



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

**IMPROVEMENTS IN ILLUMINATION COMPENSATION FOR FACE RECOGNITION
UNDER NONCONTROLLED ILLUMINATION CONDITIONS**

**TESIS PARA OPTAR AL GRADO DE DOCTOR EN
INGENIERÍA ELÉCTRICA**

LUIS ERNESTO CASTILLO FAUNE

**PROFESOR GUÍA:
DR. CLAUDIO PÉREZ FLORES**

**MIEMBROS DE LA COMISIÓN:
DR. KIM L. BOYER
DR. MARCOS DÍAZ QUEZADA
DR. GONZALO RUZ HEREDIA**

**SANTIAGO DE CHILE
2017**

Mejoras en compensación de iluminación para el reconocimiento de rostros en condiciones de iluminación no controlada

Resumen: El reconocimiento de rostros depende fuertemente de las condiciones de iluminación, especialmente en ambientes no controlados donde la iluminación del rostro no es homogénea. Por esta razón, la compensación de iluminación es crucial en esta tarea. Diversos métodos de compensación de iluminación han sido desarrollados y probados para reconocimiento de rostros utilizando bases de datos internacionales. Entre los métodos con mejores resultados están la Transformada de Coseno Discreta (DCT por sus siglas en inglés), la Normalización Local (LN por sus siglas en inglés) y la Imagen de Cuociente Propio (SQI por sus siglas en inglés). La mayoría de estos métodos han sido aplicados con gran éxito en reconocimiento de rostros usando un clasificador basado en el Análisis de Componentes Principales (PCA por sus siglas en inglés). En la última década, los clasificadores basados en el Pareo Local por Filtros Gabor (LMG por sus siglas en inglés) han mostrado mejores resultados en reconocimiento de rostros con respecto a otros clasificadores. En todos estos casos, la compensación de iluminación mejora las tasas de reconocimiento de rostros en imágenes mal iluminadas, pero afectan negativamente en algunas imágenes bien iluminadas.

El objetivo de esta tesis es desarrollar mejoras a los métodos de compensación de iluminación existentes para obtener mejores tasas de reconocimiento de rostros bajo condiciones de iluminación no homogénea. Usando algoritmos genéticos, se seleccionaron parámetros en el método SQI para mejorar el reconocimiento de rostros. Los parámetros optimizados por el algoritmo genético fueron: la fracción del valor medio dentro de la ventana en que se aplica SQI; la selección de la función de eliminación de ruido entre la arcotangente, la sigmoide, la tangente hiperbólica y la función mínimo; y los pesos para cada escala de filtrado, cuyo valor está entre 0 y 1. Los resultados obtenidos al utilizar este método fueron comparados con aquellos obtenidos sin compensación de iluminación y con otras versiones de SQI. Se utilizaron cuatro bases de datos internacionales: Yale B, CMU PIE, AR y Color FERET en escala de grises. Las primeras tres bases de datos contienen imágenes de rostros con cambios significativos en las condiciones de iluminación, mientras que la cuarta contiene imágenes de rostros con cambios pequeños en las condiciones de iluminación. El método creado tiene un mejor desempeño que SQI en imágenes con iluminación no homogénea.

Del mismo modo, se utilizaron algoritmos genéticos para optimizar los parámetros en versiones modificadas de los métodos LN y SQI, aplicados en cascada, con el fin de mejorar el reconocimiento de rostros. Al utilizar estos métodos modificados en cascada, se obtuvieron resultados del 100% de reconocimiento en rostros con iluminación no homogénea y resultados significativamente mejores que otros métodos en rostros con iluminación homogénea. También fueron optimizados los parámetros de los métodos DCT, LN y SQI, mediante algoritmos genéticos utilizando el clasificador LMG. Los resultados fueron probados en la base de datos FERET. Estos resultados muestran que el reconocimiento de rostros puede ser mejorado sustancialmente con versiones modificadas de los métodos de compensación de iluminación mencionados. Los mejores resultados fueron obtenidos con la versión genéticamente optimizada del método LN, alcanzando una reducción del 31% del número total de errores en la base de datos FERET.

Finalmente, fue propuesta una extensión del método LN utilizando estadísticos basados en Kolmogorov y Nagumo. Mediante lo anterior, se generó un marco teórico más generalizado para la normalización de iluminación y se demostró que la versión original de LN es sólo un caso particular del marco teórico generado. El método propuesto fue evaluado utilizando dos clasificadores diferentes, PCA y LMG, sobre las bases de datos estándares Extended Yale B, AR y Gray FERET. El método propuesto alcanzó mejores resultados que los publicados previamente para otras variaciones de LN en las mismas bases de datos.

Palabras clave: Reconocimiento de rostros, compensación de iluminación, transformada de coseno discreta, imagen de cuociente propio, clasificadores de rostros, normalización local, optimización genética, compensación de iluminación en cascada, media cuasi-aritmética, desviación estándar cuasi-aritmética.

Improvements in illumination compensation for face recognition under noncontrolled illumination conditions

Abstract:

Face recognition depends strongly on illumination conditions, especially in non-controlled scenarios where face illumination is not homogeneous. For this reason, illumination compensation is crucial in this task. Several methods for illumination compensation have been developed and tested on the face recognition task using international available face databases. Among the methods with best results are the Discrete Cosine Transform (DCT), Local Normalization (LN) and Self-Quotient Image (SQI). Most of these methods have been applied with great success in face recognition using a principal component classifier (PCA). In the last decade, Local Matching Gabor (LMG) classifiers have shown great success in face classification relative to other classifiers. In all cases, the illumination compensation methods improve the face recognition rates in unevenly illuminated images, but affect negatively in some well illuminated images.

The aim of this thesis is to propose improvements to the current illumination compensation methods to obtain improved face recognition rates under different illumination conditions. Using genetic algorithms (GAs), parameters of the SQI method were selected to improve face recognition. The parameters optimized by the GA were: the fraction of the mean value within the region for the SQI, selection of Arctangent, Sigmoid, Hyperbolic Tangent or Minimum functions to eliminate noise, and the weight values of each filter are selected within a range between 0 and 1. The results obtained after using the proposed method were compared to those with no illumination compensation and to those previously published for SQI method. Four internationally available face databases were used: Yale B, CMU PIE, AR, Color FERET (grayscale), where the first three contain face images with significant changes in illumination conditions, and the fourth one contains face images with small changes in illumination conditions. The proposed method performed better than SQI in images with non-homogeneous illumination.

In the same way, GAs were used to optimize parameters of the modified LN and SQI methods in cascade for illumination compensation to improve face recognition. The main novelty of this proposed method is that it applies to non-homogeneous as well as homogeneous illumination conditions. The results were compared to those of the best illumination compensation methods published in the literature, obtaining 100% recognition on faces with non-homogeneous illumination and significantly better results than other methods with homogeneous illumination. Also, the DCT, LN, and SQI illumination compensation methods were optimized using GAs to be used with the LMG face classifier. Results were tested on the FERET international face database. Results show that face recognition can be significantly improved by modified versions of the current illumination compensation methods. The best results are obtained with the optimized LN method which yields a 31% reduction in the total number of errors in the FERET database.

Finally, an extension of the LN method using Kolmogorov-Nagumo-based statistics was proposed to improve face recognition. The proposed method is a more general framework for illumination normalization and it was showed that LN is a particular case of this framework. The proposed method was assessed using two different classifiers, PCA and LMG, on the standard face databases Extended Yale B, AR and Gray FERET. The proposed method reached significantly better results than those previously published for other versions of LN on the same databases.

Keywords: Face recognition, illumination compensation, discrete cosine transform, DCT, self quotient image, SQI, face classifiers, local normalization, genetic optimization, cascaded illumination compensation, quasi-arithmetic mean, quasi-arithmetic standard deviation.



To Sofía Amanda, my beloved niece and goddaughter.

Acknowledgments

Firstly, infinite thanks to my family, especially my parents, for their unconditional love and confidence.

I would like to thank my advisor, Prof. Claudio Pérez, for his willingness, dedication, support, and for sharing his knowledge with me, beyond the academic field. I would also like to thank Prof. Kim Boyer for his warm welcome and guidance during my PhD internship at the Rensselaer Polytechnic Institute. Further, I would like to acknowledge the thesis committee for providing very useful recommendations to improve this thesis manuscript.

Thanks to my friends for being with me in difficult times and for encouraging me to finish this PhD thesis, particularly Carlos and Viole (it was difficult to choose an order, so I did it alphabetically).

Thanks to the people from the different laboratories of the Universidad de Chile for making these years an unforgettable experience: Leo×3, Pablo, Chubo, Carlonchín, Schulz, JP, Jorge, Juan, Alonso, Felipe, Jacob, Sebastián, Coté, Cristian, Gabriel.

I wish to thank the people I met in Troy, especially Pablo, Cintia, Nithya and Dami, for making me feel like at home.

Finally, I would like to thank the CONICYT National Doctoral Scholarship program, Becas Chile Scholarship for Doctoral Internship, and FONDECYT project No. 1120613 for their financial support to my thesis.

Contents

List of Figures	x
List of Tables	xii
Abbreviations And Acronyms	xiii
1 Introduction	1
1.1 General Motivation	1
1.2 Definition of the Problem	2
1.3 Objectives	3
1.3.1 General objective	3
1.3.2 Specific objectives	3
1.4 Hypotheses	3
1.5 Proposed solution	4
2 State of the Art	5
2.1 Surface models	6
2.1.1 Retinex model	6
2.1.2 Lambertian model	6
2.2 Illumination compensation methods	7
2.2.1 Discrete Cosine Transform	7
2.2.2 Local Normalization	10
2.2.3 Self-quotient Image	10
2.3 Face image databases	13
2.3.1 The AR Face Database	13
2.3.2 The CMU Pose, Illumination and Expression (PIE) Database	14
2.3.3 The Facial Recognition Technology (FERET) Database	15
2.3.4 The Yale Database	15
2.3.5 The Yale Face Database B (Yale B)	15
2.4 Face recognition	16
2.4.1 Principal Components Analysis (PCA)	16
2.4.2 Local Matching Gabor (LMG)	17
2.4.3 Gallery and testing sets	19
2.4.3.1 The AR Face Database	19
2.4.3.2 The CMU Pose, Illumination and Expression (PIE) Database	19
2.4.3.3 The Facial Recognition Technology (FERET) Database	20
2.4.3.4 The Yale Database	21

2.4.3.5	The Yale Face Database B (Yale B)	21
3	Methodology	22
3.1	Genetic optimization of illumination compensation methods	22
3.1.1	Genetic optimization of the DCT method (GO-DCT)	24
3.1.2	Genetic optimization of the LN method (GO-LN)	25
3.1.3	Genetic optimization of the SQI method (GO-SQI)	27
3.1.4	Genetic optimization of illumination compensation method applied in cascades	28
3.2	Kolmogorov-Nagumo LN	28
4	Experiments and Results	30
4.1	Genetic optimization	30
4.1.1	PCA	30
4.1.1.1	GO-SQI	30
4.1.1.2	GO-LN+SQI cascade	31
4.1.2	LMG	33
4.1.2.1	GO-DCT	35
4.1.2.2	GO-LN	35
4.1.2.3	GO-SQI	36
4.1.2.4	GO-LN+SQI	39
4.2	Kolmogorov-Nagumo LN	40
4.2.1	PCA	40
4.2.2	LMG	42
5	Discussion and Conclusions	47
5.1	Discussion	47
5.1.1	Genetic optimization of illumination compensation methods	47
5.1.2	Kolmogorov-Nagumo LN	48
5.1.3	Relation between LN and the Laplace operator	48
5.1.4	Relation between LN, its improved versions, and other illumination compensation methods	49
5.1.5	Spectral analysis based on IDCT of the Laplacian-based methods	51
5.2	Conclusions	51
5.3	Publications	53
5.3.1	Journal papers	53
5.3.2	Conference papers	53
	Bibliography	54

A Relation between the Laplacian operator and illumination compensation methods	60
---	----

List of Figures

2.1	Lambertian surface reflectance. The incident light diffuses in all directions.	7
2.2	Diffuse reflectance spectra for different ages in the skin of (a) elbows, (b) knees, and (c) wrists [Calin 2010].	7
2.3	DCT orthogonal basis $-1 \leq \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \leq 1$ for an 8×8 pixels image. The darker the pixel, the lower the value.	8
2.4	Elimination of DCT coefficient within an isosceles triangle of side D_{dis} [Chen 2006]. These coefficients represent the lower frequencies of the transformed image.	9
2.5	Effect of the illumination compensation using LN. The upper row shows the original images from the Yale B database. The lower row shows the compensated images using LN method [Xie 2006].	10
2.6	Example of anisotropic filtering in SQI method. $I_{\Omega}(i, j)$ is a block of the original image to be filtered. The central pixel of this block is $I_{\Omega}(0, 0) = 50$. \oplus represents point-wise multiplication.	12
2.7	In pairs, the original images from the Yale B database and their respective SQIs [Wang 2004].	13
2.8	Samples from AR face database. The first image shows a woman with neutral expression and homogeneous illumination. A man with sunglasses (occlusion) and heterogeneous illumination can be observed in the second image. In the third image, a woman with a non neutral expression is shown.	14
2.9	Sample images from CMU PIE database of three individuals with different illumination.	14
2.10	Samples from FERET face database. This face database exhibits low variation in illumination, but high differences in expression.	15
2.11	Samples from Yale face database. Individuals are illuminated from different angles.	16
2.12	Sample images from Yale B database. The first row contains samples from the Subset 1 and the following rows for Subsets 2, 3, 4, and 5, respectively.	17
2.13	Gallery images used in [Turk 1991] to generate Eigenfaces.	18
2.14	Seven of the Eigenfaces computed using the gallery images of the Figure 2.13 [Turk 1991].	19
2.15	Real part of the Gabor filters. Each row represents a scale. Each column represents an orientation [Zou 2007a].	20
2.16	Magnitude of the Gabor filters for 5 different scales with the same parameters of those presented in Figure 2.15 [Zou 2007a].	20

3.1	Typical behavior of the fitness function (error) in the training and the validation set of a GA. The GA is overfitting to the training set whether the error increases in the validation set [Gurney 2007].	23
3.2	Bitstring representation of the weights $\delta(i, j)$ in GO-DCTa.	24
3.3	Weights in the low frequency region (isosceles triangle) of GO-DCTa. Each element $\delta(i, j)$ represents a block of size 4×4	25
3.4	Bitstring representation of the weights $\gamma(i, j)$ in GO-DCTb.	26
3.5	Weights in the low frequency region (isosceles triangle) of GO-DCTb. Each element $\gamma(i, j)$ represents a single pixel.	26
3.6	Bitstring representation of the weights $\alpha(i, j)$ in GO-LN for a window of size $(2k + 1) \times (2k + 1)$	27
3.7	Diagram of a bitstring of the LN+SQI cascade.	28
4.1	Comparison of the face image of an individual in Yale B after optimizing SQI using GAs for recognition by a PCA-based classifier: (a) Original illumination, (b) SQI, (c) GO-SQI-YPA, (d) GO-SQI-PAY, (e) GO-SQI-APY.	33
4.2	Three original images from Yale B database (first row) and the corresponding face images (second row) using the illumination compensation GO-LN+SQI-APY. These images were obtained after the optimization of the LN+SQI cascade using GAs for recognition by a PCA-based classifier.	34
4.3	Three original images from FERET database (first row) and the corresponding face images (second row) using the illumination compensation GO-LN+SQI-APY. These images were obtained under the same conditions that the Figure 4.2.	35
4.4	Parameters of LN in the GO-LN+SQI-APY cascade obtained after optimizing using GAs for recognition by a PCA-based classifier. Parameters are multiplied by 15.	36
4.5	Face images of Gray FERET after the optimization of the illumination compensation methods using GAs for recognition by a LMG classifier: (a) Original illumination, (b) GO-DCTa, (c) GO-DCTb, (d) GO-LN, (e) GO-SQI, (f) GO-LN+SQI	37
4.6	Weights selected by the GA for the DCTa after optimizing the DCT for face recognition by a LMG classifier.	37
4.7	Weights selected by the GA for the DCTb after optimizing the DCT for face recognition by a LMG classifier.	38
4.8	Best weights found by the GA for LN after optimizing for face recognition by a LMG classifier.	39
4.9	Parameters of the optimized LN in cascade chosen by GA for face recognition by a LMG classifier.	39

4.10	Images from Gray FERET (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	41
4.11	Images from AR (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	42
4.12	Images from Extended Yale B (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	43
4.13	Images from Gray FERET (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	44
4.14	Images from AR (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	45
4.15	Images from Extended Yale B (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).	46

List of Tables

4.1	Training, validation and testing sets for optimization of the DCT, LN, SQI methods and the LN+SQI cascade using GAs.	31
4.2	Error rate in GO-SQI-YPA using a PCA-based classifier	31
4.3	Error rate in GO-SQI-PAY using a PCA-based classifier	31
4.4	Error rate in GO-SQI-APY using a PCA-based classifier	32
4.5	Parameters of GO-SQI-YPA after optimizing SQI using GAs for recognition by a PCA-based classifier.	32
4.6	Parameters of GO-SQI-PAY after optimizing SQI using GAs for recognition by a PCA-based classifier.	32
4.7	Parameters of GO-SQI-APY after optimizing SQI using GAs for recognition by a PCA-based classifier.	32
4.8	Recognition rates, using a PCA-based classifier, for different face databases published in literature for 14 different methods and the results of GO-LN+SQI.	34
4.9	Parameters of SQI in the GO-LN+SQI-APY cascade obtained after optimizing using GAs for recognition by a PCA-based classifier.	35
4.10	Results in face recognition on the training (Yale database), validation (AR) and testing (Gray FERET) database by a LMG classifier after optimizing different illumination compensation methods using GAs.	36
4.11	Parameters of GO-SQI chosen by the GA after optimizing for recognition by a LMG classifier.	38
4.12	Parameters of SQI in the GO-LN+SQI cascade chosen by the GA after optimizing for recognition by a LMG classifier.	40
4.13	Best parameters of Kolmogorov-Nagumo LN for each database using a PCA-based classifier.	40
4.14	Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over Gray FERET.	40
4.15	Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over AR.	41
4.16	Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over Extended Yale B.	42
4.17	Best parameters of Kolmogorov-Nagumo LN for each database using a LMG classifier.	43
4.18	Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over Gray FERET.	44
4.19	Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over AR.	45

4.20 Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over Extended Yale B. 46

Abbreviations and Acronyms

2-D: Two-dimensional.

CMU PIE: Carnegie Mellon University Pose, Illumination and Expression Database [[Sim 2003](#)].

DCT: Discrete Cosine Transform [[Chen 2006](#)].

FERET: The Facial Recognition Technology Database.

GA: Genetic algorithm.

IDCT: Inverse Discrete Cosine Transform [[Chen 2006](#)].

LBP: Local Binary Pattern.

LMG: Local Matching Gabor [[Zou 2007a](#)]

LN: Local Normalization [[Xie 2006](#)].

PCA: Principal Components Analysis.

SQI: Self-Quotient Image [[Wang 2004](#)].

TKEO: Teager-Kaiser Energy Operator [[Teager 1990](#), [Kaiser 1990](#)].

CHAPTER 1
Introduction

Contents

1.1	General Motivation	1
1.2	Definition of the Problem	2
1.3	Objectives	3
1.3.1	General objective	3
1.3.2	Specific objectives	3
1.4	Hypotheses	3
1.5	Proposed solution	4

1.1 General Motivation

Illumination compensation has proven to be crucial in many machine vision applications including face recognition. Face recognition is important in many applications such as security systems, human-machine interfaces, gesture based computer interfaces, automatic driver monitoring, biometric identification, and video search [Perez 2003, Perez 2005, Wang 2009, Osadchy 2010, Perez 2011]. Face detection is an essential component in human-computer interaction, video surveillance, face tracking and face recognition [Perez 2007, Phimoltares 2007, Lei 2009, Yun 2009, Kasturi 2009, Perez 2009a]. Face images are significantly changed by lighting conditions, which may cause performance degradation both in face detection and in recognition in any machine vision applications [Choi 2007, Perez 2010a, Perez 2011]. In many real situations illumination cannot be controlled and varies according to the daylight conditions, e.g. in airports, malls, train stations, etc. Therefore, it is highly desirable to develop new methods for face recognition that are robust to changes in illumination conditions.

Several methods have been proposed to compensate for variations in illumination conditions and applied to face recognition using internationally available face databases. In [Zou 2007b, Yan 2009, Makwana 2010], the illumination compensation methods have been divided into active and passive methods. Active methods consist of the use of additional devices for obtaining illumination independent images. Passive methods consist of applying

image-processing techniques to improve face recognition rates. Some of these methods attempt to eliminate the image low frequency components because those coefficients show the largest variance for images with changing illumination. In this thesis, passive methods are studied.

In [Chen 2006], a 2-D discrete cosine transform (DCT) was used for illumination compensation truncating low frequency DCT coefficients. Then, the inverse discrete cosine transform (IDCT) was applied to obtain the image with illumination compensation. In [Wang 2004], multi-scale low-pass filtering was applied to an image to obtain different illumination components. Then, the original image was divided point-to-point by each smoothed image to eliminate the illumination components. Finally, the Self-Quotient Image (SQI) was computed as the sum of the images obtained after each division. Other methods try to extract illumination invariant features. In [Xie 2006] a local normalization (LN) was applied to an image in overlapped windows. The best recognition rate was obtained when the block size was set to 7×7 . All these illumination compensation methods were developed initially for a PCA classifier.

More recently, in [Cao 2012], wavelets are used to estimate the luminance component of an image. The luminance is discarded and, therefore, reflectance is extracted. In [Baradarani 2013], images are enhanced using a logarithm-based function, and luminance components are estimated using the double-density dual-tree complex wavelet transform to be subtracted to the enhanced image. Faces are recognized using ELM. In [Han 2013], some holistic preprocessing methods such as histogram equalization, gamma intensity correction and DCT, were applied locally. Also, the partial incorporation of low frequency components to the reconstruction of images was studied. Faces were recognized using a Local Ensemble Classifier. In [Li 2013], a face compensation dictionary, with a non-necessarily orthogonal subspace, was proposed for recognition through Sparse Representation-based Classification (SRC) and Collaborative Representation-based classification with Regularized Least Square. In [Cament 2014], entropy-like weighted Gabor features and Gabor features of the image compensated through LN were combined.

1.2 Definition of the Problem

In this thesis, an image is considered to have *homogeneous illumination* if the face in the image is illuminated frontally. In the same manner, an image is considered to have *heterogeneous illumination* if the face in the image is illuminated asymmetrically, generating shadows in part of the face. Existing illumination compensation methods are not robust since they may improve face recognition in images with heterogeneous illumination, but worsen recognition rates in images with homogeneous illumination.

1.3 Objectives

In this section, general and specific objectives are described.

1.3.1 General objective

The general objective of this thesis is to improve illumination compensation methods for face recognition in images with non-controlled illumination conditions. Only frontal faces are considered.

1.3.2 Specific objectives

The specific objectives of this thesis are:

- Modify and optimize the parameters in existing methods for illumination compensation to improve face recognition rate both in controlled and non-controlled illumination conditions. This includes using illumination compensation methods in cascade.
- Propose a more general framework for LN method and assess improvements in face recognition. This implies to change the computation of some variables and to transform constants into variables. These two strategies are used to widen the search space.

1.4 Hypotheses

The hypotheses of this work are the following:

- Genetic algorithms (GAs) [Goldberg 1989, Mitchell 1996] are commonly used to optimize functions which do not have a known algebraic expression and depend on multiple variables. In this case, error must be minimized and it is considered a function of the parameters of each illumination compensation method. GAs are used to optimize parameters and variations in existing illumination compensation methods and it is expected that improved results will be obtained.
- Existing illumination compensation methods are based on different models. The application in cascades of these illumination compensation methods can be better than applying them separately, because the models can be complementary among them.
- The LN method can be improved if the mean and the standard deviation are replaced by more general statistics.

1.5 Proposed solution

The solution proposed and implemented in this thesis is the following:

- Modify the DCT, LN and SQI methods to improve their performance adjusting critical parameters that have not been modified in the past. GAs are used as an optimization tool, to search for an appropriate set of parameters for this method to improve illumination compensation parameters.
- Modify the parameters of different illumination compensation methods applied in cascade, using GAs.
- Generalize the LN method using Kolmogorov-Nagumo-based statistics.

CHAPTER 2

State of the Art

Contents

2.1	Surface models	6
2.1.1	Retinex model	6
2.1.2	Lambertian model	6
2.2	Illumination compensation methods	7
2.2.1	Discrete Cosine Transform	7
2.2.2	Local Normalization	10
2.2.3	Self-quotient Image	10
2.3	Face image databases	13
2.3.1	The AR Face Database	13
2.3.2	The CMU Pose, Illumination and Expression (PIE) Database	14
2.3.3	The Facial Recognition Technology (FERET) Database	15
2.3.4	The Yale Database	15
2.3.5	The Yale Face Database B (Yale B)	15
2.4	Face recognition	16
2.4.1	Principal Components Analysis (PCA)	16
2.4.2	Local Matching Gabor (LMG)	17
2.4.3	Gallery and testing sets	19

The aim of the illumination compensation is to separate the information of the objects in an image from the illumination information. These objects can have different textures and shapes, which determine their surfaces. Existent illumination compensation methods are focused to a certain object type, in functions of these characteristics. In this particular case, Retinex model [Jobson 1997a, Jobson 1997b] and the Lambertian model [Hallinan 1994, Bichsel 1995, Belhumeur 1997] are useful to represent the surface of frontal faces. These models are described in 2.1.1 and 2.1.2. Therefore, illumination compensation methods for face recognition are based on the Retinex model, the Lambertian model, or both of them. In 2.2, the illumination compensation methods studied in this thesis are detailed.

In 2.3, there are descriptions of face databases commonly used to compare recognition rates. The face classifiers used in this thesis are described in 2.4. In 2.4.3, gallery and testing sets for face recognition are defined.

2.1 Surface models

In this section, surface models used for frontal faces are described.

2.1.1 Retinex model

The name of this model is derived from the union of *retina* and *cortex* [Jobson 1997a, Jobson 1997b]. This model proposes that the gray intensity $I(x, y)$ is equal to the product between the reflectance $R(x, y)$ and the luminance $L(x, y)$ [Chen 2006, Wang 2004]:

$$I(x, y) = R(x, y)L(x, y) \quad (2.1)$$

Under this model, $L(x, y)$ lies on the low frequencies of the images. This fact is used for the methods based on the Retinex model to eliminate the illumination in the images.

2.1.2 Lambertian model

The Lambertian model [Hallinan 1994, Bichsel 1995, Belhumeur 1997] suggests that the gray intensity $I(x, y)$ in an image depends on the surface normal direction $\vec{n}(x, y)$, the surface albedo $\rho(x, y)$ and the illumination source direction \vec{s} [Xie 2006, Wang 2004]:

$$I(x, y) = \rho(x, y)\vec{n}(x, y)^T\vec{s} \quad (2.2)$$

The surface albedo $\rho(x, y)$ is defined as the quotient between the reflected and the incident electromagnetic radiation. For this reason, $0 \leq \rho(x, y) \leq 1$ [Weast 1989]. The magnitude of \vec{s} is the intensity of the illumination source. Figure 2.1 shows the Lambertian surface reflectance, which is uniform in all orientations.

To compensate the illumination of a face image, human skin is assumed to be a Lambertian surface [Belhumeur 1998]. Subsequent studies show that the human skin reflects diffusely much of the visible spectrum [Calin 2010]. Figure 2.2 shows diffuse reflectance spectra of human skin in several parts of the human body and for different ages.

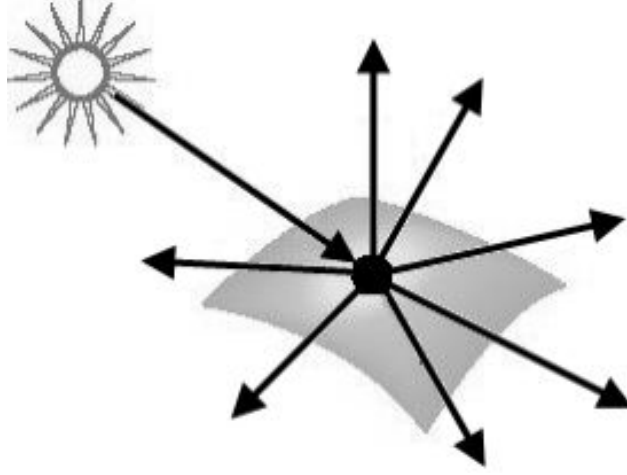


Figure 2.1: Lambertian surface reflectance. The incident light diffuses in all directions.

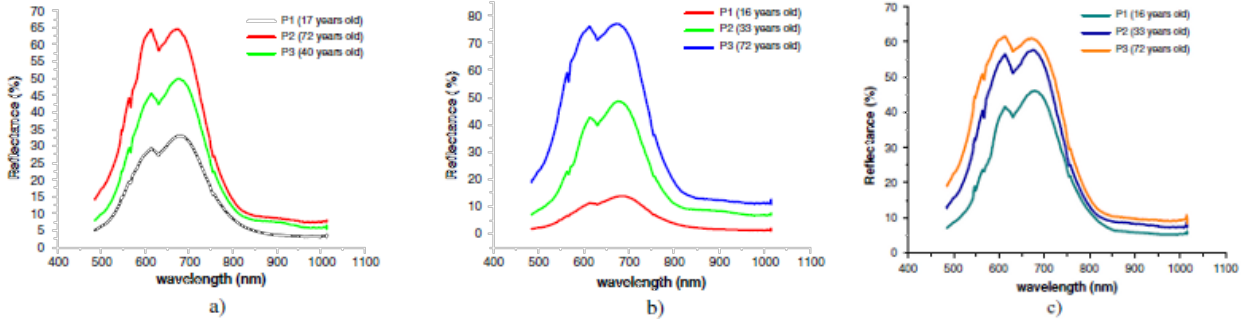


Figure 2.2: Diffuse reflectance spectra for different ages in the skin of (a) elbows, (b) knees, and (c) wrists [Calin 2010].

2.2 Illumination compensation methods

2.2.1 Discrete Cosine Transform

The DCT is used to obtain the frequency spectrum of an image, similarly to the Fourier Transform, localizing the lower frequency components in the top left corner of the spectrum. Let $f(x, y)$ be an image of $M \times N$ pixels. Its DCT $C(u, v)$ is defined by [Chen 2006]:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right), \quad (2.3)$$

where $0 \leq u \leq M-1$, $0 \leq v \leq N-1$ and:

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{M}} & \text{if } u = 0 \\ \sqrt{\frac{2}{M}} & \text{if } u \neq 0 \end{cases}, \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{if } v = 0 \\ \sqrt{\frac{2}{N}} & \text{if } v \neq 0 \end{cases} \quad (2.4)$$

Consequently, the Inverse DCT (IDCT) is defined by:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right), \quad (2.5)$$

The DCT $C(u, v)$ can be regarded as a projection of an image $f(x, y)$ onto an orthogonal basis $-1 \leq \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \leq 1$. As an example, the DCT basis for an 8×8 pixels image is illustrated in Figure 2.3. It can be observed that the further from the upper left corner, the higher the horizontal and vertical frequencies.

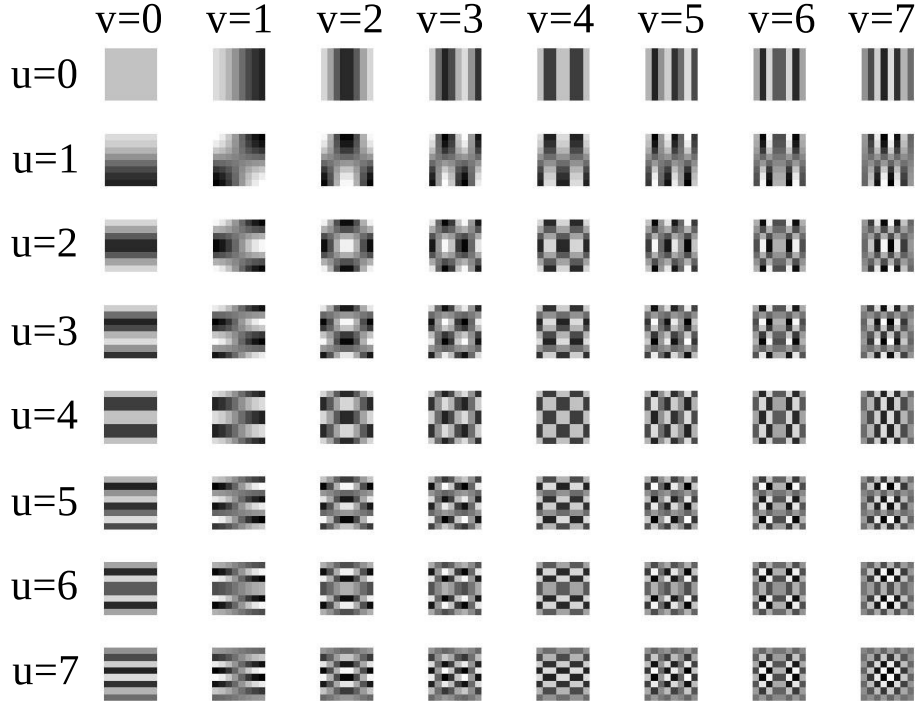


Figure 2.3: DCT orthogonal basis $-1 \leq \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \leq 1$ for an 8×8 pixels image. The darker the pixel, the lower the value.

The logarithm function is used commonly to emphasize the darker pixels in an image. Also, there is evidence that the response of the cells in the retina can be modeled as a logarithm function [Adini 1997]. If the Retinex model described in 2.1.1 is considered and if the logarithm function is applied to the Equation (2.1), then:

$$\log(I(x, y)) = \log(R(x, y)) + \log(L(x, y)) \quad (2.6)$$

It is simple to verify that the DCT is a linear transform. If the DCT $C(u, v)$ is applied over the logarithm of an image:

$$C(u, v) = C(\log(I(x, y))) = C(\log(R(x, y))) + C(\log(L(x, y))), \quad (2.7)$$

$C(\log(L(x, y)))$ corresponds to the low-frequency components of $C(u, v)$, which can be easily discarded. To compensate the illumination, the low-frequency DCT coefficients are set to zero. These components fall within an isosceles triangle of side D_{dis} in the upper left corner of the transformed image, as shown in Figure 2.4. In [Chen 2006], the value of D_{dis} is equal to 20.

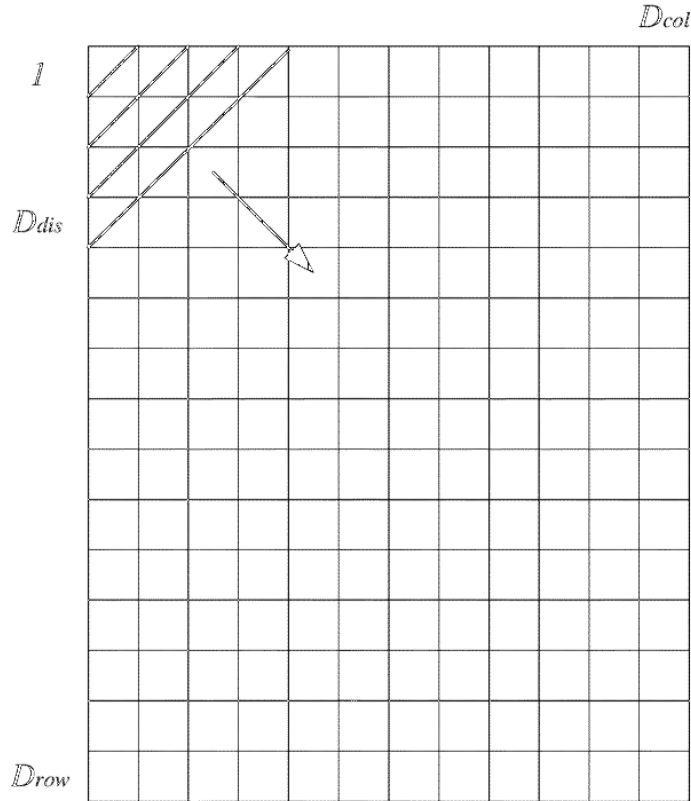


Figure 2.4: Elimination of DCT coefficient within an isosceles triangle of side D_{dis} [Chen 2006]. These coefficients represent the lower frequencies of the transformed image.

The first DCT coefficient $C(0, 0)$ is the continuous component and gather the global illumination of an image. Let μ be the average gray-scale intensity of the original image, it means, before applying the logarithm function. The continuous component $C(0, 0)$ is reconstructed as in [Chen 2006]:

$$C(0, 0) = \log(\mu)\sqrt{MN} \quad (2.8)$$

Finally, the image is reconstructed using the IDCT defined in (2.5) [Chen 2006].

2.2.2 Local Normalization

The LN method is based in the Lambertian surface model and the representation of the human face as a sequence of small and flat facets, used in graphic computing [Xie 2006]. Instead of trying to determinate the shape of each facet for each face, a normalization is applied into squared and overlapped windows. The normalized image I_{LN} is defined by:

$$I_{LN}(x, y) = \frac{I(x, y) - \mu(x, y)}{\sigma(x, y)} \quad (2.9)$$

where $\mu(x, y)$ and $\sigma(x, y)$ are, respectively, the average and the standard deviation into the window centered in (x, y) . In [Xie 2006], it is concluded that the best size of window for face recognition is 7×7 . In Figure 2.5 is shown the effect of the application of LN.

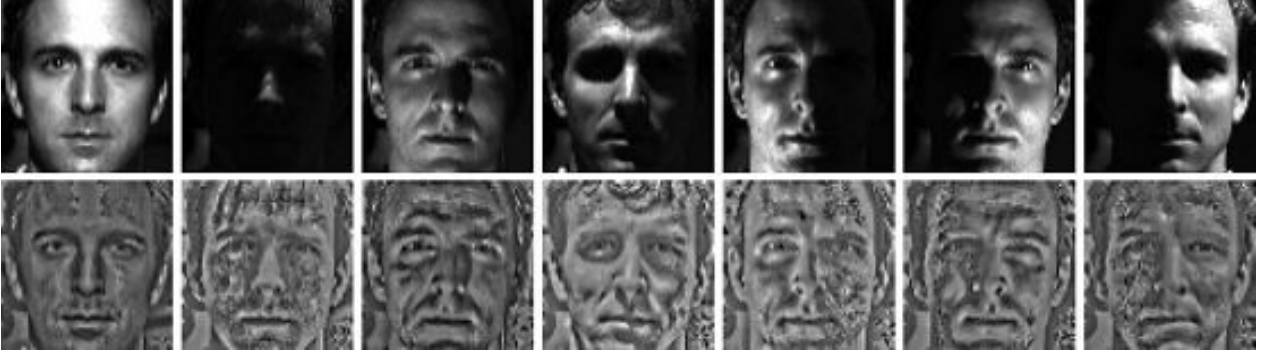


Figure 2.5: Effect of the illumination compensation using LN. The upper row shows the original images from the Yale B database. The lower row shows the compensated images using LN method [Xie 2006].

2.2.3 Self-quotient Image

The SQI method is based on the Retinex model described in (2.1) [Wang 2004]. If a low-pass filter F is applied to an image I , the luminance L is obtained as $L = F * I$, where $*$ is the convolution between matrices. Using the Retinex model described in section 2.1.1, the Equation (2.1) can be rewritten to obtain the reflectance as:

$$R = \frac{I}{L} = \frac{I}{F * I} \quad (2.10)$$

Using the above equation, the SQI Q is defined by:

$$Q = \frac{I}{\hat{I}} = \frac{I}{F * I}, \quad (2.11)$$

where \hat{I} is the smoothed (low pass filtered) version of I , F is the smoothing kernel, and the division is point-wise [Wang 2004].

A weighed Gaussian filter WG for anisotropic smoothing is applied, where W is the weight, G is the Gaussian kernel, and N is the normalization factor for which:

$$\frac{1}{N} \sum_{\Omega} WG = 1 \quad (2.12)$$

where Ω is the convolution kernel size. The convolution region is divided into two sub-regions with respect to a threshold τ calculated in [Wang 2004] as the mean of the gray intensities of the image into the convolved region:

$$\tau = \text{mean}(I_{\Omega}) \quad (2.13)$$

For the two sub-regions, W has corresponding value:

$$W_{ij} = \begin{cases} 0 & \text{if } I(i, j) < \tau \\ 1 & \text{if } I(i, j) \geq \tau \end{cases} \quad (2.14)$$

The above equation means that only the sub-region with higher gray intensities into the convolved region is smoothed.

To reduce the noise in Q caused by the division operation between I and \hat{I} , a nonlinear transformation function to change Q into D [Wang 2004]:

$$D = T(Q) \quad (2.15)$$

The SQI method is implemented by the following steps [Wang 2004]:

- **Anisotropic low-pass filtering:** Select several smoothing kernel G_1, G_2, \dots, G_n and calculate corresponding weights W_1, W_2, \dots, W_n according to image I , and then calculate the smoothed version of I by each weighed anisotropic filter $W_k G_k$. In (2.16), \oplus represents point-wise multiplication [Wang 2004]. An illustrative example for G_1 is shown in Figure 2.6.

$$\hat{I}_k = I \oplus \left(\frac{1}{N} W_k G_k \right), k = 1, 2, \dots, n. \quad (2.16)$$

- **Calculate SQIs:** Divide pixel by pixel the input image I and its smoothing versions \hat{I}_k .

$$Q_k = \frac{I}{\hat{I}_k}, k = 1, 2, \dots, n. \quad (2.17)$$

- **Noise ellimination:** Transfer SQI with a nonlinear function to filter the noise produced by the division performed above when $|\hat{I}_k|$ is small. Arctangent and Sigmoid nonlinear function are used for this effect [Wang 2004].

$$D_k = T(Q_k), k = 1, 2, \dots, n. \quad (2.18)$$

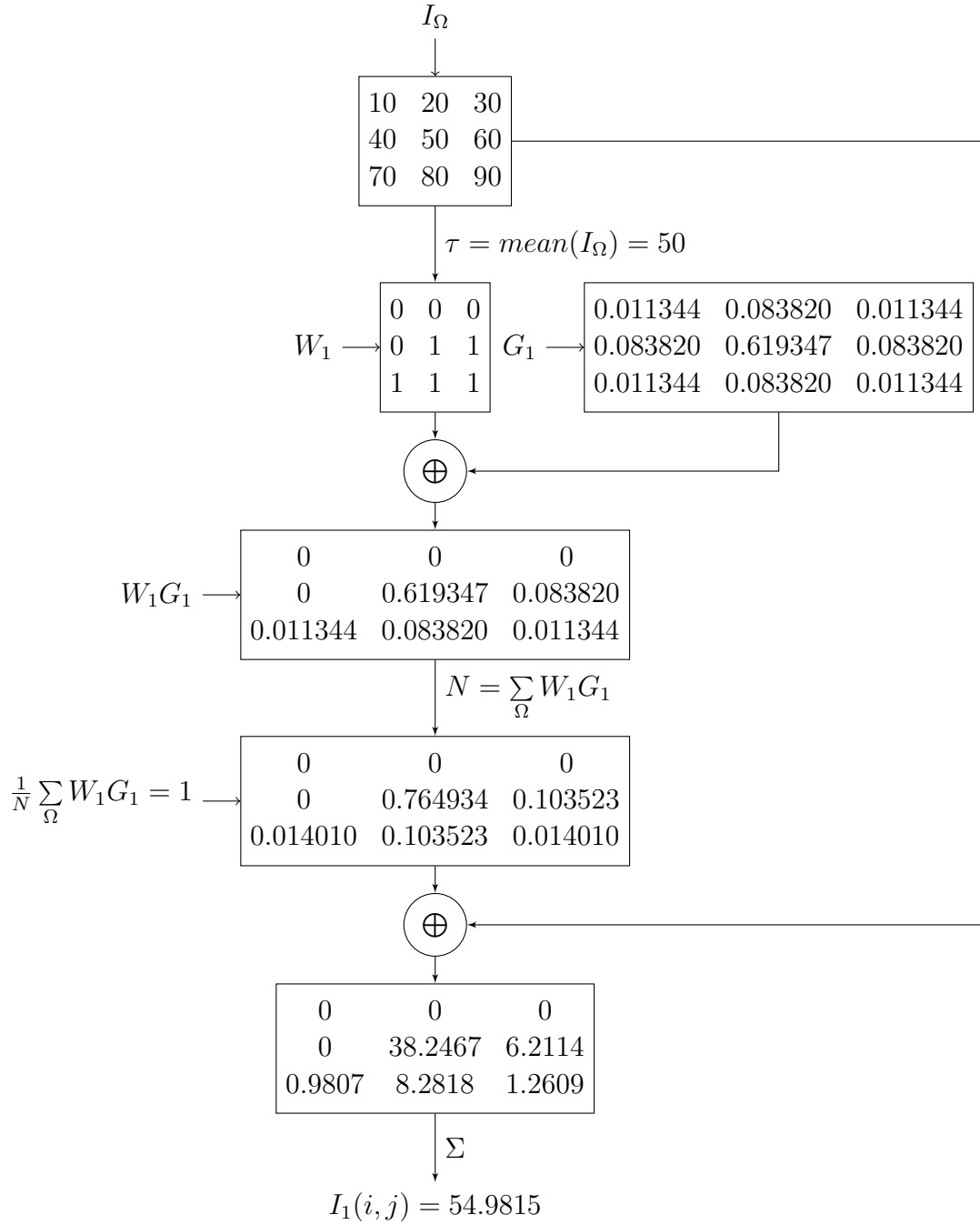


Figure 2.6: Example of anisotropic filtering in SQI method. $I_\Omega(i, j)$ is a block of the original image to be filtered. The central pixel of this block is $I_\Omega(0, 0) = 50$. \oplus represents point-wise multiplication.

- **Final SQI:** Summarize nonlinear transferred results [Wang 2004].

$$Q = \sum_{k=1}^n m_k D_k, k = 1, 2, \dots, n. \quad (2.19)$$

In [Wang 2004], $m_k = 1 \forall k$. It means that none of the SQIs is emphasized to calculate the final image. Figure 2.7 shows examples of images compensated by SQI.

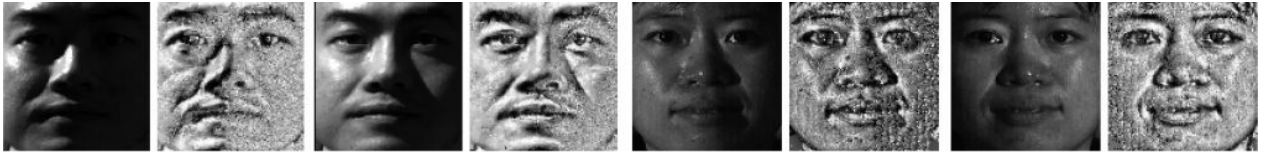


Figure 2.7: In pairs, the original images from the Yale B database and their respective SQIs [Wang 2004].

The performance of illumination compensation methods in face recognition is usually tested in face databases designed for this purpose, it means, in which lighting is the only condition that varies, and it varies greatly. In this thesis, the performance of existing and proposed methods is also tested in face databases in which not only illumination varies, but gesture, occlusion, and elapsed time between images (even years in some cases) as well. In 2.3, selected face image databases are described.

The performance measure is the face recognition rate using a certain classifier. In this thesis, two of the best current face classifiers are used. These face classifiers are described in 2.4. For each face database, some images are used to train the classifier (gallery set) and the remaining images are used to evaluate the face recognition (testing set). In 2.4.3, gallery and testing sets for each face database are described.

2.3 Face image databases

In this section, the face image databases used to improve the illumination compensation method are described. These databases are often used to evaluate the accuracy rate in face recognition. For this reason, it is possible to compare the results obtained in this thesis with those obtained in other publications. In this work, a database contains **homogeneous illumination** if the whole face can be observed in all the images. Conversely, a database contains **heterogeneous illumination** if there are images in which the face is partially or almost totally shaded.

2.3.1 The AR Face Database

This database contains images of frontal faces from 136 individuals (60 women and 76 men). These images were captured in two sessions of 13 images per individual each one. In each session, for each individual, 6 images have illumination or gesture changes, one contains neutral illumination and gesture, 3 correspond to the individual wearing sunglasses and 3 wearing a scarf that occludes the face partially. In the last two cases, the illumination source varies its position [Martinez 1998]. Figure 2.8 shows samples from the AR database.



Figure 2.8: Samples from AR face database. The first image shows a woman with neutral expression and homogeneous illumination. A man with sunglasses (occlusion) and heterogeneous illumination can be observed in the second image. In the third image, a woman with a non neutral expression is shown.

2.3.2 The CMU Pose, Illumination and Expression (PIE) Database

In this database, there are images from 68 subjects with pose, illumination and expression variations [Sim 2003]. In this thesis, only frontal face images under different lighting conditions were selected. These images were captured indoor, with and without room lights. Figure 2.9 contains samples from CMU PIE database.



Figure 2.9: Sample images from CMU PIE database of three individuals with different illumination.

2.3.3 The Facial Recognition Technology (FERET) Database

This database is organized by default in five subsets: one of them (Fa) is the gallery set and the other 4 (Fb, Fc, Dup1 and Dup2) are the testing set. Fa contains 1196 images from different individuals (one per each individual). Fb contains 1195 images with different gestures. Fc contains 194 images with different illumination intensities. Dup1 contains 722 images captured between 0 and 34 months after Fa. Dup2 contains 234 images captured at least 18 months after Fa [Phillips 1998]. Furthermore, Dup2 is a subset of Dup1. Figure 2.10 shows samples from FERET face database.



Figure 2.10: Samples from FERET face database. This face database exhibits low variation in illumination, but high differences in expression.

2.3.4 The Yale Database

This database contains 150 images of frontal faces from 15 subjects (10 images per subject). Two of the images per subject have extreme illumination conditions. Also, there are gesture changes. Figure 2.11 shows samples from Yale database.

2.3.5 The Yale Face Database B (Yale B)

This database contains faces from 10 subjects under 64 different illumination conditions and with 9 different poses [Georghiades 2001]. Figure 2.12 shows samples from Yale B database. Yale B database is divided into 5 subsets as a function of the angle of illumination between the source and the camera. The first subset is composed of 70 images with a lighting angle smaller than 12° . The second subset is composed of 120 images with lighting angle between angles 13° and 25° . The third subset is composed of 120 images with illumination angles



Figure 2.11: Samples from Yale face database. Individuals are illuminated from different angles.

between 26° and 50° . The fourth subset is composed of 140 images with illumination angles between 51° and 77° . The fifth subset is composed of 190 images with illumination angles between 78° and 190° . Extended Yale Face Database B, a lengthened version of Yale B with 37 subjects, was also used in some experiments [Lee 2005].

2.4 Face recognition

Illumination compensation is applied over the face images to enhance face recognition. For this reason, the face recognition rate is used to measure the performance of the different illumination compensation methods. Face classifiers used in this thesis, PCA and LMG, are described in sections 2.4.1 and 2.4.2, respectively. The composition of the gallery and testing sets for each database is described in 2.4.3.

2.4.1 Principal Components Analysis (PCA)

In [Turk 1991], a classifier based on Principal Components Analysis is presented as follows. First, Eigenfaces are computed. To do this, gallery images are represented as vectors. After this, these vectorized images are centered at the origin. A matrix is created using the centered vectors. Eigenvalues and Eigenvectors of this matrix are computed. These Eigenvectors, resized to the image dimension, correspond to the Eigenfaces. Eigenvalues show the variance around the Eigenvectors. To reduce the dimension, Eigenvalues are sorted from the highest to the lowest and a number of these Eigenvalues and their associated Eigenvectors are chosen. Finally, gallery images are projected over the Eigenfaces. In Figure 2.13, the gallery images used in [Turk 1991] are shown. Figure 2.14 shows seven of the resultant Eigenfaces.



Figure 2.12: Sample images from Yale B database. The first row contains samples from the Subset 1 and the following rows for Subsets 2, 3, 4, and 5, respectively.

To classify a test image, this image is vectorized, centered at the origin, and projected over the vectorized Eigenfaces. Then, the distance between the projected test image and the projected gallery image is computed to determine the nearest neighbor. The identity of the nearest neighbor is assigned to the test image. In order to compare the results obtained in this thesis with other publications, all images were scaled and rotated to have pupils positions in the same coordinates, and finally cropped to 105×120 pixels [Georghiades 2001].

2.4.2 Local Matching Gabor (LMG)

In [Zou 2007a], a LMG classifier for face recognition is proposed. This classifier is based on Gabor jets and Borda Count, achieving better results than Eigenfaces for face recognition. Gabor jets are sets of Gabor filters in 2-D with the same position and wavelength, but different orientation. A 2-D Gabor filter corresponds to a sinusoidal function modulated by a Gaussian function:

$$\Psi_{\mu,\nu} = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,\nu}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (2.20)$$



Figure 2.13: Gallery images used in [Turk 1991] to generate Eigenfaces.

where $k_{\mu,\nu} = k_{\nu}e^{i\varphi_{\mu}}$, $z = (x, y)$, μ and ν define the orientation and scale of the Gabor filter, $k_{\nu} = \frac{k_{max}}{f^{\nu}}$, $\varphi_{\mu} = \frac{\pi\mu}{N}$, k_{max} is the maximum frequency, and f is the spacing factor between filters in the frequency domain. For face recognition, in [Wang 2008a] are defined: $\sigma = 2\pi$, $k_{max} = \frac{\pi}{2}$, and $f = \sqrt{2}$. In the LMG classifier, 4172 Gabor jets are used. These jets are distributed in 5 grids with different scales expressed in function of the wavelength $\lambda \in \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$ (in pixels), and with 8 different orientations ($N = 8$ and $\mu = [0, \dots, 7]$). The distribution of the 4172 Gabor jets is: 2420 for $\lambda = 4$, 1015 for $\lambda = 4\sqrt{2}$, 500 for $\lambda = 8$, 165 for $\lambda = 8\sqrt{2}$, and 72 for $\lambda = 16$ [Zou 2007a].

In Figure 2.15, the real part of the Gabor filters are shown, for 5 different scales and 8 different orientations. Figure 2.16 shows the magnitude of the Gabor filters for 5 different scales [Zou 2007a].

All images were scaled and rotated to have pupils positions in the same coordinates, and cropped to 203×251 pixels [Zou 2007a], in order to compare the results obtained in this thesis with other publications.



Figure 2.14: Seven of the Eigenfaces computed using the gallery images of the Figure 2.13 [Turk 1991].

2.4.3 Gallery and testing sets

In this subsection, gallery and testing sets used for each face database in this thesis are described.

2.4.3.1 The AR Face Database

In the AR-564 subset, 282 images from 94 individuals (3 images per person) are selected. These frontal face images with no gestures are illuminated from center, right and left. Frontal illuminated images form the training set. The remaining images are chosen as the testing set.

If the whole database is used, images with frontal illumination and neutral expression are used as the training set. The remaining images form the testing set.

2.4.3.2 The CMU Pose, Illumination and Expression (PIE) Database

The images with frontal lighting are chosen as the training set. The remaining 20 face images with different illumination conditions are chosen as the testing set.

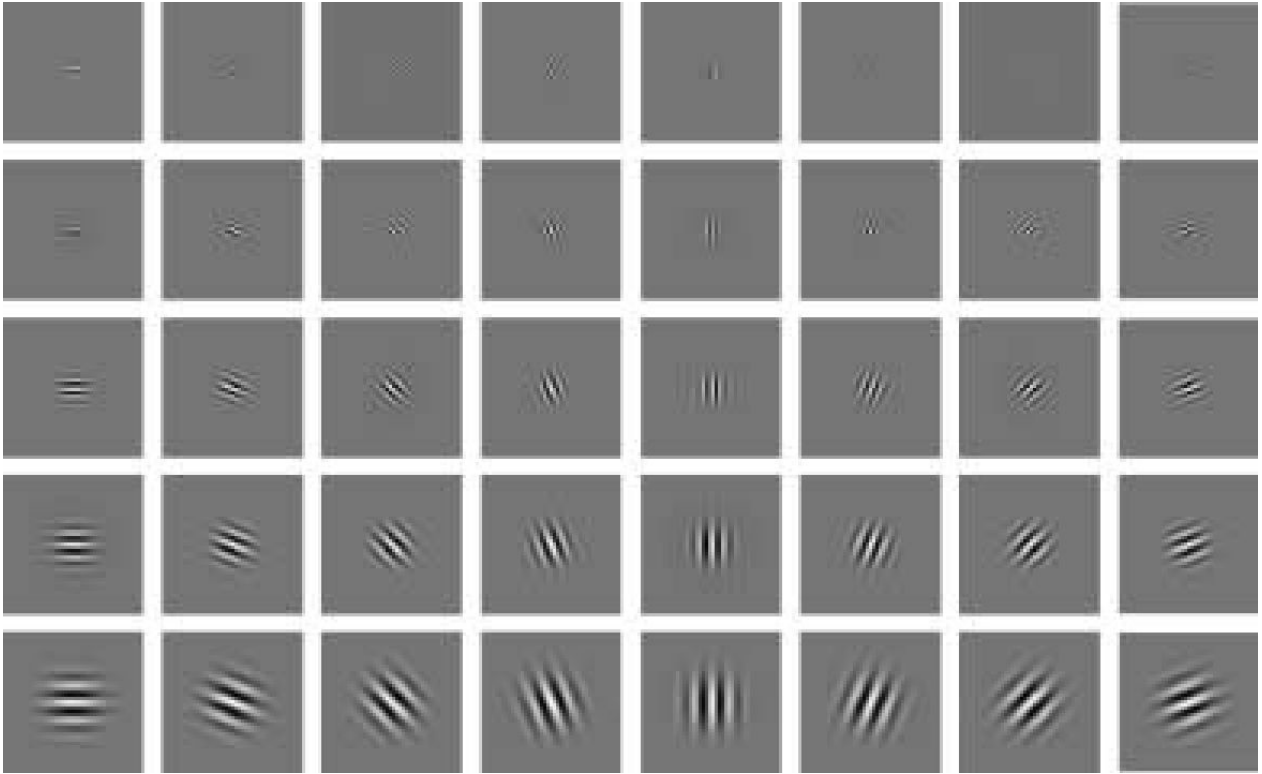


Figure 2.15: Real part of the Gabor filters. Each row represents a scale. Each column represents an orientation [Zou 2007a].

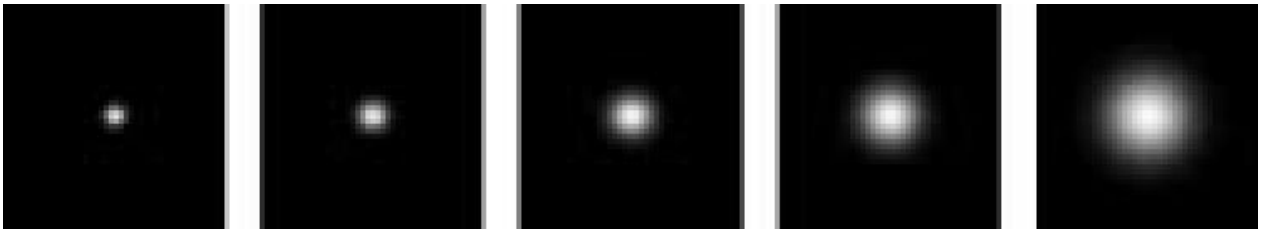


Figure 2.16: Magnitude of the Gabor filters for 5 different scales with the same parameters of those presented in Figure 2.15 [Zou 2007a].

2.4.3.3 The Facial Recognition Technology (FERET) Database

This database does not have the same number of images per individual. For this reason, in the Color FERET subset, 127 individuals with four images per person are selected. A single image per person is randomly selected for the training set and the remaining images are chosen for the testing set [Ruiz-del Solar 2008]. If the whole database is used, the default training and testing sets are considered as described in 2.3.3 [Phillips 1998].

2.4.3.4 The Yale Database

In this database, one image per person is used for training, and the remaining images are used for testing. Then, this procedure is repeated ten times choosing different images per person for training and testing [Ruiz-del Solar 2008].

2.4.3.5 The Yale Face Database B (Yale B)

In this database, the Subset 1 is chosen as the training test. The remaining subsets form the testing set [Georghiades 2001].

Methodology

Contents

3.1 Genetic optimization of illumination compensation methods	22
3.1.1 Genetic optimization of the DCT method (GO-DCT)	24
3.1.2 Genetic optimization of the LN method (GO-LN)	25
3.1.3 Genetic optimization of the SQI method (GO-SQI)	27
3.1.4 Genetic optimization of illumination compensation method applied in cascades	28
3.2 Kolmogorov-Nagumo LN	28

In this section, the contributions of this thesis are described. To enhance the illumination compensation methods, genetic optimization are used. Also, the genetic optimization of two methods applied in cascade is proposed. In 3.1, the genetic optimizations are detailed. In 3.2, a generalization of LN using Kolmogorov-Nagumo statistics is defined.

3.1 Genetic optimization of illumination compensation methods

To optimize the different illumination compensation methods using GAs, bits are interpreted as numerical values or as a choice of some qualitative parameter of the optimized illumination compensation method. To generate numerical values in $[0, 1]$, n bits are used to create 2^n levels of depth. Each level is equal to $\frac{i}{2^n-1}$, where i is the decimal representation of the n selected bits. If L is the bitstring length, the population size is chosen between L and $2L$ individuals, according to [Alander 1992]. If the bitstring is too long, the population size is limited to a maximum of 100 individuals. For greater likelihood of global maxima, two independent GAs are run. Once these GAs have finished, duplicate individuals from the last generation in both GAs are eliminated, and the remaining individuals are part of the initial population of a a third GA.

To generate new individuals, two-point crossover is used, which consists of selecting two integers m and n randomly in $[1, L]$, where L is the bitstring length. Each individual of the next generation are created from two individuals of the current generation, considering the

first m bits from the first individual, the bits between the $m+1$ -th and the n -th position from the second individual, and the bits since the $n+1$ -th from the first individual again. Also, the survival of the best two individual or *élite* of the current generation is ensured. The 80% of the population of a generation, without considering the *élite*, is created using crossover. Uniform mutation is applied. It means that every individual has the same probability of being mutated. Mutation probability is 1%. Stochastic uniform selection is used to select the parents of the new population.

Given a face database with N subsets in the testing set (see 2.4.3 for details). The fitness function F is defined by:

$$F = \sqrt{\sum_{i=1}^N err_i^2}, \quad (3.1)$$

where err_i is the error rate in the i -th subset. It is clear that the aim is to minimize the fitness function, which can be regarded as the Euclidean norm of an error vector.

For each GA, one face database is selected as the training set, a second database is chosen as the validation set (optionally), and one or more face databases are selected as the testing set. The training set is used by the GAs to evaluate the fitness of each individual in the population. The best individual is evaluated in the validation set. This procedure is repeated epoch by epoch until the fitness function reaches a minimum in the validation set [Gurney 2007]. This minimum is considered the best solution of the GA because is not overfitted to the training set. Finally, this selection is evaluated in the testing set, that has not been considered before. Figure 3.1 shows the typical behavior of the fitness function for the training and validation sets.

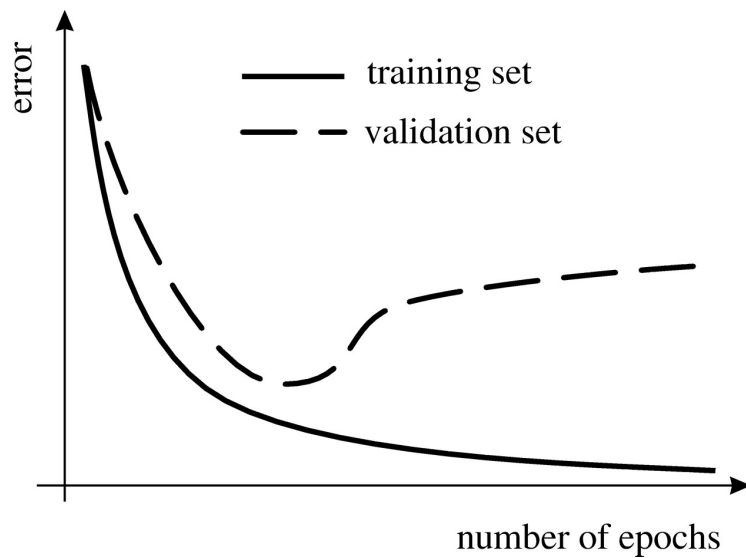


Figure 3.1: Typical behavior of the fitness function (error) in the training and the validation set of a GA. The GA is overfitting to the training set whether the error increases in the validation set [Gurney 2007].

The testing set for the GAs (entire face databases) does not have to be confused with the testing set for face recognition (subset of each face database).

In 3.1.1, 3.1.2 and 3.1.3, genetic optimizations of DCT, LN and SQI are explained, comparing the original methods described in 2.2 with their improved versions.

3.1.1 Genetic optimization of the DCT method (GO-DCT)

In this subsection, the proposed GO-DCT will be compared with the original DCT described in 2.2.1.

In the original paper of DCT [Chen 2006], the value of D_{dis} , the side of the isosceles triangle in the upper left corner of the DCT, is equal to 20. All the DCT components within this isosceles triangle were eliminated. In this thesis, the low frequency coefficients were weighted by a parameter $\gamma(u, v)$ within the isosceles triangle of side D_{dis} with values between 0 and 1. Thus, the new definition of the IDCT (see the Equation (2.5)) is:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)\gamma(u, v)C(u, v) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right). \quad (3.2)$$

Under this new definition, $\gamma(u, v) = 1$ for higher frequencies. According to 2.2.1, it is equivalent to have weighted projections of an image onto an orthogonal basis $\cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right)$.

Two different approaches were employed:

1. GO-DCTa: the isosceles triangle with side $D_{dis} = 32$ is divided into squares of 4×4 pixels with a total of 36 elements δ of 4×4 each. These elements are each one coded into 4 bits in a genetic string, with a total of 144 bits. In Equation (3.2), it implies that $\gamma(4k, 4l) = \dots = \gamma(4k + 3, 4l + 3) = \delta(k, l)$ for $k, l \in \mathbb{N}$. In Figure 3.2, a bitstring representation of the weights $\delta(i, j)$ is shown. The distribution of the weights $\delta(i, j)$ in the low frequency region is illustrated in Figure 3.3. If $\delta(k, l) = 0$ when $k + l < 5$, and the remaining elements are equal to one, a good approximation to the original DCT method is obtained.

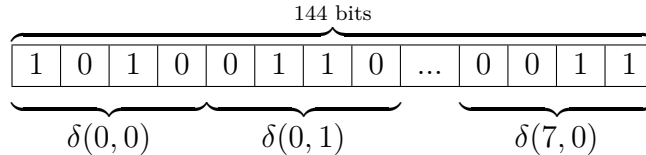


Figure 3.2: Bitstring representation of the weights $\delta(i, j)$ in GO-DCTa.

$\delta(0, 0)$	$\delta(0, 1)$	$\delta(0, 2)$	$\delta(0, 3)$	$\delta(0, 4)$	$\delta(0, 5)$	$\delta(0, 6)$	$\delta(0, 7)$
$\delta(1, 0)$	$\delta(1, 1)$	$\delta(1, 2)$	$\delta(1, 3)$	$\delta(1, 4)$	$\delta(1, 5)$	$\delta(1, 6)$	
$\delta(2, 0)$	$\delta(2, 1)$	$\delta(2, 2)$	$\delta(2, 3)$	$\delta(2, 4)$	$\delta(2, 5)$		
$\delta(3, 0)$	$\delta(3, 1)$	$\delta(3, 2)$	$\delta(3, 3)$	$\delta(3, 4)$			
$\delta(4, 0)$	$\delta(4, 1)$	$\delta(4, 2)$	$\delta(4, 3)$				
$\delta(5, 0)$	$\delta(5, 1)$	$\delta(5, 2)$					
$\delta(6, 0)$	$\delta(6, 1)$						
$\delta(7, 0)$							

Figure 3.3: Weights in the low frequency region (isosceles triangle) of GO-DCTa. Each element $\delta(i, j)$ represents a block of size 4×4 .

2. GO-DCTb: the isosceles triangle of side $D_{dis} = 20$ was divided into 210 elements γ , each one weighting a single pixel. A genetic string of 210 elements is used, where each element is coded into 4 bits with a total of 840 bits. A bitstring representation of the weights γ is illustrated in Figure 3.4. In Figure 3.5, the distribution of the weights $\gamma(i, j)$ in the low frequency region is shown. If all the weights $\gamma(i, j)$ are equal to zero in GO-DCTb, the original DCT method is reached. Therefore, the original DCT method is one of the possible solutions that can be considered by the GAs.

3.1.2 Genetic optimization of the LN method (GO-LN)

In this case, the aim is to propose variations of the statistics used in the original LN method [Xie 2006] described in section 2.2.2. The average $\mu(x, y)$ in (2.9) is generalized by a weighted average $\mu_w(x, y)$ as:

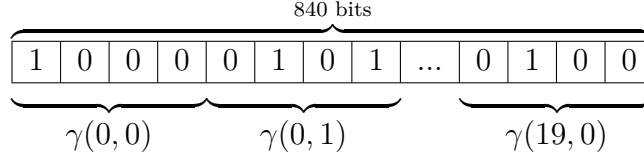


Figure 3.4: Bitstring representation of the weights $\gamma(i, j)$ in GO-DCTb.

$\gamma(0, 0)$	$\gamma(0, 1)$	$\gamma(0, 2)$...	$\gamma(0, 18)$	$\gamma(0, 19)$
$\gamma(1, 0)$	$\gamma(1, 1)$...	$\gamma(1, 17)$	$\gamma(1, 18)$	
$\gamma(2, 0)$		
...			
$\gamma(18, 0)$	$\gamma(18, 1)$				
$\gamma(19, 0)$					

Figure 3.5: Weights in the low frequency region (isosceles triangle) of GO-DCTb. Each element $\gamma(i, j)$ represents a single pixel.

$$\mu_w(x, y) = \frac{\sum_{i=-k}^k \sum_{j=-k}^k \alpha_{i,j} I(x-i, y-j)}{\sum_{i=-k}^k \sum_{j=-k}^k \alpha_{i,j}}. \quad (3.3)$$

The $\alpha_{i,j}$ are the weights of the pixels located in the position $(x-i, y-j)$ into the window centered in (x, y) . The $\alpha_{i,j}$ have values between 0 and 1, and these values are searched by the GAs. All weights $\alpha_{i,j}$ located at the same distance relative to the center of the neighborhood have the same values:

$$\alpha_{i,j} = \alpha_{-i,j} = \alpha_{i,-j} = \alpha_{-i,-j}. \quad (3.4)$$

In Figure 3.6, a bitstring representation of the weights $\alpha(i, j)$ is shown.

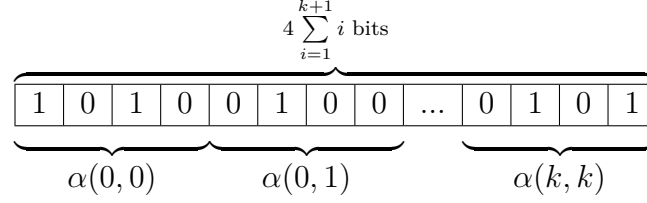


Figure 3.6: Bitstring representation of the weights $\alpha(i, j)$ in GO-LN for a window of size $(2k + 1) \times (2k + 1)$.

Consequently, the standard deviation is generalized by a deviation σ_w with respect to the weighted average μ_w as:

$$\sigma_w(x, y) = \sqrt{\frac{1}{(2k + 1)^2} \sum_{i=-k}^k \sum_{j=-k}^k (I(x - i, y - j) - \mu_w(x, y))^2}. \quad (3.5)$$

Finally, the GO-LN method [Perez 2008] for illumination compensation in human faces was performed by:

$$I_{MLN} = \frac{I(x, y) - \mu_w(x, y)}{\sigma_w(x, y)}. \quad (3.6)$$

If all the values of $\alpha_{i,j}$ are equal to 1, the original LN method is obtained. Therefore, the original LN method is a possible solution into the search space.

3.1.3 Genetic optimization of the SQI method (GO-SQI)

To optimize the SQI method [Wang 2004] described in section 2.2.3, variations in each step of this method are introduced.

- Eight Gaussian filters are used as smoothing kernels G_k , with sizes of $(2k + 1) \times (2k + 1)$ pixels. It means that the size of the smaller smoothing kernel G_1 is equal to 3×3 pixels, and the size of the greater smoothing kernel G_8 is equal to 17×17 pixels. The weights W_k presented in (2.14) were redefined by:

$$W_k(i, j) = \begin{cases} 0 & \text{if } I(i, j) < \beta_k \tau \\ 1 & \text{if } I(i, j) \geq \beta_k \tau \end{cases}, \quad (3.7)$$

where β_k is a new parameter with value between 0 and 1. This parameter changes the threshold shown in Figure 2.6. If $\beta_k = 0$, the Gaussian filter G_k is applied without changes, while the original anisotropic Gaussian filter is applied when $\beta_k = 1$.

- In (2.18), the Arctangent and Sigmoid functions are used as to eliminate noise [Chen 2006]. In GO-SQI, Hyperbolic Tangent and Minimum are proposed as well because these are also monotonically increasing and strictly concave functions in \mathbb{R}^+ . In addition, the parameter α between 0 and 1 is used to widen the search space as follows:
 - Arctangent: $T(\alpha Q_k(i, j)) = \arctan(\alpha Q_k(i, j))$
 - Hyperbolic Tangent: $T(\alpha Q_k(i, j)) = \tanh(\alpha Q_k(i, j)) = \frac{e^{\alpha Q_k(i, j)} - e^{-\alpha Q_k(i, j)}}{e^{\alpha Q_k(i, j)} + e^{-\alpha Q_k(i, j)}}$
 - Minimum: $T(\alpha Q_k(i, j)) = \min(\alpha, Q_k(i, j)) = \begin{cases} Q_k(i, j) & \text{if } Q_k(i, j) < \alpha \\ \alpha & \text{if } Q_k(i, j) \geq \alpha \end{cases}$
 - Sigmoid: $T(\alpha Q_k(i, j)) = \text{sig}(\alpha Q_k(i, j)) = \frac{1}{1 + e^{-\alpha Q_k(i, j)}}$
- Finally, the weights m_k used to summarize the smoothed images in (2.19) were assigned values between 0 and 1.

3.1.4 Genetic optimization of illumination compensation method applied in cascades

Illumination compensation methods can be also applied in cascades. It means that the output image of an illumination compensation method becomes the input image of a second method. The output of the second method is the compensated image. These method in cascade are optimized using a concatenated bitstring depending on the methods involved in the cascade, using the same strategies described in 3.1.1 (DCT), 3.1.2 (LN), and 3.1.3 (SQI). In the case of LN, a window size of 9×9 is considered, instead of the original size of 7×7 . In Figure 3.7, a diagram of a bitstring of the LN+SQI cascade is presented.

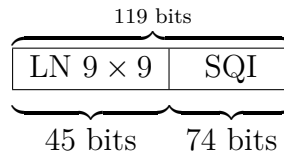


Figure 3.7: Diagram of a bitstring of the LN+SQI cascade.

3.2 Kolmogorov-Nagumo LN

Kolmogorov [Kolmogorov 1930] and Nagumo [Nagumo 1930] defined the quasi-arithmetic mean by :

$$\mu_\varphi = \varphi^{-1} \left(\frac{1}{N} \sum_{k=1}^N \varphi(x_k) \right), \quad (3.8)$$

where $\varphi(x)$ is a continuous and strictly monotone function. In [Rényi 1961], this definition of quasi-arithmetic mean was used to generalize the measure of entropy introduced by Shannon [Shannon 1949]. In this thesis, the quasi-arithmetic standard deviation is defined by:

$$\sigma_\varphi = \varphi^{-1} \left(\sqrt{\frac{1}{N} \sum_{k=1}^N (\varphi(x_k) - \varphi(\mu_\varphi))^2} \right). \quad (3.9)$$

Thus, the Kolmogorov-Nagumo LN is defined by:

$$I_{KNLN}(x, y) = \frac{I(x, y) - \mu_\varphi(x, y)}{\sigma_\varphi(x, y)}. \quad (3.10)$$

To extend the LN, $\varphi(x) = \ln(x)$ and $\varphi = x^p$ are used. The first case is equivalent to compute the geometric mean (Equation 3.11) and the geometric standard deviation (Equation 3.12).

$$\mu_{\ln(x)} = e^{\frac{1}{N} \sum_{k=1}^N \ln(x_k)} = \sqrt[N]{\prod_{k=1}^N x_k} \quad (3.11)$$

$$\sigma_{\ln(x)} = e^{\sqrt{\frac{1}{N} \sum_{k=1}^N \ln\left(\frac{x_k}{\mu_{\ln(x)}}\right)^2}} \quad (3.12)$$

The second case is equivalent to computing the power mean and the corresponding power standard deviation. The case $p = 1$, yields the original LN method [Xie 2006]. Therefore, the Kolmogorov-Nagumo LN is more general and is stated as follows:

$$\mu_{x^p} = \sqrt[p]{\frac{1}{N} \sum_{k=1}^N x_k^p} \quad (3.13)$$

$$\sigma_{x^p} = \sqrt[2p]{\frac{1}{N} \sum_{k=1}^N (x_k^p - \mu_{x^p}^p)^2} \quad (3.14)$$

The power p is the only parameter to be adjusted in this case.

Experiments and Results

Contents

4.1 Genetic optimization	30
4.1.1 PCA	30
4.1.2 LMG	33
4.2 Kolmogorov-Nagumo LN	40
4.2.1 PCA	40
4.2.2 LMG	42

In this chapter, the proposed improvements to the illumination compensation methods are tested in standard face databases, using two different classifiers: PCA and LMG.

4.1 Genetic optimization

Table 4.1 shows the training, validation, and testing sets employed in the genetic optimization of the DCT, LN, and SQI methods for two different classifiers; PCA and LMG. For example, for the GA used to optimize LN for classification using LMG, The Yale Database is the training set, the AR Face Database is the validation set, and Gray FERET is the testing set. These sets were selected to enable comparing the results to those published previously. LN+SQI was the only cascade that improved the face recognition rate. This is why the other cascades are not considered in this section.

4.1.1 PCA

4.1.1.1 GO-SQI

Tables 4.2, 4.3 and 4.4 show the results in face recognition obtained after the GO-SQI-YPA, GO-SQI-PAY and GO-SQI-APY optimization respectively, in the corresponding testing sets. Table 4.5 shows the β_k and m_k values obtained after the YPA optimization. In this case, $\min(0.91, x)$ was chosen by the GA as the nonlinear function for noise reduction. Table 4.6 shows the β_k and m_k values obtained after the PAY optimization. In this case, $\text{sig}(0.4767x)$ was chosen by the GA as the nonlinear function for noise reduction. Table 4.7 shows the

Table 4.1: Training, validation and testing sets for optimization of the DCT, LN, SQI methods and the LN+SQI cascade using GAs.

Classifier	Illumination compensation methods			
	Training set	Validation set	Testing set	
PCA [Perez 2009b]	GO-SQI-YPA	Yale B	CMU PIE	AR
	GO-SQI-PAY	CMU PIE	AR	Yale B
	GO-SQI-APY	AR	CMU PIE	Yale B
PCA [Perez 2010c]	GO-LN+SQI YPA	Yale B	CMU PIE	AR
	GO-LN+SQI APY	AR	CMU PIE	Yale B
LMG [Perez 2010b]	GO-DCTa	Yale	AR	Gray FERET
	GO-DCTb			
	GO-LN			
	GO-SQI			
	GO-LN+SQI			

β_k and m_k obtained after the APY optimization. In this case, $\min(0.98, x)$ was chosen by the GA as the nonlinear function for noise reduction. Figure 4.1 shows images of Yale B database with compensated illumination using the GO-SQI. The genetic optimization of SQI is described in the Section 3.1.3.

Table 4.2: Error rate in GO-SQI-YPA using a PCA-based classifier

Error rate %	AR-564	Color FERET	
		Mean	Standard Deviation
No compensation	40.83	22.4	0.9
SQI	2.13	25.9	1.4
GO-SQI	2.13	29.9	3.2

Table 4.3: Error rate in GO-SQI-PAY using a PCA-based classifier

Error rate %	Yale B					Color FERET	
	Subset 2	Subset 3	Subset 4	Subset 5	Fitness	Mean	Standard Deviation
No compensation	0	3.33	51.43	75.79	91.65	22.4	0.9
SQI	0	0	0	3.16	3.16	25.9	1.4
GO-SQI	0	0	0.71	0.53	0.89	28.0	1.8

4.1.1.2 GO-LN+SQI cascade

The two bottom rows of Table 4.8 show the face recognition rate obtained with this method in the four testing databases. Table 4.8 also shows results on the same databases of 14 different methods published in the scientific literature on illumination compensation. It can

Table 4.4: Error rate in GO-SQI-APY using a PCA-based classifier

Error rate %	Yale B					Color FERET	
	Subset 2	Subset 3	Subset 4	Subset 5	Fitness	Mean	Standard Deviation
No compensation	0	3.33	51.43	75.79	91.65	22.4	0.9
SQI	0	0	0	3.16	3.16	25.9	1.4
GO-SQI	0	0	0	0.53	0.53	28.9	1.9

Table 4.5: Parameters of GO-SQI-YPA after optimizing SQI using GAs for recognition by a PCA-based classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	0	2/15	2/15	1	1	2/5	0	0
m_k	0	4/15	1/5	8/15	1	4/15	0	2/15

Table 4.6: Parameters of GO-SQI-PAY after optimizing SQI using GAs for recognition by a PCA-based classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	1/5	1/15	0	4/5	2/15	1	2/15	1/5
m_k	1/15	1/15	2/15	0	0	3/5	3/5	2/3

Table 4.7: Parameters of GO-SQI-APY after optimizing SQI using GAs for recognition by a PCA-based classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	1	1/15	2/15	7/15	1	1	1	3/5
m_k	4/15	1/15	1/3	0	4/15	14/15	14/15	3/5

be observed that the best results are obtained by the cascade GO-LN+SQI-APY, reaching a perfect score (100%) recognition rate on the databases with non-homogeneous illumination and the best results on Color FERET (82.4%). It is important to note that all previous methods for illumination compensation, although they improve results in databases with non-homogeneous illumination, worsen the results with homogeneous illumination. Figures 4.2 and 4.3 show the result of the proposed transformation on three images of the Yale B face database and the FERET face database, respectively. Figure 4.4 and Table 4.9 show the parameters selected for the cascade GO-LN+SQI-APY. In this case, $\min(0.98, x)$ was chosen by the GA as the nonlinear function. The details of the genetic optimization are in the Section 3.1.4.

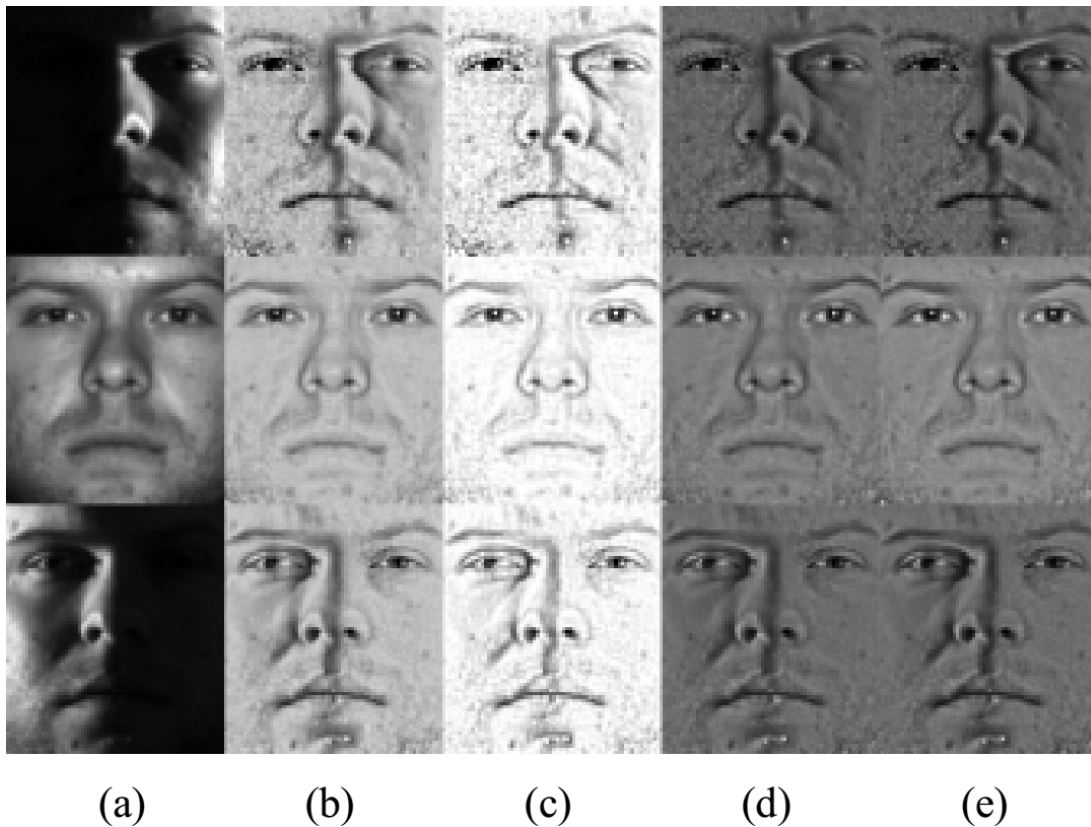


Figure 4.1: Comparison of the face image of an individual in Yale B after optimizing SQI using GAs for recognition by a PCA-based classifier: (a) Original illumination, (b) SQI, (c) GO-SQI-YPA, (d) GO-SQI-PAY, (e) GO-SQI-APY.

4.1.2 LMG

Table 4.10 shows the results in face recognition obtained after the optimization in the training set (Yale), validation set (AR) and the testing set (FERET) with the LMG classifier. In the case of Gray FERET as a testing database, it can be observed that illumination compensation with GO-LN yields one of the best results improving classification results by 5.3% for the Dup1 subset and 4.7% for the Dup2 subset. Also, small improvements were observed in Dup1 (0.9%) and in Dup2 (3%) using the GO-SQI method, and improvements of 3.8% in Dup1, and 2.1% in Dup2 when illumination compensation is performed with LN+SQI. It can also be observed in this table that the LMG classifier performs much better than the PCA-based classifier when non-homogeneous illumination is present in faces. LMG reaches better classification results than PCA for all databases without illumination compensation as well as with illumination compensation.

Figure 4.5 shows the effects of the genetically optimized method over three samples of Gray FERET.

Table 4.8: Recognition rates, using a PCA-based classifier, for different face databases published in literature for 14 different methods and the results of GO-LN+SQI.

Method	Yale B					CMU PIE [%]	AR-564 [%]	Color FERET [%]
	Subsets [%]							
	1	2	3	4	5			
No compensation	100	100	96.7	48.6	23.7	59.2	91.5	77.6 ± 0.9
9 – D QI [Wang 2008b]	100	100	100	91.7	84.3	-	-	-
MWT [Nie 2008]	-	-	0	0.7	0	-	-	-
PS+HE [Ruiz-del Solar 2008]	-	100	99.2	57.9	-	85.8	-	-
LBP [Ruiz-del Solar 2008]	-	100	99.2	92.1	-	99.1	-	-
mLBP [Ruiz-del Solar 2008]	-	100	100	98.6	-	99.8	-	-
SIP1-77-62 [Ruiz-del Solar 2008]	-	100	100	67.1	-	85.5	-	-
SIP2-100-90 [Ruiz-del Solar 2008]	-	100	100	75	-	87.8	-	-
SQI [Ruiz-del Solar 2008]	-	99.2	100	97.9	-	96.2	-	-
GIC [Du 2005]	-	-	-	-	-	80.2	-	-
Log [Du 2005]	-	-	-	-	-	78.4	-	-
HE [Du 2005]	-	-	-	-	-	76.3	-	-
HS [Du 2005]	-	-	-	-	-	79.5	-	-
SQI [Du 2005]	-	-	-	-	-	77	-	-
DCT [Perez 2008]	-	-	100	99.3	97.4	99.6	96.3	63.4 ± 1.1
LN [Perez 2008]	-	-	100	99.3	100	99.7	99.5	78.0 ± 1.9
SQI [Perez 2009b]	-	100	100	100	96.8	99.3	97.9	74.1 ± 1.4
GO-LN+SQI-YPA	-	100	100	100	100	99.9	100	81.8 ± 1.6
GO-LN+SQI-APY	-	100	100	100	100	100	100	82.4 ± 3.2

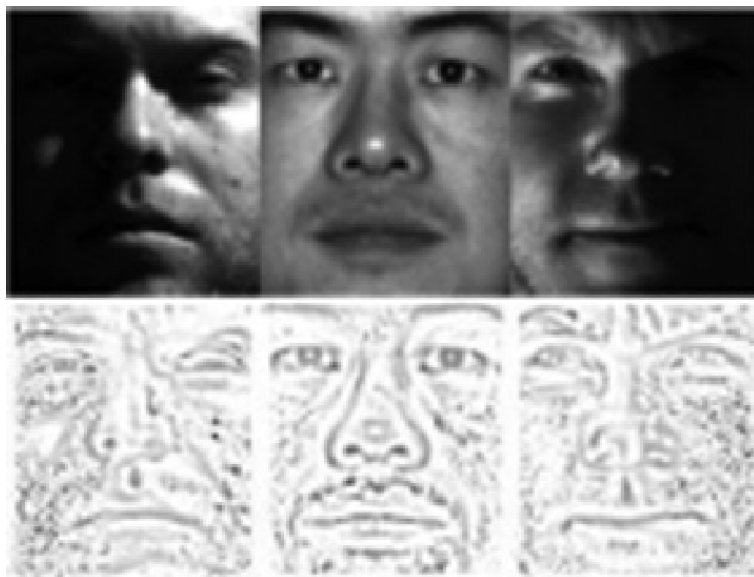


Figure 4.2: Three original images from Yale B database (first row) and the corresponding face images (second row) using the illumination compensation GO-LN+SQI-APY. These images were obtained after the optimization of the LN+SQI cascade using GAs for recognition by a PCA-based classifier.

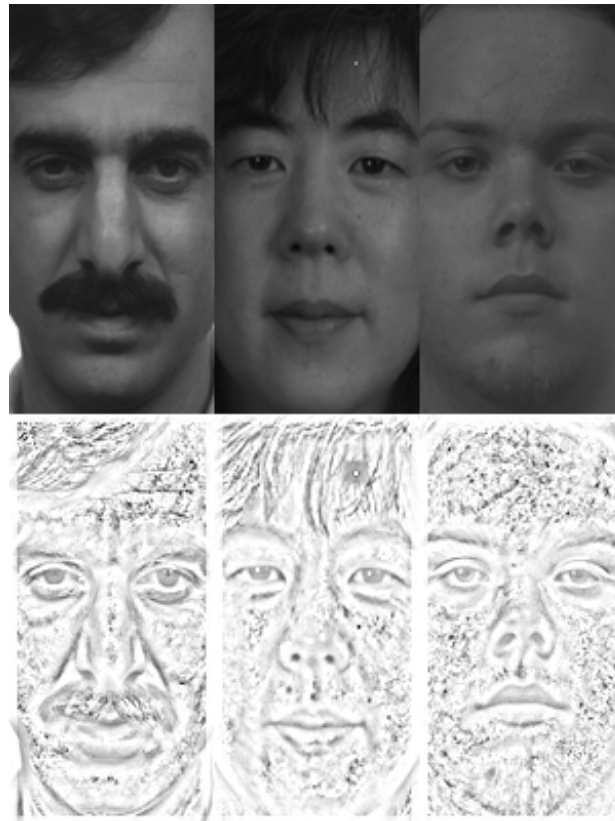


Figure 4.3: Three original images from FERET database (first row) and the corresponding face images (second row) using the illumination compensation GO-LN+SQI-APY. These images were obtained under the same conditions that the Figure 4.2.

Table 4.9: Parameters of SQI in the GO-LN+SQI-APY cascade obtained after optimizing using GAs for recognition by a PCA-based classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	1	1/15	2/15	7/15	1	1	1	3/5
m_k	4/15	1/15	1/3	0	4/15	14/15	14/15	3/5

4.1.2.1 GO-DCT

Figure 4.6 and Figure 4.7 show the best weights found by the GA for the DCTa and DCTb respectively. The genetic optimization is described in the Section 3.1.1.

4.1.2.2 GO-LN

Figure 4.8 shows the weights selected by the GA for LN. The details of the genetic optimization are in the Section 3.1.2.

$$\frac{1}{15} \times$$

10	1	11	14	15	14	11	1	10
1	6	1	6	11	6	1	6	1
11	1	12	0	15	0	12	1	11
14	6	0	4	9	4	0	6	14
15	11	15	9	7	9	15	11	15
14	6	0	4	9	4	0	6	14
11	1	12	0	15	0	12	1	11
1	6	1	6	11	6	1	6	1
10	1	11	14	15	14	11	1	10

Figure 4.4: Parameters of LN in the GO-LN+SQI-APY cascade obtained after optimizing using GAs for recognition by a PCA-based classifier. Parameters are multiplied by 15.

Table 4.10: Results in face recognition on the training (Yale database), validation (AR) and testing (Gray FERET) database by a LMG classifier after optimizing different illumination compensation methods using GAs.

Method	Database						
	Yale [%]	AR [%]	Gray FERET				
			Fb [%]	Fc [%]	Dup 1 [%]	Dup 2 [%]	No. of Errors
No	98.4	97.2	99.5	99.5	85.0	79.5	163
DCT	98.6	96.7	99.2	99.5	82.4	78.6	188
GO-DCTa	98.7	96.7	99.2	99.5	82.1	76.9	194
GO-DCTb	98.7	96.9	99.2	99.5	81.4	76.9	199
LN	99.1	98.6	99.2	99.5	82.4	78.6	188
GO-LN	99.3	98.3	99.6	100	90.3	84.2	112
SQI	98.5	98.0	99.3	99.5	85.7	81.6	156
GO-SQI	99.0	97.8	99.5	99.5	85.9	82.5	150
GO-LN+SQI	99.2	98.2	99.6	100.0	88.8	81.6	129

4.1.2.3 GO-SQI

Table 4.11 shows the β_k and m_k found by the GA for SQI. $\arctan(0.63x)$ was chosen by the GA as the nonlinear function for noise reduction. The genetic optimization is described in the Section 3.1.3.



Figure 4.5: Face images of Gray FERET after the optimization of the illumination compensation methods using GAs for recognition by a LMG classifier: (a) Original illumination, (b) GO-DCTa, (c) GO-DCTb, (d) GO-LN, (e) GO-SQI, (f) GO-LN+SQI

	1	3	15	10	8	15	3	10
	14	6	15	10	10	7	3	
	8	13	10	8	7	6		
	3	3	13	12	11			
$\frac{1}{15} \times$	5	4	11	10				
	0	0	15					
	4	5						
	10							

Figure 4.6: Weights selected by the GA for the DCTa after optimizing the DCT for face recognition by a LMG classifier.

	9	9	5	5	10	3	4	0	2	8	2	0	8	9	14	10	1	12	8	6
	10	5	1	5	13	2	9	5	14	13	7	4	15	6	3	9	0	3	13	
	3	13	4	3	6	11	0	9	7	15	12	15	10	12	7	15	3	0		
	8	4	3	9	7	1	3	6	4	15	1	4	5	10	10	1	14			
	10	2	9	11	11	5	12	6	3	2	3	9	6	7	1	5				
	15	10	4	6	12	8	11	15	3	1	0	7	9	2	13					
	10	5	0	11	8	5	11	11	2	2	15	11	3	10						
	11	1	8	9	8	6	0	10	14	5	1	15	9							
	9	5	11	11	6	15	8	6	14	15	1	5								
$\frac{1}{15} \times$	14	9	9	6	7	5	0	2	2	2	4									
	2	9	15	14	1	0	10	0	8	1										
	4	1	13	14	13	6	9	14	9											
	12	6	11	0	11	11	5	15												
	2	10	12	7	11	4	13													
	9	0	8	15	2	11														
	5	11	7	7	14															
	5	6	4	2																
	14	11	6																	
	7	14																		
	9																			

Figure 4.7: Weights selected by the GA for the DCTb after optimizing the DCT for face recognition by a LMG classifier.

Table 4.11: Parameters of GO-SQI chosen by the GA after optimizing for recognition by a LMG classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	0	$8/15$	$4/5$	$1/15$	1	$1/15$	$1/15$	$2/3$
m_k	$1/15$	0	$1/3$	1	$14/15$	$4/5$	$11/15$	$1/5$

$$\frac{1}{15} \times$$

1	2	8	14	8	2	1
2	4	10	8	10	4	2
8	10	13	15	13	10	8
14	8	15	2	15	8	14
8	10	13	15	13	10	8
2	4	10	8	10	4	2
1	2	8	14	8	2	1

Figure 4.8: Best weights found by the GA for LN after optimizing for face recognition by a LMG classifier.

4.1.2.4 GO-LN+SQI

Figure 4.9 shows the weights found by the GA for the LN stage of the LN and SQI cascade. Table 4.12 shows the β_k and m_k selected by the GA for the SQI stage of the LN and SQI cascade. $\text{sig}(0.27x)$ was chosen by the GA as the nonlinear function for noise reduction. The details of the genetic optimization are in the Section 3.1.4.

$$\frac{1}{15} \times$$

5	1	5	15	2	15	5	1	5
1	7	12	0	14	0	12	7	1
5	12	8	10	11	10	8	12	5
15	0	10	1	0	1	10	0	15
2	14	11	0	3	0	11	14	2
15	0	10	1	0	1	10	0	15
5	12	8	10	11	10	8	12	5
1	7	12	0	14	0	12	7	1
5	1	5	15	2	15	5	1	5

Figure 4.9: Parameters of the optimized LN in cascade chosen by GA for face recognition by a LMG classifier.

Table 4.12: Parameters of SQI in the GO-LN+SQI cascade chosen by the GA after optimizing for recognition by a LMG classifier.

Parameter	Size of filter							
	3×3	5×5	7×7	9×9	11×11	13×13	15×15	17×17
β_k	4/15	13/15	0	7/15	4/15	2/15	1	2/5
m_k	14/15	13/15	2/3	4/5	2/3	4/5	7/15	2/5

4.2 Kolmogorov-Nagumo LN

4.2.1 PCA

An exhaustive search is performed using odd window sizes from 3×3 to 21×21 to compute the proposed LN. Simultaneously, for $\varphi(x) = x^p$, an exhaustive search is carried out for p in $[0.5, 10]$, using a step of 0.5. Table 4.13 shows the combination of best parameters used for best results in each database.

Table 4.13: Best parameters of Kolmogorov-Nagumo LN for each database using a PCA-based classifier.

Database	Window size when $\varphi(x) = \ln(x)$	Window size when $\varphi = x^p$	p for $\varphi = x^p$
Gray FERET	7×7	11×11	10
AR	7×7	11×11	5.5
Ext. Yale B	7×7	9×9	0.5

Table 4.14 shows the results in face recognition on the Gray FERET database using the PCA-based classifier [Turk 1991]. The total number of errors was reduced from 1315 to 1132 (13.9%). In Figure 4.10, the effect of the normalization over three images from Gray FERET is shown.

Table 4.14: Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over Gray FERET.

Illumination compensation	Fb	Fc	Dup 1	Dup 2	No. of Errors
No compensation	71.3	13.4	30.6	15.4	1210
LN [Xie 2006]	53.2	45.4	33.4	27.8	1315
LN $\varphi(x) = \ln(x)$	52.0	46.4	31.2	25.2	1350
LN $\varphi(x) = x^p$	67.2	36.1	39.2	24.4	1132

In Table 4.15, the results in face recognition on the AR Face Database using the PCA-based classifier are shown. The total number of errors was reduced from 1094 to 985 (9.9%). Figure 4.11 shows the effect of the normalization over three images from AR.

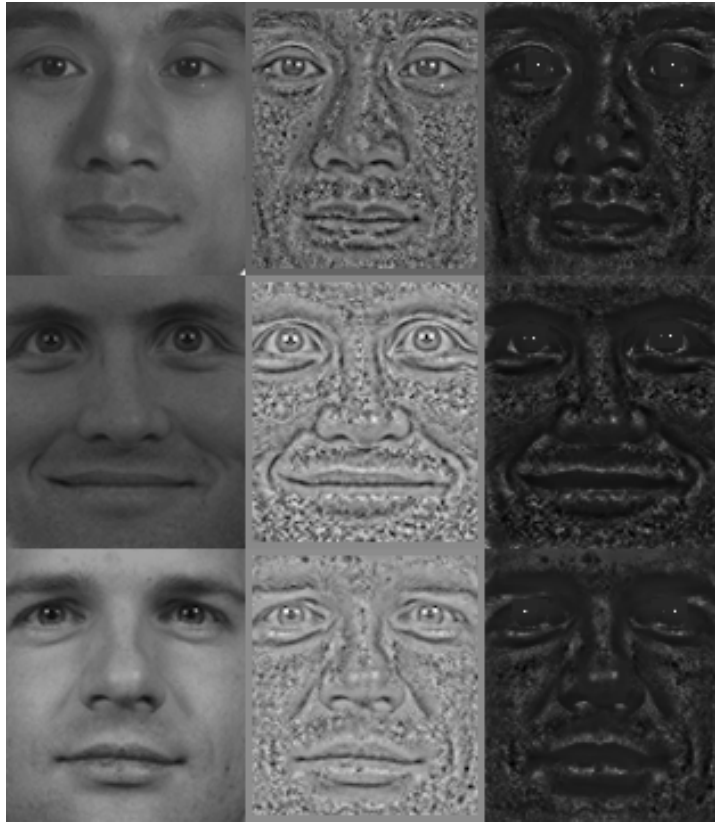


Figure 4.10: Images from Gray FERET (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Table 4.15: Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over AR.

Illumination compensation	Gest.	Light.	Sungl.	Scarf	No. of Errors
No compensation	69.8	54.4	9.7	2.0	1854
LN [Xie 2006]	56.6	85.2	27.6	74.8	1094
LN $\varphi(x) = \ln(x)$	53.9	85.0	23.8	73.8	1148
LN $\varphi(x) = x^p$	63.4	83.2	38.5	74.6	985

Table 4.16 shows the results in face recognition on the Extended Yale B Database using the PCA-based classifier. The total number of errors was reduced from 252 to 12 (95.2%). In Figure 4.12, the effect of the normalization over three images from Gray FERET is shown.



Figure 4.11: Images from AR (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Table 4.16: Recognition rates using a PCA-based classifier after applying the Kolmogorov-Nagumo LN compensation over Extended Yale B.

Illumination compensation	S2	S3	S4	S5	No. of Errors
No compensation	99.8	71.4	12.9	4.8	1248
LN [Xie 2006]	100	99.8	86.9	74.0	252
LN $\varphi(x) = \ln(x)$	100	99.8	99.4	98.9	12
LN $\varphi(x) = x^p$	100	99.8	88.8	77.0	221

4.2.2 LMG

In Table 4.17 the combination of the best parameters used for best results in each database are shown.

Table 4.18 shows the results in face recognition on the Gray FERET database using the LMG classifier [Zou 2007a]. It can be observed that the total number of errors was reduced from 111 to 97 (12.6%) using the Kolmogorov-Nagumo LN. In Figure 4.13, the effect of the

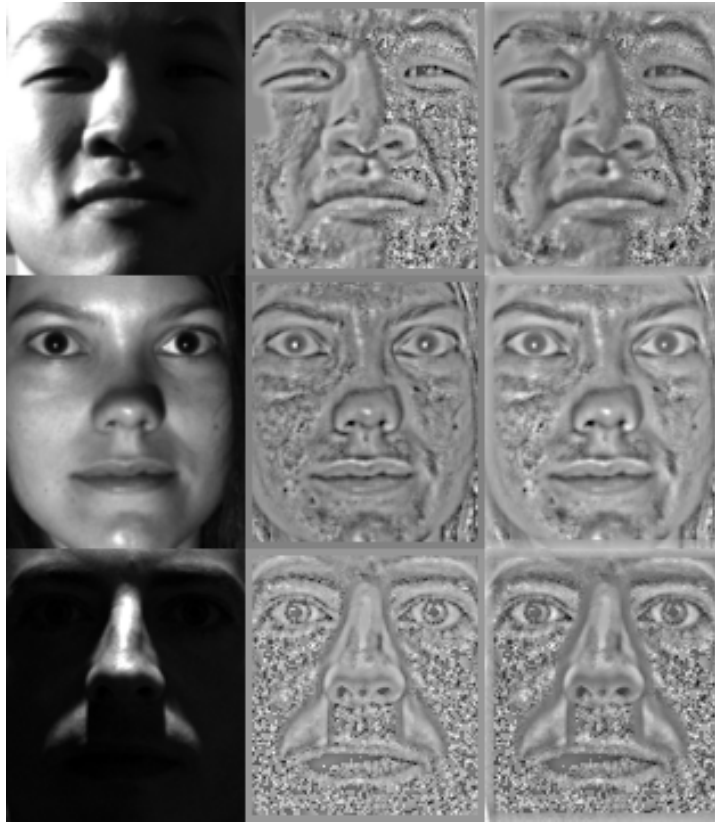


Figure 4.12: Images from Extended Yale B (first column) and corresponding normalized face images for face recognition through a PCA-based classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Table 4.17: Best parameters of Kolmogorov-Nagumo LN for each database using a LMG classifier.

Database	Window size when $\varphi(x) = \ln(x)$	Window size when $\varphi = x^p$	p for $\varphi = x^p$
Gray FERET	3×3	5×5	5
AR	3×3	7×7	5.5
Ext. Yale B	3×3	9×9	1.5

normalization over three images from Gray FERET is shown.

In Table 4.19, the results in face recognition on the AR Face database using the LMG classifier is shown. The total number of errors was reduced from 42 to 33 (21.4%). Figure 4.14 shows the effect of the normalization over three images from AR.

Table 4.20 shows the results in face recognition on the Extended Yale B database using the LMG classifier. The total number of errors was reduced from 7 to 5 (28.5%). In Figure 4.15, the effect of the proposed Kolmogorov-Nagumo LN on three images from Extended Yale B is shown.

Table 4.18: Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over Gray FERET.

Illumination compensation	Fb	Fc	Dup 1	Dup 2	No. of Errors
No compensation	99.5	99.5	85.0	79.5	163
LN [Xie 2006]	99.8	100	90.0	84.2	111
LN $\varphi(x) = \ln(x)$	99.8	100	89.9	85.5	109
LN $\varphi(x) = x^p$	99.8	100	91.3	86.3	97



Figure 4.13: Images from Gray FERET (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Table 4.19: Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over AR.

Illumination compensation	Gest.	Light.	Sungl.	Scarf	No. of Errors
No compensation	99.7	99.0	93.3	98.9	64
LN [Xie 2006]	99.7	99.6	95.2	99.6	42
LN $\varphi(x) = \ln(x)$	99.7	99.7	95.6	99.7	37
LN $\varphi(x) = x^p$	99.7	99.4	96.7	99.4	33



Figure 4.14: Images from AR (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Table 4.20: Recognition rates using a LMG classifier after applying the Kolmogorov-Nagumo LN compensation over Extended Yale B.

Illumination compensation	S2	S3	S4	S5	No. of Errors
No compensation	100	99.8	99.8	99.4	6
LN [Xie 2006]	100	99.8	99.8	99.3	7
LN $\varphi(x) = \ln(x)$	100	99.8	100	99.3	6
LN $\varphi(x) = x^p$	100	99.8	100	99.4	5

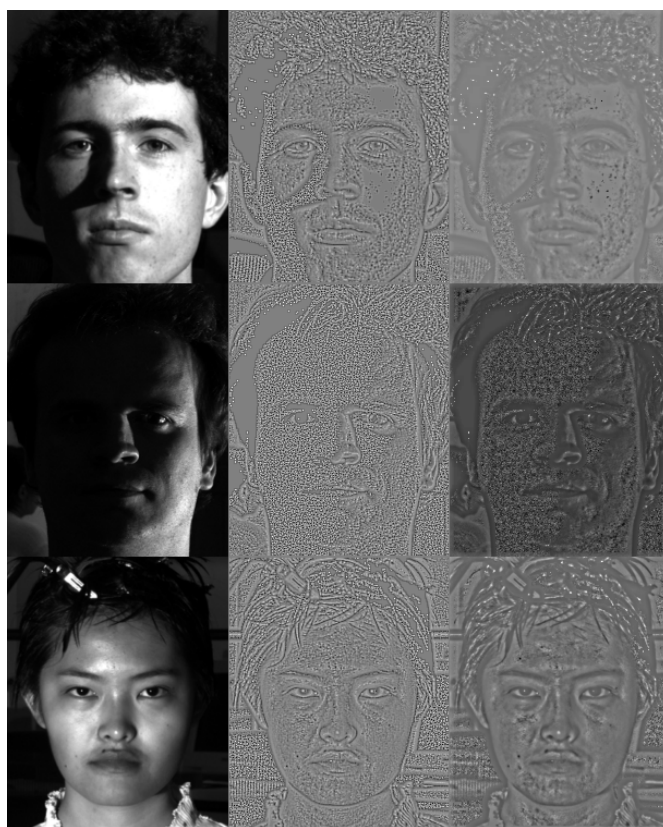


Figure 4.15: Images from Extended Yale B (first column) and corresponding normalized face images for face recognition through a LMG classifier using the Kolmogorov-Nagumo LN when $\varphi(x) = \ln(x)$ (second column) and when $\varphi(x) = x^p$ (third column).

Discussion and Conclusions

Contents

5.1 Discussion	47
5.1.1 Genetic optimization of illumination compensation methods	47
5.1.2 Kolmogorov-Nagumo LN	48
5.1.3 Relation between LN and the Laplace operator	48
5.1.4 Relation between LN, its improved versions, and other illumination compensation methods	49
5.1.5 Spectral analysis based on IDCT of the Laplacian-based methods	51
5.2 Conclusions	51
5.3 Publications	53
5.3.1 Journal papers	53
5.3.2 Conference papers	53

5.1 Discussion

5.1.1 Genetic optimization of illumination compensation methods

GAs were helpful to obtain solutions in the optimization of illumination compensation methods, considering that there are not algebraic expressions to describe the face recognition rates (or the error rates) as a function of each illumination compensation parameters. The obtained solutions were very different compared to those published in the original papers. Also, these solutions were not intuitive.

In the GO-DCT method, usually it is assumed that the illumination components are given by the low frequency coefficients of the DCT and that they must be eliminated. According to the results, it is not true that all these components must be eliminated, but weighted (see Figures 4.6 and 4.7).

In the GO-LN method, the mean in the LN definition (2.9) can be rewritten as:

$$\mu(x, y) = \frac{1}{(2k+1)^2} \sum_{i=-k}^k \sum_{j=-k}^k I(x+i, y+j)$$

Then, the LN can be redefined by:

$$I_{LN}(x, y) = \frac{\frac{(2k+1)^2}{(2k+1)^2} I(x, y) - \frac{1}{(2k+1)^2} \sum_{i=-k}^k \sum_{j=-k}^k I(x+i, y+j)}{\sigma(x, y)}$$

$$I_{LN}(x, y) = \frac{1}{(2k+1)^2} \sum_{i=-k}^k \sum_{j=-k}^k \frac{I(x, y) - I(x+i, y+j)}{\sigma(x, y)} \quad (5.1)$$

Therefore, LN depends on the differences between the central pixel and its neighbor pixels into the window where it is applied. Consequently, the weights found by the GAs can be interpreted as the relevance of directional and scaled differences. It can be observed that the weights of the corners in GO-LN had the lowest values (see Figure 4.8). Therefore, we may conclude that the pixels closer to the center are more important than those that are further away.

In the GO-SQI method, the values of β_k and m_k found by the GAs for the LMG classifier (see Table 4.11) show that the intermediate scales (between 9×9 and 15×15) are the most important to eliminate the illumination components. Also, except for the 11×11 scale, the anisotropy is almost unnecessary. It should be noted that these results may vary if image size changes.

In the GO-LN+SQI cascade method, in the LN stage (see Figure 4.9), vertical and horizontal differences are more important than the differences in other angles. In the SQI stage (see Table 4.12), scales of 13×13 and 15×15 contain most of the illumination components. Again, these scales are strongly related to the size of the images used. In contrast to GO-SQI, in this case the anisotropy is necessary.

5.1.2 Kolmogorov-Nagumo LN

For both classifiers, the results obtained using values of $p = 1$ to compute the power mean and the power standard deviation were better than the original LN method. This proves that the original LN is not necessarily the best version of LN. Considering the definition of LN expressed in (5.1), the Kolmogorov-Nagumo LN is equivalent to assign different weights to the neighbor pixels as in GO-LN, but now in function of the pixel value itself, instead of the distance from the central pixel.

5.1.3 Relation between LN and the Laplace operator

Using the definition of LN in Equation (5.1), a relation between the LN and the Laplace operator can be established.

Expanding:

$$\begin{aligned}
\sum_{i=-1}^1 \sum_{j=-1}^1 I(x, y) - I(x + i, y + j) &= 9I(x, y) - \sum_{i=-1}^1 \sum_{j=-1}^1 I(x + i, y + j) \\
&= 8I(x, y) - \sum_{\substack{i=-1 \\ i \neq 0}}^1 \sum_{\substack{j=-1 \\ j \neq 0}}^1 I(x + i, y + j)
\end{aligned} \tag{5.2}$$

is equivalent to convolve an image with the 2-D Laplacian filter:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}. \tag{5.3}$$

The above relation can be generalized for any $k \in \mathbb{N}$. Expanding:

$$\begin{aligned}
\sum_{i=-k}^k \sum_{j=-k}^k I(x, y) - I(x + i, y + j) &= (2k + 1)^2 I(x, y) - \sum_{i=-k}^k \sum_{j=-k}^k I(x + i, y + j) \\
&= \left((2k + 1)^2 - 1 \right) I(x, y) - \sum_{\substack{i=-k \\ i \neq 0}}^k \sum_{\substack{j=-k \\ j \neq 0}}^k I(x + i, y + j) \\
&= 4k(k + 1) I(x, y) - \sum_{\substack{i=-k \\ i \neq 0}}^k \sum_{\substack{j=-k \\ j \neq 0}}^k I(x + i, y + j)
\end{aligned} \tag{5.4}$$

is equivalent to convolve an image with $(2k + 1) \times (2k + 1)$ filter:

$$\begin{bmatrix} \vdots & \vdots & \vdots \\ \cdots & -1 & -1 & -1 & \cdots \\ \cdots & -1 & 4k(k + 1) & -1 & \cdots \\ \cdots & -1 & -1 & -1 & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix}. \tag{5.5}$$

This filter can be regarded as a generalization of the Laplacian filter described in 5.3, in the sense that differences in several angles and distances with respect to the center are considered.

5.1.4 Relation between LN, its improved versions, and other illumination compensation methods

Regarding the definition of LN in Equation (5.1), it can be demonstrated that LN and its improved versions belong to a set of illumination compensation methods based on the

Laplacian operator.

Weberface is an illumination compensation method based on the Weber's law, which describes how human beings perceive changes in physical stimuli [Wang 2011]. Weberface method is defined by:

$$I_W(x, y) = \arctan \left(\alpha \sum_{i=0}^{p-1} \frac{x_c - x_i}{x_c} \right), \quad (5.6)$$

where x_c is the central pixel and x_i are the neighbor pixels in a 3×3 window. Equation 5.6 can be rewritten as:

$$I_W(x, y) = \arctan \left(\alpha \sum_{i=-1}^1 \sum_{j=-1}^1 \frac{I(x, y) - I(x+i, y+j)}{I(x, y)} \right)$$

The Local Binary Pattern method (LBP) were first introduced to discriminate textures in images [Ojala 1996]. In [Heusch 2006], LBP is used to compensate illumination for face recognition. In that work, LBP was defined by:

$$LBP(x_c, y_c) = \sum_{n=0}^7 s(i_n - i_c) 2^n \quad (5.7)$$

where (x_c, y_c) is the position of the central pixel, i_c is the gray intensity of the central pixel, i_n are the neighbor pixels starting from the upper left corner and numbered in clockwise order, and $s(x)$ is defined by:

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}. \quad (5.8)$$

Equation (5.7) can be reformulated as:

$$LBP(I(x, y)) = \sum_{i=-1}^1 \sum_{j=-1}^1 s(-(I(x, y) - I(x+i, y+j))) 2^{n(i,j)} \quad (5.9)$$

where $n(i, j)$ is defined such as to be consistent with (5.7).

From (5.1), (5.7), and (5.1), it can be concluded that LBP, LN, GO-LN, KN-LN, and Weberfaces belong to a set of Laplacian-based methods \mathbb{L} defined by:

$$\mathbb{L}(I(x, y), k) = S \left(\sum_{i=-k}^k \sum_{j=-k}^k f(g(I(x, y)) - g(I(x+i, y+j))) w(i, j) \right) \quad (5.10)$$

where $S()$ is a saturation function, $f()$ is a function of the difference between the central pixels and its neighbors, $g()$ is a function of the pixel intensity, and $w(i, j)$ is a weight that depends on the position of the neighbor pixel.

5.1.5 Spectral analysis based on IDCT of the Laplacian-based methods

A spectral analysis of the Laplacian methods can be easily performed through the IDCT. The continuous 2-D Laplacian of $I(x, y)$ is defined by:

$$\nabla^2 I(x, y) = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \quad (5.11)$$

The IDCT was defined in 2.5 by:

$$I(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right),$$

Using 5.11 and 5.12, the Laplacian of the IDCT can be expressed as:

$$\nabla^2 I(x, y) = -\pi^2 \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \left[\left(\frac{u}{M}\right)^2 + \left(\frac{v}{N}\right)^2 \right] \alpha(u)\alpha(v)C(u, v) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right), \quad (5.12)$$

where the term $\left[\left(\frac{u}{M}\right)^2 + \left(\frac{v}{N}\right)^2 \right]$ assigns small weights to the lower frequency components. This agrees with the Retinex model, which proposes that the illumination of an image lies in its low frequency components (see Section 2.1.1). Furthermore, it can be concluded that applying Laplacian methods is equivalent to give the smaller weights to the lower frequencies in IDCT.

5.2 Conclusions

The aim of this thesis was to make improvements in face recognition under non-controlled illumination. These improvements were made specifically over the following illumination compensation methods: DCT, LN and SQI. The illumination compensation methods are applied in a preprocessing step before face classification. The improvements obtained are the following:

- A new method for illumination compensation has been proposed based on the SQI method. Our method selects parameters of SQI using a GA. This method used a face database with non-homogeneous illumination for training and testing was performed with a face database with homogeneous and non-homogeneous illumination. Results show improvements for face databases with non-homogeneous illumination. However, the method shows a slight performance with faces under homogeneous illumination.
- A method for illumination compensation has been proposed using a genetic optimization of the LN and SQI parameters in cascade. The new optimized method reaches

100% face recognition rate on databases with non-homogeneous illumination, which are the best published results in the literature to date. This method yields the best results among illumination compensation methods with homogeneous illumination.

- The parameters of the DCT, LN, SQI and cascade LN-SQI methods were optimized using GAs to improve face recognition rate by the LMG classifier. Results on the Gray FERET database show that face recognition performance of the LMG classifier can be significantly improved by illumination compensation methods. The best results were obtained for the GO-LN method using the LMG classifier. The best result for the optimized LN method yields a 31% reduction in the total number of errors in the FERET database.
- A new LN method is proposed using Kolmogorov-Nagumo statistics. This method is more general than the original LN [Xie 2006]. It was showed that the original LN method is a particular case of the proposed method when $\varphi = x^p$. The results show that the canonical normalization ($p = 1$) is sub-optimal for face recognition. The proposed method produced significant improvements in face recognition relative to the canonical LN method, in the six standard databases, ranging from 9.9% to 95.2%. In all but the Extended Yale B database and PCA, the best results were reached using $\varphi(x) = x^p$. In the case of Extended Yale B database and PCA, the best result was reached with $\varphi(x) = \ln(x)$. Windows sizes no larger than 11×11 are recommended for this method.

The above demonstrated that better versions of the current illumination compensation methods were obtained by relaxing their variables, beyond the optimization tool used to syntonize them (GAs in this thesis). Other tools can be used for this effect, such as Particle Swarm Optimization (PSO), or even Mutual Information Feature Selection (MIFS) and its improved versions.

In addition, the LN and its improved versions were included into a set of Laplacian-based methods which involve a filtering through the 3×3 Laplace operator or its extension to $(2k + 1) \times (2k + 1)$, with $k \in \mathbb{N}$. Also, a spectral analysis using IDCT revealed that the Laplacian-based are equivalent to give the smaller weights to the lower frequency components, which agrees with the Retinex model. All the above is a theoretical proof of the performance of the GO-LN and the KN-LN methods.

In the same context, it was found that the Teager-Kaiser Energy Operator (TKEO) $\Psi(I(x, y))$ and the Laplacian of the natural logarithm of an image $\ln(I(x, y))$ are related by:

$$\Psi(I(x, y)) = -I^2(x, y) \nabla^2(\ln(I(x, y))) \quad (5.13)$$

This identity is demonstrated in the Appendix A, and it could be considered in a future work.

Part of this work was performed during a doctoral internship at the SAMPL laboratory at the Department of Electrical, Computer and System Engineering of the Rensselaer Polytechnic Institute.

5.3 Publications

The following publications resulted from the work of this thesis:

5.3.1 Journal papers

- **L. E. Castillo**, L. A. Cament, F. J. Galdames and C. A. Perez. *Illumination normalisation method using Kolmogorov-Nagumo-based statistics for face recognition*. Electronics Letters, vol. 50, pages 940-942, 2014.
- C. A. Perez, **L. E. Castillo**, L. A. Cament, P. A. Estévez and C. M. Held. *Genetic optimisation of illumination compensation methods in cascade for face recognition*. Electronics Letters, vol. 46, no. 7, pages 498-500, 2010.

5.3.2 Conference papers

- C. A. Perez, **L. E. Castillo** and L. A. Cament. *Illumination Compensation Method for Local Matching Gabor Face Classifier*. In Int. Symp. on Optomechatronic Technologies (ISOT 2010), pages 1-5, Toronto, ON, Canada, 25-27 October 2010.
- C. A. Perez and **L. E. Castillo**. *Illumination compensation for face recognition by genetic optimization of the self-quotient image method*. In Int. Symp. on Optomechatronic Technologies (ISOT 2009), pages 322-327, Istanbul, Turkey, 21-23 September 2009.

Bibliography

- [Adini 1997] Yael Adini, Yael Moses and Shimon Ullman. *Face recognition: The problem of compensating for changes in illumination direction*. IEEE Transactions on pattern analysis and machine intelligence, vol. 19, no. 7, pages 721–732, 1997.
- [Alander 1992] J. T. Alander. *On optimal population size of genetic algorithms*. In Proceedings of CompEuro '92 "Computer Systems and Software Engineering", pages 65–70, Hague, Netherlands, 4-8 May 1992.
- [Baradarani 2013] A. Baradarani, Q. M. J. Wu and M. Ahmadi. *An efficient illumination invariant face recognition framework via illumination enhancement and DD-DTCWT filtering*. Pattern Recognition, vol. 46, no. 1, pages 57–72, 2013.
- [Belhumeur 1997] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman. *Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pages 711–720, 1997.
- [Belhumeur 1998] P. N. Belhumeur and D. J. Kriegman. *What Is the Set of Images of an Object Under All Possible Illumination Conditions?* International Journal of Computer Vision, vol. 28, no. 3, pages 245–260, 1998.
- [Bichsel 1995] M. Bichsel. *Illumination invariant object recognition*. In Proceedings of the International Conference on Image Processing, volume 3, pages 620–623, Washington, DC, 1995.
- [Calin 2010] M. A. Calin and S. V. Parasca. *In vivo study of age-related changes in the optical properties of the skin*. Lasers in Medical Science, vol. 25, no. 2, pages 269–274, 2010.
- [Cament 2014] L. A. Cament, L. E. Castillo, J. P. Perez, F. J. Galdames and C. A. Perez. *Fusion of local normalization and Gabor entropy weighted features for face identification*. Pattern Recognition, vol. 47, no. 2, pages 568–577, 2014.
- [Cao 2012] X. Cao, W. Shen, L. G. Yu, Y. L. Wang, J. Y. Yang and Z. W. Zhang. *Illumination invariant extraction for face recognition using neighboring wavelet coefficients*. Pattern Recognition, vol. 45, no. 4, pages 1299–1305, 2012.
- [Chen 2006] W. Chen, M. J. Er and S. Wu. *Illumination Compensation and Normalization for Robust Face recognition Using Discrete Cosine Transform in Logarithm Domain*. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 36, no. 2, pages 458–466, 2006.

- [Choi 2007] S. Choi, C. Kim and C. Choi. *Shadow compensation in 2D images for face recognition*. Pattern Recognition, vol. 40, no. 7, pages 2118–2125, 2007.
- [Du 2005] B. Du, S. Shan, L. Qing and L. Gao. *Empirical comparisons of several preprocessing methods for illumination*. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05), pages 981–984, Philadelphia, PA, USA, 18-23 March 2005.
- [Georghiades 2001] A. S. Georghiades, P. N. Belhumeur and D. J. Kriegman. *From few to many: Illumination cone models for face recognition under variable lighting and pose*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pages 643–660, 2001.
- [Goldberg 1989] D. Goldberg. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, 1989.
- [Gurney 2007] Kevin Gurney. *Neural networks for perceptual processing: from simulation tools to theories*. Philosophical Transactions of the Royal Society of London B: Biological Sciences, vol. 362, no. 1479, pages 339–353, 2007.
- [Hallinan 1994] P. W. Hallinan. *A low-dimensional representation of human faces for arbitrary lighting conditions*. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 995–999, Seattle, WA, 21-23 June 1994.
- [Han 2013] H. Han, S. Shan, X. Chen and W. Gao. *A comparative study on illumination preprocessing in face recognition*. Pattern Recognition, vol. 46, no. 6, pages 1691–1699, 2013.
- [Heusch 2006] Guillaume Heusch, Yann Rodriguez and Sebastien Marcel. *Local binary patterns as an image preprocessing for face authentication*. In 7th International Conference on Automatic Face and Gesture Recognition (FGR06), pages 6–pp. IEEE, 2006.
- [Jobson 1997a] D. J. Jobson, Z. U. Rahman and G. A. Woodell. *A multiscale retinex for bridging the gap between color images and the human observation of scenes*. IEEE Transactions on Image Processing, vol. 6, no. 7, pages 965–976, 1997.
- [Jobson 1997b] D. J. Jobson, Z. U. Rahman and G. A. Woodell. *Properties and performance of a center/surround retinex*. IEEE Transactions on Image Processing, vol. 6, no. 3, pages 451–462, 1997.
- [Kaiser 1990] James F Kaiser. *On a simple algorithm to calculate the energy of a signal*. In Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on, pages 381–384. IEEE, 1990.

- [Kasturi 2009] R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, J. Garofolo and R. Bowers, M. Boonstra, V. Korzhova and J. Zhang. *Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol*. IEEE T. Pattern Anal., vol. 31, no. 2, pages 319–336, 2009.
- [Kolmogorov 1930] A. Kolmogorov. *Sur la notion de la moyenne*. Atti Accad. Naz. Lincei, vol. 9, pages 221–235, 1930.
- [Kvedalen 2003] Eivind Kvedalen. *Signal processing using the Teager energy operator and other nonlinear operators*. PhD thesis, 2003.
- [Lee 2005] Kuang-Chih Lee, Jeffrey Ho and David Kriegman. *Acquiring linear subspaces for face recognition under variable lighting*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 27, no. 5, pages 684–698, 2005.
- [Lei 2009] Z. Lei, C. Wang, Q. Wang and Y. Huang. *Real-Time Face Detection and Recognition for Video Surveillance Applications*. In World Congress on Computer Science and Information Engineering, pages 168–172, Los Angeles, CA, USA, 31 March - 2 April 2009.
- [Li 2013] Y. Li, L. Meng and J. Feng. *Face illumination compensation dictionary*. Neurocomputing, vol. 101, pages 139–148, 2013.
- [Makwana 2010] R. M. Makwana. *Illumination Invariant Face Recognition: A Survey of Passive Methods*. In Proceedings of the International Conference and Exhibition on Biometrics Technology, pages 597–614, Coimbatore, India, 2-4 September 2010.
- [Martinez 1998] A. M. Martinez and R. Benavente. *The AR Face Database*. CVC Technical Report, vol. 24, 1998.
- [Mitchell 1996] M. Mitchell. Introduction to genetic algorithms. MIT Press, 1996.
- [Mitra 1991] Sanjit K Mitra, Hui Li, I-S Lin and T-H Yu. *A new class of nonlinear filters for image enhancement*. In Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on, pages 2525–2528. IEEE, 1991.
- [Nagumo 1930] M. Nagumo. *Über eine klasse der mitterwerte*. Japan. J. Math, vol. 7, pages 71–79, 1930.
- [Nie 2008] X. Nie and Z. Tan. *Face illumination compensation using multivariate wavelets transform*. In Chinese Control and Decision Conference (CCDC), pages 1853–1856, Yantai, Shandong, China, 2-4 July 2008.

- [Ojala 1996] Timo Ojala, Matti Pietikäinen and David Harwood. *A comparative study of texture measures with classification based on featured distributions*. Pattern recognition, vol. 29, no. 1, pages 51–59, 1996.
- [Osadchy 2010] M. Osadchy, P. Pinkas, A. Jarrous and B. Moskovich. *SCiFI - A System for Secure Face Identification*. In IEEE Symposium on Security & Privacy, pages 239–254, Berkeley, CA, USA, 22-25 May 2010.
- [Perez 2003] C. A. Perez, C. Salinas, P. A. Estévez and P. Valenzuela. *Genetic design of biologically inspired receptive fields for neural pattern recognition*. IEEE Transactions on Systems, Man, and Cybernetics-Part B, vol. 33, no. 2, pages 258–270, 2003.
- [Perez 2005] C. A. Perez, G. González, L. E. Medina and F. J. Galdames. *Linear Versus Non-Linear Neural Modeling for 2D Pattern Recognition*. IEEE Transactions on Systems, Man and Cybernetics-Part A, vol. 35, no. 6, pages 955–964, 2005.
- [Perez 2007] C. A. Perez, V. Lazcano and P. A. Estévez. *Real-time iris detection on coronal-axis-rotated faces*. IEEE Transactions on Systems, Man and Cybernetics-Part C, vol. 37, no. 5, pages 971–978, 2007.
- [Perez 2008] C. A. Perez and L. E. Castillo. *Genetic improvements in illumination compensation by the discrete cosine transform and local normalization for face recognition*. In SPIE Int. Symp. on Optomechatronic Technologies (ISOT 2008), pages 72661B–1–8, San Diego, CA, USA, 17-19 November 2008.
- [Perez 2009a] C. Perez, V. Lazcano, P. Estévez and C. Held. *Real-Time Template Based Face and Iris Detection on Rotated Faces*. International Journal of Optomechatronics, vol. 3, no. 1, pages 54–67, 2009.
- [Perez 2009b] C. A. Perez and L. E. Castillo. *Illumination compensation for face recognition by genetic optimization of the self-quotient image method*. In Int. Symp. on Optomechatronic Technologies (ISOT 2009), pages 322–327, Istanbul, Turkey, 21-23 September 2009.
- [Perez 2010a] C. A. Perez, C. M. Aravena, J. I. Vallejos, P. A. Estévez and C. M. Held. *Face and iris localization using templates designed by particle swarm optimization*. Pattern Recognition Letters, vol. 31, no. 9, pages 857–868, 2010.
- [Perez 2010b] C. A. Perez, L. E. Castillo and L. A. Cament. *Illumination Compensation Method for Local Matching Gabor Face Classifier*. In Int. Symp. on Optomechatronic Technologies (ISOT 2010), pages 1–5, Toronto, ON, Canada, 25-27 October 2010.
- [Perez 2010c] C. A. Perez, L. E. Castillo, L. A. Cament, P. A. Estévez and C. M. Held. *Genetic optimisation of illumination compensation methods in cascade for face recognition*. Electronics Letters, vol. 46, no. 7, pages 498–500, 2010.

- [Perez 2011] C. A. Perez, L. A. Cament and L. E. Castillo. *Methodological improvement on local Gabor face recognition based on feature selection and enhanced Borda count*. Pattern Recognition, vol. 44, no. 4, pages 951–963, 2011.
- [Phillips 1998] P. J. Phillips, H. Wechsler, J. Huang and P. Rauss. *The FERET database and evaluation procedure for face recognition algorithms*. Image Vis. Comput. J., vol. 16, no. 5, pages 295–306, 1998.
- [Phimoltares 2007] S. Phimoltares, C. L. Lursinsap and K. Chamnongthai. *Face detection and facial feature localization without considering the appearance of image context*. Image and Vision Computing, vol. 25, no. 5, pages 741–753, 2007.
- [Rényi 1961] A. Rényi. *On measures of entropy and information*. In Fourth Berkeley Symposium on Mathematical Statistics and Probability, volume 547–561, 1961.
- [Ruiz-del Solar 2008] Javier Ruiz-del Solar and Julio Quinteros. *Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches*. Pattern Recognition Letters, vol. 29, no. 14, pages 1966–1979, 2008.
- [Shannon 1949] Claude E Shannon and Warren Weaver. *University of Illinois Press*. Urbana, pages 104–107, 1949.
- [Sim 2003] T. Sim, S. Baker and M. Bsat. *The CMU pose, illumination, and expression database*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pages 1615–1618, 2003.
- [Teager 1990] HM Teager and SM Teager. *Evidence for nonlinear sound production mechanisms in the vocal tract*. In Speech production and speech modelling, pages 241–261. Springer, 1990.
- [Turk 1991] M. A. Turk and A. P. Pentland. *Face recognition using eigenfaces*. In Proceedings CVPR’91, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 586–591, Maui, HI, USA, 3-6 June 1991.
- [Wang 2004] H. Wang, S. Z. Li and Y. Wang. *Face Recognition under Varying Lighting Conditions Using Self Quotient Image*. In Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pages 819–824, Seoul, Korea, 17-19 May 2004.
- [Wang 2008a] L. Wang, Y. Li, C. Wang and H. Zhang. *2D Gaborface representation method for face recognition with ensemble and multichannel model*. Image and Vision Computing, vol. 26, no. 6, pages 820–828, 2008.

- [Wang 2008b] Y. H. Wang, X. J. Ning, C. X. Yang and Q. F. Wang. *A method of illumination compensation for human face image based on quotient image*. Inf. Sci., vol. 178, pages 2705–2721, 2008.
- [Wang 2009] X. Wang, C. Zhang and Z. Zhang. *Boosted multi-task learning for face verification with applications to web image and video search*. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 142–149, Miami, FL, USA, 20-26 June 2009.
- [Wang 2011] Biao Wang, Weifeng Li, Wenming Yang and Qingmin Liao. *Illumination normalization based on Weber’s law with application to face recognition*. IEEE Signal Processing Letters, vol. 18, no. 8, pages 462–465, 2011.
- [Weast 1989] R.C. Weast, D.R. Lide, Chemical Rubber Company and M.J. Astle. *Crc handbook of chemistry and physics: A ready-reference book of chemical and physical data*. CRC handbook of chemistry and physics. CRC Press, 1989.
- [Xie 2006] X. Xie and K. M. Lam. *An efficient illumination normalization method for face recognition*. Pattern Recognition Letters, vol. 27, no. 6, pages 609–617, 2006.
- [Yan 2009] G. Yan, J. Li, J. Li, Q. Ma and M. Yu. *Illumination Variation in Face Recognition: A Review*. In International Conference on Intelligent Networks and Intelligent Systems, pages 309–311, Barcelona, Spain, 4-6 November 2009.
- [Yun 2009] T. Yun and L. Guan. *Automatic face detection in video sequences using local normalization and optimal adaptive correlation techniques*. Pattern Recognition, vol. 42, no. 9, pages 1859–1868, 2009.
- [Zou 2007a] J. Zou, Q. Ji and G. Nagy. *A Comparative Study of Local Matching Approach for Face Recognition*. IEEE Transactions on Image Processing, vol. 16, no. 10, pages 2617–2628, 2007.
- [Zou 2007b] X. Zou, J. Kittler and K. Messer. *Illumination Invariant Face Recognition: A Survey*. In IEEE International Conference on Biometrics: Theory, Applications, and Systems, pages 1–8, Washington DC, USA, 27-29 September 2007.

Relation between the Laplacian operator and illumination compensation methods

The continuous 2-D Laplacian of $f(I(x, y))$ is defined by:

$$\nabla^2(f(I(x, y))) = \frac{\partial^2 f(I(x, y))}{\partial x^2} + \frac{\partial^2 f(I(x, y))}{\partial y^2} \quad (\text{A.1})$$

Expanding the first term of the right-hand side:

$$\begin{aligned} \frac{\partial^2 f(I(x, y))}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial f(I(x, y))}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial f(I(x, y))}{\partial I} \frac{\partial I(x, y)}{\partial x} \right) \\ &= \frac{\partial}{\partial x} \left(\frac{\partial f(I(x, y))}{\partial I} \right) \frac{\partial I(x, y)}{\partial x} + \frac{\partial f(I(x, y))}{\partial I} \frac{\partial^2 I(x, y)}{\partial x^2} \\ &= \frac{\partial}{\partial I} \left(\frac{\partial f(I(x, y))}{\partial x} \right) \frac{\partial I(x, y)}{\partial x} + \frac{\partial f(I(x, y))}{\partial I} \frac{\partial^2 I(x, y)}{\partial x^2} \\ &= \frac{\partial}{\partial I} \left(\frac{\partial f(I(x, y))}{\partial I} \frac{\partial I(x, y)}{\partial x} \right) \frac{\partial I(x, y)}{\partial x} + \frac{\partial f(I(x, y))}{\partial I} \frac{\partial^2 I(x, y)}{\partial x^2} \\ \frac{\partial^2 f(I(x, y))}{\partial x^2} &= \frac{\partial^2 f(I(x, y))}{\partial I^2} \left(\frac{\partial I(x, y)}{\partial x} \right)^2 + \frac{\partial f(I(x, y))}{\partial I} \frac{\partial^2 I(x, y)}{\partial x^2} \end{aligned} \quad (\text{A.2})$$

Regarding that the second term of the right-hand side can be expanded in the same manner, (A.1) can be rewritten as:

$$\begin{aligned} \nabla^2(f(I(x, y))) &= \frac{\partial^2 f(I(x, y))}{\partial x^2} + \frac{\partial^2 f(I(x, y))}{\partial y^2} \\ &= \frac{\partial^2 f(I(x, y))}{\partial I^2} \left[\left(\frac{\partial I(x, y)}{\partial x} \right)^2 + \left(\frac{\partial I(x, y)}{\partial y} \right)^2 \right] \\ &\quad + \frac{\partial f(I(x, y))}{\partial I} \left[\frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \right] \end{aligned}$$

$$\nabla^2(f(I(x, y))) = \frac{\partial^2 f(I(x, y))}{\partial I^2} \|\nabla I(x, y)\|^2 + \frac{\partial f(I(x, y))}{\partial I} \nabla^2(I(x, y)) \quad (\text{A.3})$$

If $f(I(x, y)) = \ln(I(x, y))$, as proposed in DCT:

$$\nabla^2(\ln(I(x, y))) = -\frac{1}{I^2(x, y)} \|\nabla I(x, y)\|^2 + \frac{1}{I(x, y)} \nabla^2(I(x, y)) \quad (\text{A.4})$$

It is interesting to notice that the second term of the right-hand side of the above equation is equivalent to the Weberface method before been saturated by the $\arctan()$ function (see 5.1.4).

The Teager-Kaiser Energy Operator (TKEO) [Teager 1990, Kaiser 1990] was first proposed to explain the nonlinearities of speech generation in the vocal tract, but also has been used in AM-FM demodulation[Kvedalen 2003]. The TKEO is defined by:

$$\Psi(x(t)) = \dot{x}^2(t) - x(t)\ddot{x}(t) \quad (\text{A.5})$$

In [Mitra 1991], the 2-D TKEO was applied to image processing. This was defined by:

$$\Psi(I(x, y)) = \|\nabla I(x, y)\|^2 - I(x, y)\nabla^2 I(x, y) \quad (\text{A.6})$$

Using (A.4) and (A.6), the following identity can be established:

$$\Psi(I(x, y)) = -I^2(x, y)\nabla^2(\ln(I(x, y))) \quad (\text{A.7})$$