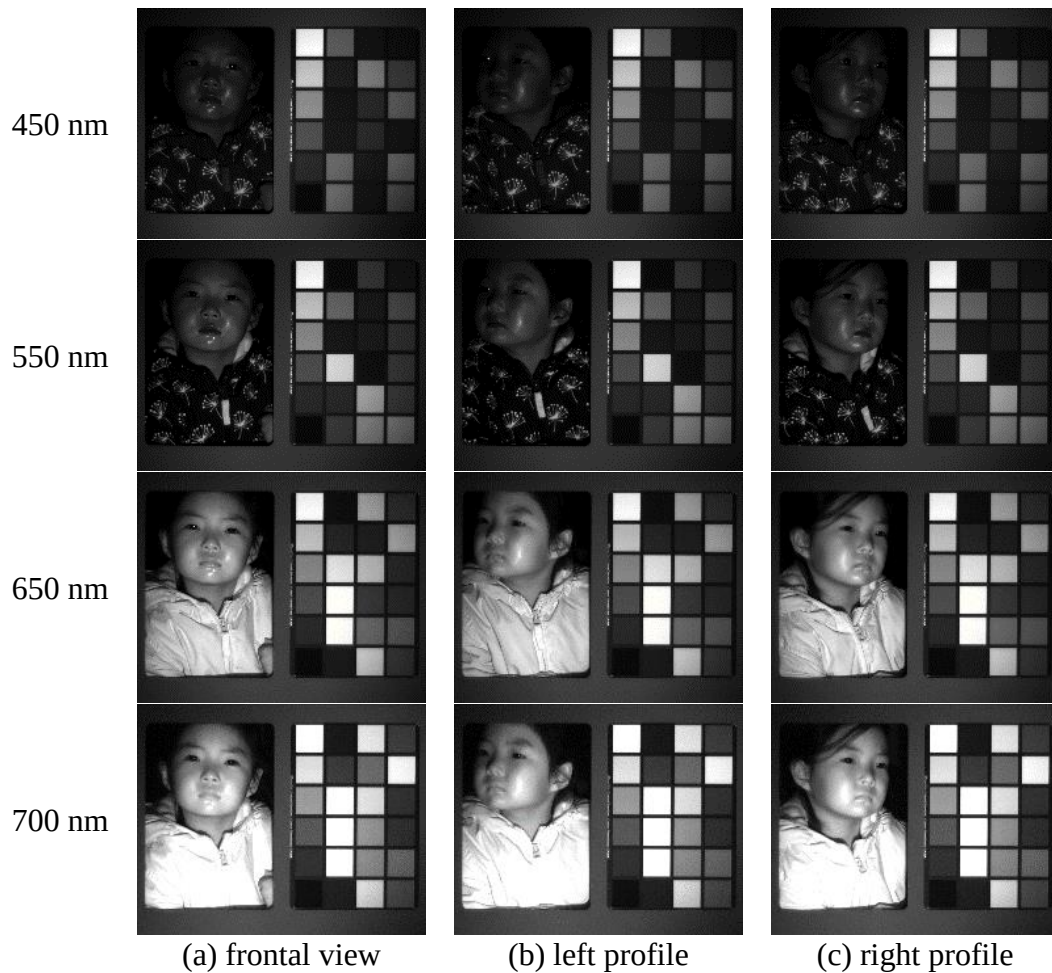


Figure 4.21: Sample HFI sequences for each pose in IRIS-HFD-2014-V2, after data calibration.

4.4 Summary

In this Chapter, we introduced the new large hyperspectral face database, named IRIS-HFD-2014, which can serve as a benchmark for comprehensively and statistically evaluating the performance of current and future algorithms for hyperspectral face recognition. This database consists of a total of 19,157 facial images of 115 males and 64 females (179 individuals) of diverse ethnic backgrounds and appearance. IRIS-HFD-2014 was acquired by using the LCTF covering a range from 420 nm to 700 nm in 10 nm steps

over multiple sessions in four different poses (frontal view, 45 degree left and right profiles, and structural feature, i.e., glasses). IRIS-HFD-2014 will be publicly available.

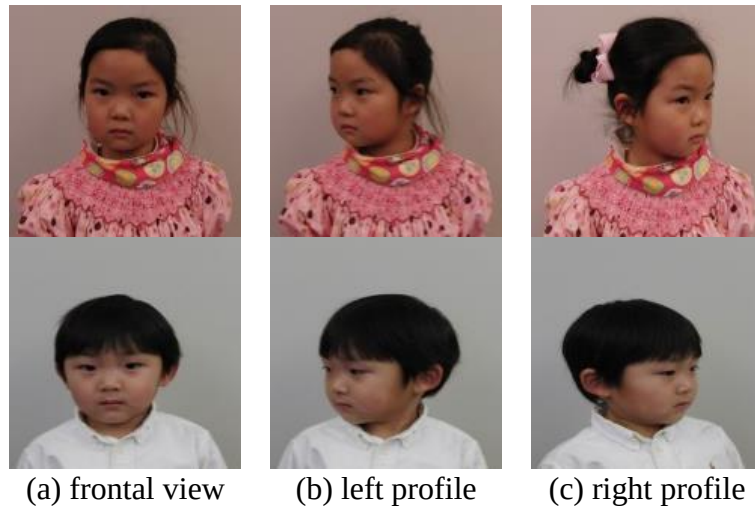


Figure 4.22: Sample color images representing each pose in IRIS-HFD-2014-V2.

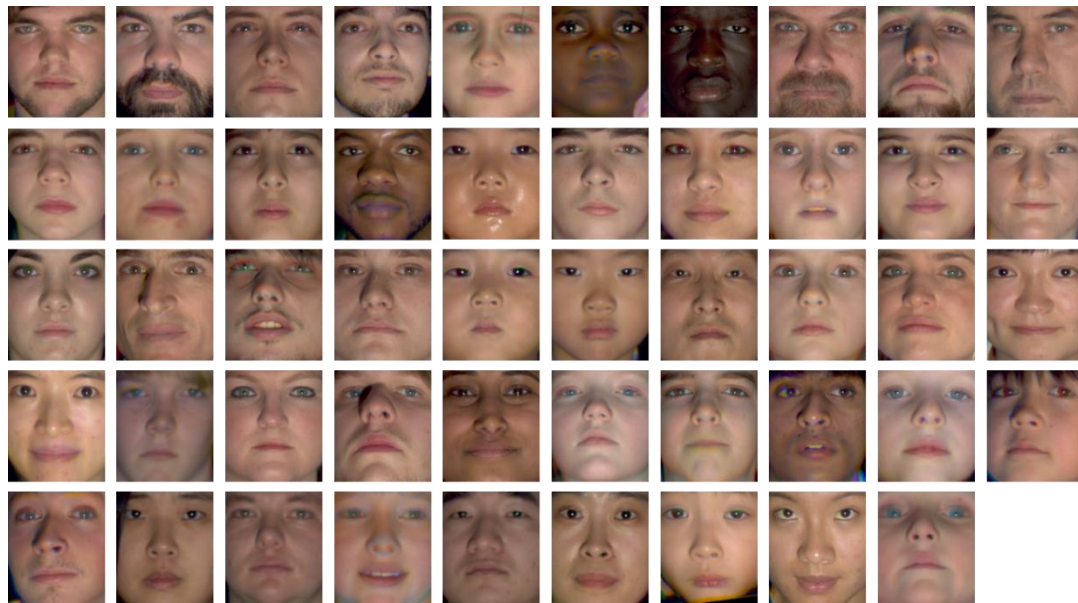


Figure 4.23: Entire aligned datasets of IRIS-HFD-2014-V2 in the frontal view where full sub-bands (420 nm to 700 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.

5 Proposed Spectral Data Compression Method

Spectral imaging typically generates a large amount of high-dimensional data that are acquired in different bands for each spatial location of interest. The high dimensionality of spectral data imposes limitations on numerical analysis. As such, there is an emerging demand for robust data compression techniques; specifically post-processing techniques that can cope with real spectral data.

In this Chapter, we describe a reduced-order spectral data modeling (ROM) technique based on local proper orthogonal decomposition (LPOD), which can be used as an efficient procedure to compute low-dimensional models by projecting high-dimensional clusters onto subspaces spanned by local reduced-order bases (LROB). Experimental results are reported on three public domain databases and an in-house database. Comparisons with three leading spectral recovery techniques, three decomposition techniques used for hyperspectral imaging, and two baseline techniques show that the proposed method leads to promising improvement on spectral and colorimetric accuracy corresponding to the reconstructed spectral reflectance.

5.1 Background on proper orthogonal decomposition

Proper orthogonal decomposition (POD) is perhaps one of the most assuring techniques for model reduction. POD has received an increasing amount of attention due to its potential for optimally extracting empirical information from numerical simulations or experimental data.

POD was first developed in the context of turbulence by Lumley in [Lumley67]. As stated in [Lumley07], POD can be seen as a way to replace the traditional Fourier decomposition in nonhomogeneous directions. POD was established by a variety of researchers according to [Holmes98]: Kosambi [Kosambi43], Loève [Loève55], and Obukhov [Obukhov42]. The procedure of POD has been used in different fields: signal analysis [Nobach07],

classification of human faces [Kirby90], data compression [Freund03] and, recently, optimal control [Djouadi08], [Camphou.05].

In various applications, POD has been used to extract coherent structures which are distinct patterns that are highly correlated to neighboring ensembles in space and time. Hence, POD plays a strong role in building low-dimensional models for a given set of computational data by truncating an optimal set of basis functions. In principle, the method of snapshots proposed by Sirovich [Sirovich87] is a procedure for extracting the POD modes without explicitly computing the kernel needed for the POD, where the eigenfunctions can be represented as the summation of snapshots. The optimality and geometric interpretation of POD in terms of linear n -widths of the corresponding compact operators are discussed in [Djouadi08]. The n -widths measure the inherent errors which are generated in the information-collecting stage of simulation or identification. According to [Djouadi12], optimal approximation without the usual POD integral and inner product constraints should be attained by solving a nonlinear optimization problem where the partial differential equation (PDE) solution is estimated by operators of a given finite rank in the respective trace class 2-norm. As mentioned in [Efe03], there are a variety of many powerful analytical techniques for solving diverse linear PDEs, particularly Galerkin projection. For multiple dimensional processes, the POD should be employed with the method of snapshots. Otherwise, for single dimensional processes, the equivalent modeling procedure can be followed by using the SVD technique. We refer the interested reader to [Holmes98] and the references therein for more detail.

In this dissertation, we use POD to find a relevant set of basis functions from spectral data. By means of a few basis functions, we can identify a low-dimensional subspace to construct a fidelity model by projection of the reconstruction equation that will be shown in the next section. In the preceding section, we present the theory behind the canonical POD and we then show the close link between the singular value decomposition (SVD) and POD by exploring SVD with respect to POD.

5.2 Theory of proper orthogonal decomposition

In this section, we introduce the proper orthogonal decomposition (POD) technique in the Euclidean space $\mathbb{R}^m$. The fundamental purpose behind POD is to yield an optimally ordered, orthonormal basis for the modal decomposition of an ensemble of functions [Holmes98].

Suppose that we have a set of the sampled data $A = \{a_1, a_2, \dots, a_n\} \subset [0,1]$ in subspace V of $\mathbb{R}^m$. In finding the proper representations of vectors $\{a_j\}_{j=1}^n$, we need to project each a_j

onto a candidate basis $\{\varphi_d\}$. We represent $\{a_j\}_{j=1}^n$ as a linear combination of basis functions:

$$a_j = \sum_{d=1}^r c_{dj} \varphi_d, \text{ for } 1 \leq r \leq m. \quad (5.1)$$

The coefficients c_{dj} are computed as:

$$c_{dj} = \langle a_j, \varphi_d \rangle_{\mathbb{R}^m} \quad (5.2)$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}^m}$ denotes the Euclidean inner product in $\mathbb{R}^m$. The coefficients c_{dj} correspond to the basis functions φ_d . The optimal subspace for representing the data is $V_r \subset V$ that is specified by $V_r = \text{span}\{\varphi_1, \dots, \varphi_r\}$. The vectors $\{\varphi_d\}_{d=1}^r$ are referred to as POD modes. To determine the orthonormal basis function $\{\varphi_d\}_{d=1}^r, r \leq m$, we arrange the data set A in an $m \times n$ matrix $\mathbb{A}$:

$$\mathbb{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \mathbb{A} \in \mathbb{R}^{m \times n}. \quad (5.3)$$

Then, we derive the orthonormal basis function $\{\varphi_d\}_{d=1}^r, r \leq m$ by truncating the SVD of $\mathbb{A}$ to the length r because POD is closely associated with SVD. Specifically, let $A = [a_1, \dots, a_n]$ be a real-valued $m \times n$ matrix of rank $d \leq \min(m, n)$ with columns $a_j \in \mathbb{R}^m, 1 \leq j \leq n$. Essentially, SVD guarantees the existence of real singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_d \geq 0$ and orthogonal matrices $U \in \mathbb{R}^{m \times m}$ with columns $\{u_i\}_{i=1}^m$ and $V \in \mathbb{R}^{n \times n}$ with columns $\{v_i\}_{i=1}^n$ such that

$$U^T A V = \begin{pmatrix} \Sigma_d & 0 \\ 0 & 0 \end{pmatrix} \triangleq \Sigma \in \mathbb{R}^{m \times n}, \quad (5.4)$$

where $\Sigma_d = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \in \mathbb{R}^{d \times d}$ and the rank of A equals the number of nonzero singular values. From SVD, we can derive that the vectors $\{u_i\}_{i=1}^d$ and $\{v_i\}_{i=1}^d$ are the eigenvectors for AA^T and $A^T A$, respectively, with eigenvalues $\lambda_i = \sigma_i^2$ on the diagonal. The approximation of the columns a_j in A for $d < n$ can be optimized by the first d eigenvectors $\{u_i\}_{i=1}^d$. Furthermore, an analogous result to the matrix $A^T A$ holds. The POD basis of $A^T A \in \mathbb{R}^{n \times n}$ with rank r can be achieved by solving the $n \times n$ eigenvalue problem as follows:

$$A^T A v_i = \lambda_i v_i, \text{ for } 1 \leq i \leq r. \quad (5.5)$$

By multiplying A on both sides in (5.5), we have the eigenvector u_i of AA^T due to $u_i = A v_i$ and so we can obtain the POD basis of $A^T A$ such that

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i. \quad (5.6)$$

Finally, we can directly construct the approximated $\tilde{A}$ with a few basis functions u_i from (5.6) that is used as the orthonormal basis functions $\{\varphi_d\}$ in (5.1) and (5.2).

In the following section, we introduce a robust reduced-order modeling (ROM) method which can be carried out using fewer dominant eigenvalues but significantly improves the performance of ROM for spectral data.

5.3 Local proper orthogonal decomposition

This section illustrates local proper orthogonal decomposition (LPOD) based on k-means clustering, which is a reduced-order modeling (ROM) technique. The proposed LPOD broadly consists of two steps: 1) clustering a set of the given data A and 2) constructing local reduced-order bases (LROB). The difference between global POD and local POD is illustrated in Fig. 5.1. In LPOD, A is grouped into k subsets (or clusters) based on a k-means clustering method. Next, for each subset, we construct LROBs to project the full solution space. Finally, we find the approximated solution with the constructed LROBs.

5.3.1 Clustering a set of the sample data

In the problem to classify a set of k points in $\mathbb{R}^m$ [Kanungo02], three issues should be considered: 1) the preferred data points to cluster, 2) the centroid of a cluster, and 3) the criterion of clustering. In our case, A is what we expect to group into k clusters; the centroid of a cluster is closest to the current state; and the clustering criterion is evaluated by assigning a_j to the cluster if the relative distance between a_j and the corresponding cluster centroid is smallest.

A solution, which contains markedly different features, relies on subdividing the dataset A into sub-regions. Accordingly, we first group together nearby analogous snapshots to attain a partition of A to construct the local reduced-order bases (LROB). In this dissertation, we use the k-means clustering method for grouping a_j into k clusters because of its simplicity and computational efficiency [Kanungo02], [Laszlo12]. In addition, k-means clustering is one of the most popular clustering techniques that has been significantly applied to many different applications and is freely available in many open libraries, including open source computer vision library (OpenCV)²³. Note that whereas the k-means method plays a fundamental role in clustering A , alternative clustering methods could be used for LPOD if some modifications are considered to obtain a steady partition and a consistent selection procedure from the data space.

²³ <http://opencv.org/>.

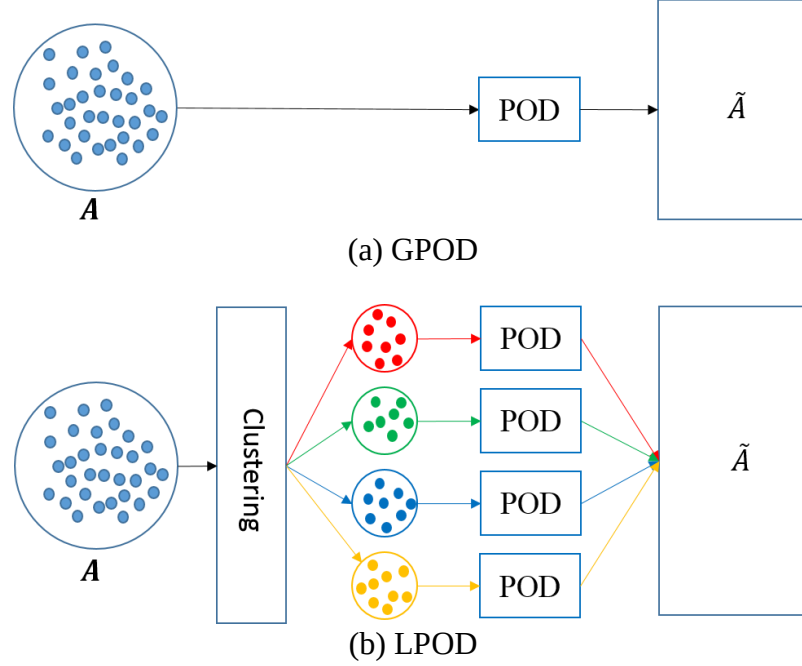


Figure 5.1: Comparison for overviews of (a) global POD and (b) local POD proposed in this dissertation. As the difference between global and local POD illustrated in (a) and (b), a clustering method in local POD splits A into k groups. For each group, the POD approximation is performed to construct local reduced-order bases (LROB). Then, we find the approximated $\tilde{A}$.

As mentioned in several earlier papers [Aldahd.13], [Kanungo02], [Laszlo12], the performance of the k-means clustering is significantly dependent on the initial centroid. Thus, we randomly allocate k initial locations to the cluster centroids by means of leading to a good set of initial centers and, then, iteratively find an optimal partition of A into $\chi_1 \cup \chi_2 \cup \dots \cup \chi_k$ so as to minimize the objective function (5.7) until either the cluster centroids have converged or a maximum iteration $t_{\max}$ is reached. The objective function of the k-means method yields the least within-cluster sum of squares (WCSS):

$$\arg \min_{\chi} \sum_{i=1}^k \sum_{a_j \in \chi_i} d(a_j, a_c^i)^2, \text{ for } i \in [1, k], j \in [1, n], \quad (5.7)$$

where a_j and a_c^i denote the given data and cluster centroids, and $d(a_j, a_c^i)$ is the relative distance between a_j and a_c^i . The distance metric is defined as:

$$d(a_j, a_c^i) = \|a_j - a_c^i\|_M = \sqrt{\|a_j\|_M^2 - 2a_c^{iT} M a_j + \|a_c^i\|_M^2} \quad (5.8)$$

where $M \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix. In this dissertation, we consider $M = I_n$ to get the 2-norm, $\|\cdot\|_2$.

While solving the optimization problem, if all a_j has been allocated to candidate clusters, new cluster centroids are updated by averaging a_j assigned to each cluster:

$$a_c^i = \frac{1}{|\chi_i|} \sum_{a_j \in \chi_i} a_j. \quad (5.9)$$

As a result, each a_j is assigned to the respective cluster χ_i that contains the same cluster index. For further details we refer to [Gordon99], [Jain99].

5.3.2 Construction of local reduced-order bases

In constructing local reduced-order bases (LROB), we first build the correlation matrix with each of the clusters $\{\chi_i\}_{i=1}^k$ obtained from the previous section:

$$R_i = \langle \chi_i, \chi_i^T \rangle, \text{ for } i \in [1, k], \quad (5.10)$$

where R_i is a $p \times p$ positive semi-definite matrix. Next, we perform a SVD of the matrix R_i to derive LROB by truncating SVD of length $d \leq \text{rank}(R_i)$ such that

$$R_i = U_i^d \Sigma_i V_i^{dT}, \quad (5.11)$$

where U_i^d and V_i^d are orthogonal matrices with reduced rank d and Σ_i is the diagonal matrix with singular values sorted in descending order. The columns of the matrix U_i^d then form a LROB which can be seen as the optimal POD basis of rank d . We then normalize the singular vectors of the SVD as $\|U_i^d\|_2 = 1/\Sigma_i^d$. The final step is to project the LROB onto the subspaces spanned by the empirical orthonormal basis functions. The POD basis functions $\{\varphi_i^d\}_{d=1}^r$ are computed by linearly combining the LROBs with the clusters χ_i as:

$$\varphi_i^d = \sum_{l=1}^p U_{i,l}^d \chi_{i,l}, \quad (5.12)$$

where $U_{i,l}^d$ and $\chi_{i,l}$ are the l th components U_i^d and χ_i . Once the uncorrelated POD coefficients c_i^d are derived by solving the linear system (5.2) with the corresponding orthonormal function $\{\varphi_i^d\}$, every member of the given data can be reconstructed as:

$$\{\tilde{A}\}_{i=1}^r \approx \{\sum_{d=1}^r c_i^d \varphi_i^d\}_{i=1}^k. \quad (5.13)$$

Note that for implementation purposes the index of the given data should be stored to properly recover a high-fidelity model $\tilde{A}$ approximated by (5.13), before and after we perform the k-means.

5.3.3 Choice of parameter k

In k-means clustering [Laszlo12], the number of clusters k (NC) must be selected as an initial parameter before the clustering process. However, it would be rather difficult to determine the valid number of clusters without yielding any empty clusters. One way to deal with the problem of NC is to analyze the quality of clustering results with internal indices [Maulick02], [Arbelaitz13].

The internal index is to assess the goodness of a clustering structure with objective information without any *a priori* knowledge about the data. Indeed, according to [Arbelaitz13], there are 30 distinct ways to confirm the cluster validity index (CVI). However, we are interested in the Calinski-Harabasz index (CH) [Maulick02], which was one of the most recommended indices and performed the best possible results to achieve the optimal NC with k-means clustering method in [Arbelaitz13]. The CH index aims to find the maximum hierarchy level (MHL) used to determine the appropriate NC in the preferred data. This CVI is computed as [Maulick02]:

$$CH(k) = \frac{|A| - k}{k - 1} \frac{\sum_{i=1}^k |\chi_i| d(a_c^i, \bar{A})}{\sum_{i=1}^k \sum_{a_j \in \chi_i} d(a_j, a_c^i)}, \quad (5.14)$$

where $\bar{A}$ is the centroid of A . On both Munsell and NCS databases, the optimal NCs estimated by CH index are two as shown in Fig. 5.2. Figure 5.2 demonstrates the criterion scores assessed by CH index across different k from 2 to 20 where the optimal NC is marked in red on each database. However, the reconstruction accuracy with optimal NC corresponding to the five error metrics is less discriminated than other methods in Section 5.4.3. Hence, instead of searching the optimal NC, we construct a set of multi-scale NCs, $\rho = \{\rho_1, \rho_2, \rho_3\}$, with scale factors: 0.01, 0.03, and 0.05. More specifically, we increase NC by steps of 1 percent up to 5 percent of the spatial dimension of the input data. In the cases of hyperspectral imaging, we set the multi-scale NC to be proportional to the shorter spatial dimension. The primary consideration specific to the limitation of NC (5 percent of the original spatial dimension), is to avoid too large k values as achieving a global optimum becomes much more difficult as NC increases [Krishna99].

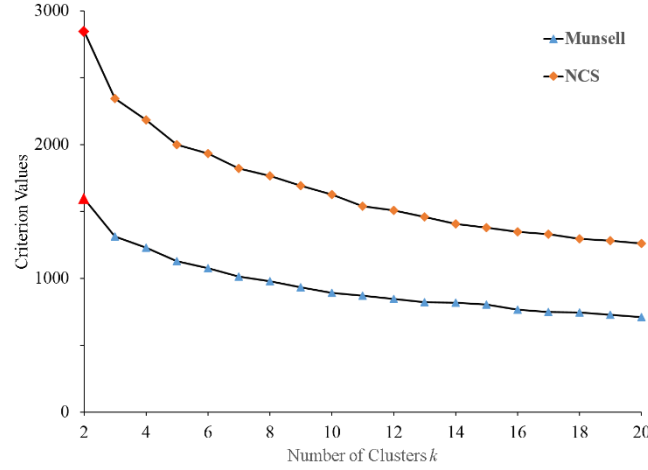


Figure 5.2: Results for the cluster validity indices in Calinski-Harabasz (CH) criterion on the Munsell and Natural color system (NCS) databases. In both databases, the optimal number of clusters k is two.

5.4 Experimental Results

In this section, we test our reduced-order data modeling approach compared with eight leading techniques on four databases. To assess the effectiveness of the proposed method, five error metrics are employed. These error metrics have been commonly used for evaluating the performance of reconstruction of reflecting specimens. Experimental results on each database are reported as average and standard deviations of the outcomes of the testing approaches for each error metric in the following sections.

5.4.1 Test databases

For a comprehensive study on the performance of the proposed method, we test four databases: 1) the Munsell database [SCRG], 2) a natural color system (NCS) [Hård81] database measured by Peyvandi [Peyvandi11], 3) CAVE database [Yasuma10], and 4) an in-house database including 17 different subjects recently built. The first two databases are used for evaluation of the hyperspectral signal and the latter two databases are used for testing the compression of hyperspectral image and the reconstruction of reflecting specimens.

The Munsell database [SCRG], which is composed of 1269 color chips measured by spectrophotometers ranging from 380nm to 800nm with a 1 nm interval, has been used to test whether a certain basis function is well-suited for data compression and reconstruction.

The NCS database consists of 1,950 color patches measured by a Datacolor 550TM Spectrophotometer ranging from 360nm to 700nm with intervals of 10 nm. The NCS color patches [Hård81] are intended to aid in measuring colors by visual comparison and have additionally been used to evaluate the performance for compression and reconstruction.

In this dissertation we only consider the range of the visible spectrum from 400 to 700 nm with 10 nm intervals. In the case under which the dataset was measured over our preferred wavelength range, we rearrange the dataset to match the desired wavelength range. The wavelength interval is set to the maximum interval among the datasets in order to avoid having to use any interpolation techniques on the selected datasets that were measured in different ranges and intervals. Thus, we choose $m = 1269$ and $n = 31$ for the Munsell database throughout our experiments.

CAVE [Yasuma10] consists of 32 static scenes that include 31-band hyperspectral images ranging from 400 to 700 nm at 10 nm steps. It was acquired using a cooled charge-coupled device (CCD) camera (Apogee Alta F260, providing a resolution of 512×512 pixels) with a VariSpec liquid crystal tunable filter (LCTF) under CIE illuminant D65. The 32 static scenes are classified into five sections: 1) stuff, 2) skin and hair, 3) paints, 4) food and drinks, and 5) real and fake. It is suitable for researchers who work in hyperspectral imaging to estimate the effectiveness of their proposed methods broadly used for scene segmentation and material recognition.

For the in-house database, the face images of 17 individuals were captured with the X-rite ColorChecker Classic that was placed to the side of the individuals to allow for color calibration. We utilized the LCTF covering the visible range from 420 to 700 nm in steps of 10 nm (29 narrow bands) under a Lumia 5.1 Reef as a light source.

At this point, we remark that for hyperspectral image databases m denotes the number of pixels of an image and n indicates the number of sub-bands. As mentioned in Section 5.3.3, we utilize a set of multiscale NCs, $\rho = \{\rho_1, \rho_2, \rho_3\}$, as increasing NC by steps of 1 to 5% of the spatial dimension of the input data. For hyperspectral images, we set the multiscale NC to be proportional to the shorter spatial dimension. In our experiments, we construct the LROBs to all of the clusters using the same number of subspace (NS) as increasing NS from 3 to 6.

5.4.2 Error metrics

For spectral and colorimetric accuracy, the spectral data compressed by the proposed method are analyzed using spectral reflectance. We consider five different error metrics: 1) the CIE color difference [Berns00], ΔE_{ab} , under CIE illuminant D65 with 2-degree defined as CIE 1931 standard observer for colorimetric accuracy, 2) the goodness of fit coefficient (GFC) [García.98], [Hernandez.01], 3) the peak signal-to-noise ratio (PSNR) [Bianco10], 4) the root mean square (RMS) difference [Amirshahi10], [Berns05], [Drew11], and 5) an index of metamerism [Berns00], [Berns05], [Fairman87] between CIE illuminants A and D65 for spectral accuracy.

CIELAB, which is the standard perceptual uniform color space, is used for the specification of surface colors. Computing a color difference ΔE_{ab} in CIELAB space allows one to compare equally perceptual color differences between pairs of samples. Let o and $\tilde{o}$ be the original and estimated reflectance, respectively. The color difference ΔE_{ab} can be computed as in [Westland04]:

$$\Delta E_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}, \quad (5.15)$$

where $\Delta L^* = L_o^* - L_{\tilde{o}}^*$, $\Delta a^* = a_o^* - a_{\tilde{o}}^*$, and $\Delta b^* = b_o^* - b_{\tilde{o}}^*$.

GFC [García.98], [Hernandez.01], based on the inequality of Schwartz, is for measuring the likelihood between o and $\tilde{o}$ that is intended to evaluate the accuracy of the mathematical reconstruction:

$$GFC = \frac{\langle o, \tilde{o} \rangle}{\langle o, o \rangle^{1/2} \langle \tilde{o}, \tilde{o} \rangle^{1/2}}. \quad (5.16)$$

The score of GFC ranges from zero to one where one indicates a perfect reconstruction.

PSNR [Bianco10] and RMS [Amirshahi10], [Berns05], [Drew11] are adopted for objective measures of the accuracy of spectral reconstruction, whereas ΔE_{ab} is for the evaluation of color reproducibility. MSE indicates the cumulative squared error between o and $\tilde{o}$, whereas PSNR is intended to estimate the peak error computed as

$$PSNR = 10 \log_{10} \frac{m_f^2}{M}, \quad (5.17)$$

where m_f^2 is the maximum fluctuation in the input data and M denotes MSE:

$$M = \frac{1}{n} \sum_{i=1}^n (o_i - \tilde{o}_i)^2, \quad (5.18)$$

where o_i and $\tilde{o}_i$ denotes i th samples of n -dimensional o and $\tilde{o}$. The root mean squared error (RMS) can be readily derived from (5.18) as follows:

$$RMS = \sqrt{M}. \quad (5.19)$$

Finally, the metamer index (MI) [Berns00], [Berns05], [Fairman87] is to evaluate the color stability in illumination changes. In other words, HI indicates how well two samples match under two different illuminants. The CIE color difference ΔE_{00} is used for calculating the MI such that

$$MI = \sqrt{\left(\frac{\Delta L}{s_L}\right)^2 + \left(\frac{\Delta C_{ab}}{s_C}\right)^2 + \left(\frac{\Delta H_{ab}}{s_H}\right)^2}, \quad (5.20)$$

where $\Delta L = L_o - L_{\tilde{o}}$ for the trial condition. ΔC_{ab} and ΔH_{ab} denote the chroma and hue differences. s_L , s_C , and s_H denote positional functions that correct the lack of visual uniformity of CIELAB for the same set of reference conditions. As mentioned in [Fairman87], the HI first corrects each pair of spectra to have zero color difference for a reference illuminant; ΔE_{00} is then calculated for a test illuminant wherein properties of test illuminant differ characteristically from those of test reference illuminant. As the spectral match gets poorer, the MI increases. In our experiments, the metamer index is calculated with CIE illuminant D65 as the reference and CIE illuminant A as the test illuminant, and vice versa.

Table 5.1: Summary of the performance results MSE-GIA, MINIMAX, GMM, GPOD, and LPOD on the MUNSELL database.

Methods	ΔE_{ab}		RMS		PSNR (dB)		GFC		MI D65 to A		MI A to D65	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
<i>Sub. = 3</i>												
MSE_GIA	0.3587	0.3214	0.0229	0.0173	34.7	5.7	0.9940	0.0115	1.1713	0.7989	1.3319	0.9640
MINIMAX	0.3587	0.3214	0.0229	0.0173	34.7	5.7	0.9940	0.0115	1.1713	0.7989	1.3319	0.9640
GMM	0.2482	0.2699	0.0169	0.0144	37.7	6.3	0.9952	0.0124	0.7628	0.7485	0.8586	0.8549
GPOD	2.9954	3.3759	0.0195	0.0139	36.1	6.1	0.9949	0.0105	0.8722	0.6981	1.0046	0.8963
$LPOD_{k=\rho_1}$	2.1568	2.7746	0.0119	0.0082	40.2	5.5	0.9976	0.0053	0.6018	0.5305	0.7102	0.6307
$LPOD_{k=\rho_2}$	1.4273	1.5643	0.0078	0.0059	44.1	5.8	0.9992	0.0016	0.3372	0.2906	0.3741	0.3151
$LPOD_{k=\rho_3}$	1.0916	1.2087	0.0063	0.0050	46.2	6.1	0.9995	0.0007	0.2510	0.2139	0.2790	0.2379
<i>Sub. = 4</i>												
MSE_GIA	0.3601	0.3244	0.0224	0.0174	34.7	5.3	0.9940	0.0114	1.1903	0.8340	1.3611	1.0346
MINIMAX	0.4573	0.3731	0.0292	0.0215	33.0	6.6	0.9912	0.0142	1.4944	1.0559	1.7130	1.2314
GMM	0.2269	0.2370	0.0156	0.0142	38.7	6.6	0.9961	0.0123	0.7379	0.6390	0.8487	0.8070
GPOD	1.8887	2.0182	0.0136	0.0096	39.1	5.6	0.9973	0.0068	0.4018	0.4271	0.3839	0.4052
$LPOD_{k=\rho_1}$	1.1594	1.3071	0.0073	0.0047	44.3	5.4	0.9991	0.0028	0.2770	0.2744	0.3132	0.3354
$LPOD_{k=\rho_2}$	0.6519	0.7710	0.0044	0.0027	48.6	5.0	0.9997	0.0005	0.1417	0.1338	0.1661	0.1620
$LPOD_{k=\rho_3}$	0.4956	0.5816	0.0033	0.0022	51.0	5.0	0.9998	0.0003	0.1053	0.0973	0.1206	0.1133
<i>Sub. = 6</i>												
MSE_GIA	0.3587	0.3224	0.0223	0.0175	34.8	5.3	0.9940	0.0116	1.1868	0.8249	1.3559	1.0148
MINIMAX	0.4651	0.3936	0.0332	0.0245	32.0	6.7	0.9887	0.0188	1.2943	0.8897	1.3984	0.9823
GMM	0.1991	0.2093	0.0146	0.0133	39.3	6.7	0.9966	0.0115	0.6187	0.5607	0.6982	0.6667
GPOD	0.7438	0.6856	0.0101	0.0058	41.2	4.8	0.9987	0.0030	0.1836	0.2057	0.1999	0.1969
$LPOD_{k=\rho_1}$	0.2980	0.2892	0.0034	0.0020	50.4	4.6	0.9998	0.0005	0.0692	0.0783	0.0726	0.0847
$LPOD_{k=\rho_2}$	0.1997	0.2013	0.0021	0.0011	54.4	4.1	0.9999	0.0001	0.0387	0.0399	0.0435	0.0435
$LPOD_{k=\rho_3}$	0.1483	0.1391	0.0015	0.0007	57.0	4.1	1.0000	0.0001	0.0318	0.0340	0.0350	0.0381

5.4.3 Test on hyperspectral signal databases

As mentioned above, we first evaluate the proposed methods compared with global POD (GPOD) and the spectral recovery methods, including minimum of averaging root mean square error based on the generalized inverse approach model (MSE_GIA) [Peyvandi13], minimax of the estimation error based on GIA (MINIMAX) [Peyvandi13], and Gaussian mixture model (GMM) clustering [Peyvandi12] on two databases consisting of the Munsell and NCS. Note that the performance of LPOD without clustering the given data is the same as GPOD. For compression and reconstruction in all our experiments, we increase the number of dimensional subspaces from 3 to 6 corresponding to the length r in (5.1). Due to the space limitation, we do not include the results from 5-dimensional subspace experiments. Furthermore, the results in terms of the compression ratio for each of the testing methods, including the proposed method, are not provided as there is no significant difference.

Table 5.2: Summary of the performance results MSE-GIA, MINIMAX, GMM, GPOD, and LPOD on the NCS database.

Methods	ΔE_{ab}		RMS		PSNR (dB)		GFC		MI D65 to A		MI A to D65	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
<i>Sub. = 3</i>												
MSE_GIA	0.3463	0.3480	0.0244	0.0192	34.0	5.2	0.9944	0.0121	1.0244	0.9143	1.1670	1.0375
MINIMAX	0.3463	0.3480	0.0244	0.0192	34.0	5.2	0.9944	0.0121	1.0244	0.9143	1.1670	1.0375
GMM	0.2512	0.2613	0.0195	0.0167	36.1	5.4	0.9961	0.0105	0.6965	0.6655	0.7956	0.7447
GPOD	3.2753	4.1828	0.0199	0.0151	35.8	5.5	0.9955	0.0116	0.6826	0.6948	0.9104	0.8794
$LPOD_{k=\rho_1}$	2.2792	3.3205	0.0126	0.0080	39.5	5.3	0.9982	0.0048	0.4689	0.4676	0.5505	0.5296
$LPOD_{k=\rho_2}$	1.5638	1.7583	0.0093	0.0056	42.1	5.3	0.9992	0.0017	0.3105	0.2876	0.3659	0.3197
$LPOD_{k=\rho_3}$	1.3626	1.5375	0.0082	0.0055	43.6	6.0	0.9995	0.0008	0.2279	0.1835	0.2830	0.2302
<i>Sub. = 4</i>												
MSE_GIA	0.3426	0.3454	0.0238	0.0193	34.2	5.1	0.9944	0.0120	1.0235	0.9290	1.1783	1.0921
MINIMAX	0.8741	0.8049	0.0542	0.0263	26.5	5.1	0.9742	0.0480	3.3029	3.7035	3.7197	4.0498
GMM	0.2705	0.2895	0.0193	0.0155	36.2	5.7	0.9951	0.0131	0.7365	0.7949	0.8312	0.8616
GPOD	1.0984	1.0591	0.0145	0.0101	38.3	5.1	0.9977	0.0066	0.4713	0.5947	0.4978	0.6694
$LPOD_{k=\rho_1}$	1.0567	1.3722	0.0082	0.0048	43.1	5.1	0.9992	0.0021	0.2662	0.3222	0.3065	0.3731
$LPOD_{k=\rho_2}$	0.7456	0.8100	0.0062	0.0037	45.8	5.4	0.9997	0.0007	0.1659	0.1462	0.1973	0.1957
$LPOD_{k=\rho_3}$	0.6649	0.7322	0.0052	0.0033	47.3	5.8	0.9998	0.0004	0.1366	0.1187	0.1632	0.1571
<i>Sub. = 6</i>												
MSE_GIA	0.3414	0.3447	0.0237	0.0193	34.2	5.1	0.9944	0.0121	1.0198	0.9260	1.1727	1.0830
MINIMAX	0.8042	0.7166	0.0554	0.0264	26.3	5.0	0.9728	0.0504	2.4176	2.7676	2.6473	2.9543
GMM	0.2196	0.2249	0.0164	0.0145	37.8	5.9	0.9971	0.0104	0.6082	0.5802	0.6998	0.6824
GPOD	0.6424	0.6605	0.0090	0.0047	42.0	4.3	0.9993	0.0011	0.0631	0.0520	0.0765	0.0558
$LPOD_{k=\rho_1}$	0.3555	0.3804	0.0041	0.0026	49.2	5.5	0.9999	0.0003	0.0676	0.0605	0.0728	0.0655
$LPOD_{k=\rho_2}$	0.2547	0.2687	0.0030	0.0019	52.2	5.6	0.9999	0.0002	0.0504	0.0501	0.0567	0.0591
$LPOD_{k=\rho_3}$	0.2161	0.2294	0.0023	0.0015	54.4	5.6	1.0000	0.0001	0.0452	0.0439	0.0534	0.0573

The summary of the performance results in terms of five error metrics are shown in Table 5.1 and 5.2. For ΔE_{ab} under CIE Illuminant D65, MSE_GIA, MINIMAX, and GMM perform better than GPOD and LPOD in 3 and 4 dimensional subspace. However, when

the dimension of a subspace is six, $LPOD_{k=\rho_3}$ becomes more accurate than three spectral recovery methods. For the four remaining error metrics, GPOD and LPOD should be able to yield not just comparable, but better results. As we increase NC, we can find that the accuracy of GPOD and LPOD gradually increases for each of the considered metrics since RIC approaches one, whereas it is noticeable that the accuracy of the spectral recovery methods in higher dimensional subspace slightly increases. Furthermore, as reflected in the results, particularly with respect to the MI, three spectral recovery methods are sensitive to changing illuminants. The results establish that those methods fit poorly at long wavelengths when altering CIE illuminants D65 to A, and at short wavelengths when altering CIE illuminants A to D65 as mentioned in [Berns05]. As a result, the overall performance of LPOD on the test databases is consistent with the results in five considered merits. More specifically, LPOD is little affected by the choice of the spectral signal data and is quite stable.

Table 5.3: Summary of the performance results ICA, NMF, RSNMU, RSPNMU, GPOD, and LPOD on the CAVE database.

Methods	ΔE_{ab}		RMS		PSNR (dB)		GFC		MI D65 to A		MI A to D65	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
<i>Sub. = 3</i>												
ICA	6.4644	4.9820	0.0148	0.0121	41.0	6.5	0.9480	0.0642	1.1602	0.8478	1.1124	0.7906
NMF	10.883	8.5361	0.0194	0.0193	38.6	6.6	0.9181	0.0861	1.2048	0.8656	1.4380	1.0257
RSNMU	7.9239	7.0451	0.0205	0.0172	37.6	5.8	0.8959	0.1083	0.7466	0.5635	1.0163	0.7429
RSPNMU	7.6389	6.9547	0.0213	0.0182	37.4	5.9	0.8944	0.1097	0.7479	0.5908	1.0084	0.7680
GPOD	6.1354	4.8444	0.0117	0.0084	41.9	5.1	0.9421	0.0708	1.1933	0.8987	1.1372	0.8298
$LPOD_{k=\rho_1}$	3.6316	3.5568	0.0062	0.0063	48.6	6.8	0.9880	0.0373	0.4765	0.4646	0.5123	0.5121
$LPOD_{k=\rho_2}$	2.1014	2.3684	0.0042	0.0049	53.0	7.8	0.9950	0.0334	0.2760	0.3173	0.3219	0.3714
$LPOD_{k=\rho_3}$	1.7888	2.0724	0.0036	0.0045	54.1	7.8	0.9959	0.0326	0.2342	0.2759	0.2805	0.3349
<i>Sub. = 4</i>												
ICA	3.0133	2.4445	0.0105	0.0099	44.6	7.5	0.9765	0.0442	0.7637	0.5849	0.9033	0.7216
NMF	9.0275	7.5026	0.0158	0.0151	39.9	6.3	0.9267	0.0828	1.1019	0.8432	1.2952	0.9722
RSNMU	6.0576	5.8707	0.0150	0.0131	40.9	6.4	0.9386	0.0806	0.6177	0.5293	0.8226	0.6942
RSPNMU	5.9715	5.6794	0.0163	0.0137	39.8	6.1	0.9256	0.0915	0.6322	0.5441	0.8588	0.7271
GPOD	2.9689	2.4299	0.0077	0.0062	45.8	5.7	0.9729	0.0481	0.7632	0.6254	0.8878	0.7579
$LPOD_{k=\rho_1}$	1.7780	1.9164	0.0041	0.0044	52.2	7.1	0.9944	0.0317	0.3228	0.3318	0.3714	0.4055
$LPOD_{k=\rho_2}$	1.2221	1.4414	0.0029	0.0034	55.5	7.4	0.9967	0.0308	0.2139	0.2550	0.2618	0.3178
$LPOD_{k=\rho_3}$	1.0503	1.2716	0.0026	0.0031	56.4	7.3	0.9970	0.0302	0.1876	0.2308	0.2271	0.2836
<i>Sub. = 6</i>												
ICA	0.9417	0.7595	0.0059	0.0059	49.9	7.8	0.9945	0.0273	0.3510	0.2886	0.3908	0.3141
NMF	6.0819	6.1987	0.0117	0.0115	42.8	6.6	0.9497	0.0683	0.9108	0.7847	1.0580	0.8972
RSNMU	4.2112	4.1512	0.0100	0.0093	44.2	6.8	0.9719	0.0540	0.4673	0.4298	0.5858	0.5571
RSPNMU	3.8262	3.7149	0.0100	0.0096	44.3	6.8	0.9728	0.0556	0.4435	0.4466	0.5670	0.5782
GPOD	0.8392	0.7685	0.0037	0.0036	52.1	6.0	0.9943	0.0276	0.3074	0.2690	0.3469	0.2936
$LPOD_{k=\rho_1}$	0.4663	0.5719	0.0023	0.0024	57.0	6.7	0.9974	0.0279	0.1651	0.1923	0.1920	0.2287
$LPOD_{k=\rho_2}$	0.4122	0.5453	0.0018	0.0019	59.0	6.7	0.9978	0.0271	0.1209	0.1626	0.1419	0.1941
$LPOD_{k=\rho_3}$	0.3894	0.5667	0.0016	0.0017	59.6	6.5	0.9979	0.0258	0.1101	0.1522	0.1302	0.1816

5.4.4 Test on hyperspectral image databases

Following the quantitative validation with two hyperspectral signal datasets, in order to verify the usefulness of the proposed method on higher dimensional datasets, the proposed workflow is applied to two sets of hyperspectral image databases including CAVE and our in-house face database by comparing the performance with GPOD, ICA [Bianco10], [Hyvärinen00], NMF [Gillis14], RSNMU [Gillis13], and RSPNMU [Gillis12].

Table 5.4: Summary of the performance results ICA, NMF, RSNMU, RSPNMU, GPOD, and LPOD on the in-house database.

Methods	ΔE_{ab}		RMS		PSNR (dB)		GFC		MI D65 to A		MI A to D65	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
<i>Sub. = 3</i>												
ICA	5.2314	5.2818	0.0195	0.0185	36.8	7.3	0.9395	0.1418	1.0971	1.0050	1.5468	1.2902
NMF	9.3288	12.886	0.0354	0.0449	33.9	9.2	0.9148	0.1661	1.2454	1.0721	1.6029	1.3371
RSNMU	9.6341	12.532	0.0336	0.0444	33.9	8.8	0.9306	0.1592	0.9359	0.9051	1.2152	1.2309
RSPNMU	9.7917	12.410	0.0363	0.0471	34.1	12.8	0.9369	0.1481	0.9860	1.0193	1.2455	1.3264
GPOD	4.3573	4.9381	0.0184	0.0172	37.3	7.2	0.9413	0.1406	1.0456	0.9495	1.3579	1.1436
$LPOD_{k=\rho_1}$	2.6896	3.5335	0.0106	0.0091	41.1	5.8	0.9725	0.0738	0.6247	0.6481	0.7942	0.8730
$LPOD_{k=\rho_2}$	2.3705	3.3125	0.0088	0.0073	42.5	5.6	0.9793	0.0609	0.4665	0.5362	0.5983	0.7200
$LPOD_{k=\rho_3}$	2.1721	2.9949	0.0084	0.0068	43.0	5.9	0.9830	0.0541	0.3946	0.4838	0.5149	0.6557
<i>Sub. = 4</i>												
ICA	1.8990	2.0042	0.0142	0.0126	38.9	6.4	0.9483	0.1235	0.9220	1.1235	1.0178	1.3377
NMF	7.7031	10.843	0.0313	0.0383	34.5	8.7	0.9207	0.1597	1.2298	1.0897	1.5782	1.3343
RSNMU	7.4452	8.9995	0.0268	0.0306	35.0	8.0	0.9338	0.1531	0.8594	0.9098	1.1780	1.2505
RSPNMU	7.2814	8.5673	0.0283	0.0313	35.3	12.4	0.9405	0.1483	0.9405	1.0407	1.2059	1.3552
GPOD	1.8204	1.8930	0.0140	0.0124	39.1	6.3	0.9485	0.1286	0.9306	1.1306	1.0280	1.3544
$LPOD_{k=\rho_1}$	1.9284	2.6498	0.0086	0.0068	42.5	5.3	0.9781	0.0633	0.4332	0.4952	0.5964	0.6985
$LPOD_{k=\rho_2}$	1.7406	2.4828	0.0076	0.0058	43.6	5.4	0.9831	0.0528	0.3428	0.4222	0.4612	0.5878
$LPOD_{k=\rho_3}$	1.6181	2.3292	0.0073	0.0055	44.1	6.0	0.9866	0.0453	0.2987	0.3863	0.4080	0.5438
<i>Sub. = 6</i>												
ICA	0.7652	0.8261	0.0094	0.0066	41.7	5.5	0.9751	0.0677	0.2135	0.2430	0.2386	0.2560
NMF	7.0359	10.957	0.0248	0.0346	36.8	8.6	0.9245	0.1597	0.8674	0.9219	1.1282	1.1590
RSNMU	5.5414	6.7945	0.0214	0.0230	36.3	7.4	0.9414	0.1414	0.6775	0.6324	0.8274	0.8014
RSPNMU	5.2329	6.0331	0.0219	0.0228	37.3	12.7	0.9479	0.1331	0.7031	0.7665	0.8623	0.8974
GPOD	0.7448	0.8067	0.0091	0.0063	42.0	5.4	0.9757	0.0666	0.2163	0.2445	0.2416	0.2573
$LPOD_{k=\rho_1}$	0.9100	1.4119	0.0066	0.0046	44.6	4.9	0.9841	0.0506	0.2280	0.3131	0.2867	0.4228
$LPOD_{k=\rho_2}$	0.8794	1.3388	0.0061	0.0042	45.3	5.4	0.9878	0.0421	0.1926	0.2692	0.2386	0.3472
$LPOD_{k=\rho_3}$	0.8714	1.3617	0.0060	0.0041	45.9	6.6	0.9912	0.0333	0.1843	0.2557	0.2303	0.3344

As shown in Table 5.3 and 5.4, increasing the subspace dimension for NMF, RSNMU, and RSPNMU marginally affects the results of the spectral and colorimetric performances with regard to the reconstructed reflectance spectra. However, the results of those methods cannot be superior to the achievements of ICA with respect to all of five error metrics for each database, even when the increased subspace dimension is considered. The results achieved establish GPOD as a suitable approach to compress hyperspectral image databases as it is more accurate in terms of spectral approximation and also provides lower colorimetric differences when changing illuminants. Whereas GPOD and ICA in ΔE_{ab} yield slightly better results in the six dimensional subspace on in-house face database, the

conducted results indicate that the proposed method, $LPOD_{k=\rho_3}$, is able to reliably accomplish significant improvement according to the metrics considered with two real databases among all the testing methods.

5.5 Summary

In this Chapter, we have proposed local-based proper orthogonal decomposition (LPOD) methods by employing the k-means clustering approach, efficiently reducing the dimension of the given spectral data that were highly correlated to neighboring spectral samples. Whereas we applied k-means clustering to the set of snapshots for achieving good partitions, the proposed approach is flexible enough to accommodate other clustering methods such as hierarchical clustering, self-organizing maps (SOM), partitioning around medoids (PAM), etc. Note that the main reason for using k-means clustering was its simplicity and computational efficacy. To avoid imposing the limitations of the choice of the parameter k as a constant for each dataset, we first analyzed the quality of the clustering results based on the cluster validity index such as Calinski-Harabasz that can determine the optimal number of clusters k . However, we found that the determination of the optimal number of clusters by Calinski-harabasz index did not bring critical improvements on our scheme. Subsequently, we further devised multi-scale approach for choosing the number of clusters to enhance the spectral and colorimetric accuracy.

For a comprehensive study on the effectiveness of our reduced-order data modeling approach where six leading and two conventional techniques on three public domain databases and one our initial database in five error metrics were established, experimental results showed that the proposed method achieved highly competitive accuracy in terms of five considered error metrics on four different database.

6 Design of New Framework for Face Alignments

Hyperspectral imaging (HI) applied to non-rigid objects introduces new challenges such as inter-band misalignments (IBMs) resulting from subject motion during data acquisition. To deal with IBMs in HFIs, we investigate four different individual techniques: 1) manual alignment approaches (AAs) based on selecting regions of interest, such as fixed boundary box-based AA (FBB) [Denes02], [Di10], [Szeliski10], [Uzair13], [Uzair15] and eye coordinate-based AA (EC); 2) iterative convex optimization (ICO) processes for face alignment, such as RASL [Peng12] and ORIA [Wu12]; 3) landmark-based AAs, such as DRMF [Asthana13], IPCM [Asthana14], SDM [Xiong12], CDM [Yu13], TSPM [Zhu12]; and 4) two popular image AAs, such as Lucas-Kanade (LK) [Baker04] and SIFTFlow [Liu11]. In the following sections, we analyze those approaches on our database towards experimental determination of promising AAs.

6.1 Two conventional alignment approaches (FBB and EC)

We suppose that we are given a set of HFIs $I = \{I_1, I_2, \dots, I_n\}$ in $\mathbb{R}^{w \times h \times n}$ where w and h are width and height of HFIs, respectively. In aligning m -pixel region of interest (ROI) from $\{I_i\}_{i=1}^n$, $i \in [1, n]$, we define a function of the reshaping operation (i.e., vectorization) $f: \mathbb{R}^{w \times h \times n} \rightarrow \mathbb{R}^{m \times n}$, typically for $m \gg n$. To approximate a set of the aligned HFIs, $A = \{a_1, a_2, \dots, a_n\} \in \mathbb{R}^{m \times n}$, using fixed bounding boxes (FBB) and eye coordinates (EC) approaches, we define the canonical frame of the ROI within I_i as being $w' \times h'$ pixels and the distance between the selected eye points is normalized to ξ pixels. We note that whereas FBB [Denes02], [Di10], [Szeliski10], [Uzair13] is mostly retained to align HFDs by selecting the initial location of the ROI from sequential images, we adopt it for a task to build the ground truth in order to compare the studied metrics used to predict alignment quality (AQ) in Chapter 7.

In this dissertation, the canonical points of FBB to select the fixed ROI in I_i are initially obtained from the κ th eye coordinate set in EC. The overview of FBB is illustrated in Fig.

6.1 where we choose $\kappa = 1$ in our experiments. The selection of the ROI using the canonical points is the same as EC approach.

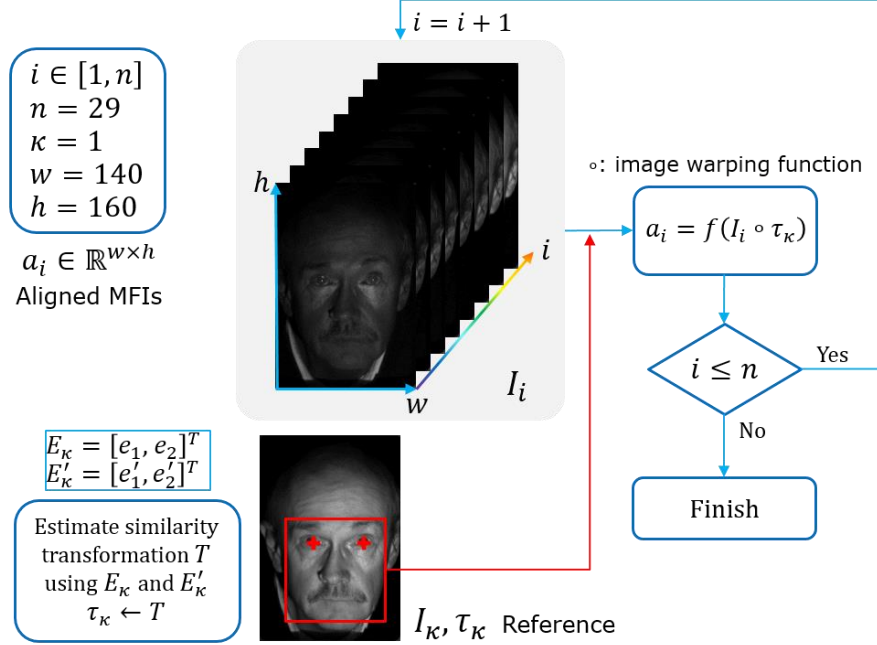


Figure 6.1: Overview of the fixed bounding box (FBB) approach where $\kappa = 1$ in our experiments.

The approach using the manual selection of eye coordinates is the most straightforward approach to address IBMs in HFIs. However, it should be hard to consistently select two eye coordinates at the same positions by hand over a large HFI set such as IRIS-HFD-2014. As shown in Fig. 6.2, the eye coordinates $E_i = [e_1, e_2]^T$ were manually selected to middle of the eyes for each I_i where a point e represents a pair of the coordinates (x, y) in $\mathbb{R}^2$.

To find a set of the domain transformation $\{\tau_i\}_{i=1}^n: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ in a Lie group $\mathbb{G}$ [Peng12], we define $\tau_i \leftarrow T$ and the canonical points $E' = [e'_1, e'_2] \in \mathbb{R}^{w' \times h'}$ where T denotes a 3×3 similarity transformation matrix [Hartley00]. The similarity transformation T is defined as:

$$T = \begin{bmatrix} t_1 & t_2 & t_3 \\ t_4 & t_5 & t_6 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6.1)$$

To estimate six unknown coefficients of T , we need to find one more points e_3 and e'_3 for E and E' , respectively, which can be defined as:

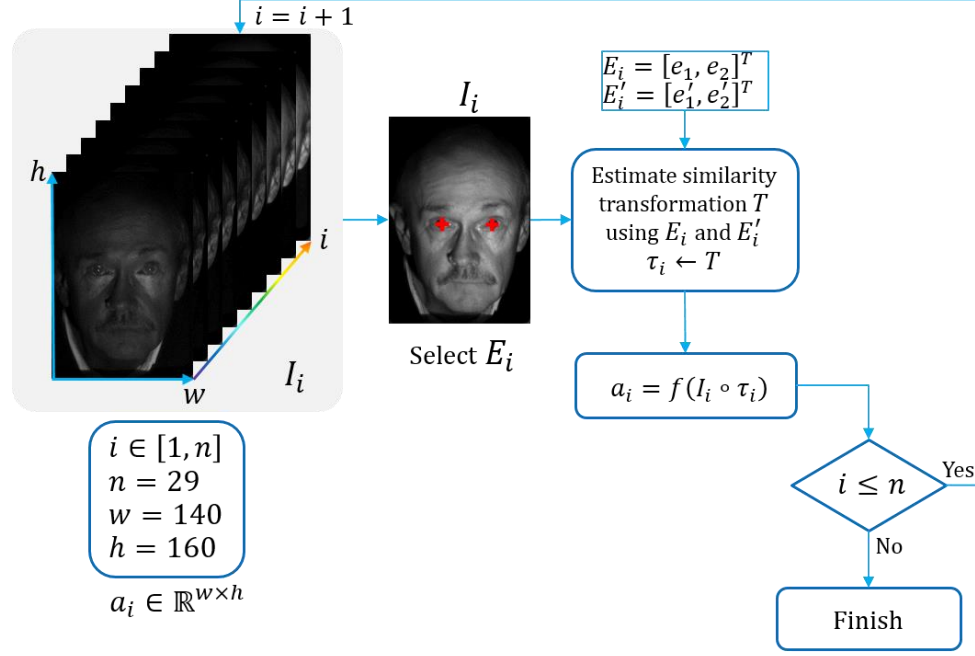


Figure 6.2: Overview of the eye coordinates (EC) approach.

$$e_3 = e_1 + \gamma^T(e_2 - e_1), e'_3 = e'_1 + \gamma^T(e'_2 - e'_1) \text{ for } \gamma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (6.2)$$

Then, we denote $e = (x, y, 1)$ as the homogeneous coordinates of a point [Hartley00] to solve $e'_\ell = T e_\ell, \ell \in [1, 2, 3]$ such that:

$$\begin{bmatrix} e'_{\ell,1} \\ e'_{\ell,2} \\ 1 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 & t_3 \\ t_4 & t_5 & t_6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{\ell,1} \\ e_{\ell,2} \\ 1 \end{bmatrix} \rightarrow \begin{cases} e'_{\ell,1} = t_1 e_{\ell,1} + t_2 e_{\ell,2} + t_3 \\ e'_{\ell,2} = t_4 e_{\ell,1} + t_5 e_{\ell,2} + t_6 \\ 1 = 1 \end{cases} \quad (6.3)$$

Without loss of generality, Equation (6.3) can be rewritten in block form as:

$$\begin{bmatrix} e'_{1,1} \\ e'_{1,2} \\ e'_{2,1} \\ e'_{2,2} \\ e'_{3,1} \\ e'_{3,2} \end{bmatrix} = \begin{bmatrix} \Omega_{3 \times 3} & 0 \\ 0 & \Omega_{3 \times 3} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix}, \text{ for } \Omega_{3 \times 3} = \begin{bmatrix} e_{1,1} & e_{1,2} & 1 \\ e_{2,1} & e_{2,2} & 1 \\ e_{3,1} & e_{3,2} & 1 \end{bmatrix}. \quad (6.4)$$

Finally, we can derive T by solving (6.4) as:

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix} = \begin{bmatrix} \Omega_{3 \times 3} & 0 \\ 0 & \Omega_{3 \times 3} \end{bmatrix}^{-1} \begin{bmatrix} e'_{1,1} \\ e'_{1,2} \\ e'_{2,1} \\ e'_{2,2} \\ e'_{3,1} \\ e'_{3,2} \end{bmatrix}. \quad (6.5)$$

After finding τ_i from T , we can compute the aligned images as:

$$a_i = f(I_i \circ \tau_i), \text{ for } i \in [1, n], \quad (6.6)$$

where $I_i \circ \tau_i$ denotes an image warping function such that $I_i \circ \tau_i(x, y) \triangleq I_i(\tau_i(x, y))$.



Figure 6.3: Results of FBB approach on ID: F009_02 in IRIS-HFD-2014-V1 where we select the canonical frame to be 140×160 pixels and $\xi = 80$ pixels for the distance between two eye coordinates throughout this dissertation.



Figure 6.4: Results of EC approach on ID: F009_02 in IRIS-HFD-2014-V1.

Figure 6.3 and 6.4 show the results of FBB and EC approaches, respectively. Note that the primary consideration specific to FBB approach is to establish an alignment-baseline to specify subject's movement during data acquisition as mentioned above. Therefore, there is no improvement in the results of FBB as evidenced in Fig. 6.3.

6.2 Iterative convex optimization approaches (RASL and ORIA)

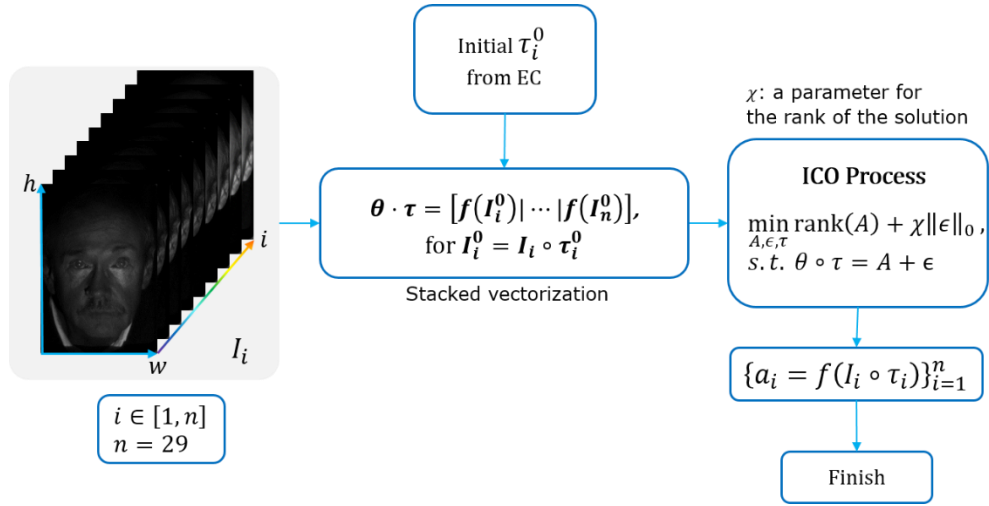


Figure 6.5: Overview of RASL approach.

Robust alignment by sparse and low-rank decomposition (RASL) [Peng12] and online robust image alignment (ORIA) [Wu12] based on iterative convex optimization (ICO) approach could be considered promising approaches to overcome IBMs in HFI. RASL seeks an optimal set of image domain transformations to minimize the rank of pairs of the transformed images:

$$\min_{A, \epsilon, \tau} \text{rank}(A) + \chi \|\epsilon\|_0 \text{ s.t. } \theta \circ \tau = A + \epsilon, \quad (6.7)$$

where $\chi > 0$ represents a parameter for the rank of the solution against the sparsity of the error, τ indicates a set of transformations, ϵ is the error, and $\|\cdot\|_0$ stands for ℓ_0 -norm. A is a matrix of stacking vectors as:

$$A \triangleq [f(I_1)|f(I_2)|\cdots|f(I_n)] \in \mathbb{R}^{m \times n}. \quad (6.8)$$

$\theta \circ \tau \triangleq [f(I_1^0)|f(I_2^0)|\cdots|f(I_n^0)] \in \mathbb{R}^{m \times n}$ where $I_i^0 = I_i \circ \tau_i$ in [Peng12].

Whereas RASL achieves impressive performance with high accuracy and consistency of face alignment for the entire in-house database, it requires eye coordinates for each sub-band of the input HFI set and has high computational cost and memory-demand. Thus, it becomes increasingly impractical for dynamically growing the total number of image sets. Figure 6.5 shows the overview of RASL alignment approach on a HFI set. The results of RASL on ID: F009_02 of IRIS-HFD-2014-V1 are shown in Fig. 6.6.



Figure 6.6: Results of RASL approach on ID: F009_02 in IRIS-HFD-2014-V1.

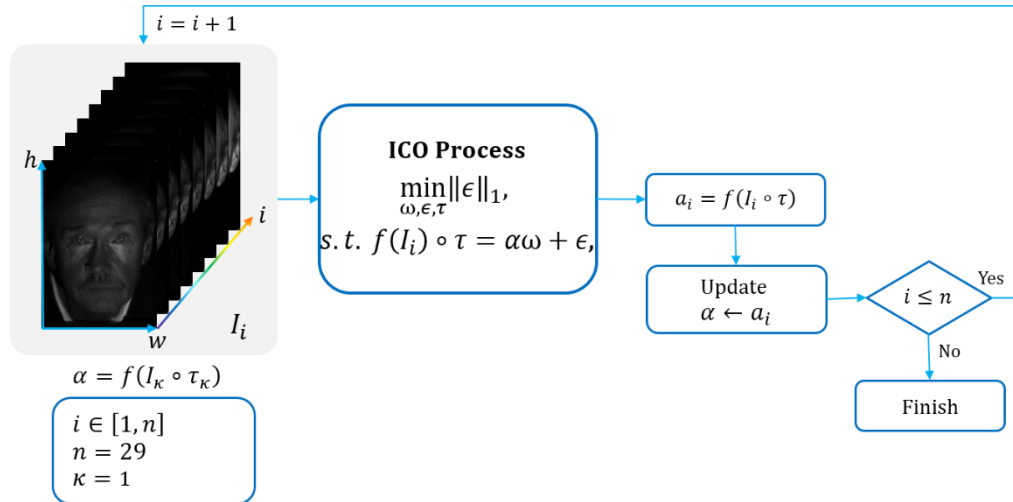


Figure 6.7: Overview of ORIA approach.

To address the limitations of RASL while maintaining the comparable accuracy of face alignment with low computational cost and memory-demand, ORIA was applied to solve a sequence of convex optimization to minimize an l_1 -norm and update previously well-aligned images at the same time. ORIA aims to find an optimal transformation $\tau: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by solving the sparsely regularized alignment [Wu12]:

$$\min_{\omega, \epsilon, \tau} \|\epsilon\|_1 \text{ s.t. } f(I_i \circ \tau) = \alpha\omega + \epsilon, \text{ for } i \in [1, n] \quad (3.9)$$

where $\|\cdot\|_1$ denotes the l_1 -norm, ω stands for the reconstruction coefficients, ϵ is the reconstruction errors, and α is initially defined as $\alpha = f(I_\kappa \circ \tau_\kappa)$ where τ_κ is achieved from EC. When ORIA finds τ , α is simultaneously updated as being $\alpha = a_i$ such that $a_i = f(I_i \circ \tau)$. Figure 6.7 illustrates the overview of ORIA approach on an input HFI set.



Figure 6.8: Results of ORIA approach on ID: F009_02 in IRIS-HFD-2014-V1.



Figure 6.9: Example of a failure of ORIA on ID: F129_01 in IRIS-HFD-2014-V1 at 600 nm, 630 nm, 660 nm, and 690 nm.

Although ORIA has lower computational cost than RASL and achieves better alignment images than FBB in most of the HFI sets in our database as shown in Fig. 6.8, it is more sensitive to the substantial subject motion compared to RASL. This is because ORIA only utilizes one EC set as the canonical points instead of using the entire EC set as RASL does. Furthermore, if ORIA fails to align one sub-band of an input HFI set, all other sub-bands are affected and, accordingly, will be misaligned. For example, on ID: F129_01 in our

database where it innately contains substantial subject motion from 630 nm to 700 nm, ORIA fails to establish correct alignment at those sub-bands shown in Fig. 6.9.

6.3 Landmark-based alignment approaches (LAs)

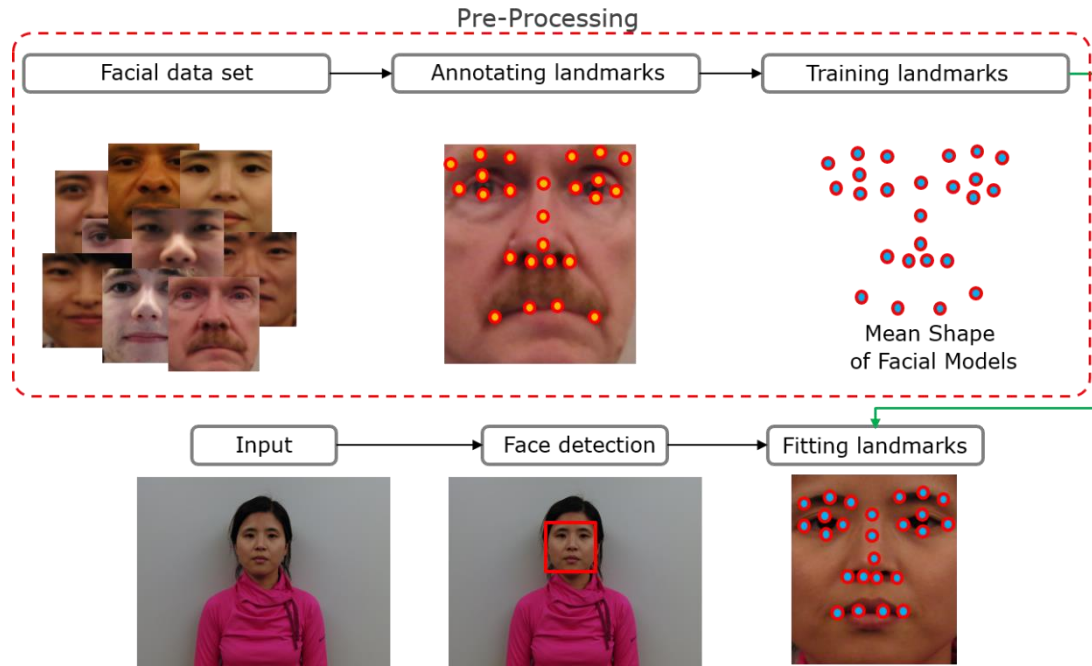


Figure 6.10: A framework of landmark-based alignment approach.

A viable way to deal with the complexity of aligning HFIs can be landmark-based AAs. Detecting and Fitting landmarks have been typically used to address the problem of tracking and registering deformable facial models. Landmarks are a distinguishable point set to represent an appearance or shape of the facial model. To perform a framework of detecting and localizing facial features, facial models are initially predefined by manually annotating landmarks throughout the entire data set. The predefined facial models are then trained to obtain a canonical reference frame that is typically defined by a mean shape [Asthana13]. Figure 6.10 illustrates an overall framework of the landmark-based alignment approach.

Whereas face detectors in [Asthana13], [Asthana14], [Xiong12], [Yu13] widely used for color or gray images yield high accuracy of the performance of face detection, they yielded

unconvincing results for HFIs because HI introduces high variations of intensity at different wavelengths. This observation implies that those landmark-based alignment approaches are heavily sensitive to HFIs with high intensity variations. However, we heuristically found that tree structure part model (TSPM) [Zhu12] is capable of accurately detecting faces over the entire HFI sets in our database. Thus, we utilize TSPM to incorporate other landmark-based AAs in the initialization of the facial models.

For experimental verification of fitting landmarks to align HFIs, we examine five state-of-the-art landmark-based AAs: 1) TSPM [Zhu12], 2) supervised descent methods (SDM) [Xiong12], 3) cascaded deformable model (CDM) [Yu13], 4) discriminative response map fitting (DRMF) [Asthana13], and 5) incremental parallel cascade model (IPCM) [Asthana14]. The initial position of each facial models of the studied landmark-based AAs is determined by the estimated boundary boxes involved in the results of TSPM on our database.

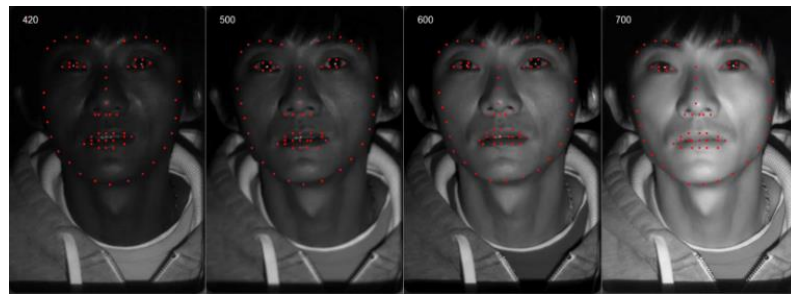


Figure 6.11: Sample results of TSPM [Zhu12] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.

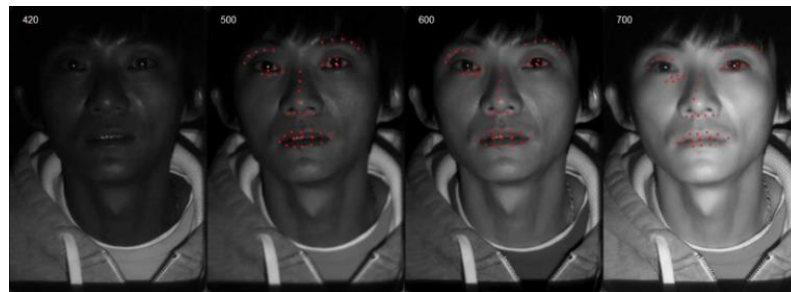


Figure 6.12: Sample results of SDM [Xiong12] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.



Figure 6.13: Sample results of CDM [Yu13] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.



Figure 6.14: Sample results of DRMF [Asthana13] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.



Figure 6.15: Sample results of IPCM [Asthana14] at 420 nm, 500 nm, 600 nm, and 700 nm tested on ID: F048_01 in IRIS-HFD-2014-V1.

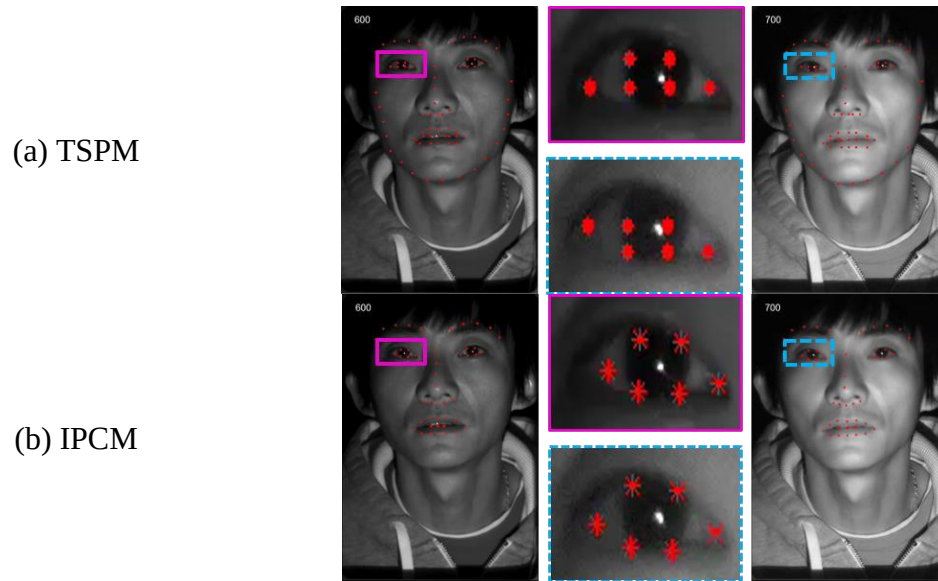


Figure 6.16: Examples of inaccurate localizations of landmark points: (a) the results of TSPM [Zhu12] and (b) the results of IPCM [Asthana14] at 600 nm and 700 nm. As illustrated in (a) and (b).

Landmark-based AAs, which are able to offer fully automatic processes, have already achieved remarkable performance over color or gray images captured by traditional cameras in uncontrolled environments. Nevertheless, as shown in Fig. 6.11 to Fig. 6.16, the landmark-based AAs often failed either to update correct facial model parameters related to controlling the shape of the facial model particularly at short wavelengths or to localize consistent landmark points over the HFI sets, even though we properly assigned the initial locations to each facial model of the conducted landmark-based AAs as mentioned above.

According to our observations we suspect two main reasons: 1) lack of the photon energy at short wavelengths near blue ranges in the visible spectrum deteriorate the performance of landmark-based AAs; and 2) the landmark models trained with dissimilar facial databases cannot be directly applied to HFIs without training them. Indeed, for annotating the facial models, it is generally necessary to manually and steadily select more than 30 landmarks for each image. This becomes impractical when applied to a large HFI set.

6.4 Image alignment approaches (LK and SIFTFlow)

An alternative way to directly handle HFI alignment can be adopting general image alignment approaches: Lukas-Kanade (LK) [Baker04] and SIFTFlow [Liu11]. The image alignment techniques are frequently used for image stitching and stereo matching [Liu11]. The challenging problems in image alignment are typically solved using gradient decent. Baker and Matthew [Baker04] presented comprehensive review of LK approaches to align a reference image I_r to an input image I_t by computing optical flow. LK methods based on Gauss-Newton optimization aim to minimize the sum of squared error between template and target images which can be formulated as [Baker04]:

$$\sum_p \left(I_t(W(p; \tau)) - I_r(p) \right)^2 \quad (3.10)$$

where p denotes pixel coordinates of an image, $W(p; \tau)$ indicates the parameterized set of the warps. LK methods iteratively update the parameter $\tau \leftarrow \tau + \Delta\tau$ until τ converges. Figure 6.17 shows the overview of LK approach. We note that to estimate the optimal transformation τ in RASL and ORIA approaches was directly inspired by the primary consideration of LK algorithms associated with the warp update rule.

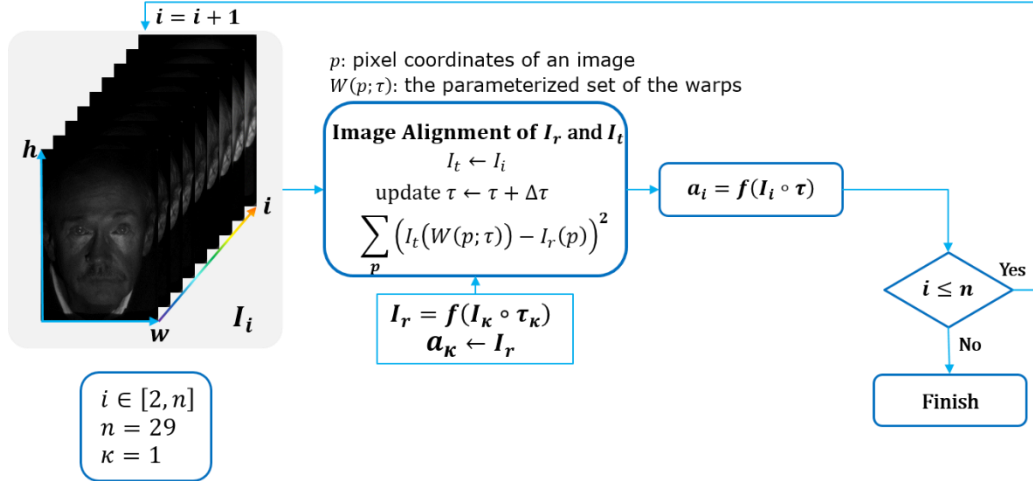


Figure 6.17: Overview of LK approach.

For SIFTFlow [Liu11], Liu et al. utilized scale-invariant feature transform (SIFT) to characterize local gradient information on reference and target images. The overview of SIFTFlow approach is illustrated in Fig. 6.18. Liu et al. proposed an objective function containing a data term, small displacement term, and smoothness term for spatial

regularization. The data term is to match SIFT descriptors along the flow vectors from two images. The small displacement term includes the flow vectors. The smoothness term is used to deal with matching outlier and flow discontinuities. We refer the interested reader to [Liu11] and the references therein for more detail.

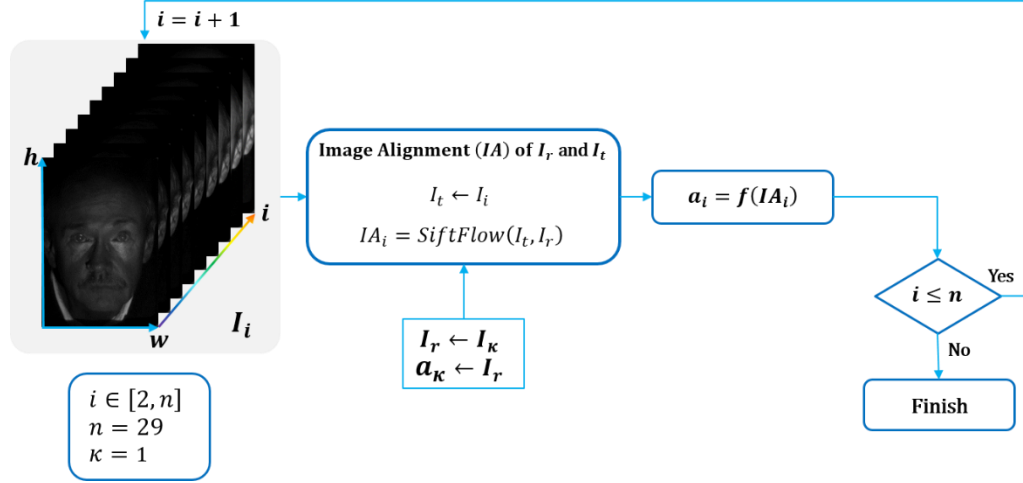


Figure 6.18: Overview of SIFTFlow approach.

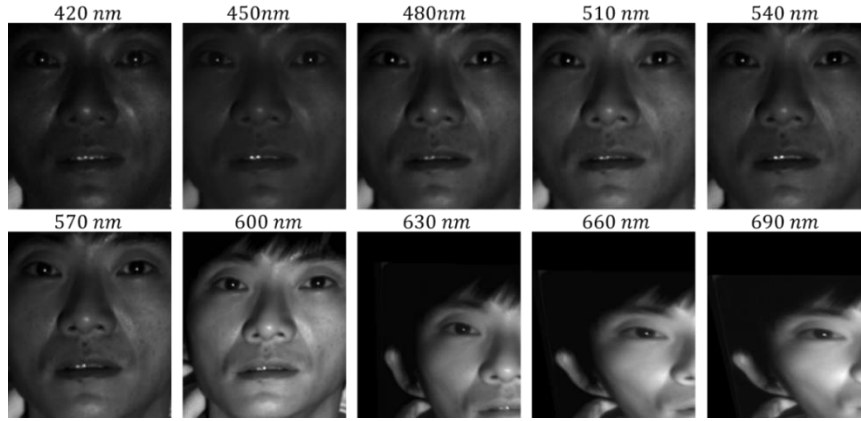


Figure 6.19: An example of a failure of LK approach on ID: F048_01 in IRIS-HFD-2014-V1.

For experimental determination of image alignment approaches, we investigate the accuracy of HFI alignment on our database. For the reference image, we select the first sub-band of the input HFI set as mentioned in the previous sections. We iteratively update the target images from 430 nm to 700 nm for increments to 10 nm. Figure 6.19 and 6.20

show the results of LK and SIFTFlow, respectively. In most of our HFI sets, both image alignment approaches are unsuccessful to handle alignment problems in HFI including high variations of intensity at different wavelengths.

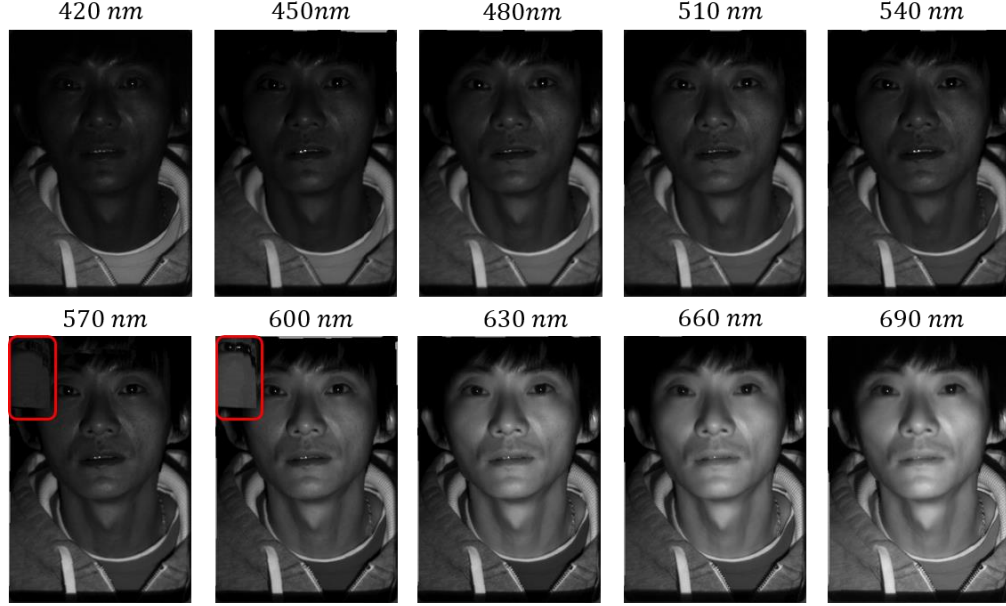


Figure 6.20: An example of a failure of SIFTFlow approaches on ID: F048_01 in IRIS-HFD-2014-V1.

Therefore, for hyperspectral face alignment approaches, we utilize two manual approaches and two ICO approaches corresponding to FBB, EC, RASL, and ORIA, which can deal with IBMs in HFI throughout this dissertation. In Fig. 6.21, we show the aligned HFIs and the corresponding sRGB images rendered under CIE D65 where ROIs marked in rectangles are magnified in the last column. To visually compare the aligned HFIs, we connected the left eye at 420 nm to the right eyes at 700 nm. As illustrated with the red line in Fig. 6.21 (a), the subject tended to move toward the bottom right-hand corner during data acquisition. Thus, this HFI set cannot be sufficiently aligned by FBB approach due to the presence of the subject's movement whereas EC, RASL, and ORIA explicitly achieved well-aligned results, see Fig. 6.21(b)-(d). Nevertheless, on the enlarged ROIs in the last column of Fig. 6.21, we can observe that there are still distorted colors near the right jaw particularly in the result of EC. Though vital to produce visually acceptable alignment results from RASL and ORIA, such AAs are nevertheless imperfect. It is thus strategic to predict the better-quality alignment using an objective score over a large HFI set as will be mentioned in Chapter 7.

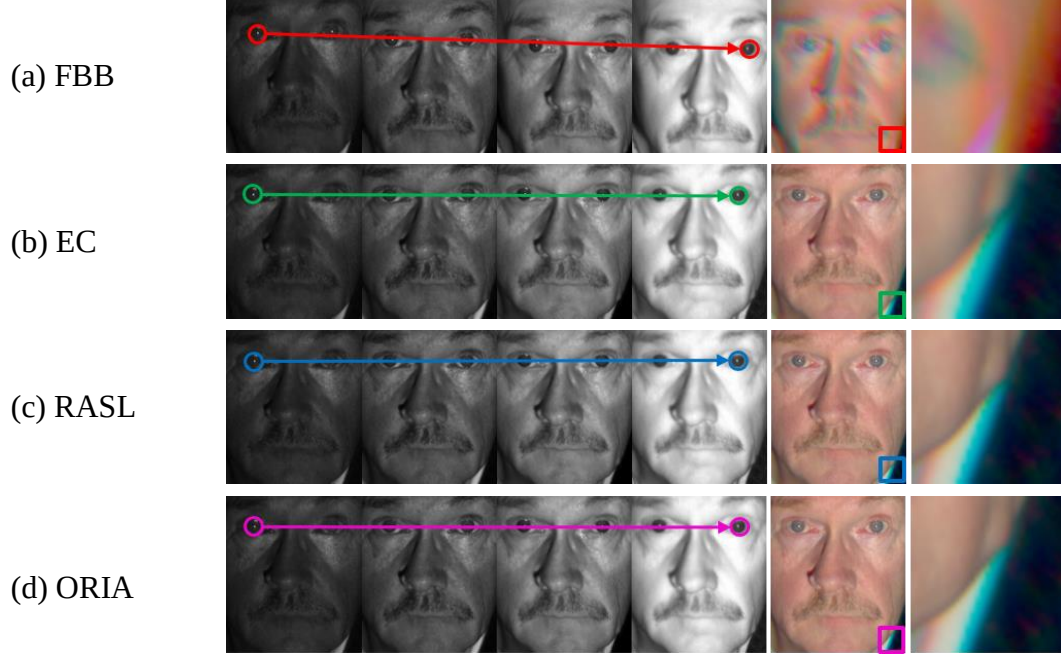


Figure 6.21: Examples of the aligned HFIs at 420 nm, 500 nm, 600 nm, and 700 nm according to (a) FBB, (b) EC, (c) RASL, and (d) ORIA. The color images in the fifth column are displayed using sRGB rendered with full sub-bands under CIE D65 where the regions of interest (ROIs) marked with rectangles are magnified and displayed in the last column.

6.5 Summary

In this Chapter, to address IBMs involved in HFI, we employed four different techniques: 1) manual approaches based on selecting regions of interest, such as fixed bounding box-based AA (FBB) [Szeliski10] and eye coordinate-based AA (EC) [Denes02], [Di10], [Uzair13]; 2) iterative convex optimization (ICO) processes for face alignment, such as RASL [Peng12] and ORIA [Wu12]; 3) landmark-based AAs, such as DRMF [Asthana13], IPCM [Asthana14], SDM [Xiong12], CDM [Yu13], TSPM [Zhu12]; and, 4) image AAs, such as Lucas-Kanade (LK) [Baker04] and SIFTFlow [Liu11]. For experimental determination of HFI alignment, we examined all of the studied techniques on our database. However, landmark-based AAs and image AAs were unsuccessful to align our HFIs in the experiments. Therefore, for hyperspectral face AAs, we employ two manual AAs and two ICO-based AAs corresponding to FBB, EC, RASL, and ORIA to address IBMs in HFIs throughout this dissertation.

7 Proposed Qualitative Assessments of Alignment Methods

In this Chapter, we propose two qualitative prediction models based on: 1) a principal curvature map for evaluating the similarity index between sequential target bands and a reference band in the HFI using average pooling as a full-reference metric; and 2) the cumulative probability of target colors in the HSV color space for evaluating the alignment index of a single sRGB image rendered using the entire bands of the HFI as a no-reference metric. Figure 7.1 illustrates the overview of the proposed framework.

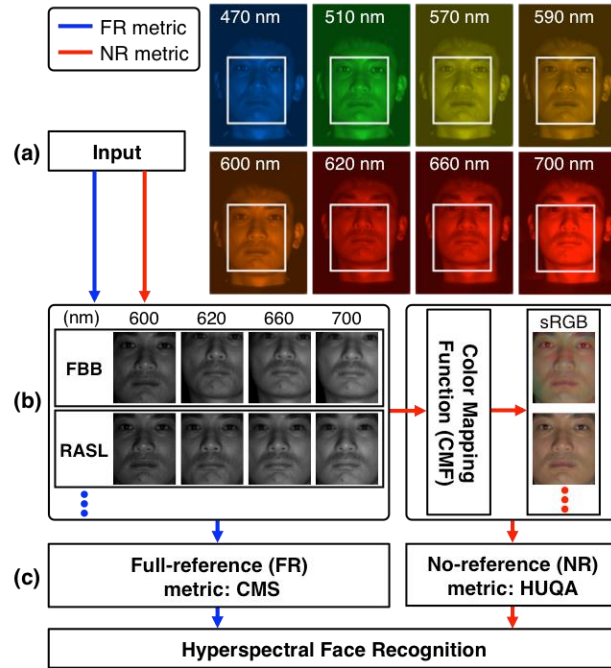


Figure 7.1: Overview of the proposed framework (best viewed in color): (a) input, (b) alignment, and (c) assessing improved alignment via two proposed qualitative prediction models.

7.1 Curvature-based alignment quality assessment

We present the proposed method for a full reference alignment quality assessment (FR-AQA) based on principal curvature map [Deng07], [Stokes98] derived by computing the maximum or minimum eigenvalues of a 2×2 Hessian matrix where the Hessian matrix describes the local curvature of the image. A major difficulty for FA-AQA with HFIs is the existence of high intensity changes at different wavelengths. To handle varying intensity levels depending on the wavelength in HFIs, we adopt the basic framework of recently published gradient map-based image quality assessments (IQAs) [Liu12], [Xue14]. A primary difference in our approach compared to two existing GM-based approaches is that we utilize the principal curvature map as a local quality map in order to evaluate the pixel-wise gradient similarity index between the reference and the target images in each sub-band of the HFI set. The proposed method can provide more consistent and sufficient information for describing facial features as shown in Fig. 7.2.

To obtain the principal curvature map, we first convolve an input aligned image A at the point p with a variable-scale Gaussian $g(p, \sigma)$ as a function $d(p, \sigma) = g(p, \sigma) \otimes A(p)$ where $\otimes$ is the convolution operation at the point p . We then form a 2×2 Hessian matrix as in [Deng07]:

$$HM(p) = \begin{bmatrix} D_{xx}(p) & D_{xy}(p) \\ D_{xy}(p) & D_{yy}(p) \end{bmatrix}, \quad (7.1)$$

where D_{xx} , D_{xy} , and D_{yy} denote the second-order derivatives of the input image at point p .

After building the Hessian matrix $HM(p)$, we compute the maximum and minimum eigenvalues, λ_1 and λ_2 , respectively, from $HM(p)$. Next, we obtain the principal curvature map (PCMP) of the input image as $PCMP = |\lambda_1|$ where $|\cdot|$ refers to the absolute value. Finally, a curvature model (CM) can be given by normalizing $PCMP$ as:

$$CM(p) = \frac{PCMP(p)}{\max(PCMP(p))}. \quad (7.2)$$

Note that to achieve high accuracy of AQA, we only utilize the maximum eigenvalue λ_1 of the Hessian matrix because we heuristically found that the CM built from the minimum eigenvalue λ_2 tends to be unpredictable where inconsistent curvature lines or edges result in reductions in the accuracy of AQA.

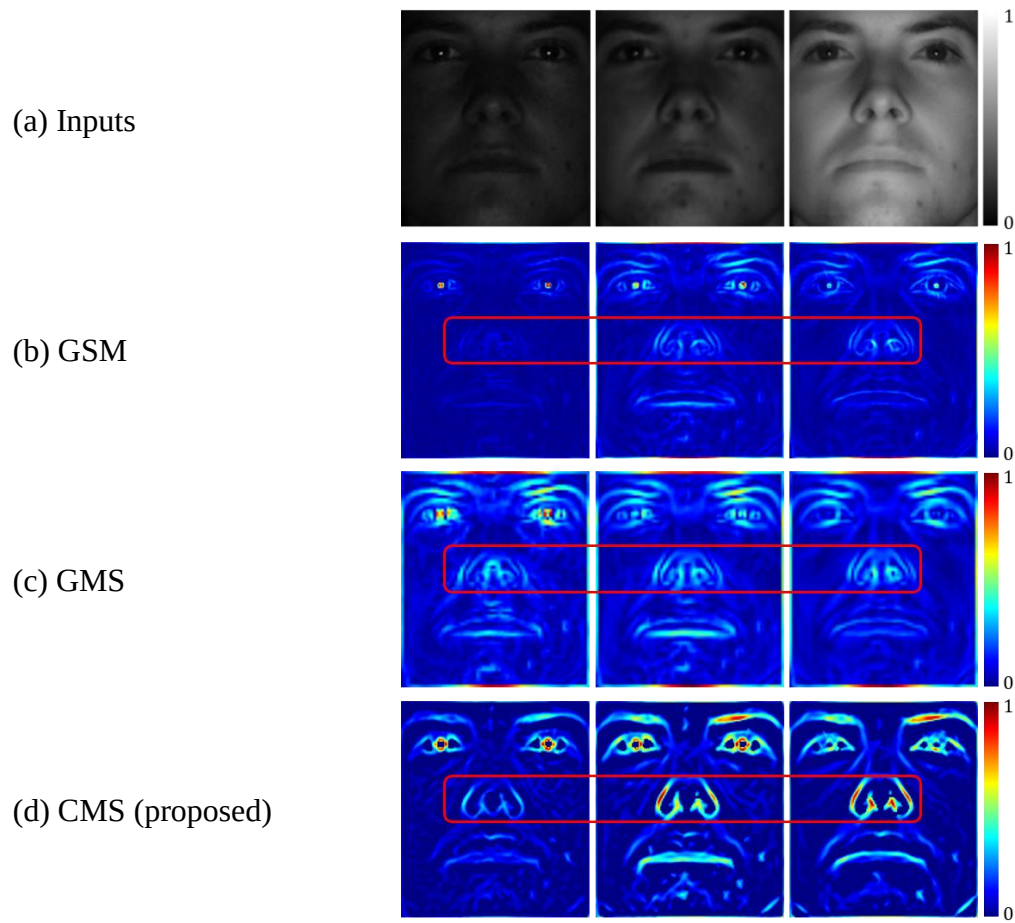


Figure 7.2: A comparison of the results of (b) gradient similarity model (GSM) [Liu12], (c) gradient magnitude similarity (GMS) [Xue14], and (d) the proposed method based on curvature model similarity (called CMS) at 420 nm, 550 nm, and 700 nm on the results of RASL [Peng12] applied to ID: F067_01 in IRIS-HFD-2014-V1 as shown in (a).

Given n aligned grayscale HFIs A_i , $i \in [1, n]$, in an input set, we choose the first band image (i.e. 420 nm) as a reference image $A_{i=1}$ and the remaining bands as target images A_j , $j \in [2, n]$. We first compute CM_r for $A_{i=1}$ once and then iteratively compute CM_j for A_j until $j = n$. The CM similarity (CMS) score is defined as:

$$CMS(j) = \frac{1}{k} \sum_{p=1}^k \frac{2 \cdot CM_r(p) \cdot CM_j(p) + C}{CM_r(p)^2 + CM_j(p)^2 + C}, \quad (7.3)$$

where C denotes a small positive constant that supplies numerical stability by keeping the denominator from being zero. We set C to 10^{-5} in our experiments. m is the total number of pixels in the image.

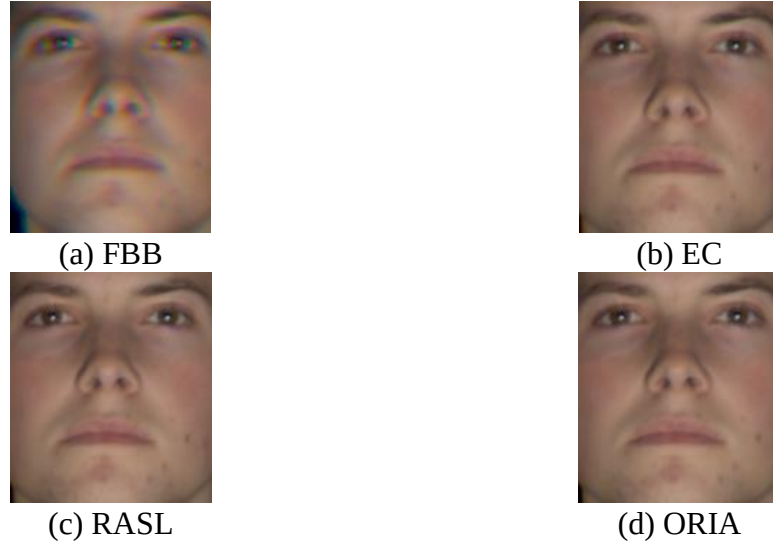


Figure 7.3: A comparison of the results of four alignment approaches with the corresponding sRGB images rendered under CIE D65. The HFIs are taken from ID: F067_01 in IRIS-HFD-2014-V1.

A single overall score for the AQ in all of the bands of the input HFI set using average pooling is computed as:

$$CMSM = \frac{1}{n-1} \sum_j CMS(j). \quad (7.4)$$

In Table 7.1, we show the scores of the proposed CMS over the aligned HFI sets produced by FBB, EC, RASL, and ORIA by comparing two existing approaches: 1) gradient

similarity model (GSM) [Liu12] and 2) gradient magnitude similarity (GMS) [Xue14]. Figure 7.3 shows that the AQ of RASL is visually better than FBB. The scores of CMS in Table 7.1 agree with the visual assessment. However, both GSM and GMS incorrectly indicate that the AQ of FBB is better than RASL as opposed to our observation.

Accordingly, the proposed CMS offers better prediction for the AQ among the four aligned HFI sets. Note that higher scores indicate better AQ on an HFI set. These examples validate that the proposed method can accurately predict the better-quality alignment among four aligned HFI sets. Despite the efficacy of the proposed CMS in predicting AQ, CMS needs to iteratively evaluate the similarity scores over full sub-bands in the input HFI set. Thus, CMS has high computational cost. To improve the computational cost, we propose a no-reference AQA based on the cumulative probability of target colors in the HSV color space [Jayaraman10] for an overall AQA with a single sRGB image.

Table 7.1: A comparison of the reported scores for the sample sets in Fig. 7.2.

Metrics	FBB	EC	RASL	ORIA
GSM	0.6581	0.6870	0.6364	0.6789
GMS	0.6171	0.6418	0.5807	0.6307
CMS (proposed)	0.5742	0.7574	0.7672	0.7693

7.2 Hue-based alignment quality assessment

In this section, we propose a no-reference alignment quality assessment (NR-AQA) based on the cumulative probability of target colors in the HSV color space in Fig. 7.4 for evaluating the alignment quality of a single sRGB image rendered with full sub-bands under CIE D65 [Moan14], instead of evaluating the sharpness or blurriness of a target image [Hassen13], [Liu12], [Mittal12], [Narvekar11], [Vu12].

The proposed method was inspired by the presence of distorted colors caused by inter-band misalignments (IBMs) as illustrated in Fig. 3.12. In general, the distorted colors cannot be seen in natural color facial images. More specifically to analyze the effects of IBMs in HFIs, both of the misaligned and aligned sRGB images produced by the selected alignment approaches such as FBB and RASL are projected into the HSV color space [Jayaraman10]. As shown in Fig. 7.5, the color distribution of the misaligned image is more widely spread over the HSV color space compared to the aligned image, which appears to be concentrated on the red color of the hue component. Notice that in Fig. 7.5(d), the dark colors with low V values are produced by the subject’s clothes where the colors are not presented in the

misaligned image in Fig. 7.5(a). These observations imply that we can predict the alignment quality (AQ) of a single sRGB image by exploring the distribution of the distorted colors caused by IBMs.

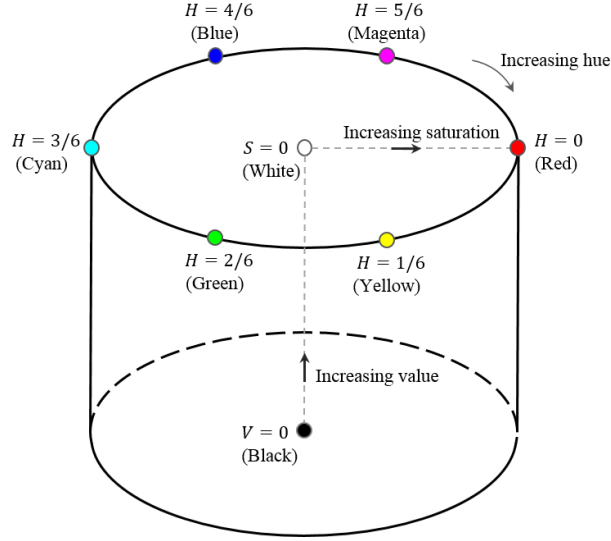


Figure 7.4: HSV color space.

To convert the color space from sRGB [Stokes98] to HSV [Jayaraman10] shown in Fig. 7.6, which is specified by three components: hue (H), saturation (S), value (V), we first find the maximum and minimum component values among R , G , and $B \in [0, 1]$ as:

$$C_{max} := \max(R, G, B), \quad C_{min} := \min(R, G, B), \quad (7.5)$$

where R , G , and B stands for the red, green, and blue components in the sRGB color space, respectively. Next, we compute the saturation S as:

$$S = \begin{cases} \frac{\Delta}{V} & \text{if } V > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (7.6)$$

where the value $V = C_{max}$ and $\Delta := C_{max} - C_{min}$. Note that if $S = 0$, then H is undefined. The preliminary hue, $\tilde{H} \in [-1, 5]$ is then computed as:

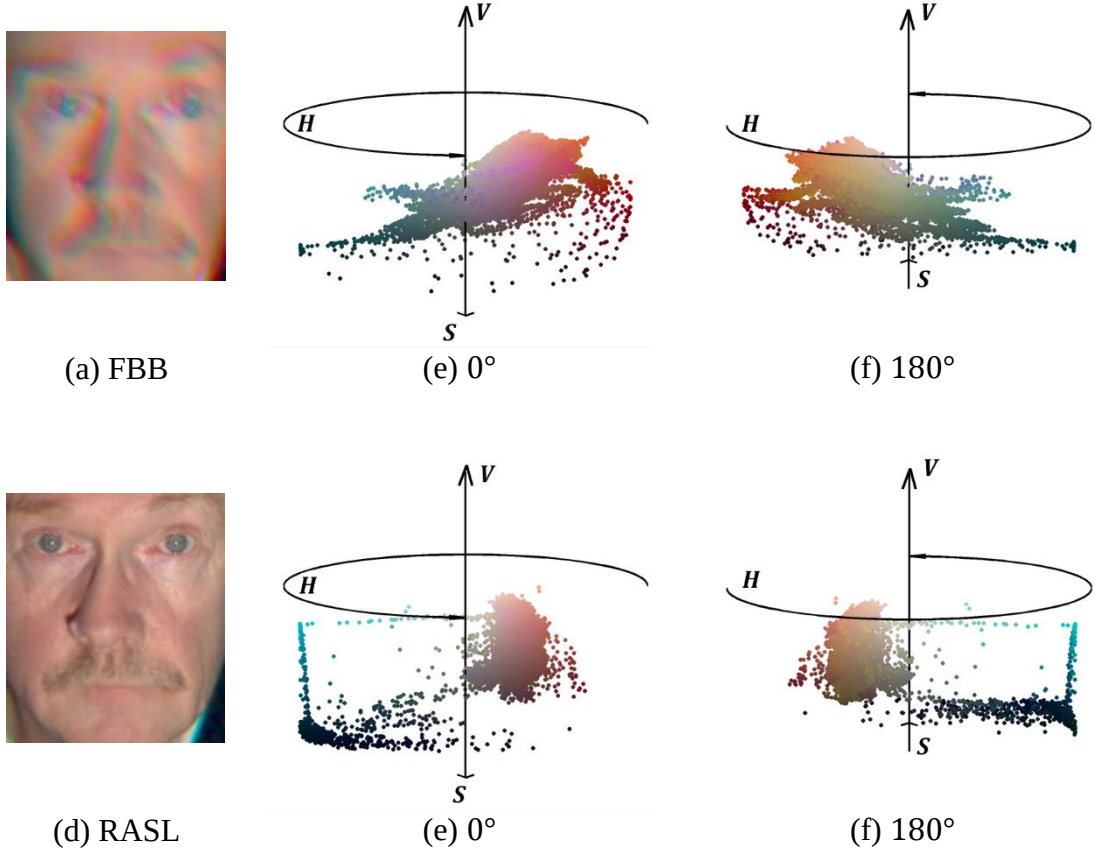


Figure 7.5: Analyses of the effects of IBMs in the HSV color space. For the misaligned sRGB image in (a), we can observe that the distribution of the colors is more extensively spread over the HSV color space as shown in (b) and (c) in different views where the vertical axis is the V value, the horizontal distance from the axis is the S value, and the angle is the H value. However, the color distribution of the aligned sRGB image in (d) is concentrated near the red color of the hue component as illustrated in (e) and (f). Note that the dark colors with low V values in (d) are induced by the subject's clothes where the colors do not appear in (a).

$$\tilde{H} = \begin{cases} (G - B)/\Delta & \text{if } R = C_{max} \\ (B - R)/\Delta + 2 & \text{if } G = C_{max} \\ (R - G)/\Delta + 4 & \text{if } B = C_{max}. \end{cases} \quad (7.7)$$

Next, the hue $H \in [0,1]$ can be given by normalizing $\tilde{H}$ as:

$$H = \frac{1}{6} \cdot \begin{cases} (\tilde{H} + 6) & \text{if } \tilde{H} < 0 \\ \tilde{H} & \text{otherwise.} \end{cases} \quad (7.8)$$

Finally, we remove undesired hue values with excessively low and high saturation by assigning β values as:

$$H'(p) = \begin{cases} H(p) & \text{if } \alpha \leq S(p) < 1 - \alpha \\ \beta & \text{otherwise.} \end{cases} \quad (7.9)$$

where $\alpha = 0.05$ in our experiments and p denotes pixel coordinates on an image. We set $\beta = -1$, which is a value out of range in H . In Fig. 7.7, we illustrate the hues on a color wheel that is divided into six sectors related to three primary colors (red, green, and blue) and three mixed color (cyan, magenta, and yellow).

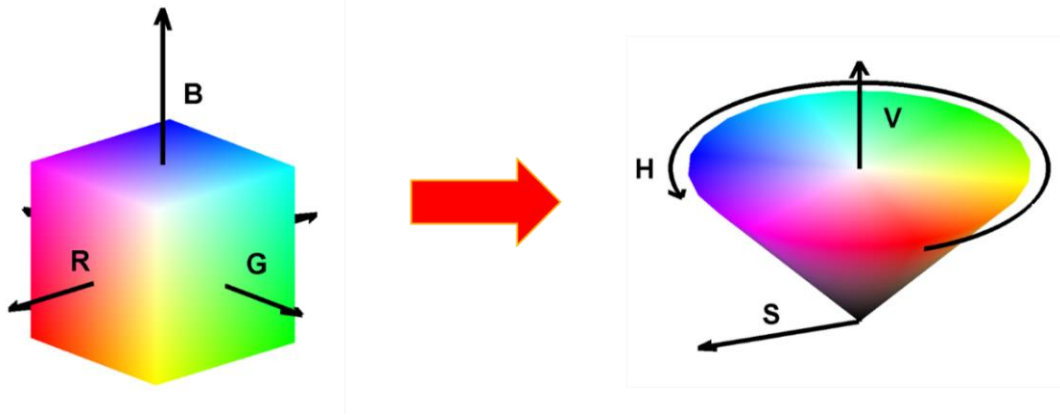


Figure 7.6: Converting the color space from sRGB to HSV.

To assess the AQ of an sRGB image, we compute the probability of the distorted colors, $\phi \in [2, 6]$ of the hue component as:

$$P_{\phi}(x) = \Pr \left(x \mid \frac{2(\phi - 1) - 1}{12} < x \leq \frac{2\phi - 1}{12} \right), \quad (7.10)$$

where we define a random variable x as $0 \leq x \leq 1$, $x \in H'$ in the interval $[-1, 1]$. Note that the probability of the red color of the hue component can be given as:

$$P_{\phi=1}(x) = \Pr \left(x \mid \frac{11}{12} < x \cup x \leq \frac{1}{12} \right). \quad (7.11)$$

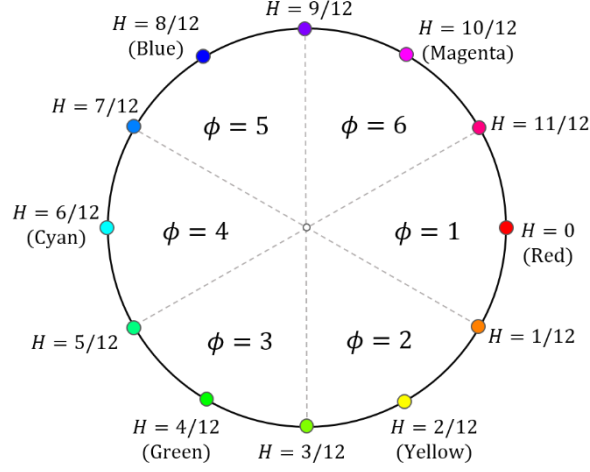


Figure 7.7: An illustration of H on a color wheel divided into six sectors associated with three primary and three pairwise mixed colors, ϕ , including red, yellow, green, cyan, blue, and magenta.

In Fig. 7.8, we show the difference of the probability of six representative colors such as red, yellow, green, cyan, blue, and magenta in the hue component between the misaligned and aligned sRGB images. As mentioned in the observations involved in the color distribution of the sRGB images, the probabilities of distorted colors on the misaligned image in Fig. 7.8(a) are observably higher than the case of the aligned image in Fig. 7.8(d). These results imply that the proposed method can be used for assessing an overall single AQ to determine the better-aligned image among four aligned image sets formed by FBB, EC, RASL, and ORIA.

7.3 Evaluation of alignment quality assessment

We demonstrate the efficacy of the proposed metrics on two cases shown in Fig. 7.9: 1) rigid objects such as mannequins in Fig. 7.9(b) and 2) non-rigid subjects in the frontal HFIs taken from the first session in our database, consisting of 3,770 HFIs collected from 86 males and 44 females (130 individuals) of diverse ethnic backgrounds and appearance. The performance of the proposed methods is evaluated by prediction accuracy. Notice that the developed metrics are called curvature model similarity (CMS) for a FR-AQA and hue-based alignment quality assessment (HUQA) for a NR-AQA.

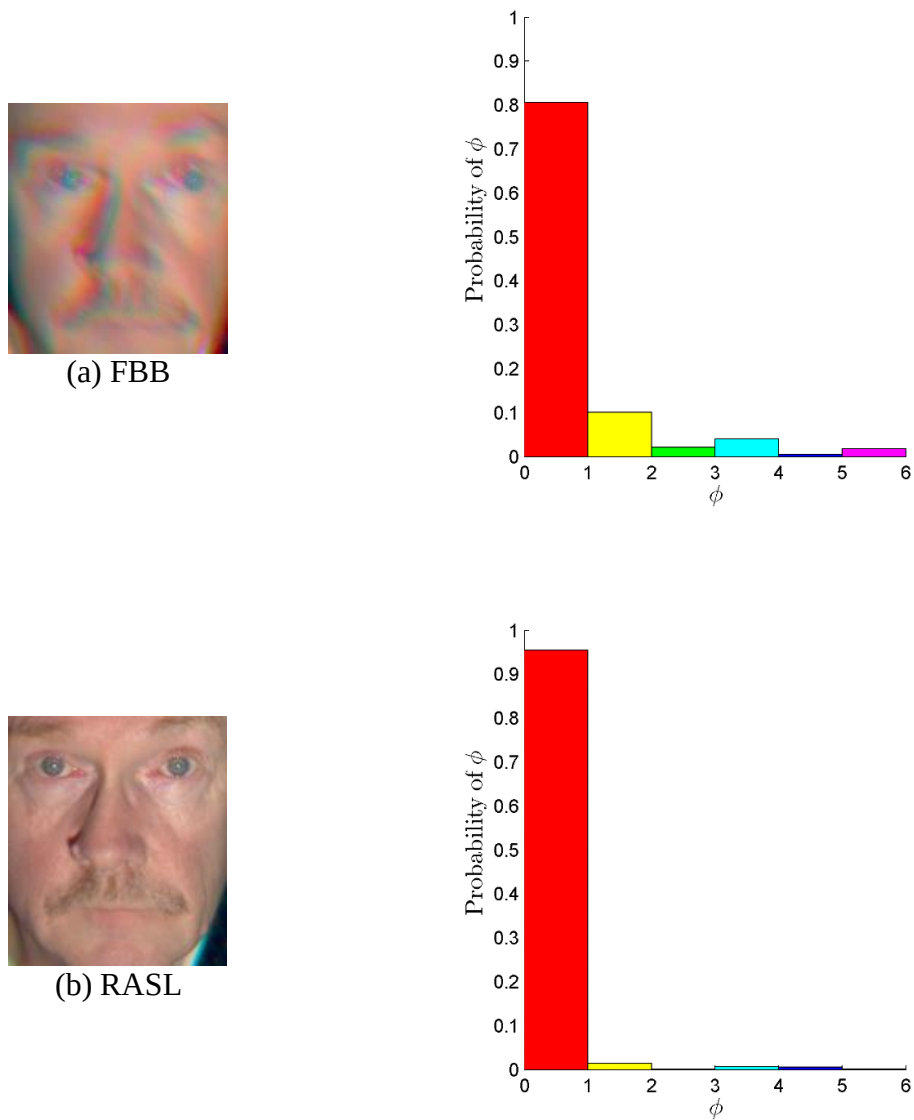


Figure 7.8: A comparison of the probability of the six representative colors (red, yellow, green, cyan, blue, and magenta) between (a) the misaligned and (b) the aligned sRGB images.

We first explain the choice of a parameter for CMS in the following subsection. We then verify the correction of CMS metric on the rigid object sets shown in Fig. 7.9(b). Next, we further test the CMS approach on the non-rigid subject sets that include more challenging problems associated with IBMs in IRIS-HFD-2014-V1. The performance of CMS is compared with two existing gradient map-based image quality assessments (IQAs) [Liu12], [Xue14]. For HUQA compared with five state-of-the-art NR-IQA metrics [Hassen13], [Liu14], [Mittal12], [Narvekar11], [Vu12] typically used to evaluate the sharpness or blurriness of an input grayscale image, we repeatedly follow the same tasks as the verification of the CMS metric. All of our experimental results are obtained by using the original software provided by the competing approaches. In our experiment results, we note that the higher predicted scores represent better AQ throughout this Chapter and the lower slope coefficients (SCs) of a linear regression line indicate the higher prediction consistency where the predicted scores are sorted in ascending order.



(a) Original images of rigid objects

(b) Canonical images of rigid objects

Figure 7.9: Examples of input HFI sets for rigid objects in (b). Note that these sRGB images are rendered with full sub-bands under CIE D65.

7.3.1 Choice of σ for curvature model similarity

We only have one parameter σ to define the Gaussian scale in CMS metric. For experimental determination of the optimal σ value on IRIS-HFD-2014-V1, we examine the effects of varying σ values from 1.0 to 2.0 on a task for predicting the improved alignment among four selected alignment approaches. As shown in Fig. 7.10, the highest accuracy of CMS is obtained when $\sigma \geq 1.4$. There is no impact on further accuracy even when considering substantial differences in scale. Thus, for efficiency, we define the Gaussian scale for CMS metric as $\sigma = 1.4$, for all other experiments throughout this dissertation. Note that the errors are computed by counting the number of the subject IDs where the scores of FBB are higher than other scores of EC, RASL, or ORIA. In Fig. 7.11, we show the effects of varying σ values on the results of RASL.

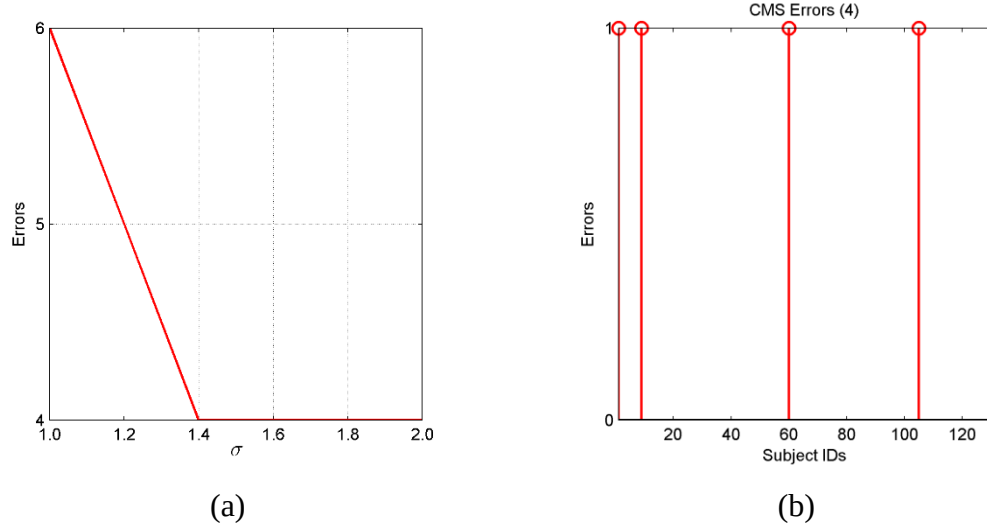


Figure 7.10: (a) shows the number of errors on IRIS-HFD-2014-V1 as increasing σ values from 1.0 to 2.0. The σ denotes the Gaussian scale and is a parameter for the proposed CMS metric. (b) indicates the subject IDs associated with the errors where a value of ‘1’ represents the error. In (a), we can observe that CMS appears not to enhance the accuracy of the prediction to determine the most improved alignment by which the σ value is greater than 1.4. Note that whereas CMS obtain four errors in this experiment where we compare the scores of FBB with EC, RASL, and ORIA, there is no error if we associate the errors with only RASL and ORIA.

7.3.2 Results of CMS on rigid object sets

On rigid object sets, we investigate the correction of CMC compared with two recent FR-IQA approaches based on the gradient map which consist of gradient similarity model (GSM) [Liu12] and gradient magnitude similarity (GMS) [Xue14]. GSM uses four 5×5 kernels predefined with weighting coefficients to compute gradient values according to two versions: 1) a block-wise version was used for image blocks x and y ; and 2) a pixel-wise version was used for the central pixels of image blocks x and y . GMS utilizes Prewitt filters along horizontal and vertical directions to obtain gradient map. Note that to properly examine the performance of GSM, GMS, and the proposed CMS, we define the parameter C to prevent the denominator from being zero in the function used for measuring gradient similarity in both GSM and GMS, such as $C = 10^{-5}$. We refer the interested reader to their papers and the references therein for more detail. For rigid objects there is no difference between the aligned HFI sets produced by FBB and EC because an eye coordinate set was selected on the first band (i.e. 420 nm) for all sub-bands. Hence, we examine only three alignment approaches: FBB, RASL, and ORIA. Furthermore, since there is explicitly no movement of object on rigid sets, the alignment quality of FBB can be better than the others.

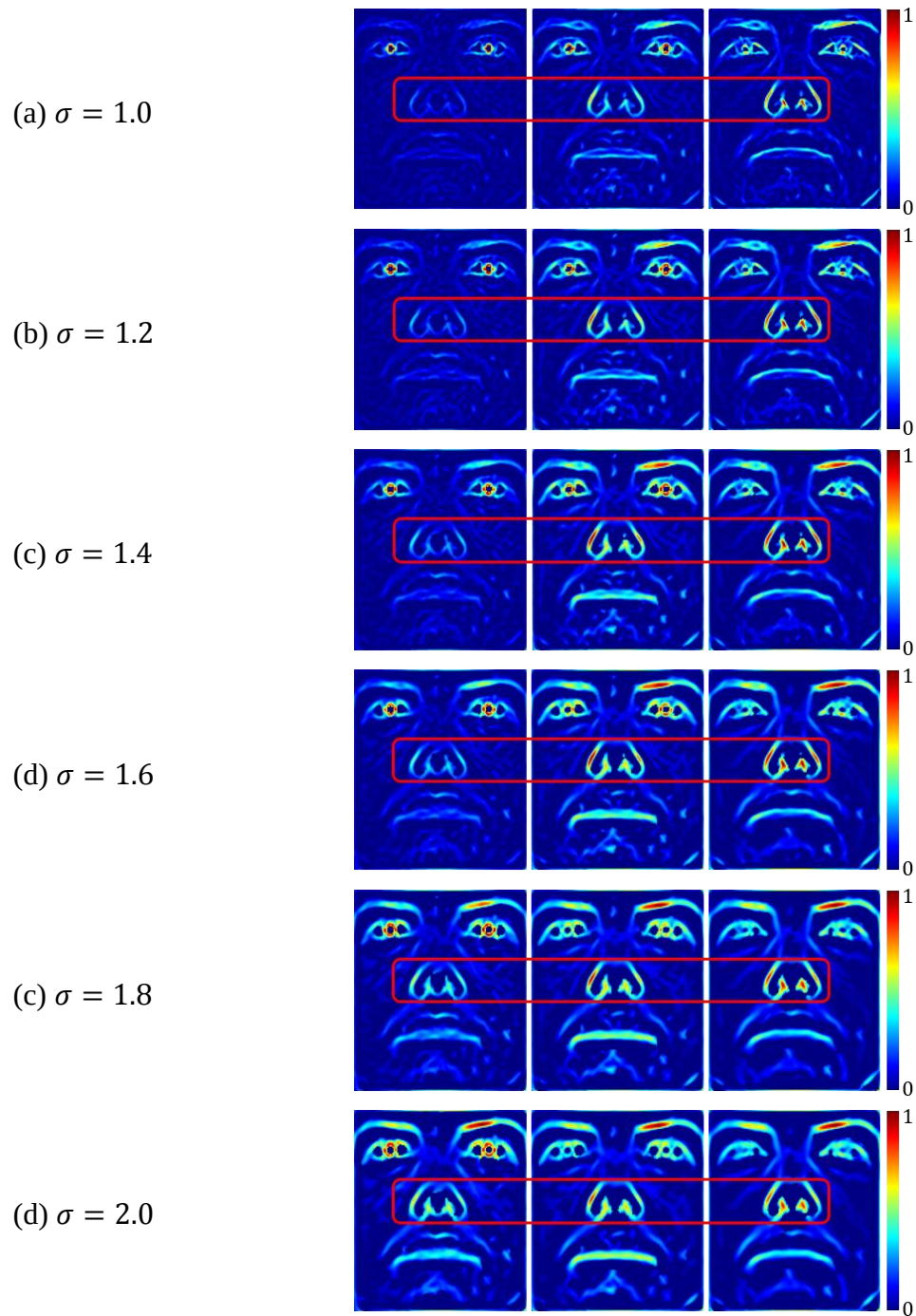


Figure 7.11: Effects of varying σ values at 420 nm, 550 nm, and 700 nm on the results of RASL.

Table 7.2: Results of two FR-IQA approaches and the proposed CMS on two rigid object sets.

Input Sets	Metrics	FBB	RASL	ORIA
Mannequin 1	GSM	0.8563	0.8337	0.8614
	GMS	0.8663	0.8375	0.8729
	CMS (proposed)	0.8311	0.8298	0.8296
Mannequin 2	GSM	0.8194	0.8085	0.8183
	GMS	0.8097	0.8007	0.8095
	CMS (proposed)	0.8049	0.8002	0.8042

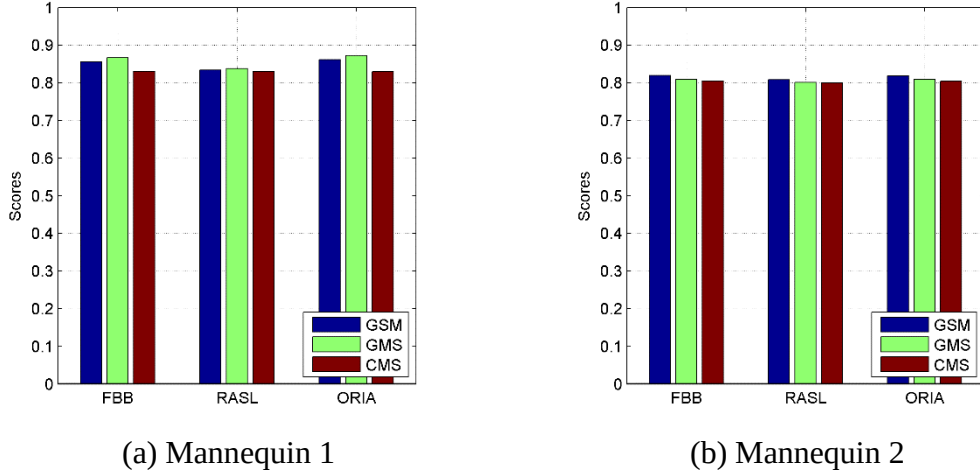


Figure 7.12: Results of two FR-IQAs and the proposed CMS on rigid object sets.

In Fig. 7.12 and Table 7.2, we show the results of two FR-IQA approaches and CMS on rigid object sets. Compared to GSM and GMS, the proposed CMS metric consistently predicts all of the scores for evaluating the AQ of the reference and target images in each sub-band of the rigid HFI sets created by three alignment approaches. In addition, although all CMS scores are relatively close in range, CMS correctly indicates the improved alignment such as FBB in both rigid object sets.

7.3.3 Results of CMS on IRIS-HFD-2014-V1

We further demonstrate the efficacy of the proposed CMC metric on the non-rigid subject sets that include more challenging problems associated with IBMs. In this task we examine

the prediction consistency and accuracy involved in determining the better-quality alignment among four selected alignment approaches: FBB, EC, RASL, and ORIA.

FBB is employed as the ground truth to evaluate the performance of the considered assessment metrics for AQ. For example, if subject movement occurs during data acquisition, the AQ of FBB is typically worse than the other AAs. However, although FBB is not able to compete with RASL and ORIA in most cases, if there is insignificant subject movement during data acquisition, FBB yields better alignment results than EC as EC requires continual and repetitive manual data input of the eye coordinates. Therefore, under these experimental observations we additionally examine the scores of FBB with RASL and ORIA to compute the considered errors.

To validate the effectiveness of the proposed CMS metric with respect to prediction accuracy, we compare the number of the errors on IRIS-HFD-2014-V1 where we examine the scores of FBB with RASL and ORIA in Fig. 7.13. For instance, GSM incorrectly predicts the AQ for 9 HFI sets in Fig. 7.13(a). 15 errors were found in GSM metric in Fig. 7.13(b). However, compared to GSM and GMS metrics, the proposed CMS more precisely predicts the alignment quality without any error over 130 HFI sets in our database in Fig. 7.13(c). We compute the prediction accuracy associated with the errors found in GSM, GMS, and CMS as:

$$Accuracy = 1 - \varepsilon/N \quad (7.12)$$

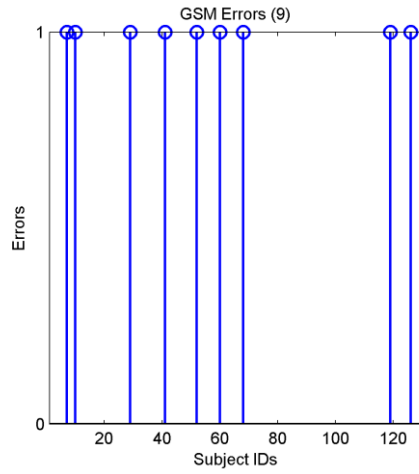
where ε denotes the number of errors and N represents the number of the HFI sets such as $N = 130$ in IRIS-HFD-2014-V1. We report the prediction accuracies of two IQAs and the proposed CMS on Table 7.3 where the prediction accuracies are computed by (7.12) in two cases: 1) comparing FBB with EC, RASL, and ORIA; and, 2) comparing FBB with RASL and ORIA.

Table 7.3: A comparison of prediction accuracies of two FR-IQA approaches and the proposed CMS on IRIS-HFD-2014-V1.

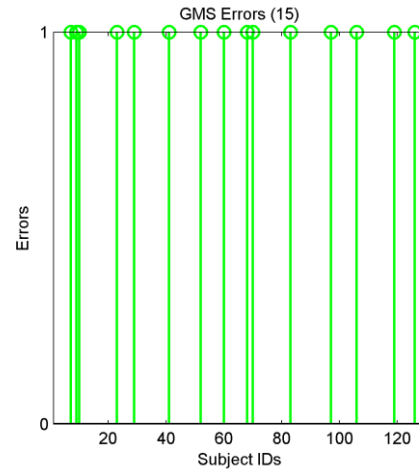
Cases	GSM	GMS	CMS
Case 1	0.93	0.88	0.97
Case 2	0.93	0.88	1.00

Table 7.4: Parameters of the estimated regression lines for FBB, EC, RASL, and ORIA via the proposed CMS.

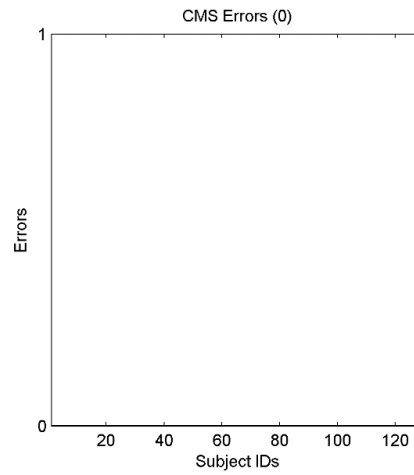
Parameters	FBB	EC	RASL	ORIA
Intercept	0.45	0.68	0.71	0.70
Slope Coefficient	0.0020	0.0010	0.0009	0.0009



(a) GSM



(b) GMS



(c) The proposed CMS

Figure 7.13: Errors of two FA-IQAs and the proposed CMS on IRIS-HFD-2014-V1. Note that the errors are computed by counting the number of the subject IDs where the scores of FBB are higher than either the scores of RASL or ORIA. As can be shown in (c), there is no error in the CMS metric compared to GSM and GMS where the number of errors of (a) GSM and (b) GMS is 9 and 15, respectively.

As shown in Table 7.3, the proposed CMS achieves highly accurate prediction in determining improved alignment in both cases. Next, we examine the performance of the selected alignment approaches using the proposed CMS metric.

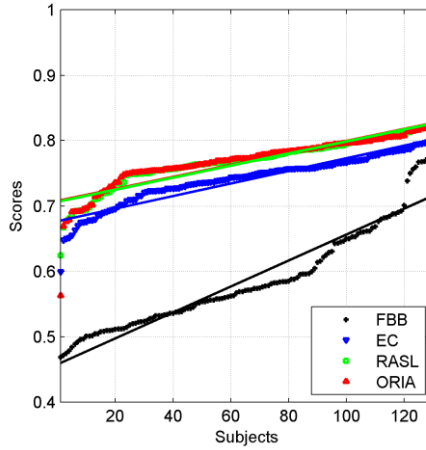


Figure 7.14: Scatter plots with linear regression lines corresponding to results of the proposed CMS on IRIS-HFD-2014-V1.

In Fig. 7.14, we depict all of the predicted scores of the proposed CMS using scatter plots and regression lines according to four selected alignment approaches where the predicted scores are sorted in ascending order. To investigate trends in the predicted scores computed by the proposed CMS metric, we utilize the least-squares method [Press92] to estimate a linear regression. As shown in Table 7.4, the intercept of the regression line according to RASL is higher than other alignment approaches, implying that the alignment quality of RASL is better than the others. Furthermore, the estimated slope coefficients of RASL and ORIA are lower than that of FBB and EC, indicating that RASL and ORIA consistently yield well-aligned results because the trend lines have lower correlation between subjects and the predicted scores. Therefore, these experimental results demonstrate that RASL achieves highly accurate alignment with high consistency as compared to FBB, EC, and ORIA.

7.3.4 Results of HUQA on rigid object sets

For the experimental verification of the correction of the proposed HUQA, we examine the prediction accuracy of HUQA on rigid object sets. To evaluate the alignment quality of a

single sRGB image, we first render the entire sub-bands in the HFI set using CIE 1931 2° Standard Observer data under CIE D65 [Moan14].

To validate the performance of HUQA, we employ five state-of-the-art NR-IQA metrics that are typically used to evaluate the sharpness or blurriness of an input image. Five competing approaches are composed of spatial-spectral entropy-based quality index (SSEQ) [Liu14], local phase coherence-based sharpness index (LPC-SI) [Hassen13], blind/referenceless image spatial quality evaluator (BRISQUE) [Mittal12], spectral and spatial sharpness (S3) [Vu12], cumulative probability of blur detection (CPBD) [Narvekar11]. We refer to their papers for further detail.

Table 7.5: Results of five NR-IQA approaches and the proposed HUQA on two rigid object sets.

Input Sets	Metrics	FBB	RASL	ORIA
Mannequin 1	SSEQ	0.7740	0.8109	0.7441
	LPC-SI	0.8807	0.8832	0.8960
	BRISQUE	0.6771	0.6729	0.6849
	S3	0.2703	0.3290	0.2611
	CPBD	0.1211	0.1062	0.1161
	HUQU (proposed)	0.9911	0.9883	0.9909
Mannequin 2	SSEQ	0.7300	0.7483	0.7291
	LPC-SI	0.8848	0.8691	0.8866
	BRISQUE	0.5978	0.5771	0.5941
	S3	0.8208	0.8756	0.8004
	CPBD	0.1000	0.0996	0.0938
	HUQU (proposed)	0.9599	0.9582	0.9591

In Fig. 7.15 and Table 7.4, we show the experimental results of five NR-IQA metrics and the proposed HUQA on the rigid object sets. As aforementioned in Chapter 7.3.2, we investigate the determination accuracy of predicting the enhanced alignment among FBB, RASL, and ORIA.

In this task, HUQA and CPDB produce higher prediction accuracy of selecting the better-quality alignment on both input sets as shown in Fig. 7.15 and Table. 7.5. Although BRISQUE correctly identifies that the AQ of FBB is better than the other AAs for the second input set, it fails in predicting the first rigid object set. SSEQ, LPC-SI, and S3 are unsuccessful in these experiments as shown in Table 7.5. Note that to fix a range of y-axis involved in the scores of the alignment quality on figures, the scores of both SSEQ and CPBD are rescaled by adding 0.3 and 0.5, respectively, for all of the experiments throughout this dissertation.

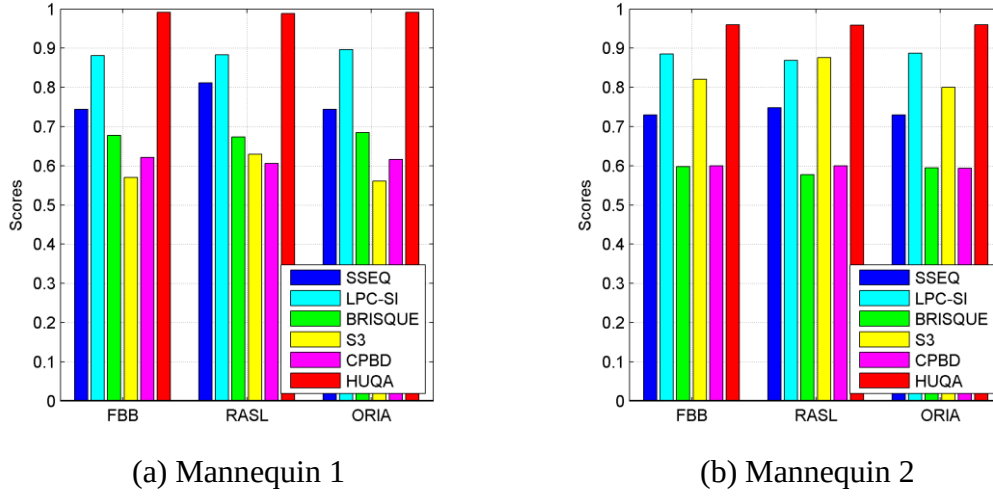


Figure 7.15: Results of five state-of-the-art NR-IQAs and the proposed HUQA on rigid object sets. Note that the scores of SSEQ and CPBD are rescaled by adding 0.3 and 0.5, respectively.

7.3.5 Results of HUQA on IRIS-HFD-2014-V1

For a comprehensive study on the effectiveness of the proposed HUQA, we further examine the prediction consistency and accuracy of the proposed HUQA on our database including more challenges associated with IBMs caused by subject's movement during data acquisition. We notice that HUQA only requires a single sRGB image as an input, instead of the entire sub-bands in the HFI set. Furthermore, whereas FR-AQA metrics assess the AQ of the input HFI set using average pooling schemes by iteratively measuring the similarity between the reference and target HFIs in each sub-band, there is neither an iterative process nor a pooling scheme in the proposed HUQA metric.

We validate the prediction accuracy of five competing IQA approaches and the proposed HUQA in terms of the estimation of the errors where the scores of FBB are compared with RASL and ORIA. HUQA achieves the highest prediction accuracy with only one error caused by ID: F009_01 shown in Fig. 7.16 where the predicted scores corresponding to FBB, EC, RASL, and ORIA are 0.9802, 0.9788, 0.9792, and 0.9793, respectively as illustrated in Fig. 7.17. Even though HUQA yields one error, the scores are quite close together. Table 7.6 shows the prediction accuracy computed by (7.12) in two cases: 1) comparing FBB with EC, RASL, and ORIA and 2) comparing FBB with RASL and ORIA, additionally. In case 1, the number of errors of SSEQ, LPC-SI, BRISQUE, S3, and CPBD are 30, 14, 34, 55, and 26, respectively. In case 2, those of SSEQ, LPC-SI, BRISQUE, S3, and CPBD are 9, 13, 15, 41, and 7, respectively. Compared to other metrics, HUQA in case

1 achieves only 4 errors. Therefore, HUQA yields the highest accuracy of predicting the better-quality alignment in both cases.

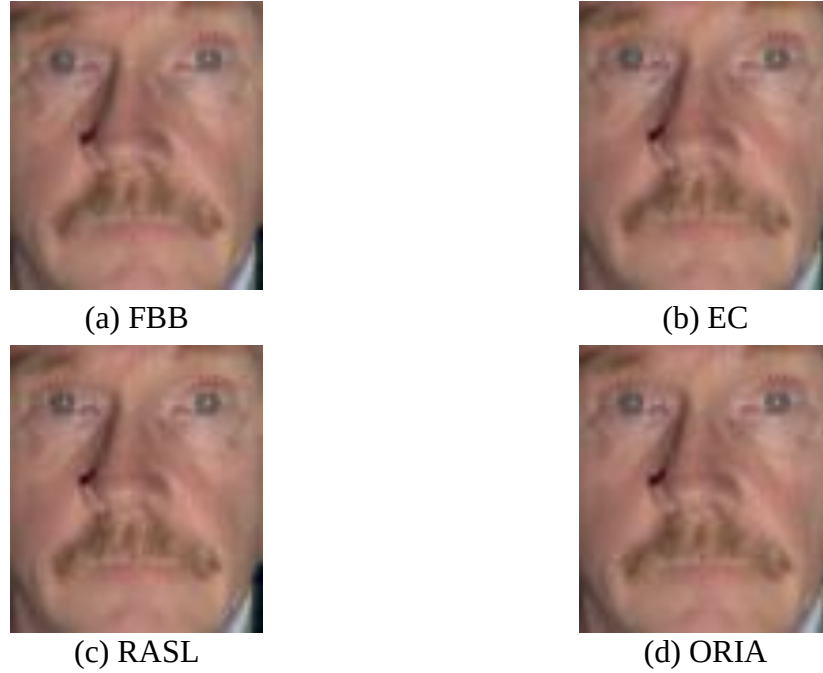
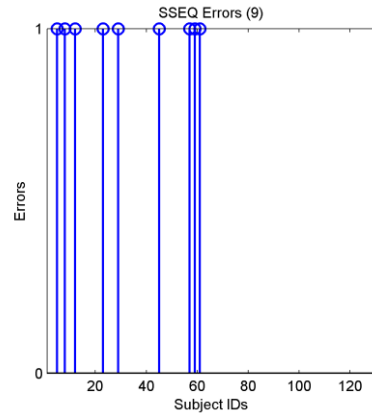


Figure 7.16: One error HFI set found in HUQA. The HFI set is taken from ID: F_009_01 of IRIS-HFD-2014-V1.

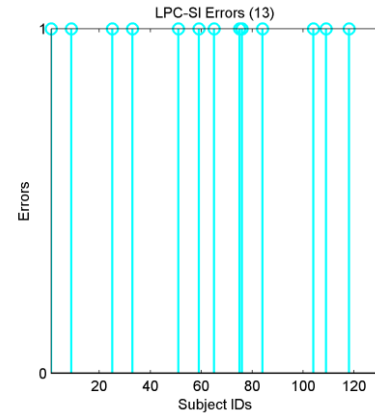
Table 7.6 A comparison of prediction accuracies of five NR-IQA approaches and the proposed HUQA on IRIS-HFD-2014-V1.

Metrics	SSEQ	LPC-SI	BRISQUE	S3	CPBD	HUQA
Case 1	0.77	0.89	0.74	0.58	0.80	0.97
Case 2	0.93	0.90	0.88	0.68	0.95	0.99

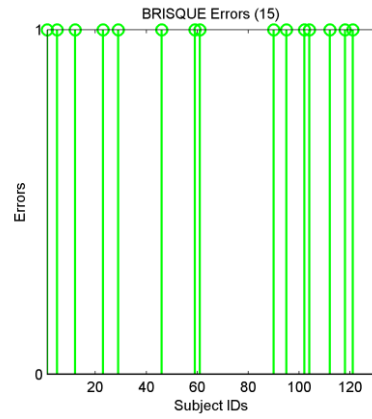
In Fig. 7.18, the experimental results show the predicted scores in the proposed HUQA corresponding to the conducted alignment approaches. The intercepts and slope coefficients of the regression lines according to FBB, EC, RASL, and ORIA are shown in Table 7.7. The proposed HUQA also shows that RASL yields improved alignment with higher consistency compared to the competing alignment approaches as illustrated in Table VIII. This determination can also be visually verified by interpreting the slopes and intercepts of the regression lines of the conducted alignment approaches in Fig. 7.18.



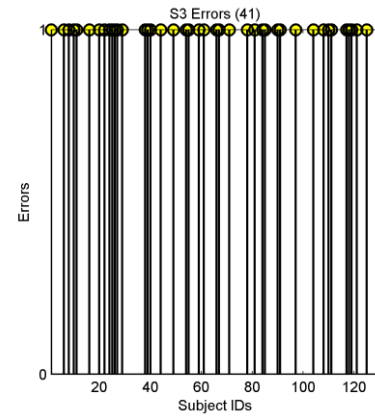
(a) SSEQ



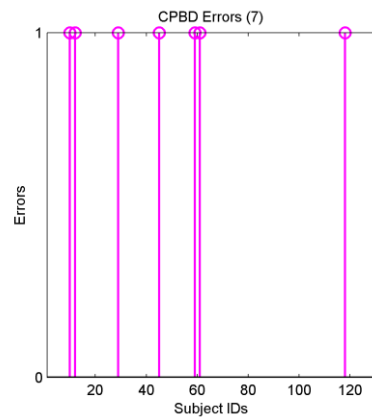
(b) LPC-SI



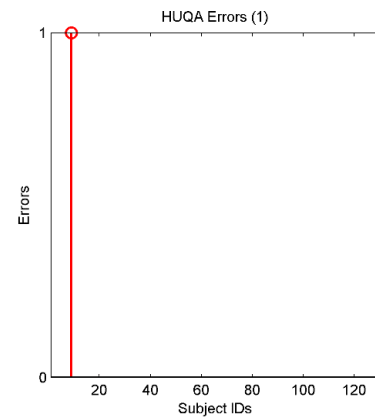
(c) BRISQUE



(d) S3



(e) CPBD



(f) HUQA (Proposed)

Figure 7.17: Errors of five NR-IQA approaches and the proposed HUQA.

Table 7.7: Parameters of the estimated regression lines for FBB, EC, RASL, and ORIA via the proposed HUQA.

Parameters	FBB	EC	RASL	ORIA
Intercept	0.81	0.95	0.96	0.95
Slope Coefficient	0.0015	0.0004	0.0003	0.0004

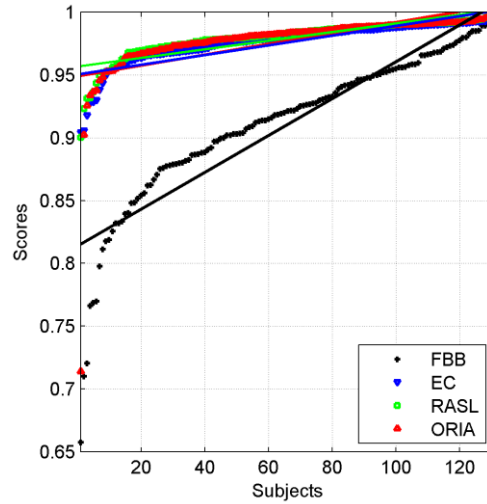


Figure 7.18: Scatter plots with the regression lines corresponding to the results of the proposed HUQA on non-rigid subject sets in IRIS-HFD-2014-V1.

7.4 Summary

In this Chapter, we proposed a novel framework to determine the better-quality alignment among four selected alignment approaches, such as fixed bounding box-based, eye coordinate-based, RASL, and ORIA, in order to account for inter-band misalignments in hyperspectral face images. As a result, the proposed framework is able to deal with reducing inter-band misalignment artifacts in hyperspectral face images.

To evaluate the alignment quality of the input hyperspectral face image sets, we developed two different metrics. First, the curvature model similarity for a full-reference alignment quality assessment was used to investigate the alignment quality by iteratively evaluating the similarity index between the reference and target images in each band of an input hyperspectral face image set. Second, the hue-based alignment quality assessment for a no-

reference alignment quality assessment was used to evaluate the alignment quality of a single sRGB image that was generated using all of the bands of an input hyperspectral face image set.

Comparisons with seven state-of-the-art image quality assessments on our new hyperspectral face database (called IRIS-HFD-2014-V1) showed that both the proposed metrics led to promising accuracy in determining the most improved alignment among four selected alignment approaches. Therefore, the proposed metrics can be used to verify the alignment quality of both current and future face alignment algorithms tested on hyperspectral face image sets, which particularly contain inter-band misalignments caused by subject motion during data acquisition.

8 Proposed Accurate Alignment Approach (A₃) Based on a Mixture Model

Existing methods based on manual inputs of eye coordinates in each band have tackled how to align hyperspectral face images, which contain significant intensity variation over the visible spectrum. However, such a method relied on eye coordinates becomes impractical as the total number of image sets grows. In addition, eye coordinate-based alignment approaches are particularly limited for partial faces in profile views or structural features (e.g., glasses) where one of the two eyes is partially occluded. In this chapter, we propose an accurate alignment approach (A₃) for hyperspectral face images that only requires one manual input of the two eyes to obtain a template image. The proposed method, A₃, is developed by incorporating the strengths of point correspondence used to match representative features between image pairs and a low-rank model used to seek an optimal set of inter-band domain transformations.

8.1 Accurate alignment approach using a mixture model

As mentioned in [Liao13], an edge-based feature detector is superior to extract many more keypoints on facial images than the scale invariant feature transform (SIFT), which is one of the most popular keypoint detector. This is because there are more edges than blobs on a face. Therefore, we develop a new edge-based feature detector based on the Laplacian of Gaussian (LoG) operator, which calculates the second spatial derivative of an image with pixel intensity values, $I(x, y)$. The LoG function centered on zero and with Gaussian standard deviation σ (also called scale factor) is given by:

$$LoG(x, y, \sigma) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}}. \quad (8.1)$$

The LoG filter as $\sigma = 2.0$ is illustrated in Fig. 9.1. The LoG filter can be typically used to produce extrema at corners and blobs [Miao13]. Compared to existing feature detectors, we employ the LoG filter to detect feature edges on images. For an input image $I(x, y)$, the output of the LoG filter is obtained by:

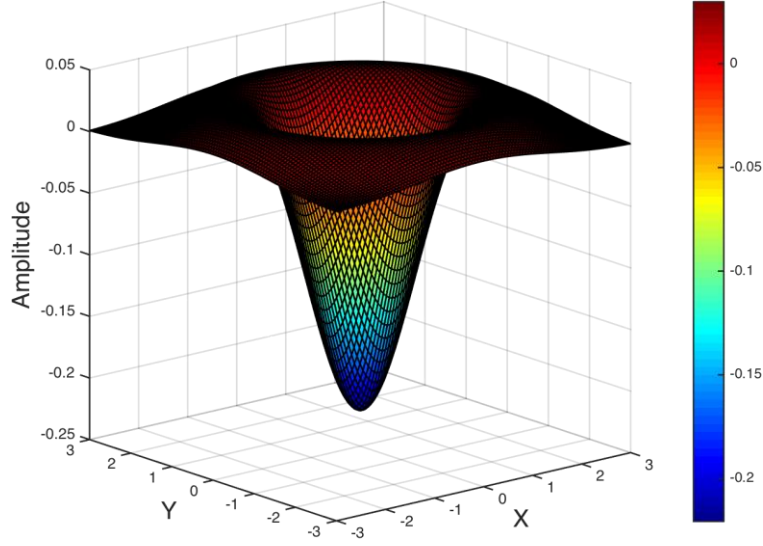


Figure 8.1: The shape of the LoG filter as $\sigma = 2.0$.

$$I_{LoG}(x, y, \sigma) = LoG(x, y, \sigma) * I(x, y), \quad (8.2)$$

where $*$ is the convolution operation in x and y . Figure 8.2 shows the response of a 1-D LoG filter to a step edge function.

In order to extract the interest points, *IP*s, we find zero-crossings, $y_{LoG}(x, y, \sigma)$, in $I_{LoG}(x, y, \sigma)$, which are stronger than a sensitivity threshold, t_{LoG} :

$$IP(x, y, \sigma) = y_{LoG}(x, y, \sigma) > t_{LoG}. \quad (8.3)$$

where $y_{LoG}(x, y, \sigma)$ occurs in which both x and y gradients change signs. In our experiments, we empirically set $t_{LoG} = 0.002$ and $\sigma = 1.2$.

Once we detect the interest points, we represent each image with a set of feature descriptors. In this dissertation, we employ a speed-up robust feature descriptor (SURF) [Bay08]. The point correspondences can be directly established by comparing the local feature descriptors, such as SURF, which can provide robustness in scale changes and rotation and produce descriptors half the size of previous descriptors such as SIFT [Lowe04] while maintaining the same matching performance.

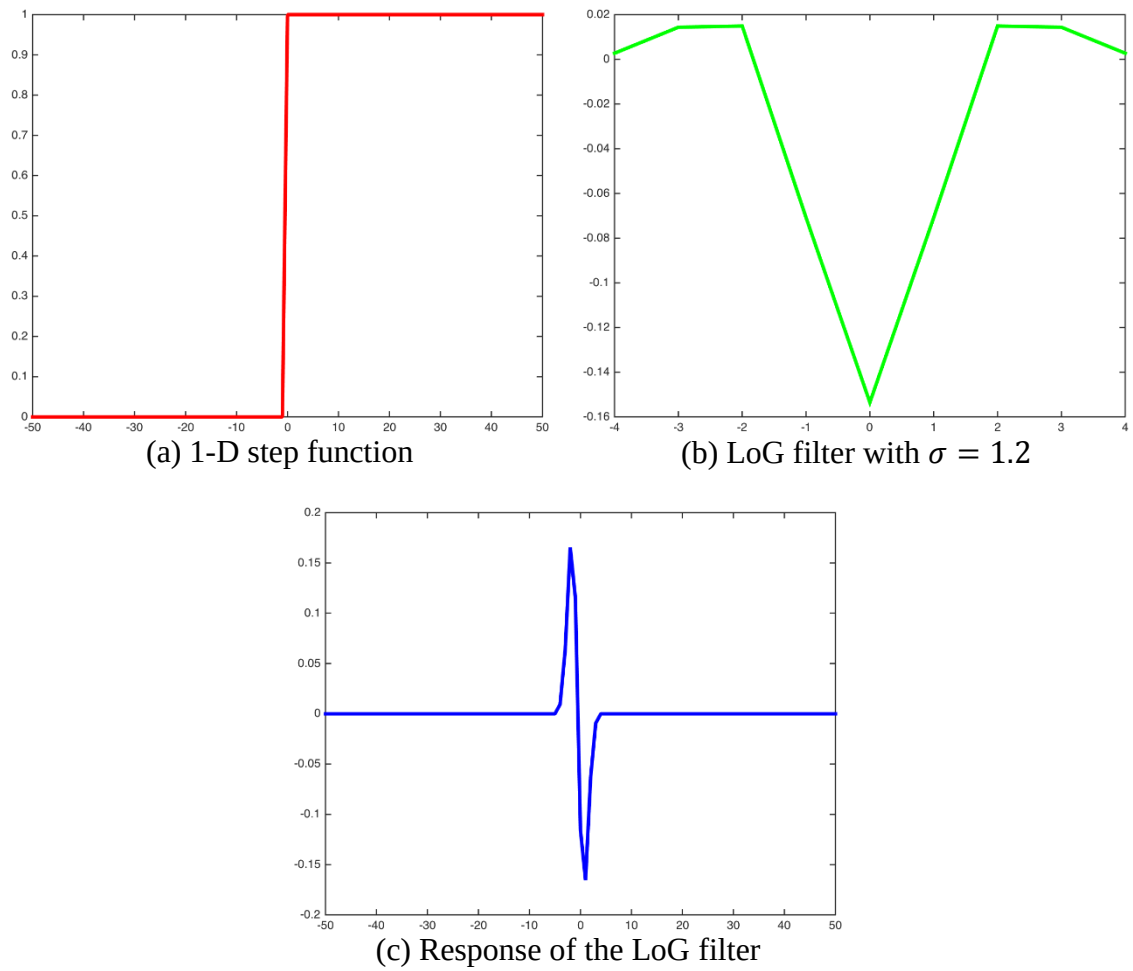


Figure 8.2: Response of a 1-D LoG filter to a step edge function.

In order to achieve invariance to rotation, we first assign a dominant orientation to the interest point by calculating the sum of all the Haar wavelet responses in x and y directions within a sliding orientation window of a size $\frac{\pi}{3}$. Figure 8.3 shows the Haar wavelet filters in each of x and y directions.

Next, a square region is split up regularly into smaller 4×4 square sub-regions. For each sub-region, we compute a few simple features at 5×5 regularly spaced sample points. The horizontal and vertical Haar wavelet responses dx and dy are calculated and summed up over each sub-region and form a first set of entries to the feature vector. The absolute values of the responses $|dx|$ and $|dy|$ are also calculated, and together with the sum of vector to form a four-dimensional descriptor. Therefore, for all 4×4 sub-regions, it results in a vector of the length $4 \times 4 \times 4 = 64$. Figure 8.4 illustrates how to build the SURF descriptor.



Figure 8.3: Haar wavelet filters.

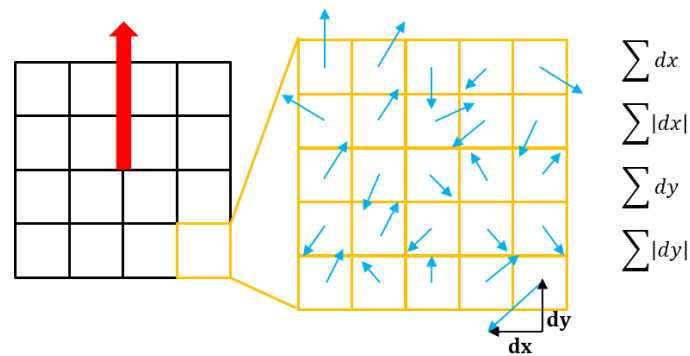


Figure 8.4: Example of SUFT descriptor with the oriented 4×4 square sub-regions. For each sub-region, we compute the wavelet responses from 5×5 regularly spaced sample points.

Once we have extracted features and their descriptors from template and input images, we find corresponding features across two images as shown in Fig. 8.5. In our experiments, we utilize a sum of squared differences (SSDs) to match the set of the descriptors. The similarity scores between two feature descriptors is computed:

$$SSD(IP_1, IP_2) = \sum_{p=1}^{N-1} (IP_1(p) - IP_2(p))^2. \quad (8.4)$$

where N denotes the total number of interest points. After we find the corresponding features across two images, we estimate a similarity transformation model, τ_{i-1} as defined in Chapter 6.1, by using a maximum likelihood estimation sample consensus (MLESAC) algorithm [Torr00]. The basic scenario of estimating τ via the MLESAC is shown in Fig. 8.6. MLESAC is referred to as a generalization of the RANSAC estimator that robustly estimates multiple view relations from point correspondences. Compared to RANSAC, MLESAC estimates the solution that maximizes the likelihood rather than just the number of inliers.

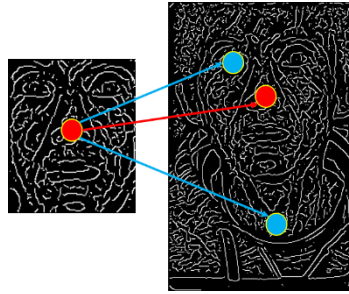


Figure 8.5: Example of feature matching across two images such as the template and input images.

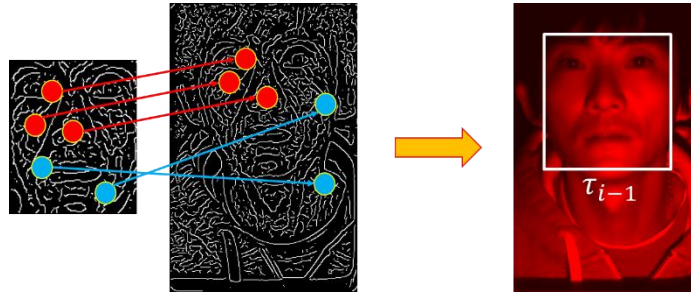


Figure 8.6: Basic scenario of estimating a similarity transformation model, τ , from the matched features.

To optimize the estimated similarity transformation model τ_{i-1} , we employ ORIA as explained in Chapter 6.2. The process pipeline of the proposed alignment approach for HFI sets is shown in Fig. 8.7.

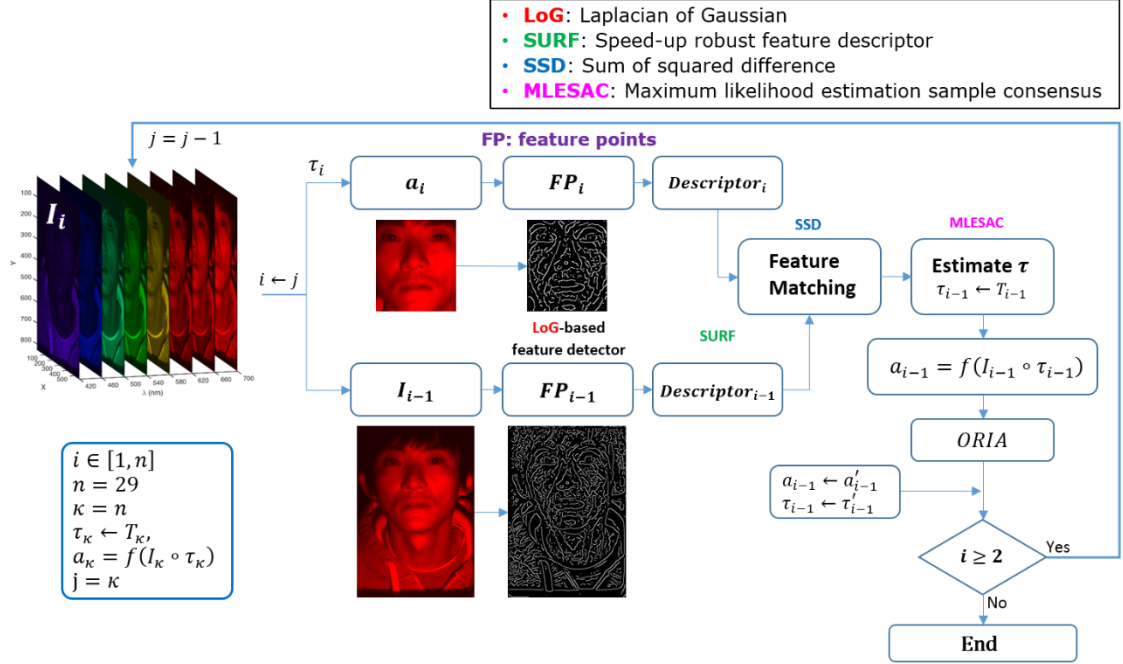


Figure 8.7: Process pipeline of the proposed alignment approach.

8.2 Experimental results

We demonstrate the effectiveness of the proposed alignment approach (A_3) to two cases: 1) rigid objects such as mannequins (Fig. 8.8) and 2) non-rigid subjects such as UWA-HSFD and IRIS-HFD-2014. We compare A_3 with FBB, LoG, and ORIA approaches where the LoG approach is performed without the optimization process.

8.2.1 Results on rigid object sets

For rigid object, we verify the correctness of the alignment algorithms because the ground truth alignment is known. We synthetically perturb each of the input images by Euclidean transformation $eucl[\theta, tr]$ defined as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} tr_x \\ tr_y \end{bmatrix}. \quad (8.5)$$

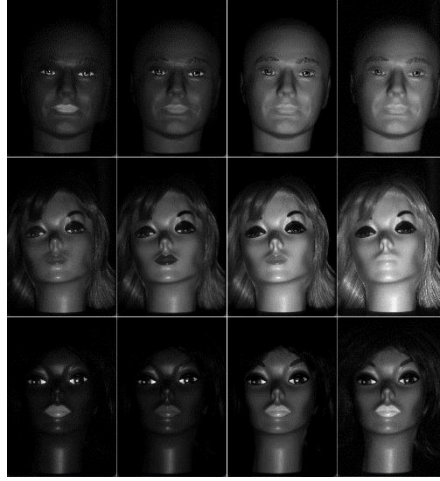


Figure 8.8: Sample HFI sequences of rigid objects taken from 420nm, 500nm, 600nm, and 700nm.



Figure 8.9: Sample artificial HFI images under controlled conditions in each row from top to bottom: translations, rotations, and translations and rotations, respectively.

For example of x and y -translations, given N input images, $I_i, i \in [1, N]$, I_i is translated by $euc[0, tr(i)]$ for $tr(i) = (j - i) \times \Delta, j = N, \Delta \in \{1, 2, 4\}$. For θ rotations, I_i is rotated by

$auc[\theta(i), 0]$ for $\theta(i) = (j - i) \times \phi, \phi \in \{1, 2\}$. For a mixture of varying levels of translations and rotations, each image I_i is artificially perturbed by $auc[\theta(i), tr(i)]$. Figure 8.9 shows sample artificial images under controlled conditions in order to examine the ability of the proposed method to deal with varying levels of misalignment.

We evaluate the alignment errors as:

$$AQ = \frac{1}{N} \sum_{i=1}^N \|G(i) - A(i)\|_2, \quad (8.6)$$

where G represent the ground truth and A denotes the aligned image. We show the results of the alignment errors of LoG, ORIA, and A_3 on artificially translated rigid sets (Figs. 8.10 and Table 8.1), rotated rigid sets (Figs. 8.11 and Table 8.2), and translated and rotated rigid sets (Figs. 8.12 and Table 8.3) in the frontal view. Although occlusions in HFIs nullify the benefits of hyperspectral imaging associated with improving face identification performance, we examine the ability of the proposed A_3 to deal with varying levels of misalignments. As shown in Figs. 8.12, we found that the low-rank model in ORIA is more sensitive to large-scale occlusions involved in Mannequins 1 and 2 compared to the proposed A_3 . Therefore, A_3 , overall, achieves slightly higher accuracies of alignment performance than that of ORIA as shown in Tables 8.1 to 8.3.

Table 8.1: Results of averaging alignment errors of the conducted alignment approaches on translated rigid sets.

Input Sets	LoG	ORIA	A_3
Mannequin 1	2.95	1.35	1.31
Mannequin 2	2.35	1.51	1.51
Mannequin 3	2.42	1.40	1.40

Table 8.2: Results of averaging alignment errors of the conducted alignment approaches on rotated rigid sets.

Input Sets	LoG	ORIA	A_3
Mannequin 1	3.49	1.37	1.37
Mannequin 2	2.77	1.67	1.66
Mannequin 3	2.31	1.52	1.51

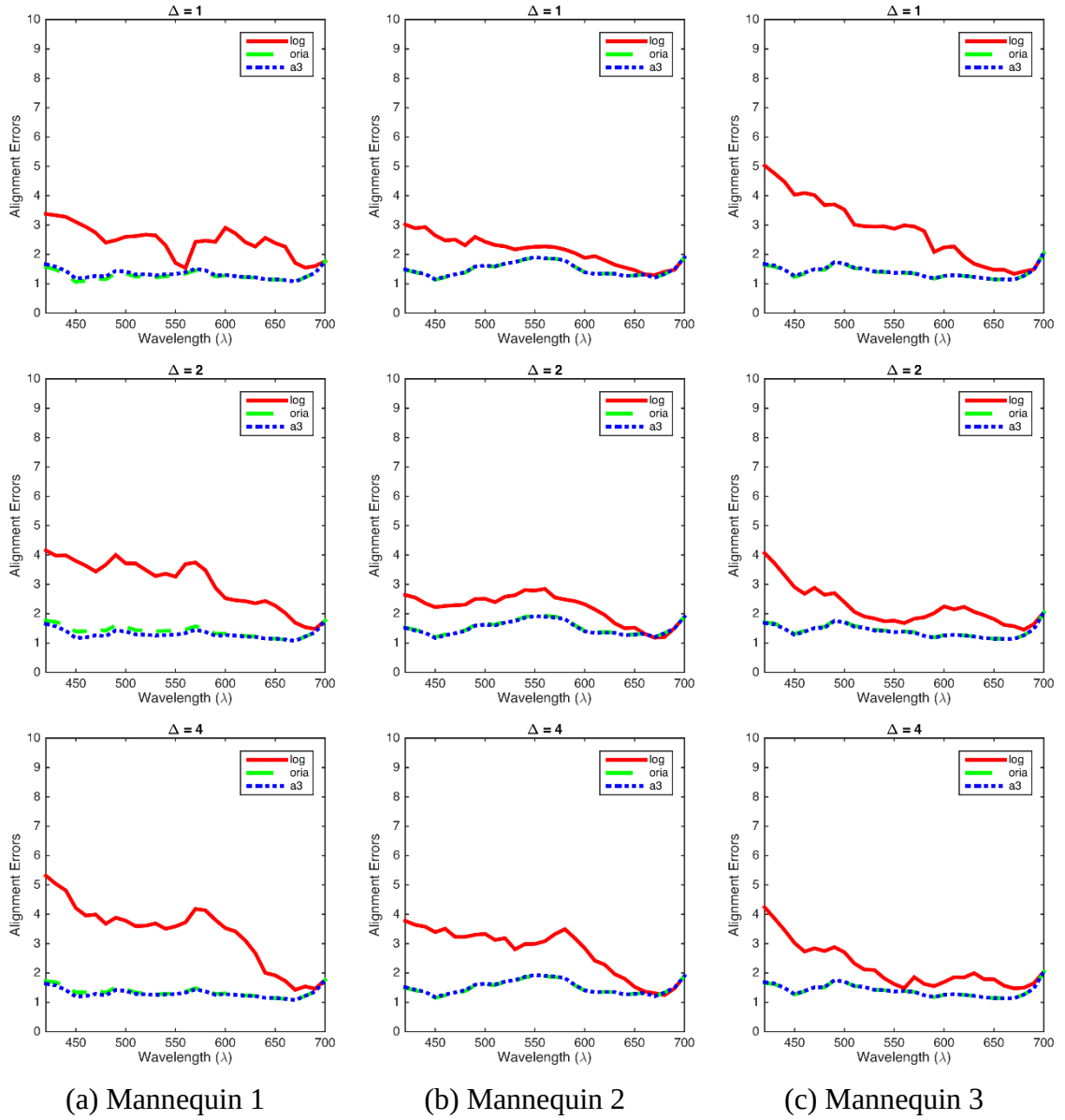


Figure 8.10: Comparison of the alignment errors of LoG, ORIA, and the proposed A_3 on the translated rigid sets where each rigid set is translated along with both x - and y -directions in the interval Δ .

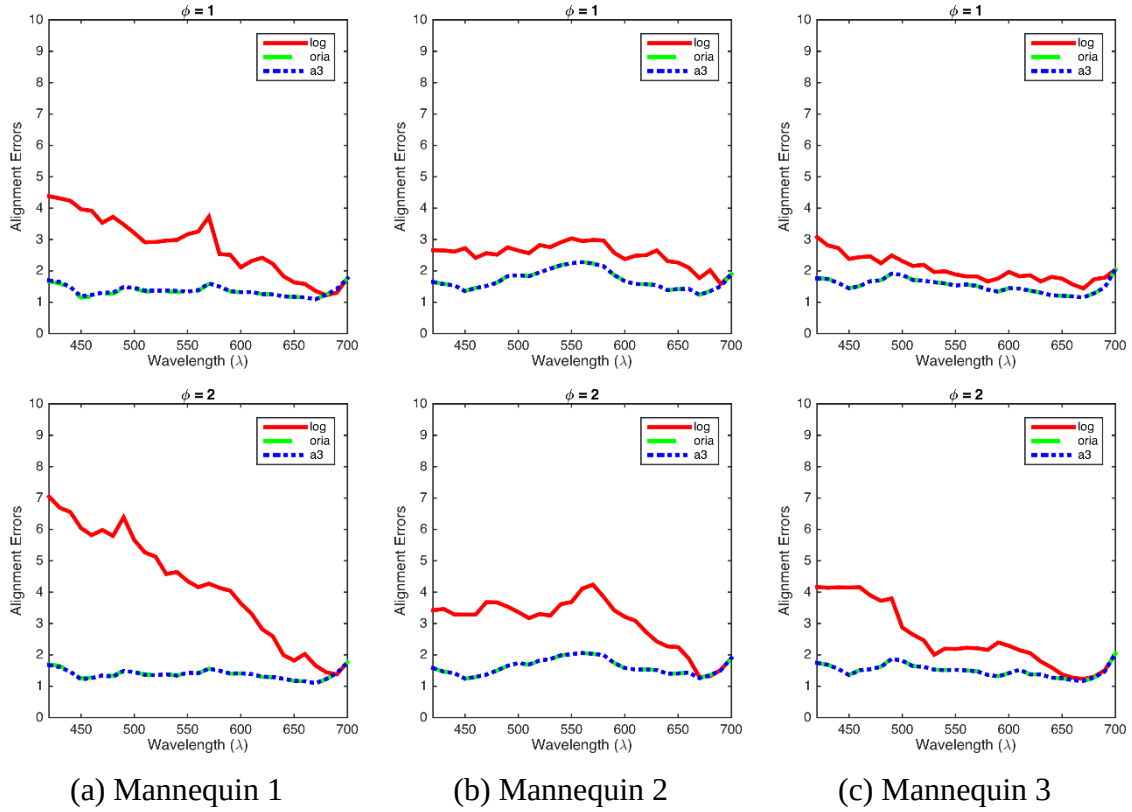


Figure 8.11: Comparison of the alignment errors of LoG, ORIA, and the proposed A_3 on the rotated rigid sets where each rigid set is rotated as $\phi \in \{1, 2\}$.

Table 8.3: Results of averaging alignment errors of the conducted alignment approaches on the rigid sets artificially disturbed by Δ and ϕ .

Input Sets	LoG	ORIA	A_3
Mannequin 1	3.59	2.48	1.88
Mannequin 2	2.48	1.99	1.80
Mannequin 3	2.78	1.51	1.50

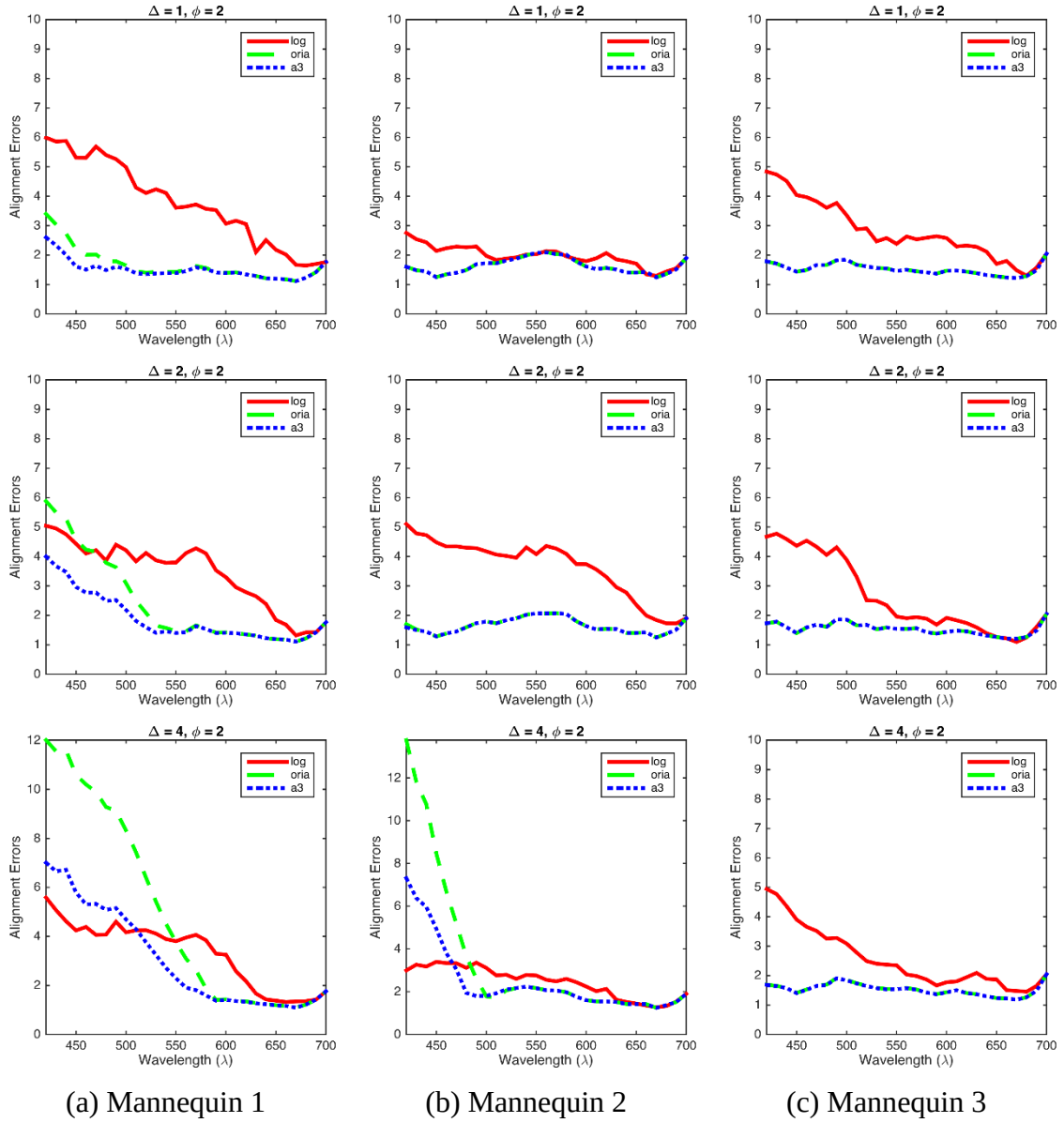


Figure 8.12: Comparison of the alignment errors of LoG, ORIA, and the proposed A_3 on the rigid sets disturbed by artificially disturbed by Δ and ϕ in the front (F), left profile (L), and right profile (R) views for $\Delta = 1$ and $\phi = 2$.

8.2.2 Results on UWA-HSFD

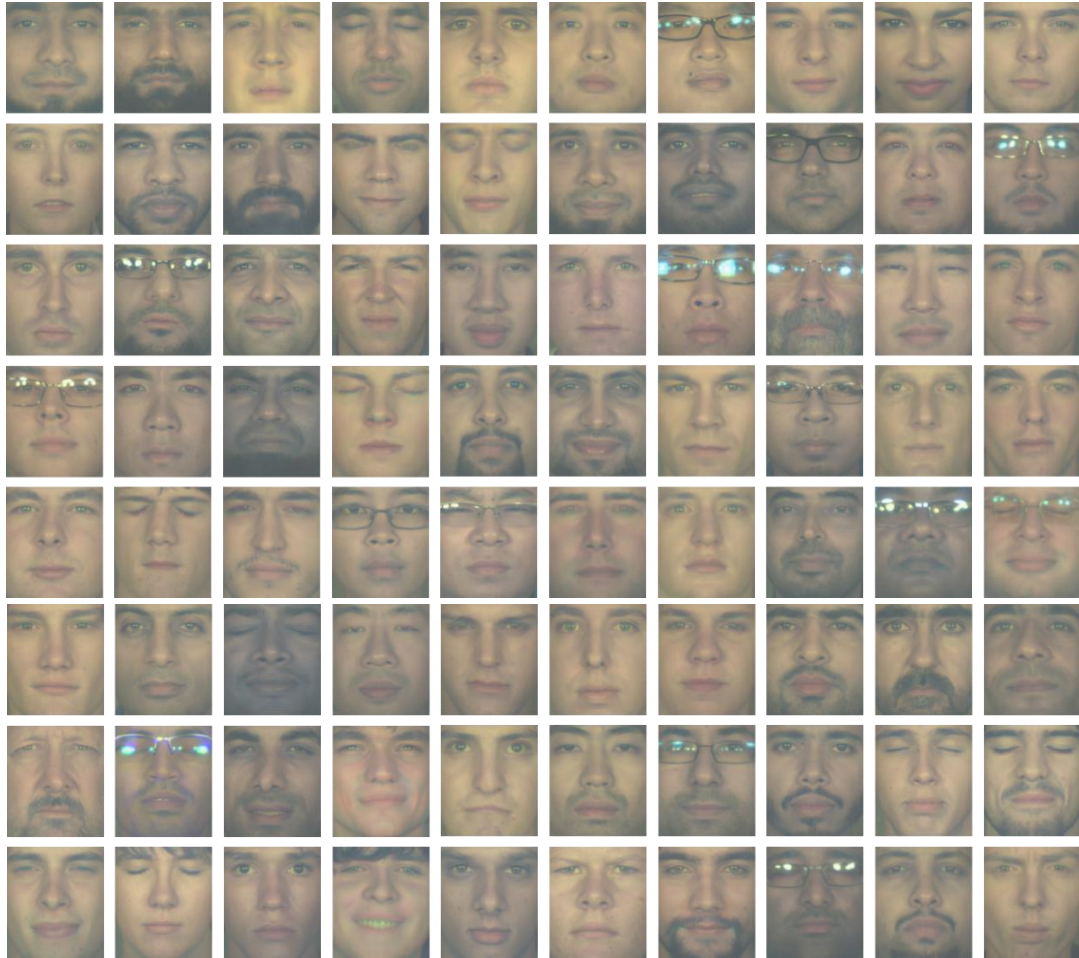


Figure 8.13: The aligned datasets of UWA-HSFD where full sub-bands (400 to 720 nm in 10 nm steps) are displayed using the sRGB values rendered under a CIE illuminant D65.

To further validate the alignment performance of the proposed method, we investigate the experimental evaluation on a large-scale hyperspectral face database, UWA-HSFD, including realistic inter-band misalignments. In this task, the effectiveness of the proposed method is compared to FBB, LoG, and ORIA which only require one manual input set of two eye coordinates on the last band (720 nm) in UWA-HSFD where one manual input set is used to crop each band in the HFIs. We employ two proposed different metrics (CMS and HUQA in Chapter 7) to predict the alignment quality of the HFI produced by the considered alignment approaches. We note that UWA-HSFD contains less inter-band misalignment compared to IRIS-HFD-2014. In addition, since UWA-HSFD considered the effects of spectral radiance for face identification, there was no effort to measure the

reflectance spectrum, or spectral signature. Thus, when we display the entire sub-bands using an sRGB image rendered under a CIE illuminant D65, an imaged subject seems to be yellowish and oversaturated as shown in Fig. 8.13. Table 8.4 demonstrates the results of averaging alignment errors of the considered alignment approaches on UWA-HSFD, which are estimated by the CMS and HUQA metrics.

Table 8.4: Results of averaging alignment errors of the conducted alignment approaches on UWA-HSFD.

Input Sets	Metrics	FBB	LoG	ORIA	A_3
Session 1	CMS	0.1857	0.1326	0.1324	0.1323
	HUQA	0.1076	0.1031	0.1017	0.1016
Session 2	CMS	0.1934	0.1320	0.1319	0.1318
	HUQA	0.1409	0.1319	0.1280	0.1283

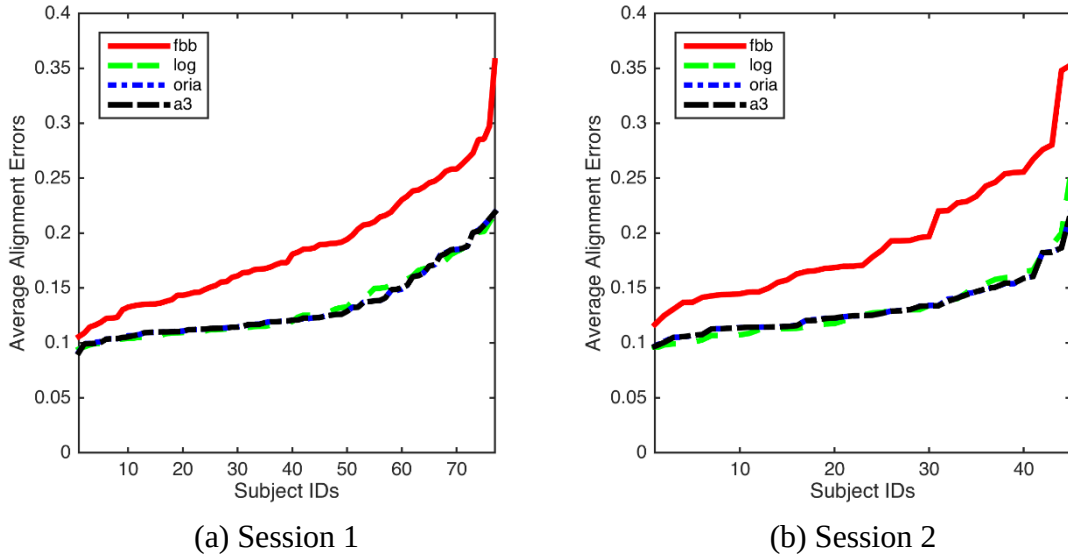


Figure 8.14: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the CMS metric on UWA-HSFD.

We note that the predicted scores are sorted in ascending order. From the results of FBB in Table 8.4 and Fig. 14, we can observe that there are IBMs in UWA-HSFD. ORIA and A_3 yield better-quality alignment compared to FBB and LoG methods. As shown in Fig. 8.15, the estimated scores of HUQA corresponding to FBB, LoG, ORIA, and A_3 are quite close together because the HFIs in UWA-HSFD are yellowish and oversaturated as mentioned

before. Therefore, we found that the highest prediction accuracy of HUQA should be achieved on the recovered reflectance images.

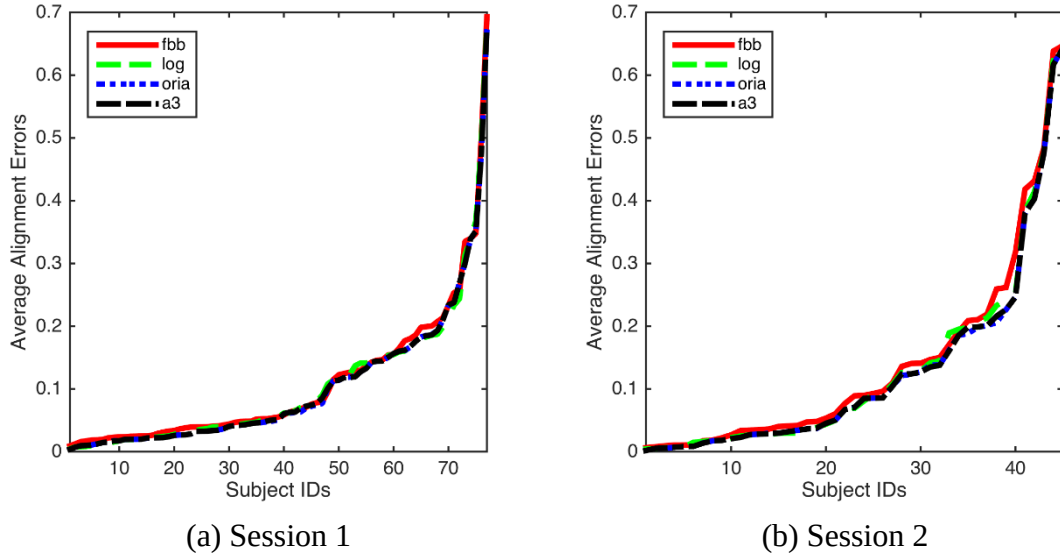


Figure 8.15: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the HUQA metric on UWA-HSFD.

8.2.3 Results on IRIS-HFD-2014

We further demonstrate the effectiveness of the proposed alignment method on IRIS-HFD-2014 that contains more significant IBMs in HFIs compared to UWA-HSFD. Specifically, IRIS-HFD-2014-V2 is rife with serious IBMs as shown in Fig. 8.16 because of the participants in age (4 to 15). Tables 8.5 to 8.7 summarize the results of alignment quality of the conducted alignment approaches on IRIS-HFD-2014 including a total of 644 hyperspectral face cubes.

In this experiment, we evaluate a comprehensive study on the performance of varying levels of practical misalignments in the HFIs. As we demonstrated previous experiments, if there is insignificant subject motion, the performances of ORIA and A_3 are quite close together. However, as shown in Table 8.7, we observed that ORIA is sensitive to align the HFIs including large-scale subject motion in each band. Compared to ORIA, the proposed method, A_3 , can handle significant subject motion by combining the strengths of LoG-based point correspondence and the low-rank model of ORIA. In addition, as shown in Figs. 17 to 22, HUQA yields highly consistent prediction accuracy on the reflectance images in IRIS-HFD-2014 compared to the prediction performance of HUQA on UWA-

HSFD. Therefore, Comparisons of FBB, LoG, and ORIA validate that A_3 is superior to reducing IBM artifacts by aligning HFIs including large-scale subject motion.

Table 8.5: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V1 in session 1.

Input Sets	Metrics	FBB	LoG	ORIA	A_3
Front view	CMS	0.2750	0.1569	0.1480	0.1461
	HUQA	0.0737	0.0222	0.0184	0.0176
Glass view	CMS	0.2811	0.1957	0.1781	0.1778
	HUQA	0.1010	0.0540	0.0498	0.0491
Left profile	CMS	0.2744	0.1656	0.1590	0.1584
	HUQA	0.0589	0.0184	0.0172	0.0165
Right profile	CMS	0.2719	0.1657	0.1618	0.1606
	HUQA	0.0575	0.0172	0.0163	0.0147

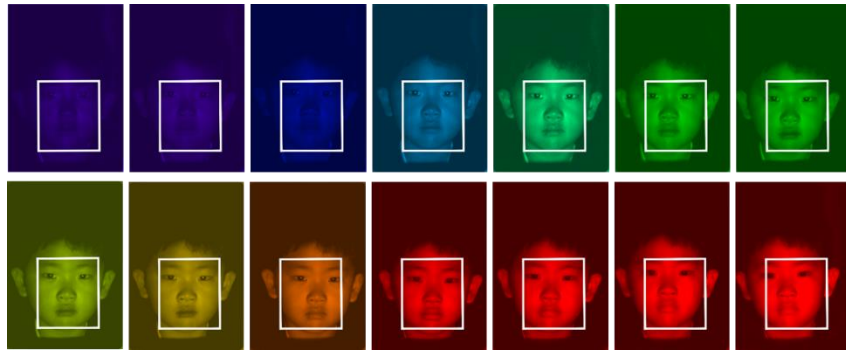
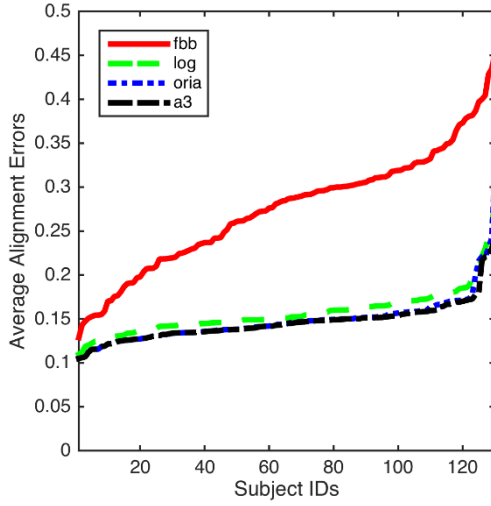


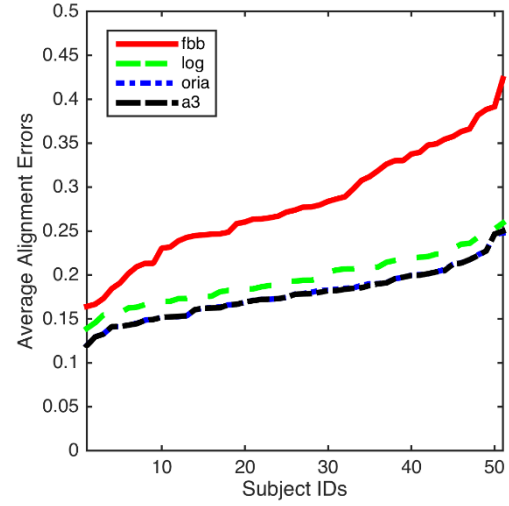
Figure 8.16: Examples of IBMs in IRIS-HFD-2014-V2. Each of the estimated warps from the FBB approach is depicted by a rectangle in each band.

Table 8.6: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V1 in sessions 2 and 3.

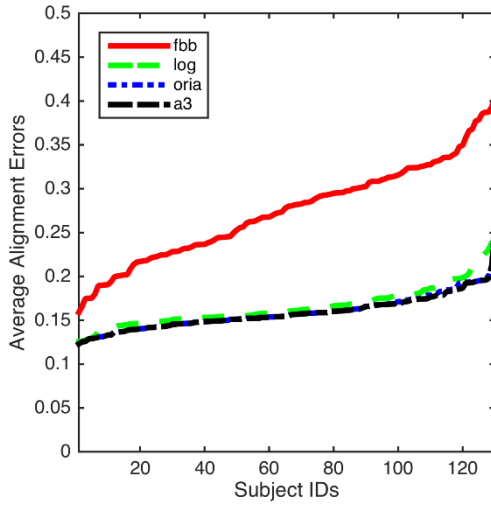
Input Sets	Metrics	FBB	LoG	ORIA	A_3
Session 1	CMS	0.2948	0.1626	0.1578	0.1532
	HUQA	0.0826	0.0224	0.0186	0.0185
Session 2	CMS	0.2841	0.1560	0.1438	0.1437
	HUQA	0.0661	0.0148	0.0122	0.0121



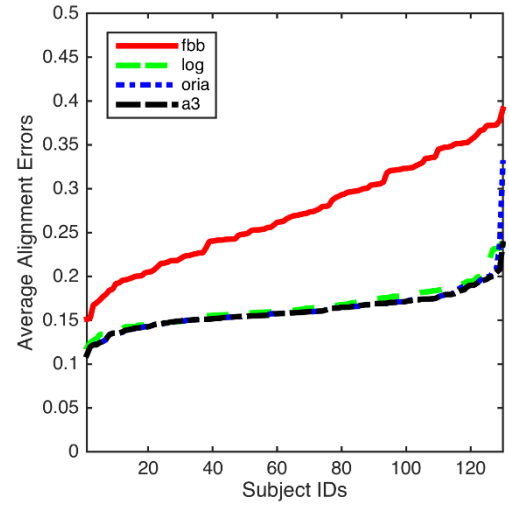
(a) Front view



(b) Glass view



(c) Left profile view



(d) Right profile view

Figure 8.17: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the CMS metric on IRIS-HFD-2014-V1 in session 1.

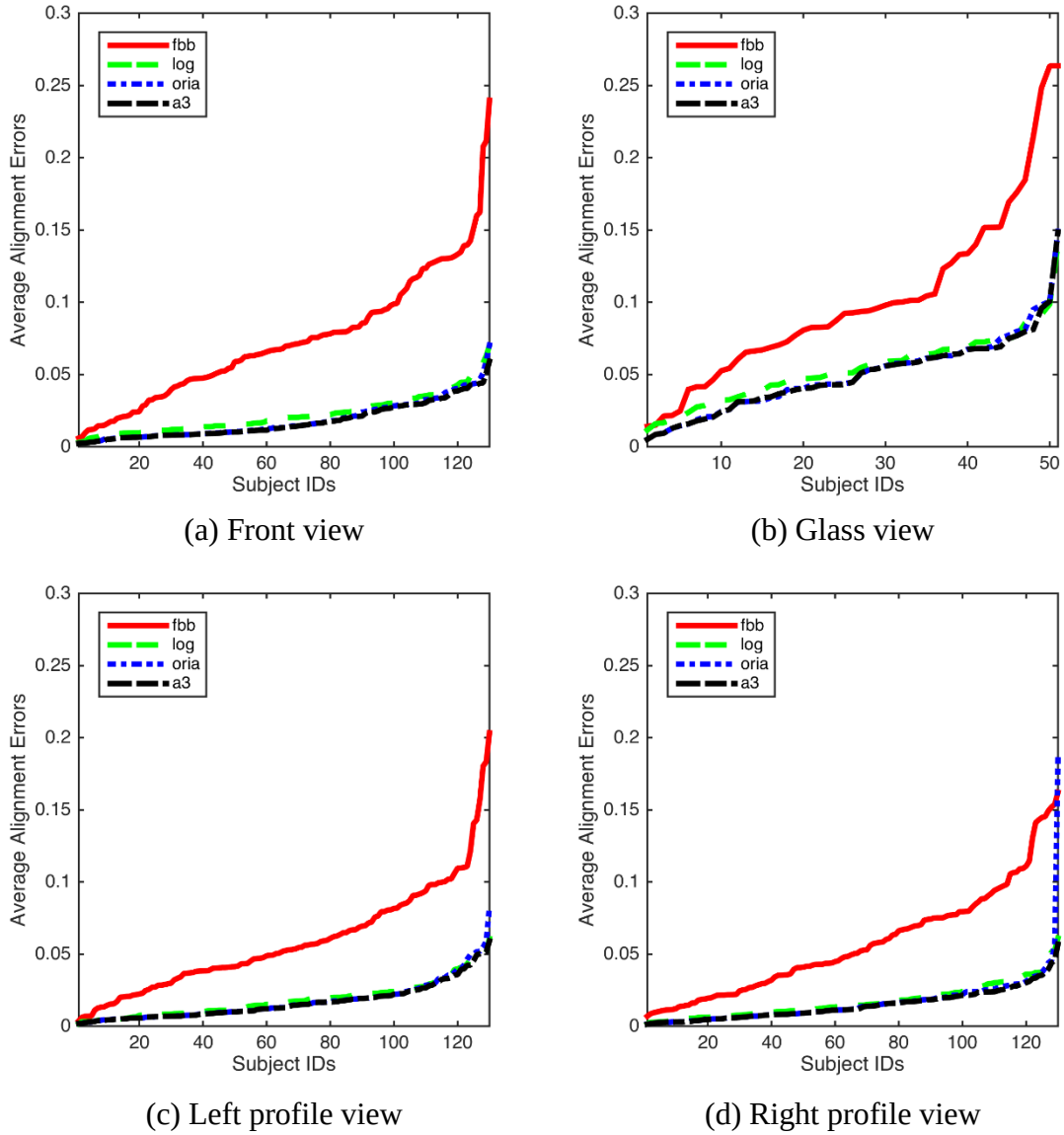
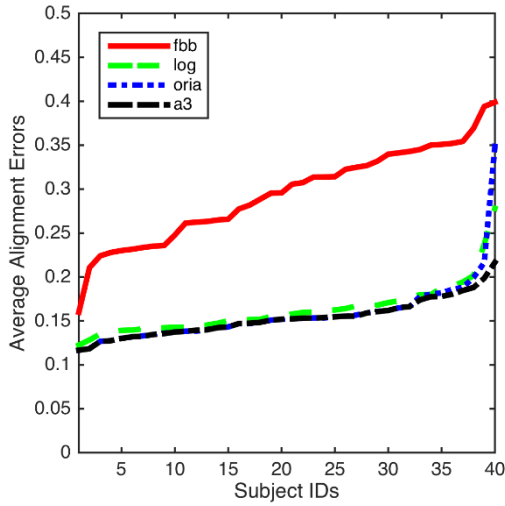
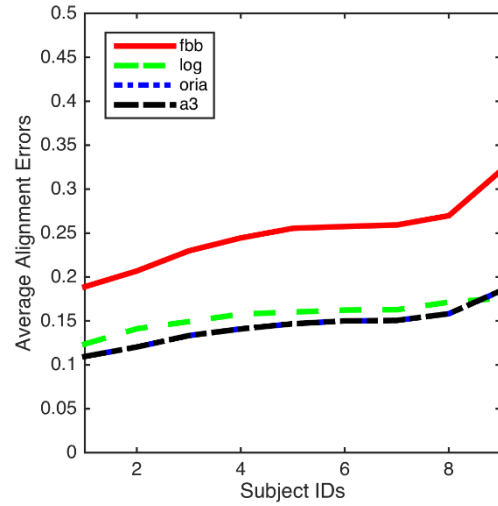


Figure 8.18: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the HUQA metric on IRIS-HFD-2014-V1 in session 1.

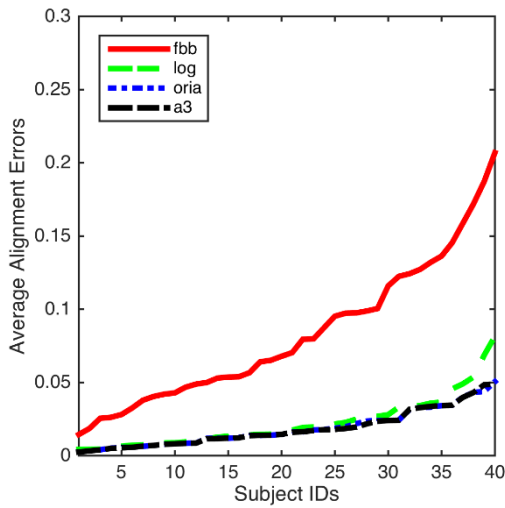


(a) Front view in Session 2

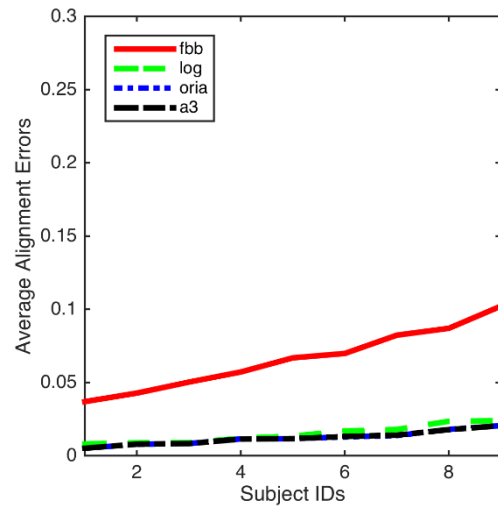


(b) Front view in Session 3

Figure 8.19: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 the CMS metric on IRIS-HFD-2014-V1.

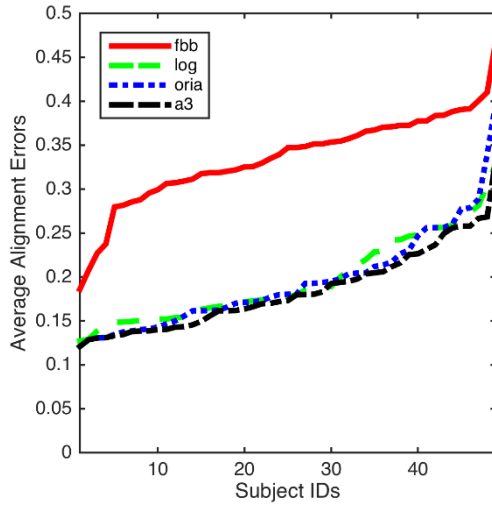


(a) Front view in Session 2

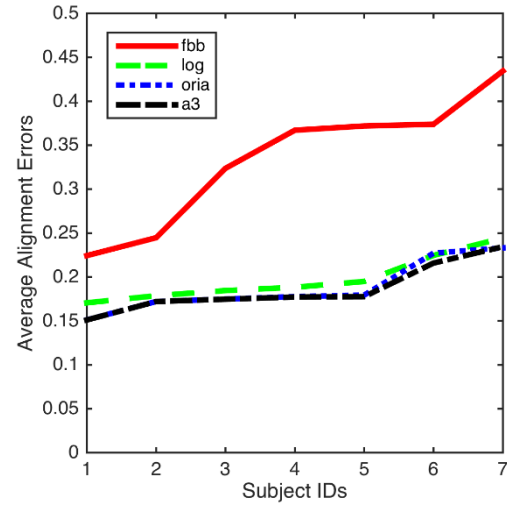


(b) Front view in Session 3

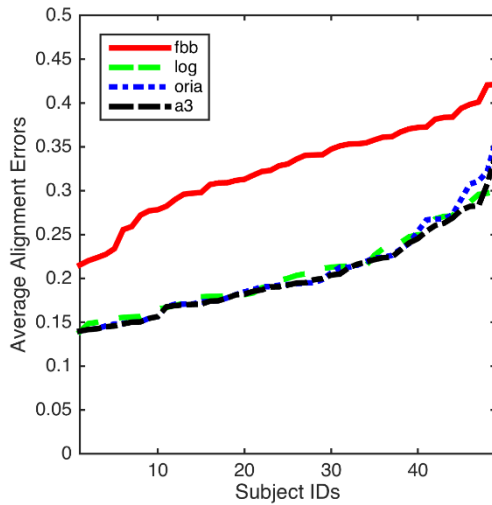
Figure 8.20: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the HUQA metric on IRIS-HFD-2014-V1.



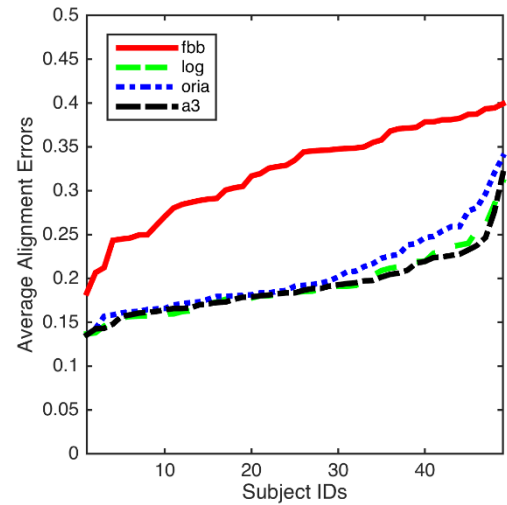
(a) Front view



(b) Glass view

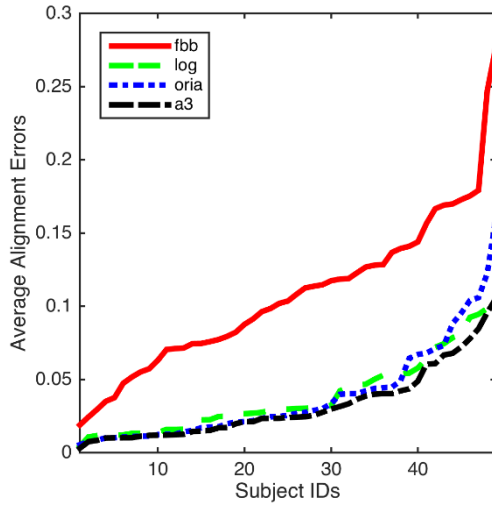


(c) Left profile view

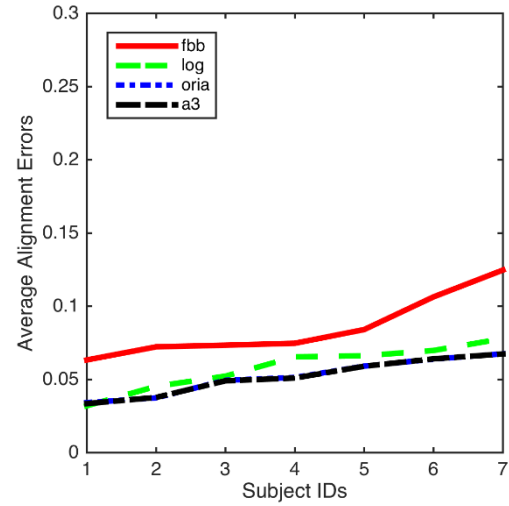


(d) Right profile view

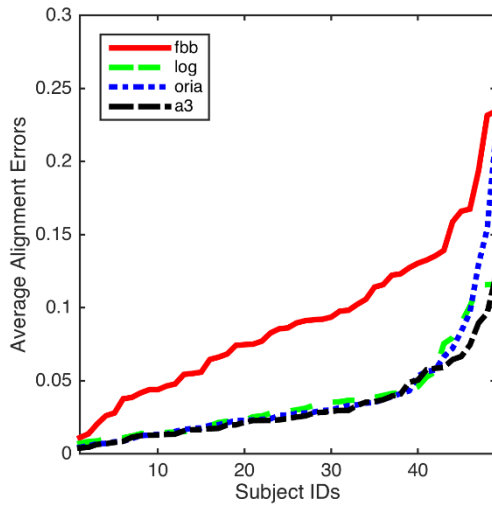
Figure 8.21: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the CMS metric on IRIS-HFD-2014-V2.



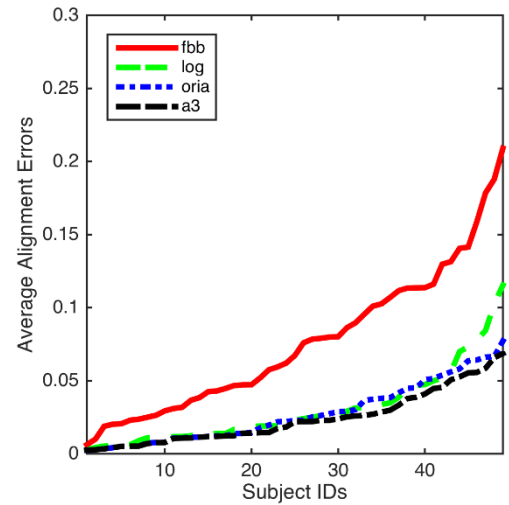
(a) Front view



(b) Glass view



(c) Left profile view



(d) Right profile view

Figure 8.22: Comparison of the alignment quality of FBB, LoG, ORIA, and the proposed A_3 via the HUQA metric on IRIS-HFD-2014-V2.

Table 8.7: Results of averaging alignment errors of the conducted alignment approaches on IRIS-HFD-2014-V2.

Input Sets	Metrics	FBB	LoG	ORIA	A_3
Front view	CMS	0.3355	0.1965	0.1951	0.1850
	HUQA	0.1074	0.0386	0.0388	0.0323
Glass view	CMS	0.3343	0.1980	0.1879	0.1861
	HUQA	0.0856	0.0584	0.0519	0.0517
Left profile	CMS	0.3258	0.2066	0.2059	0.2026
	HUQA	0.0913	0.0367	0.0380	0.0322
Right profile	CMS	0.3222	0.1923	0.2046	0.1911
	HUQA	0.0762	0.0306	0.0278	0.0241

8.3 Summary

We proposed a mixture model combining the strengths of a Laplacian of Gaussian (LoG)-based point correspondence to handle large-scale subject motion during data acquisition and a low-rank model to optimize image domain transformations. In the experiment results, we employed two proposed metrics (CMS and HUQA introduced in Chapter 7) to evaluate the alignment quality of the HFIs produced by FBB, LoG, ORIA, and the proposed alignment method. The efficacy of the proposed alignment approach is verified with extensive experiments including two large-scale hyperspectral face databases (UWA-HSFD and IRIS-HFD-2014) where two databases contain a wide range of realistic inter-band misalignments.

9 Evaluation of Face Recognition

In this Chapter, we demonstrate that the improved alignment leads to better accuracies of face recognition performance. To verify the improvements to face recognition performance, we employ a technique known as probabilistic linear discriminant analysis (PLDA) [Prince07], [Li12], which models intraclass and interclass variance as multidimensional Gaussian to seek maximum facial discriminability. As mentioned in [Shafey13], PLDA has achieved the state-of-the-art performance in face and speaker recognition fields.

9.1 Probabilistic linear discriminant analysis (PLDA)

As described in [Shafey13], the PLDA model introduced by Prince and Elder [Prince07] has been shown to achieve state-of-the-art performance for both face and speaker recognition. The PLDA model assumes that the j th image of the i th individual can be described as:

$$x_{i,j} = \mu + Fh_i + Gw_{i,j} + \epsilon_{i,j}. \quad (9.1)$$

In this process, the input signal $x_{i,j}$, of dimensionality D_x , is considered: 1) the identity part given by $\mu + Fh_i$ and 2) the noise component given by $Gw_{i,j} + \epsilon_{i,j}$. The matrices F and G are subspaces that represent the bases for the between-class variation and within-class variation, respectively. h_i and $w_{i,j}$ indicate the position in these subspaces for $x_{i,j}$ and are of size D_F and D_G , respectively. As a final point, the term $\epsilon_{i,j}$ denotes a stochastic noise term to be Gaussian with zero mean and diagonal covariance Σ . Figure 9.1 shows a graphical representation of the latent identity model.

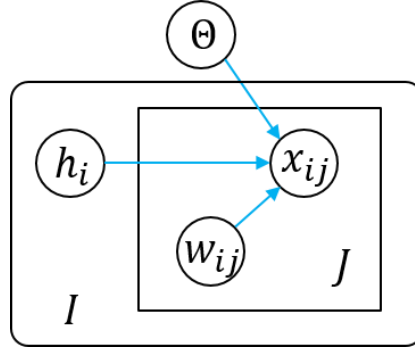


Figure 9.1: Graphical model of the PLDA: the j th image of the i th individual, x_{ij} , is modeled using a linear latent model with identities h_i , noise variables $w_{i,j}$, and parameter θ .

We can describe the process above in terms of a conditional probability:

$$\Pr(x_{i,j} | h_i, w_{i,j}, \theta) = \mathcal{N}[\mu + Fh_i + Gw_{i,j}, \Sigma], \quad (9.2)$$

and prior probabilities:

$$\Pr(h_i) = \mathcal{N}[0, I], \quad (9.3)$$

$$\Pr(w_{i,j}) = \mathcal{N}[0, I], \quad (9.4)$$

where I denotes the identity matrix and the parameters of the model are $\theta = [\mu, F, G, \Sigma]$.

In (9.3) and (9.4), we can define simple priors on the latent variables, h_i and $w_{i,j}$, to be Gaussian. The equations above can be written in a compact form by redefining $A = [F, G]$:

$$y_{i,j} = [h_i^T, w_{i,j}^T]^T. \quad (9.5)$$

We can rewrite the process as:

$$x_{i,j} = \mu + Ay_{i,j} + \epsilon_{i,j}, \quad (9.6)$$

and the compound model in terms of probabilities as:

$$\Pr(x_{i,j} | y_{i,j}, \theta) = \mathcal{N}[\mu + Ay_{i,j}, \Sigma], \quad (9.7)$$

$$\Pr(y_{i,j}) = \mathcal{N}[0, I]. \quad (9.8)$$

For example, given that $J_i = 2$ observations for identity i , we would have

$$\tilde{A} = \begin{bmatrix} F & G & 0 \\ F & 0 & G \end{bmatrix}. \quad (9.9)$$

Accordingly, we would set that $\tilde{x}_i = [\tilde{x}_{i,1}^T, \tilde{x}_{i,2}^T]^T$, $\tilde{\epsilon}_i = [\epsilon_{i,1}^T, \epsilon_{i,2}^T]^T$, $\tilde{w}_i = [w_{i,1}^T, w_{i,2}^T]^T$, $\tilde{y}_i = [y_{i,1}^T, y_{i,2}^T]^T$,

$$\tilde{\Sigma} = \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix}, \quad (9.10)$$

where

$$\bar{x}_{i,j} = x_{i,j} - \mu \quad (9.11)$$

and the symbol ‘ $\sim$ ’ denotes that the size of a variable depends on the number of samples J_i for i individuals.

For the general case of a class i with J_i samples, the model could be defined as:

$$\tilde{x}_i = \tilde{A}\tilde{y}_i + \tilde{\epsilon}_i. \quad (9.12)$$

There are two main tasks to estimate the probabilistic model: 1) training the model by using an expectation-maximization (EM) algorithm and 2) using the trained model to perform recognition by calculating the likelihood that a set of observations, $[x_{i,1}, x_{i,2}, \dots, x_{i,J_i}]$, share the same latent identity variable h_i .

To train the PLDA model, we employ an EM algorithm. All of the M-Steps are performed on a per sample basis once the latent variables have been estimated. In the E-Step, we need to calculate the first-order and second-order moments of the latent variables, we reproduce the equations as follows:

$$E[\tilde{y}_i | \tilde{x}_i, \Theta] = (\tilde{I} + \tilde{A}^T \tilde{\Sigma}^{-1} \tilde{A})^{-1} \tilde{A}^T \tilde{\Sigma}^{-1} (\tilde{x}_i), \quad (9.13)$$

$$E[\tilde{y}_i \tilde{y}_i^T | \tilde{x}_i, \Theta] = (\tilde{I} + \tilde{A}^T \tilde{\Sigma}^{-1} \tilde{A})^{-1} + E[\tilde{y}_i | \tilde{x}_i, \Theta] E[\tilde{y}_i | \tilde{x}_i, \Theta]^T. \quad (9.14)$$

From (9.13) and (9.14), we can observe that the E-Step is how to deal with the matrix $(\tilde{I} + \tilde{A}^T \tilde{\Sigma}^{-1} \tilde{A})^{-1}$ efficiently as it should be recomputed for each iteration of EM.

The likelihood for the PLDA is to solve several problems for face identification and authentication. To calculate the likelihood that a set of samples, $\tilde{x}_i$, we share the same latent identity variable. This can be performed on a probabilistic way by integrating out over all the latent variables. Therefore, we tie together the latent identity variable, h_i , of the samples to have the same identity and we then consider each observation, $x_{i,j}$, to have a separate latent session variable, $w_{i,j}$. Next, we integrate over h_i and all of the individual $w_{i,j}$ s. For the case of $J_i = 2$ we define this as follows:

$$\begin{aligned} & \Pr(x_{i,1}, x_{i,2}) \\ &= \int \left[\int \Pr(x_{i,1} | h_i, w_{i,1}) \Pr(w_{i,1}) dw_{i,1} \int \Pr(x_{i,2} | h_i, w_{i,2}) \Pr(w_{i,2}) dw_{i,2} \right] \Pr(h_i) dh_i, \end{aligned} \quad (9.15)$$

where for simplicity we have dropped the reference to Θ . This problem can be rewritten as:

$$\Pr(\tilde{x}_i | \Theta) = \mathcal{N}[0, \tilde{\Sigma} + \tilde{A} \tilde{A}^T]. \quad (9.16)$$

Equivalently, we can estimate the log-likelihood as:

$$\ln[\Pr(\tilde{x}_i | \Theta)] = -\frac{J_i D_x}{2} \ln[2\pi] - \frac{1}{2} \ln[\det(\tilde{\Sigma} + \tilde{A} \tilde{A}^T)] - \frac{1}{2} \tilde{x}_i^T (\tilde{\Sigma} + \tilde{A} \tilde{A}^T)^{-1} \tilde{x}_i. \quad (9.17)$$

We refer the interested reader to the respective papers [Li12], [Shafey13] for efficient calculation methods of the likelihood for the PLDA model to reduce the complexity of the computational cost.



(a) Mean, μ



(b) Noise, Σ

Figure 9.2: PLDA model of IRIS-HFD-2014: (a) mean face and (b) noise covariance sRGB images.

Figure 9.2 shows the mean face and noise covariance images of the PLDA model trained from IRIS-HFD-2014. Figure 9.3 presents examples of the sRGB images reconstructed by the PLDA model once performing the training stage.

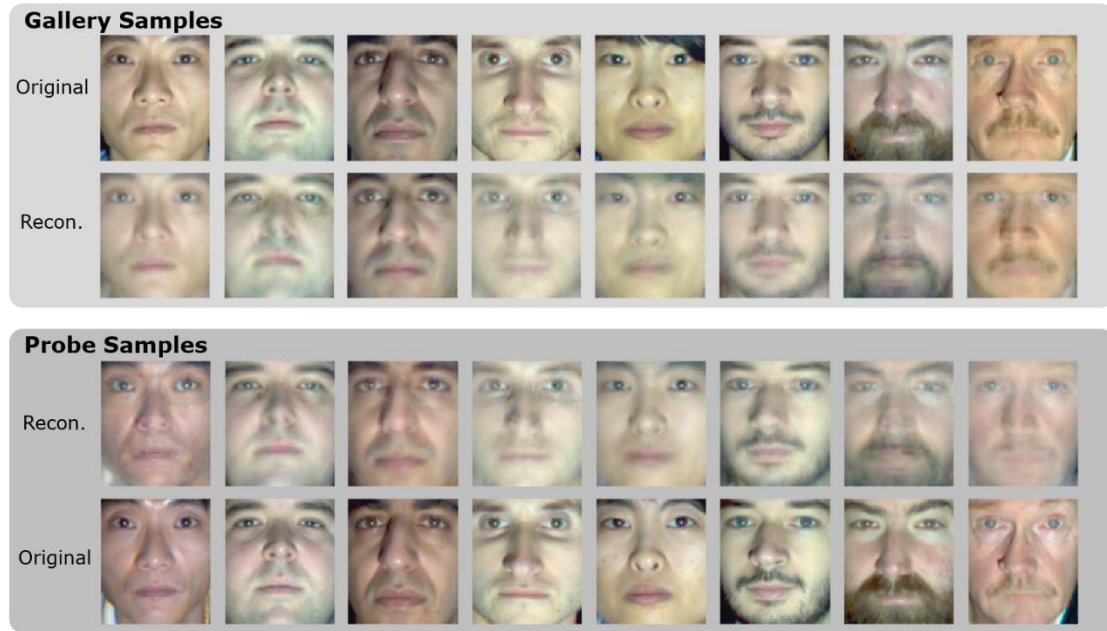


Figure 9.3: Examples of the reconstructed sRGB images by the PLDA model trained from IRIS-HFD-2014.

9.2 Experimental Results

In Chapter 8, the proposed alignment approach, called accurate alignment approach (A_3) consistently and accurately yielded the most improved alignment on the considered experiments, compared to the selected alignment approaches (AAs): fixed bounding box-based AA (FBB), Laplacian of Gaussian-based AA (LoG), online robust image alignment (ORIA). In this section, we validate how the alignment accuracies affect the performance of face recognition of the PLDA on two large-scale databases such as UWA-HSFD and IRIS-HFD-2014.

9.2.1 Results on UWA-HSFD

The UWA-HSFD contains a total of 142 hyperspectral face cubes of 79 data subjects in the frontal view. For the verification of the effect of face recognition performance, we select each cube from the UWA-HSFD. For the training set, we use 105 cubes of 72 subjects over 4 sessions. The gallery set in the testing set is constructed by choosing 30 cubes from the training set and 30 cubes as the probe set are selected from the remaining cubes over 4 sessions, but not included in the training stage. For computational efficiency,

the face cubes are resized to $40 \times 35 \times 33$. Figure 9.6 shows the identification result of 20 iterations of training from HFIs as an increase in the signal and noise subspace size from 16 to 48 dimension with a step of 8.

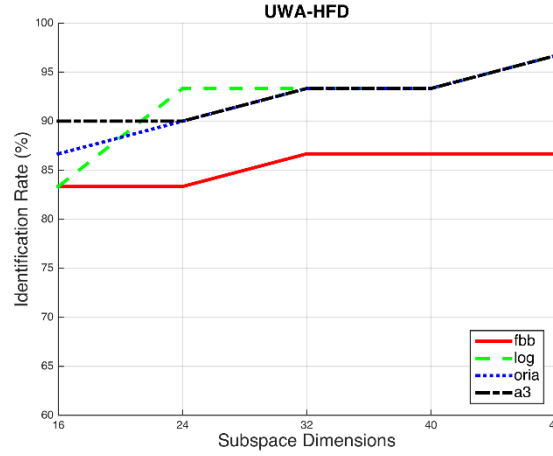


Figure 9.4: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on UWA-HSFD.

As seen in Fig. 9.4, the optimal subspace dimension of PLDA on the aligned HFIs of the proposed A_3 is 48. Therefore, PLDA based on 48 factors using the aligned HFIs produced by LoG, ORIA, A_3 achieves better accuracies of face recognition than the misaligned HFIs of FBB, which are 10 percentage points higher than those of the misaligned HFIs as summarized in Table 9.1.

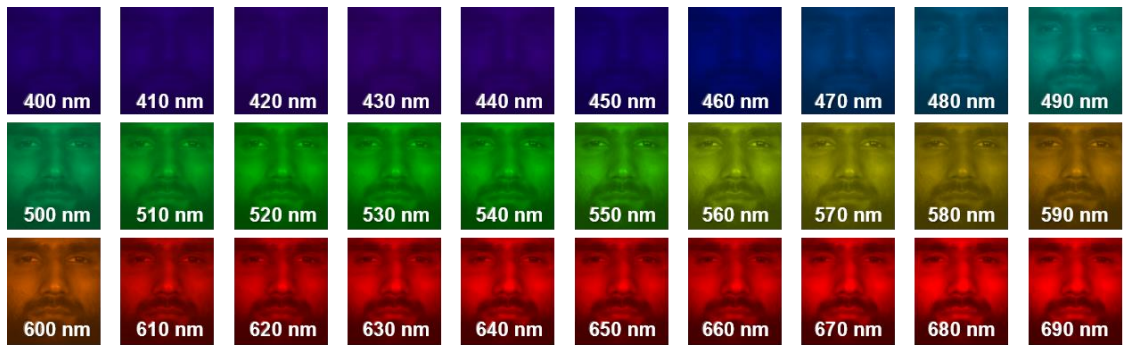


Figure 9.5: HFI sequences of UWA-HSFD ranging from 400 to 720 nm with 10 nm steps.

Table 9.1 Comparison of the first rank identification rate based on 48 factors of PLDA on UWA-HSFD.

Input sets	FBB	LoG	ORIA	A_3 (Proposed)
UWA-HSFD (%)	87	97	97	97

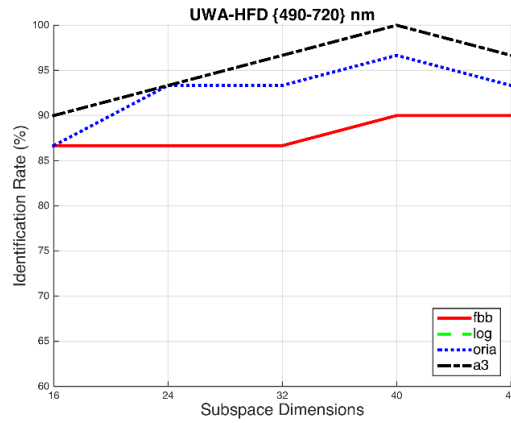


Figure 9.6: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on UWA-HSFD once resizing the face cubes to $30 \times 30 \times 24$ ranging from 490 to 720 nm.

Table 9.2 Comparison of the first rank identification rate based on 40 factors of PLDA on UWA-HSFD once resizing the face cubes to $30 \times 30 \times 24$ ranging from 490 to 720 nm.

Input sets	FBB	LoG	ORIA	A_3 (Proposed)
UWA-HSFD (%)	90	100	97	100

As shown in Fig. 9.5, we observed that lower bands ranging from 400 to 480 nm in UWA-HSFD are considerably blurred. Thus, once we reject the blurred bands in the input sets resized to $30 \times 30 \times 24$, we further investigate the performance of PLDA for the same experimental protocol as in Fig. 9.4.

Figure 9.6 shows the improved face recognition performance by removing the blurred bands in UWA-HSFD. When the subspace of PLDA is 40, the highest accuracies of PLDA on the aligned HFIs of LoG and A_3 are equivalently achieved as illustrated in Fig. 9.6. With

better quality alignments of LoG and A_3 , PLDA improves the accuracy of identification by 3 percentage points for ORIA and 10 percentage points for FBB as shown in Table 9.2.

9.2.2 Results on IRIS-HFD-2014

We further verify the improvements to face recognition performance on IRIS-HFD-2014, which contains more significant subject motion than the UWA-HSFD. For the training set, we select 290 hyperspectral face cubes over the entire datasets including the glass view in the version 1 and the profile views in the version 2. Each of 47 cubes for the gallery and probe set are chosen from the frontal view in the version 1. Note that since there is no overlap between the training and testing sets (gallery and probe sets) the probabilistic model generalizes from the training set to new data subjects. For computational efficiency, the hyperspectral face cubes are resized to $40 \times 35 \times 29$.

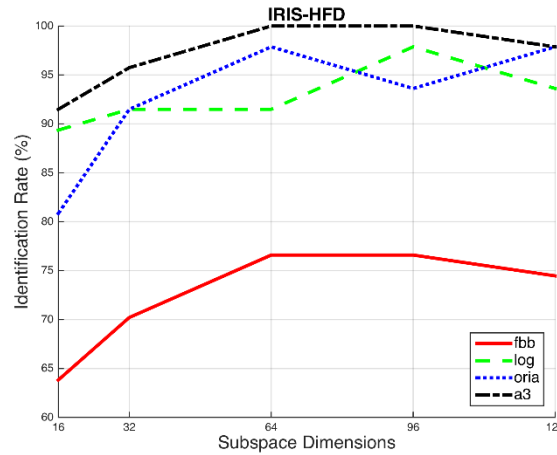


Figure 9.7: Result of the first rank identification rate of PLDA versus the signal and noise subspace size on IRIS-HFD-2014.

Figure 9.9 shows the identification results of the PLDA for different subspace size from 16 to 128 dimension once 20 iterations of training from HFIs derived from FBB, LoG, ORIA, and A_3 . The optimal subspace dimension of PLDA on the aligned HFIs of the proposed A_3 is 64. As shown in Table 9.3, the PLDA with 64 factors improves the accuracy of identification by 23 percentage points for FBB as the misaligned HFIs. Therefore, experimental results in Fig. 9.8 and Table 9.3 demonstrate that the proposed alignment approach A_3 achieves superior performance for face recognition compared to four studied alignment approaches.

Table 9.3 Comparison of the first rank identification rate based on 64 factors of PLDA on IRIS-HFD-2014.

Input sets	FBB	LoG	ORIA	A_3 (Proposed)
IRIS-HFD-2014 (%)	77	91	98	100

9.3 Summary

In this Chapter, we demonstrated the efficacy of the proposed alignment approach (A_3) on two large-scale public hyperspectral face databases such as UWA-HSFD and IRIS-HFD-2014. To verify the improvements to face recognition performance, we employed the state-of-the-art face recognition approach known as probabilistic linear discriminant analysis (PLDA). Experimental results showed highly competitive abilities of the proposed alignment method to improve hyperspectral face recognition.

10 Conclusion

This dissertation was motivated by a need of consistent and robust algorithms for current face recognition system. Current face recognition systems were rife with multiple, varied, and significant challenges. Hyperspectral imaging (HI) was typically employed to counter many of these challenges, by incorporating the spectral information within different sub-bands. Primarily, two factors contribute to this enhanced sub-band spectral information: 1) the invariance of changes in illumination conditions resulting from the recovery of spectral properties of objects; and 2) the ability to detect distinct patterns contained in human faces where such discriminative patterns cannot be captured by trichromatic (RGB) color or monochromatic (grayscale) cameras. Although numerous methods based on hyperspectral imaging have been developed for face recognition with promising results, there are three challenges specific to HI for practical use of hyperspectral face recognition: 1) low signal to noise ratios and low intensity values in the bands of the hyperspectral image specifically near blue bands; 2) high dimensionality of hyperspectral data; and 3) inter-band misalignments when hyperspectral imaging is applied to non-rigid objects. Our work concentrated mainly on those challenges in this dissertation.

In Chapter 2, we introduced fundamentals of HI relevant to this dissertation, including HI technologies and hyperspectral image formations. We presented an overview of the theoretical and practical issues to exploit hyperspectral face images.

In Chapter 3, we reviewed three challenges in HI towards providing information on the key points of each of the challenges for practical use of hyperspectral face recognition. First, we provided a review in terms of five publicly available hyperspectral face databases (HFDs): CMU, IRIS-M, PolyU-HSFD, Stanford, and UWA-HSFD. Second, we offered the historical review of data compression techniques to address the high dimensionality of the spectral data. Third, we presented a study of inter-band misalignments through an extensive analysis of hyperspectral face images collected by the liquid crystal tunable filter in our experiments.

In Chapter 4, we introduced herein a new large HFD, IRIS-HFD-2014, which can serve as a benchmark for comprehensively and statistically evaluating the performance of current and future algorithms for hyperspectral face alignment and recognition. IRIS-HFD-2014 was captured by adjusting appropriate exposure time in each band. IRIS-HFD-2014 is designed to address several challenging problems in face recognition research, including

variations in time, pose (both frontal and profile views), and structural features (i.e., glasses). In addition, the database contains RGB color images of 142 participants captured by a traditional color camera under varying illuminant conditions and blurring in uncontrolled settings. IRIS-HFD-2014 consists of a total of 19,346 facial images of 115 males and 64 females of diverse ethnic backgrounds and appearance.

In Chapter 5, we proposed local proper orthogonal decomposition (LPOD) method to address the dimensionality of hyperspectral data by employing the k-means clustering approach efficiently reducing the dimension of the given spectral data that were highly correlated to neighboring spectral samples. Whereas we applied k-means clustering to the set of snapshots for achieving good partitions, the proposed approach is flexible enough to accommodate other clustering methods, e.g., hierarchical clustering, self-organizing maps (SOM), partitioning around medoids (PAM), and so on. For a comprehensive study on the effectiveness of our reduced-order data modeling approach, six leading and two conventional techniques on three public domain databases and one in-house hyperspectral face database in five error metrics were established. Experimental results showed that the proposed method achieved highly competitive accuracy in terms of five considered error metrics on four different databases.

In Chapter 6, we introduced the new framework to reduce inter-band misalignment artifacts in hyperspectral face images (HFIs). Towards addressing IBMs we individually employed four different alignment techniques: 1) two conventional alignment approaches (AAs) based on manually selecting the region of interest (ROI) such as fixed bounding box-based (FBB) and eye coordinate-based (EC); 2) two iterative convex optimization (ICO) approaches such as RASL and ORIA; 3) five state-of-the-art landmark-based AAs including DRMF, IPCM, SDM, CDM, and TSPM; and, 4) two popular image AAs such as LK and SIFTFlow.

In Chapter 7, we developed two different metrics to automatically predict the improved alignment among the selected alignment approaches to address inter-band misalignments in HFIs. We proposed 1) a full-reference alignment quality assessment based on principal curvature map for evaluating the alignment quality of each sub-band in the HFI set; and 2) a no-reference alignment quality assessment based on the cumulative probability of target colors in the HSV color space for evaluating the alignment quality of a single sRGB image generated by the entire sub-bands in the HFI set. Experimental results were reported on IRIS-HFD-2014 compared with two recent full-reference approaches based on the gradient map and five state-of-the-art no-reference approaches. For experimental verifications, we examined the proposed metrics in two aspects: prediction consistency and prediction accuracy on rigid object sets and non-rigid subject sets. Comparisons with 7 state-of-the-art IQA metrics showed that both the proposed metrics led to promising accuracy of determination of the improved alignment among FBB, EC, RASL, and ORIA.

In Chapter 8, we proposed a mixture model combining the strengths of a Laplacian of Gaussian (LoG)-based point correspondence to handle significant subject motion and a low-rank model to optimize image domain transformations. The efficacy of the proposed alignment approach is verified with extensive experiments on two large-scale hyperspectral face databases (UWA-HSFD and IRIS-HFD-2014) that include a wide range of realistic inter-band misalignments.

In Chapter 9, we employed the probabilistic linear discriminant analysis technique to investigate how the alignment accuracy affects face recognition performance. The experimental results on two large-scale hyperspectral face databases (UWA-HSFD and IRIS-HFD-2014) demonstrated that the proposed alignment approach led to better accuracies of face recognition performance of the probabilistic linear discriminant analysis, compared to three existing alignment approaches.

10.1 Directions for future research

In this dissertation, our new database mainly considered how variations in pose, accessory, and time delay in indoor environments affect facial appearance via the LCTF. However, in practical use of hyperspectral imaging for face recognition, a future database should be developed in two different environments: 1) indoor environment under several custom light sources that have relatively more radiant power at short wavelengths; and 2) outdoor environments in the consideration of the state of the atmosphere such as temperature, wind, cloud, and rain. To acquire the hyperspectral images in two environments, it is necessary to design a new hyperspectral imaging system mounted on a mobile platform. In addition, a new large-scale face database including hyperspectral image cubes (HICs) obtained from more than 1000 participants puts into devising a benchmark for more comprehensively and statistically evaluating the performance of current and future algorithms for hyperspectral face recognition.

In evaluation of the identification performance, we further investigate performance of PLDA with the sRGB images instead of selecting maximally discriminative bands of the visible spectrum. For example, we select 117 cubes of 99 individuals from the first and third sessions for the training set and we test with 31 individuals, using the two cubes from the first and second sessions as the gallery and probe sets, respectively. Next, for computational efficiency, the face cubes are resized to $30 \times 30 \times 29$. Table 10.1 demonstrates the identification results of 20 iterations of training from HICs. The optimal subspace dimension of PLDA on the aligned HICs of RASL is 20. Therefore, PLDA based on 20 factors using the aligned HICs achieves higher accuracies of face recognition performance than the misaligned HICs, which are 26 percentage points higher than those of the misaligned HICs as shown in Table 10.1. We test performance of PLDA for the same

experimental protocol on the results of FBB and RASL with the corresponding sRGB images. The sRGB images are resized to $30 \times 30 \times 3$. When the subspace dimension of PLDA is 20, the highest accuracy of PLDA on the sRGB images of RASL is achieved. Using the aligned sRGB images, PLDA improves the accuracy of identification by 20 percentage points compared to the misaligned images. Therefore, the recognition rates on the HICs of RASL are increased by 6 percentage points compared to the aligned sRGB images as shown in Table 10.1, but it is unwise to draw a strong conclusion because of the difference in the number of used bands. We believe that a technique for band selection to seek maximally discriminative bands of the visible spectrum for face recognition can accomplish a similar recognition performance while we only use fewer bands by developing a robust band selection technique based on cross-validatory procedures for face recognition. The cross-validation is one of the simplest ways to seek interesting patterns and then to reduce data redundancy.

Table 10.1 Comparison of the identification rate of PLDA based on 20 factors using misaligned and aligned sets on IRIS-HFD-2014-V1.

Input sets	RASL (aligned)	FBB (misaligned)
IRIS-HFD-2014-V1 (%)	100	74
sRGB (%)	94	74

In this dissertation, although we mainly focused on demonstrating how alignment accuracy affects the face identification rate of PLDA, we can further consider face verification performance using receiver operating characteristics (ROC) curves that plot the verification rate against the false accept rate where the false accept rate provides the performance when a probe face is not of a face in the gallery.

Bibliography

-
- [Abate07] A. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2d and 3d face recognition: a survey," *Pattern Recognition Letters*, vol. 28, pp. 1885-1906, Oct. 2007.
 - [Ahonen06] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 2037-2041, Dec. 2006.
 - [Aldahd.13] R. Aldahdooh and W. Ashour, "DIMK-means—distance-based initialization method for k-means clustering algorithm," *Int. Jour. Intelligent Systems and Applications*, vol. 5, pp. 41, 2013.
 - [Altmann13] Y. Altmann, N. Dobigeon, S. McLaughlin, and J. Tournieret, "Nonlinear spectral unmixing of hyperspectral images using gaussian processes," *IEEE Trans. Signal Processing*, vol. 61, pp. 2442-2453, May 2013.
 - [Amirshahi10] S. Amirshahi and S. Amirshahi, "Adaptive non-negative bases for reconstruction of spectral data from colorimetric information," *Optical Review*, vol. 17, pp. 562-569, Nov 2010.
 - [Arbelaitz13] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. Pérez, and I. Perona, "An extensive comparative study of cluster validity indices," *Pattern Recognition*, vol. 46, pp. 243-256, Jan. 2013.
 - [Ashok10] R. Ashok and S. Nousath, "Subspace methods for face recognition," *Computer Science Review*, vol. 4, pp. 1-17, 2010.
 - [Asthana13] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Robust discriminative response map fitting with constrained local models," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2013, pp. 3444-3451.
 - [Asthana14] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2014, pp. 1859-1866.
 - [Baker04] S. Baker and I. Matthews, "Lucas-kanade 20 years on: a unifying framework," *Int. J. Computer Vision*, vol. 56, pp. 221-255, 2004.
 - [Barakzahi13] M. Barakzahi, S. Amirshahi, S. Peyvandi, and M. Afjeh, "Reconstruction of total radiance spectra of fluorescent samples by means of nonlinear principal component analysis," *J. Opt. Soc. Am. A*, vol. 30, pp. 1862-1870, 2013.
 - [Bartlett98] M. Bartlett, H. Lades, and T. Sejnowski, "Independent component representations for face recognition," in *Proc. SPIE Symposium Electronic Imaging: Sci. and Tech., Conf. Human Vision and Electronic Imaging III*, 1998.
 - [Bay08] H. Bay, A. Ess, T. Tuytelaar, and L. Gool, "Speeded-up robust features," *Computer Vision and Image Understanding*, vol. 110, pp. 346-359, 2008.

-
- [Belhum.97] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711-720, July 1997.
 - [Berns00] R. Berns, *Billmeyer and Saltzman's Principles of Color Technology*, 3rd ed. New York: J. Wiley and sons, 2000.
 - [Berns05] R. Berns, L. Taplin, F. Imai, E. Day, and D. Day, "A comparison of small-aperture and image-based spectrophotometry of paintings," *Studies in Conservation*, vol. 50, pp. 253-266, 2005.
 - [Bianco10] S. Bianco, "Reflectance spectra recovery from tristimulus values by adaptive estimation with metameric shape correction," *J. Opt. Soc. Am. A*, vol. 27, pp. 1868-1877, 2010.
 - [Brunelli93] R. Brunelli and T. Poggio, "Face recognition - features versus templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1042-1052, Oct. 1993.
 - [Camphou.05] R. Camphouse, "Boundary feedback control using proper orthogonal decomposition models," *J. Guidance Control and Dynamics*, vol. 28, pp. 931-938, Sept. 2005.
 - [Cao11] X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin, "A Prism-Mask System for Hyperspectral Video Acquisition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2423-2435, Dec. 2011.
 - [Chang06] H. Chang, H. Harishwaran, M. Yi, A. Koschan, B. Abidi, and M. Abidi, "An indoor and outdoor, multimodal, hyperspectral and multi-illuminant database for face recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshop*, 2006, pp. 54-54.
 - [Chang08] H. Chang, Hyperspectral imaging for face recognition over varying illumination, Ph.D. dissertation, Dept., Elec. Eng. Computer Sci., Univ. of Tenn., TN, Dec. 2008.
 - [Chellappa95] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, vol. 83, pp. 705-741, May 1995.
 - [Chen04] S. Chen, J. Liu, and Z. Zhou, "Making flda applicable to face recognition with one sample per person," *Pattern Recognition*, vol. 37, pp. 1553-1555, July 2004.
 - [Chen10] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "WLD: a robust local image descriptor," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1705-1720, Sept. 2010.

-
- [Chi10] C. Chi, H. Yoo, and M. Ben-Ezra, "Multi-spectral imaging by optimized wide band illumination," *Int. Jour. Computer Vision*, vol. 86, pp. 140-151, Jan. 2010.
 - [Cohen64] J. Cohen, "Dependency of the spectral reflectance curves of the munseel-color chips," *Psychonomic Science*, vol. 1, pp. 369-370, 1964.
 - [Denes02] L. Denes, P. Metes, and Y. Liu, "Hyperspectral face database," Tech. Rep., Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, CMU-RI-TR-02-25, 2002.
 - [Deng07] H. Deng, W. Zhang, E. Mortensen, T. Dietterich, and L. Shapiro, "Principal curvature-based region detector for object recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
 - [Di10] W. Di, L. Zhang, D. Zhang, and Q. Pan, "Studies on hyperspectral face recognition in visible spectrum with feature band selection," *IEEE Trans. Systems Man and Cybernetics A, Systems and Humans*, vol. 40, pp. 1354-1361, Nov. 2010.
 - [Djouadi08] S. Djouadi, "On the optimality of the proper orthogonal decomposition and balanced truncation," in *Proc. IEEE Conf. Decision and Control*, 2008, pp. 4221-4226.
 - [Djouadi12] S. Djouadi and S. Sahyoun, "On a generalization of the proper orthogonal decomposition and optimal construction of reduced order models," in *Proc. American Control Conf.*, 2012, pp. 1436-1441.
 - [Drew11] M. Drew and H. Reza, and V. Joze, "Planckian regression temperature for least spectral error and least cielab error," *J. Opt. Soc. Am. A*, vol. 28, pp. 1954-1961, 2011.
 - [Duda12] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd eds, Wiley-Interscience, 2012.
 - [Efe03] M. Efe and H. Ozbay, "Proper orthogonal decomposition for reduced order modeling: 2D heat flow," in *Proc. IEEE Conf. Control Applications*, 2003, pp. 1273-1277.
 - [Fairman87] H. Fairman, "Metameric correction using parameric decomposition," *Color Research and Application*, vol. 12, pp. 261-265, 1987.
 - [Fairman04] H. Fairman and M. Brill, "The principal components of reflectances," *Color Research and Application*, vol. 29, pp. 104-110, 2004.
 - [Farajikhah11] S. Farajikhah, F. Madanchi, and S. Amirshahi, "Nonlinear principal component analysis for compression of spectral data," in *Proc. Conf. Data Mining and Data Warehouses*, 2011

-
- [Fleming90] M. Fleming and G. Cottrell, "Categorization of faces using unsupervised feature extraction," in *Proc. Int. Joint Conf. Neural Networks*, 1990, pp. 65-70.
- [Foster06] D. Foster, K. Amano, S. Nascimento, and M. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A*, vol. 23, pp. 2359-2372, 2006.
- [Freund03] R. Freund, "Model reduction methods based on Krylov subspaces," *Acta Numerica*, vol. 12, pp. 267-319, 2003.
- [Gao02] Y. Gao and M. Leung, "Face recognition using line edge map," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 764-779, June 2002.
- [Gao05] Y. Gao and Y. Qi, "Robust visual similarity retrieval in single model face databases," *Pattern Recognition*, vol. 38, pp. 1009-1020, July 2005.
- [García.98] A. García-Beltrán, J. Nieves, J. Hernández-Andrés, and J. Romero, "Linear bases for spectral reflectance functions of acrylic paints," *Color Research and Application*, vol. 23, pp. 39-45, 1998.
- [Gillis12] N. Gillis, R. Plemmons, and Q. Zhang, "Priors in sparse recursive decompositions of hyperspectral images," in *Proc. SPIE 8390, Algorithms and Tech. Hyperspectral, Hyperspectral, and Ultraspectral Imagery XVIII*, 2012.
- [Gillis13] N. Gillis and R. Plemmons, "Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis," *Linear Algebra and its Applications*, vol. 438, pp. 3991-4007, 2013.
- [Gillis14] N. Gillis and S. Vavasis, "Fast and robust recursive algorithms for separable nonnegative matrix factorization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, pp. 698-714, Apr. 2014.
- [Gillis15] N. Gillis, D. Kuang, and H. Park, "Hierarchical clustering of hyperspectral images using rank-two nonnegative matrix factorization," *IEEE Trans. Geoscience and Remote Sensing*, vol. 53, pp. 2066-2078, Apr. 2015.
- [Gordon99] A. Gordon, *Classification*, Chapman and Hall, 1999.
- [Gu08] C. Gu and J. Roychowdhury, "Model reduction via projection onto nonlinear manifolds, with applications to analog circuits and biochemical systems," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, 2008, pp. 85-92.
- [Harde.01] J. Hardeberg, *Acquisition and reproduction of color images: colorimetric and hyperspectral approaches*: Dissertation. com, 2001.

-
- [Hartley00] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 1st ed. Cambridge Univ. Press, 2000.
 - [Hassen13] R. Hassen, Z. Wang, and M. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, pp. 2798-2810, July 2013.
 - [Hård81] A. Hård and L. Sivik, "NCS—natural color system: a swedish standard for color notation," *Color Research and Application*, vol. 6, pp. 129-138, 1981.
 - [He05] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, pp. 328-340, Mar. 2005.
 - [Hernande.01] J. Hernandez-Andres, J. Romero, and R. Lee, "Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain," *J. Opt. Soc. Am. A*, vol. 18, pp. 412-420, Feb. 2001.
 - [Holmes98] P. Holmes, J. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, 1st ed. Cambridge Univ. Press, 1998
 - [Hu12] Y. Hu, A. Mian, and R. Owens, "Face Recognition Using Sparse Approximated Nearest Points between Image Sets," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, pp. 1992-2004, Oct. 2012.
 - [Hua11] G. Hua, M. Yang, E. Learned-Miller, Y. Ma, M. Turk, D. Kriegman, and T. Huang, "Introduction to the special section on real-world face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, pp. 1921-1924, Oct. 2011.
 - [Hyvärinen00] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, pp. 411-430, 2000.
 - [Imai02] F. Imai, L. Taplin, and E. Day, "Comparison of the accuracy of various transformations from multi-band images to reflectance spectra As part of end-to-end color reproduction from scene to reproduction using spectral imaging," *Technical Report*, Munsell Color Science Laboratory, Rochester Institute of Technology, 2002.
 - [Jafri09] R. Jafri and H. Arabnia, "A survey of face recognition techniques," *Jour. Information Processing Systems*, vol. 5, pp. 41-68, 2009.
 - [Jain99] A. Jain, M. Murty, and P. Flynn, "Data clustering: a review," *ACM computing surveys*, vol. 31, pp. 264-323, 1999.
 - [Jayaraman10] S. Jayaraman, S. Esakkirajan, and T. Veerakumar, *Digital image processing*, 3rd ed., New Delhi, India: Tata McGraw-Hill Education, 2010.

-
- [Jiang10] R. Jiang, D. Crookes, and N. Luo, "Face recognition in global harmonic subspace," *IEEE Trans. Information Forensics and Security*, vol. 5, pp. 416-424, Sept. 2010.
 - [Kanade73] T. Kanade, Picture Processing System by Computer Complex and Recognition of Human Faces, Ph.D Dissertation, Kyoto University, 1973.
 - [Kanungo02] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 881-892, July 2002.
 - [Khan13] R. Khan, J. Weijer, D. Karatzas, and D. Muselet, "Towards hyperspectral data acquisition with hand-held devices," in *Proc. IEEE Int. Conf. Image Processing*, 2013, pp. 2053-2057.
 - [Kim95] Y. Kim and J. Ra, "Adaptive learning-method in self-organizing map for edge-preserving vector quantization," *IEEE Trans. Neural Networks*, vol. 6, pp. 278-280, Jan. 1995.
 - [Kim07] T. Kim, J. Kittler, and R. Cipolla, "Discriminative learning and recognition of image set classes using canonical correlations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1005-1018, June 2007.
 - [Kim11] J. Kim and H. Park, "Fast nonnegative matrix factorization: an active-set-like method and comparisons," *SIAM Jour. Sci. Computing*, vol. 33, pp. 3261-3281, 2011.
 - [Kim12] B. Kim, J. Han, and S. Park, "Spectral reflectivity recovery from the tristimulus values using a hybrid method," *J. Opt. Soc. Am. A*, vol. 29, pp. 2612-2621, 2012.
 - [Kirby90] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 103-108, Jan. 1990.
 - [Kittler98] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, pp. 226-239, Mar. 1998.
 - [Kong05] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition—a review," *Computer Vision and Image Understanding*, vol. 97, pp. 103-135, Jan. 2005.
 - [Kosambi43] D. Kosambi, "Statistics in function space," *J. Indian Math. Soc*, vol. 7, pp. 76-88, 1943.

-
- [Koschan11] A. Koschan, Y. Yao, H. Chang, and M. Abidi, "Hyperspectral face imaging and analysis," in *Handbook of Face Recognition*, S. Li and A. Jain, Eds., ed: Springer London, 2011, pp. 401-428.
 - [Krishna99] K. Krishna and M. Murty, "Genetic k-means algorithm," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 29, no. 3, pp. 433-439, June 1999.
 - [Lades93] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, C. Vandermalsburg, R. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Computers*, vol. 42, pp. 300-311, Mar. 1993.
 - [Laszlo12] M. Laszlo and S. Mukherjee, "A genetic algorithm using hyper-quadtrees for low-dimensional k-means clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 533-543, Apr. 2012.
 - [Le04] H. Le and H. Li, "Recognizing frontal face images using hidden markov models with one training image per person," in *Proc. Int. Conf. Pattern Recognition*, 2004, pp. 318-321.
 - [Li12] P. Li, Y. Fu, U. Mohammed, J. Elder, and S. Prince H. Le, "Probabilistic models for inference about identity," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, pp. 144-157, Jan. 2012.
 - [Lin97] S. Lin, S. Kung, and L. Lin, "Face recognition/detection by probabilistic decision-based neural network," *IEEE Trans. Neural Networks*, vol. 8, pp. 114-132, Jan. 1997.
 - [Liao13] S. Liao, A. Jain, and S. Li, "Partial face recognition: alignment-free approach," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, pp. 1193-1205, May 2013.
 - [Liu06] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 863-876, June 2006.
 - [Liu11] C. Liu, J. Yuen, A. Torralba, "SIFT flow: dense correspondence across scenes and its applications," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, pp. 978-994, May 2011.
 - [Liu12] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, pp. 1500-1512, Apr. 2012.
 - [Liu14] L. Liu, B. Liu, H. Huang, and A. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Process.: Image Communication*, vol. 29, pp. 856-863, Sept. 2014.

-
- [Loève55] M. Loève, *Probability Theory, Foundations, Random Sequences*, New York: D. Van Nostrand Company, 1955.
 - [Lowe04] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Jour. Computer Vision*, vol. 60, pp. 91-110, Nov. 2004.
 - [Lu13] C. Lu, H. Min, J. Gui, L. Zhu, and Y. Lei, "Face recognition via weighted sparse representation," *Jour. Visual Communication and Image Representation*, vol. 24, pp. 111-116, Feb. 2013.
 - [Lumley68] J. Lumley, "The structure of inhomogeneous turbulent flows," in *Atmospheric turbulence and radio propagation*, A. Yaglom and V. Tatarski, Eds., ed: Nauka, 1967, pp. 166-178.
 - [Lumley07] J. Lumley, *Stochastic Tools in Turbulence*, Dover Publications, 2007.
 - [Maloney86] L. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Opt. Soc. Am. A*, vol. 3, pp. 1673-1683, 1986.
 - [Manj.92] B. Manjunath, R. Chellappa, and C. Malsburg, "A feature based approach to face recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1992, pp. 373-378.
 - [Mansouri08] A. Mansouri, T. Sliwa, J. Hardeberg, and Y. Voisin, "An adaptive-pca algorithm for reflectance estimation from color images," in *Proc. Int. Conf. Pattern Recognition*, 2008, pp. 2941-2944.
 - [Martinez02] A. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 748-763, June 2002.
 - [Maulick02] U. Maulick and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1650-1654, Dec. 2002.
 - [Meng12] Y. Meng, Z. Lei, S. Shiu, and D. Zhang, "Monogenic binary coding: an efficient local feature extraction approach to face recognition," *IEEE Trans. Information Forensics and Security*, vol. 7, pp. 1738-1751, May 2012.
 - [Meng13A] Y. Meng, L. Gool, and Z. Lei, "Sparse variation dictionary learning for face recognition with a single training sample per person," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 689-696.
 - [Meng13B] Y. Meng, Z. Lei, Y. Jian, and D. Zhang, "Regularized robust coding for face recognition," *IEEE Trans. Image Processing*, vol. 22, pp. 1753-1766, 2013.
 - [Miao13] Z. Miao and X. Jiang, "Interest point detection using rank order log filter," *Pattern Recognition*, vol. 46, pp. 2890-2901, 2013.

-
- [Mittal12] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, pp. 4695-4708, Dec. 2012.
- [Mittelman12] R. Mittelman, N. Dobigeon, and A. Hero, "Hyperspectral image unmixing using a multiresolution sticky HDP," *IEEE Trans. Signal Processing*, vol. 60, pp. 1656-1671, Apr. 2012.
- [Moan14] S. Moan and P. Urban, "Image-difference prediction: from color to spectral," *IEEE Trans. Image Process.*, vol. 23, pp. 2058-2068, May 2014.
- [Mogh.97] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 696-710, 1997.
- [Narvekar11] N. Narvekar and L. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (cpbd)," *IEEE Trans. Image Process.*, vol. 20, pp. 2678-2683, Sept. 2011.
- [Nascimen.05] J. Nascimento and J. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, pp. 898-910, Apr. 2005.
- [Naseem10] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, pp. 2106-2112, Nov. 2010.
- [Nobach07] H. Nobach, C. Tropea, L. Cordier, J. Bonnet, J. Delville, J. Lewalle, M. Farge, K. Schneider, and R. Adrian, "Review of some fundamentals of data processing," in *Springer Handbook of Experimental Fluid Mechanics*, C. Tropea, A. Yarin, and F. Foss, Eds., ed: Springer Berlin Heidelberg, 2007, pp. 1337-1398.
- [Obukhov42] A. Obukhov, "Statistical description of continuous fields," *Trudy Geophys. In-ta AN SSSR*, no. 24, pp. 3-42, 1954.
- [Ojala02] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, July 2002.
- [Pan03] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1552-1560, Dec. 2003.
- [Pan07] Z. Pan, G. Healey, and B. Tromberg, "Hyperspectral face recognition under unknown illumination," *Optical Engineering*, vol. 46, July 2007.

-
- [Pan09] Z. Pan, G. Healey, and B. Tromberg, "Comparison of spectral-only and spectral/spatial face recognition for personal identity verification," *EURASIP Jour. Advances Signal Processing*, vol. 2009, pp. 1-6, 2009.
 - [Park07] J. Park, M. Lee, M. Grossberg, and S. Nayar, "Hyperspectral imaging using multiplexed illumination," in *Proc. IEEE Int. Conf. Computer Vision*, 2007, pp. 2049-2056.
 - [Peng12] Y. Peng, A. Ganesh, J. wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, pp. 2233-2246, Nov. 2012.
 - [Peyvandi11] S. Peyvandi and S. Amirshahi, "Generalized spectral decomposition: a theory and practice to spectral reconstruction," *J. Opt. Soc. Am. A*, vol. 28, pp. 1545-1553, 2011.
 - [Peyvandi12] S. Peyvandi, S. Amirshahi, J. Hernández-Andrés, J. Nieves, and J. Romero, "Spectral recovery of outdoor illumination by an extension of the Bayesian inverse approach to the Gaussian mixture model," *J. Opt. Soc. Am. A*, vol. 29, pp. 2181-2189, 2012.
 - [Peyvandi13] S. Peyvandi, S. Amirshahi, J. Hernández-Andrés, J. Nieves, and J. Romero, "Generalized inverse-approach model for spectral-signal recovery," *IEEE Trans. Image Process.*, vol. 22, pp. 501-510, Feb. 2013.
 - [Phillips99] P. Phillips, "Support vector machines applied to face recognition," *Advances in Neural Information Processing Systems 11*, vol. 11, pp. 803-809, 1999.
 - [Press92] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, *Numerical recipes in c: the art of scientific computing*, 2nd ed., New York, NY, USA: Cambridge Univ. Press, 1992.
 - [Prince07] S. Prince and J. Elder, "Probabilistic linear discriminant analysis for inferences about identity," in *Proc. IEEE Int. Conf. Computer Vision*, 2007, pp. 1-8.
 - [Robles.13] A. Robles-Kelly and C. Huynh, *Imaging spectroscopy for scene analysis*. London, England: Springer London, 2013.
 - [Ryer11] D. Ryer, Quest hierarchy for hyperspectral face recognition, Ph.D Dissertation, Air Force Institute of Tech., 2011.
 - [Samal92] A. Samal and P. Iyengar, "Automatic recognition and analysis of human faces and facial expressions - a survey," *Pattern Recognition*, vol. 25, pp. 65-77, Jan 1992.
 - [Samaria94] F. Samaria and S. Young, "HMM-based architecture for face identification," *Image and Vision Computing*, vol. 12, pp. 537-543, Oct. 1994.

-
- [Scholz08] M. Scholz, M. Fraunholz, and J. Selbig, "Nonlinear principal component analysis: neural network models and applications," in *Principal Manifolds for Data Visualization and Dimension Reduction*, A. Gorban, B. Kégl, D. Wunsch, A. Zinovyev Eds., ed: Springer Berlin Heidelberg, 2008, pp. 44-67.
- [SCRG] Spectral Color Research Group, University of Eastern Finland, [Online]. Available: <https://www.uef.fi/spectral/spectral-database>.
- [Shafey13] L. Shafey, C. McCool, R. Wallace, and S. Marcel, "A scalable formulation of probabilistic linear discriminant analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, pp. 1788-1794, July 2013.
- [Shen12] L. Shen and S. Zheng, "Hyperspectral face recognition using 3d gabor wavelets," in *Proc. Int. Conf. Pattern Recognition*, 2012, pp. 1574-1577.
- [Shimano06] N. Shimano, "Recovery of spectral reflectances of objects being imaged without prior knowledge," *IEEE Trans. Image Process.*, vol. 15, pp. 1848-1856, July 2006.
- [Sirovich87] L. Sirovich, "Turbulence and the dynamics of coherent structures," *Quarterly of applied mathematics*, vol. 45, pp. 561-571, 1987.
- [Skauli13] T. Skauli and J. Farrell, "A collection of hyperspectral images for imaging systems research," in *Proc. SPIE 8660, Digital Photography IX*, 2013.
- [Steger98] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, pp. 113-125, Feb. 1998.
- [Stokes98] M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta, "Multimedia systems and equipment – colour measurement and management, part 2.1: colour management in multimedia systems – default rgb colour space – srgb," *Int'l Electrotech. Commission*, IEC 61966-2-1, 1998.
- [Sun14] Z. Sun, H. Zhang, T. Tan, and J. Wang, "Iris image classification based on hierarchical visual codebook," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, pp. 1120-1133, June 2014.
- [Szeliski10] R. Szeliski, *Computer vision: algorithms and applications*, 1st ed., New York, NY, USA: Springer, 2010.
- [Tan05] X. Tan, S. Chen, Z. Zhou, F. Zhang, "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble," *IEEE Trans. Neural Networks*, vol. 16, pp. 875-886, July 2005.

-
- [Tan06] X. Tan, S. Chen, Z. Zhou, and F. Zhang, "Face recognition from a single image per person: a survey," *Pattern Recognition*, vol. 39, pp. 1725-1745, Sept. 2006.
- [Torr00] P. Torr and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, pp. 138-156, 2000.
- [Turk91] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Jour. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.
- [Uzair13] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition using 3d-dct and partial least squares," in *Proc., British Mach. Vision Conf.*, 2013.
- [Uzair15] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatio-spectral information fusion and pls regression," *IEEE Trans. Image Process.*, vol. 24, pp. 1127-1137, 2015.
- [Vu12] C. Vu, T. Phan, and D. Chandler, "S3: a spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, pp. 934-945, Mar. 2012.
- [Wandell95] B. Wandell, *Foundations of vision*. Sunderland, MA, US: Sinauer Associates, 1995.
- [Westland04] S. Westland and C. Ripamonti, *Computational Colour Science Using MATLAB*, 1st ed. New York: John Wiley and Sons, Inc., 2004.
- [Wiskott97] L. Wiskott, J. Fellous, N. Krüger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775-779, July 1997.
- [Wright09] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, pp. 210-227, Feb. 2009.
- [Wu12] Y. Wu, B. Shen, and H. Ling, "Online robust image alignment via iterative convex optimization," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 1808-1814.
- [Xiong12] X. Xiong and F. Torre, "Supervised descent method and its applications to face alignment," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 532-539.
- [Xue14] W. Xue, L. Zhang, X. Mou, and A. Bovik, "Gradient magnitude similarity deviation: a highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, pp. 684-695, Feb. 2014.

-
- [Yang02] M. Yang, D. Kriegman, N. Ahuja, "Detecting faces in images: a survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 34-58, Jan. 2002.
 - [Yang04] J. Yang, D. Zhang, A. Frangi, and J. Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, pp. 131-137, Jan. 2004.
 - [Yang07] J. Yang, D. Zhang, J. Yang, and B. Niu, "Globally maximizing, locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, pp. 650-664, Apr. 2007.
 - [Yasuma10] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, pp. 2241-2253, Sept. 2010.
 - [Yoon12] S. Yoon, J. Feng, A. Jain, "Altered fingerprints: analysis and detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, pp. 451-464, Mar. 2012.
 - [Yu13] X. Yu, J. Huang, S. Zhang, W. Yan, and D. Metaxas, "Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model," in *Proc. IEEE Int'l Conf. Computer Vision*, 2013, pp. 1944-1951.
 - [Zhang05] D. Zhang and Z. Zhou, " $(2D)^2$ PCA: two-directional two-dimensional PCA for efficient face representation and recognition," *Neurocomputing*, vol. 69, pp. 224-231, Dec. 2005.
 - [Zhang08] X. Zhang and H. Xu, "Reconstructing spectral reflectance by dividing spectral space and extending the principal components in principal component analysis," *J. Opt. Soc. Am. A*, vol. 25, pp. 371-378, 2008.
 - [Zhang09] X. Zhang and Y. Gao, "Face recognition across pose: a review," *Pattern Recognition*, vol. 42, pp. 2876-2896, Nov. 2009.
 - [Zhang13] H. Zhang, Y. Zhang, and T. Huang, "Pose-robust face recognition via sparse representation," *Pattern Recognition*, vol. 46, pp. 1511-1521, May 2013.
 - [Zhao03] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, pp. 399-459, Dec. 2003.
 - [Zheng11] Y. Zheng and A. Elmagbraby, "A brief survey on hyperspectral face recognition and multimodal score fusion," *IEEE Int. Symp. Signal Processing and Information Technology*, pp. 543-550, 2011.

- [Zhu12] X. Zhu and D. Ramanan, “Face detection, pose estimation, and landmark localization in the wild,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 2879-2886.

Vita

Woon Cho was born in Korea. He received the M.S. degree in image processing from Chung-Ang University, Korea, in 2008. He joined the Imaging, Robotics, and Intelligent Systems Laboratory (IRIS) in 2009. He completed his Doctor of Philosophy degree in the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville in 2015.

His research interests include feature extraction and description for object recognition, image dehazing and enhancement, face detection and recognition, and hyperspectral data acquisition, data calibration, and data alignment.