Gender Classification Using Local Binary Patterns and its Variants

Parichehr Behjati Ardakani

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Engineering

Eastern Mediterranean University
September 2016
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Mustafa Tümer
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Computer Engineering.

Prof. Dr. Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Computer Engineering.

Assoc. Prof. Dr. Önsen Toygar
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Önsen Toygar
2. Asst. Prof. Dr. Yıltan Bitirim
3. Asst. Prof. Dr. Ahmet Ünveren

ABSTRACT

Many social interactions and services are dependent on gender today, so, gender classification is appearing as an active research area. Most of the existing studies are based on face images acquired under controlled conditions. In our work, we used different databases such as FERET, AR and ORL for controlled conditions and Labeled Faces in the Wild (LFW) database as real-life faces for uncontrolled conditions. Local Binary Patterns (LBP) and its variants such as Uniform LBP, Completed LBP and Rotation - Invariant LBP are employed to describe faces by extracting features from the region of interests. Manhattan distance measure is used to compare difference between test and training images for gender recognition. Based on the results reported as the state-of-the-art, we have achieved satisfactory results.

Keywords: Gender recognition, feature extraction, Local Binary Patterns (LBP)

ÖZ

Günümüzde birçok sosyal etkileşim ve hizmetler cinsiyete bağlı olduğu için cinsiyet sınıflandırma aktif bir araştırma alanıdır. Literatürde varolan birçok çalışma, denetimli durumlardan elde edilen yüz resimlerini kullanmaktadır. Bu çalışmada, denetimli ortamlarda elde edilen FERET, AR ve ORL yüz veritabanları ve denetimsiz ortamlar için de doğal yaşamda çekilen yüz resimlerini içeren LFW veritabanı kullanılmıştır. Yüz resimlerinin özniteliklerini elde etmek için Yerel İkili Örüntü (LBP) yaklaşımı ve bu yaklaşımın Birbiçimli LBP, Tamamlanmış LBP, Dönme Değişimsiz LBP isimli değişik varyantları kullanılmıştır. Cinsiyet tanımada, test ve eğitilmiş yüz resimlerinin farkını karşılaştırmak için Manhattan uzaklık ölçüsü kullanılmıştır. Literatürde bildirilen cinsiyet sınıflandırma sonuçlarıyla karşılaştırıldığında bu tezde elde edilen sonuçlar memnuniyet vericidir.

Anahtar Kelimeler: Cinsiyet tanıma, öznitelik çıkarma, Yerel İkili Örüntü (LBP).

To My Family

To My Little Princess: Shaina Safaie

ACKNOWLEDGMENT

I would like to express my deep-felt to my advisor, Assoc. Prof. Dr. Önsen Toygar for giving me an opportunity to work with and for her advice, encouragement and constant support. She always gave me her time, in spite of anything else that was going on. I want to thank the faculty members of Computer Engineering department and the staff for all their hard work and dedication, providing me the means to complete my degree and prepare for a career as a computer scientist. And finally, I must thank my family for supporting me during the development of this work with no complaints and continuing, loving support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Biometrics is the use of physical characteristics like face, fingerprints, iris etc. of an individual for personal identification. Some of the challenging problems of face biometrics are face detection, face recognition, and face identification. These problems are being researched by the computer vision community for the last few decades. Considering the large population, the authentication process of an individual usually consumes a significant amount of time. One of the possible solutions is to divide the population into two halves based on gender. This will help to reduce the search space of authentication to almost half of the existing data and save substantial amount of time. Gender identification through face demands use of strong discriminative features and robust classifiers to separate the female and male faces without any ambiguity.

Applications such as smart human-computer interface, biometrics, security industry, and surveillance would gain greatly from knowledge of the attribute of the human subjects under scrutiny [1]. The gender is one such substantial demographic attribute. Gender classification estimates person's gender as female or male on their visual information and is a basic task for human beings, as many social functions critically based on the correct gender conception. Gender classification output can increase the performance of many applications such as face recognition, age classification and so on.

Gender classification using facial images has become an important area of research during past several years. It is easy for human to identify male or female by seeing a face, but it is a difficult task for the computer. Machines need some meaningful data to perform the identification. There exist some distinguishable features between male and female which are used by machine to classify a face image based on gender. Gender recognition is a pattern recognition problem. Pattern recognition can be divided into two classes, one and two stage pattern recognition systems. One stage pattern recognition system classifies input data directly. Two stage pattern recognition systems consist of feature extractor, followed by some form of classifier. Gender classification is a binary classification problem therefore machine needs an appropriate data (feature) and a classifier for gender classification.

Generally facial images provide important clues about the identity, gender, age and ethnicity of people and are used to extract features. In order to learn a gender, a feature extractor is applied to extract features and a classifier is used to classify the gender. Based on this pipeline, many approaches have been proposed, which have reached promising accuracy on various data sets in the last few decades. Most of the existing studies have focused on face images acquired under controlled conditions. Howsoever, real-world applications need gender classification on real-life environments, which is more challenging due to considerable appearance variations in unconstrained scenarios.

In this thesis, gender recognition is investigated for face images under controlled and uncontrolled conditions. Gender recognition on face images acquired under controlled conditions uses the FERET database, which usually includes frontal, with clean background, occlusion free, consistent lighting, and limited facial expression. On the other hand, there are significant appearance variations on real life faces, which include facial expression, illumination changes, head pose variations, occlusion or makeup, poor image quality and so on. Additionally, real-world applications need real-life faces, so for this purpose, Labeled Faces in the Wild (WLF) database is used. We used different databases for gender recognition and applied Local Binary Patterns (LBP) and its variants as feature extractors in the experiments.

In this study, LBP histograms are extracted from local facial regions as the region-level description, where the n-bin histogram is taken as a whole. In addition, we had done our experiments under illumination conditions, expressions and occlusions separately for each database and based on the results obtained, the performances achieved for gender recognition for images captured under controlled environments and under real-life environments.

The rest of this thesis is organized as follows. A literature review is presented in Chapter 2 and in Chapter 3 gender recognition steps using facial images are discussed. Chapter 4 describes Local Binary Patterns and its variants. Chapter 5 presents the databases used in this study. The experimental results for gender classification on female and male facial images are demonstrated in Chapter 5. Finally, Chapter 6 concludes our work.

Chapter 2

LITERATURE REVIEW

Gender recognition is an essential module for many computer vision applications such as human-robot interaction, visual surveillance and passive demographic data collections. More recently, the advertising industry's growing interest in the launching demographic-specific marketing and targeted advertisements in public places has attracted the attention of more and more researchers specialized in the field of computer vision. In this section, we will take a look at the different techniques proposed in the field of gender recognition. A detailed survey of studies on gender recognition can be found [2], [3].

Among the early algorithms in the field of gender recognition, Cottrel and Empath [4] extracted features from the whole face, called Holons, which were fed into a backpropagation network model to classify females and males. Golomb et al. [5] proposed a Neural Network model for gender classification. This network compresses faces using faces' raw pixels and then determines their gender in subsequent layers of their proposed network. Brunelli and Poggio [6] achieved a 79% accuracy for gender classification by using the HyperBF network on a set of geometrical features extracted from faces and, shortly after, Abdi et al. [7] employed the RBF network and achieved a 91% accuracy on data preprocessed by Principal Component Analysis (PCA), in 1995. In 1997, wavelet components were used by Wiskott et al. [8] in order to describe face features and build Elastic Bunch Graph models.

Lyons et al. [9] applied Linear Discriminant Analysis and Gabor wavelets to create a neuro-fuzzy system for gender recognition, which gave them a much more accurate system when compared with previously used methods. In 2002, Sun et al. [10] proposed a feature selection method by using Genetic Algorithms to select features extracted by Principal Component Analysis (PCA). They compared different classifiers such as NN, LDS, Bayesian and Support Vector Machine (SVM) and demonstrated that using an SVM classifier is a better method to classify gender. Moghaddam and Yang [11] contended that the Support Vector Machine (SVM) generates a stronger classifier than those previously used in gender recognition, when using an RBF kernel. They have done their experiments on both good quality images (64×72) and small images (21×12) of the FERET database and they achieved a 96.6% recognition rate on the second image data. In addition to this, they proved that the difference between the two different qualities is just 1%.

In 2004, Jain and Huang [12] suggested an approach using an Independent Component Analysis (ICA) as one of the feature-based methods to extract features and as classifier they employed LDA. Costen et al. [13] proposed a sparse SVM for gender classification and claimed a recognition rate of 94.42% on Japanese face images.

Local Binary Patterns (LBP) was utilized as a method for extracting texture features by Sun et al. [14] in 2006. LBP, which builds up the spatial structure of each pixel by comparing its intensity with that of its neighbors, was used with both Adaboost classifier [15] and Support Vector Machine (SVM) classifier to simplify gender classification.

Classifying facial expressions prior to gender classification was utilized to improve the classification accuracy by Saatci and Town, in 2006 [16]. Though Saatci and Town investigated the interdependency of gender recognition upon expression, they demonstrated that the gender classification accuracy reduced even with using separate gender data for different expression classes. In 2011, for gender classification, a color descriptor relying on the construction of histograms with 4 bins per color channel in the RGB color space was proposed [17]. The merger approach, 2D PCA and the centralized Gabor gradient histogram (CGGH) were other methods that were applied for extracting features. Fu et. al [18] combined Gabor gradient magnitudes with CBP to extract more discriminative features at multiple scales and orientation [18]. By using these features as input to a nearest center-based neighbor classifier, they got a 96.56% accuracy recognition on FERET and a 95.25% rate on CAS-PEAL. Meantime, in 2010, Lin and Zhao [19] suggested a color-based method on SVM for gender classification.

In 2011, compressive sensing framework was applied by Chen and Hsieh [20] to show face images in a sparse frequency domain. For each gender they extracted two feature sets as follows; the first set showing features that are common between all the images in the corresponding gender classes and the other one showing each individual face in that class. Spatial Weber Local Descriptor as a texture descriptor for gender recognition was used by Ullah et. al in 2012 [21]. They partitioned the image into a number of blocks (sub-region), computed the WLD descriptor for each block and concatenated them (as a vector). Tapia and Perez [22] suggested a method in 2013, which uses feature selection based on mutual information and the fusion of intensity, texture features and shape for gender classification. They then computed LBP features with different spatial orientation and radial, and then selected features using mutual information. They experimented three different techniques to estimate mutual information as follows: minimum redundancy and maximum relevance [23], and conditional mutual information-based [24], normalized mutual information [25].

In general, most of the efforts made to optimize in gender recognition from a face object attempt to best represent the face object. While some methods choose to use raw pixels without any modification, the majority of the existing methods use local visual descriptors to produce stronger and, often, more compact representations of face images. Examples of visual information commonly used for gender recognition are shape information (e.g., used in [24], [25]) color information (e.g., used in [21] and [26]) and texture information (e.g., used in [27], [28]), In these approaches, local descriptors are extracted from a dense regular grid placed over the entire image and the face representation is built by concatenating these extracted descriptors into a single vector. A key issue in this framework is to determine the optimal grid parameters such as number of grids in multi-resolution/pyramid approaches, spacing, size and etc.

Yu et. al [29] suggested an approach to use raw pixels, Gabor jets and LBPs on Gallaghers database to classify genders. They decreased the size of extracted features by using PCA and showed that, by using Gabor jets followed by SVM high gender classification, a better accuracy is acquired. While the aforementioned methods for gender recognition used fixed settings and performed trial-and-error to determine the right grid parameters, a better approach consists of using feature selection to allow only the most informative image regions to contribute to the face representation, i.e., those that can best separate face images that belong to different demographic classes. This approach further facilitates the integration of different types of descriptors like color based, shape based, texture based and etc. and allow for more compact representations by preventing redundant features from contributing to the face representation.

Chapter 3

GENDER RECOGNITION STEPS USING FACIAL IMAGES

In this chapter, we explain the steps of gender recognition systems. Generally, the framework of a gender recognition system can be seen as consisting of pre-processing, feature extraction and classification steps. Every face database needs some pre-processing such as face detection, normalization and etc. As the next step, we need to use feature extractors to extract features. In general, two types of features are extracted namely appearance-based features (local feature) and geometric features (global features). Classification is the last step of gender classification in which the gender is estimated as either female or male. General framework of gender recognition is shown in Figure 1. Sometimes, feature extraction and classification may be integrated in some framework such as in neural networks.

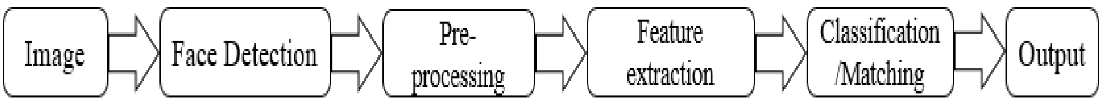

Figure 1. General framework for gender recognition system

3.1 Preprocessing Using Face Detection

Images of subjects captured by camera will normally be out of phase with each other because of difference in background, pose of head, contrast of images, number of intensity levels, size of the images and other causes. Most of the programs cannot automatically solve all of these problems. Therefore, some of the deficiencies of these images should be solved before defining them to the program as input and some of them should be solved during process of classification. Figure 2 shows the processes that took place to prepare images for being used in recognition systems.

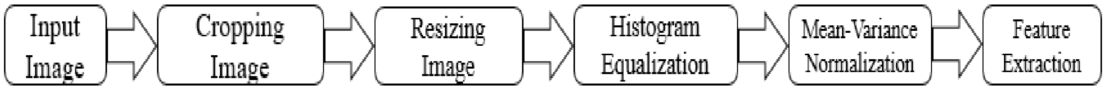

Figure 2. Preprocessing steps on images

Once an image is considered, and the region containing the person's face is detected. The image is cropped (cropping can be done using different face detection techniques) to return the region of interest, generally having the form of a bounding box. The reason of cropping image is some part of face like hairs and neck are sources of failure in the classification, therefore they should be removed from the image. The other reason of cropping image is to decrease memory consumption and increase the speed of detecting gender because vast useless data has removed from the image.

In gender recognition systems the cropped face image is obtained first, it may be followed by some preprocessing before using it as input to the feature extractor. The aim of this step is that classifiers are sensitive to variations such as illumination, poses and detection inaccuracies. In order to reduce this sensitivity of the system, some pre-processing steps are performed. There are some basic pre-processing steps that may be applied to the images such as facial portion detection and removal of background region such as hair and neck area. Face portion alignment (either manually or using automatic methods) is best done before downsizing.

The next step is to normalize for contrast and brightness using histogram equalization function. Histogram in image processing represents the relative frequency of occurrence of various gray levels in the image. Images may have different number of intensity levels. Congestion of intensities in different levels might be different and these differences among images will decrease the efficiency of facial gender classification. Histogram Equalization increases the range of intensity and spreads the intensity distributions which are better than having flattened peaks and valleys for an image in terms of a histogram [30]. This operation increases contrast of the low contrast areas without affecting the overall contrast of the image. Histogram equalization technique increases the facial gender classification rate by equalizing the levels of intensities of different images that should be equalized as much as possible. Implementing the histogram equalization technique almost equalizes the distribution of intensity levels in different images.

Downsizing to reduce the number of pixels (i.e. number of features) is another operation performed in preprocessing. For instance, each image that is given to the algorithm as an input may have different size therefore in this step it is necessary to unify size of all images. There are three types of interpolation techniques such as nearest neighbor interpolation, bilinear interpolation and bicubic interpolation that are used frequently to perform downsizing in images.

Normalizing (rescaling) the pixel values, for example to unit variance and zero mean is the next step of the preprocessing operation. Mean and variance normalization technique (MVN) is commonly used to increase the robustness of recognition features. Feature normalization techniques are able to largely reduce the actual mismatch between training and testing conditions. Both histogram equalization (HE) and mean and variance normalization (MVN) [31] have been used to process feature vectors in order to significantly improve the classification performance.

3.2 Feature Extraction

In machine learning, pattern recognition and image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is sometimes used for dimensionality reduction [32].

Extraction of various facial features probably is the most important sub-task of gender classification that contributes to improve classification accuracy. In order to detect a face, it is required to extract some features from the given image and learn classification from the given dataset which can be labeled or not. In computer vision, a feature is defined as a piece of information which is related to solve the computational task. Features can be specific structures in the image such as objects, points, edges and region of interest.

Based on feature extraction, gender classification approaches are categorized into two classes as local facial features and global facial features. Geometric-based feature extraction uses local features. In geometric-based approach, features are extracted from some facial points such as nose, eyes and face. By using this method some useful information is lost. Appearance-based feature extraction uses global features. In appearance-based approach, features are extracted from the face as a whole instead of extracting features from facial points.

One way to achieve the texture classification in gray scale images is to use Local Binary Patterns texture descriptors to build several local descriptions of the facial image and concatenate them into a global description. The LBP operator [33] is one of the best performing texture descriptors and it has been widely used in various applications.

In this study, facial images are used for the classification of person's gender and Local Binary Patterns (LBP) and its variants are used for feature extraction. Then classification strategy is used to obtain the classification accuracy. More detailed review of feature extraction approaches is described in the following chapters.

3.3 Matching

Male and female faces differ in both local features and shape. Men's faces on average have thicker eyebrows and greater texture in the bread region. In women's faces, the distance between the eyes and brows is greater, the protuberance of the nose is smaller, and the chin is narrower than in men's.

Classification or matching is the last step of gender classification in which the subject will be classified as either male or female. For this purpose, different types of classifiers are used. In our work, we used Manhattan Distance Measure to decide whether the subject is female or male by finding the distance between test images and all the training images. In this process, the most resembling image to a test image is the one with the smallest distance.

Chapter 4

LOCAL BINARY PATTERNS AND ITS VARIANTS AS FEATURE EXTRATORS

In this chapter, Local Binary Patterns Approach and its variants such as Completed LBP, Uniform LBP and Rotational Invariant LBP are described and we also present how to extract features of a given image by using these feature extractors.

4.1 Local Binary Patterns

Local Binary Patterns (LBP) was introduced by Ojala et al. [33] and used for gray scale images for texture description in computer vision. LBP has since been found to be a powerful feature for texture description and is also known as a very strong texture classification method [33] [34]. The LBP creates a descriptor or texture model using a set of histograms of the local texture neighborhood surrounding each pixel and also LBP can be used as an image processing operator.

The original version of Local Binary Pattern operator works in a 3×3 pixel block of an image. The LBP operator labels each pixel of an image, which thresholds the pixel's local neighborhood at its gray scale value into a binary number. The local neighborhood is a circular symmetric set of any number of pixels and radius. Then histogram of the labels can be used as a texture descriptor. The basic LBP operator is illustrated in Figure 3.

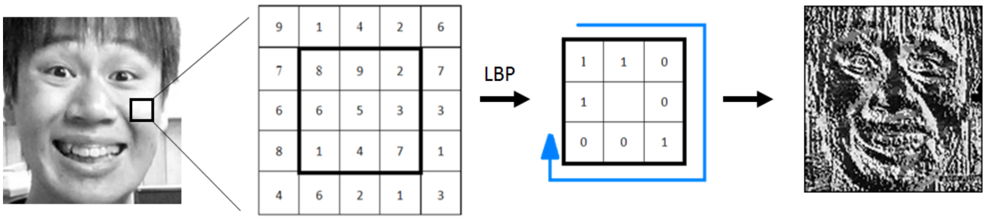

Figure 3. Example of an input image and the corresponding LBP image [34].

LBP generates a binary 1 if the neighbor of the center pixel has larger value than the center pixel. The operator generates a binary 0 if the neighbor is less than the center. The eight neighbors of the center can then be represented with an 8-bit number such as an unsigned 8-bit integer, making it a very compact description. Figure 4 shows an example of an LBP operator.

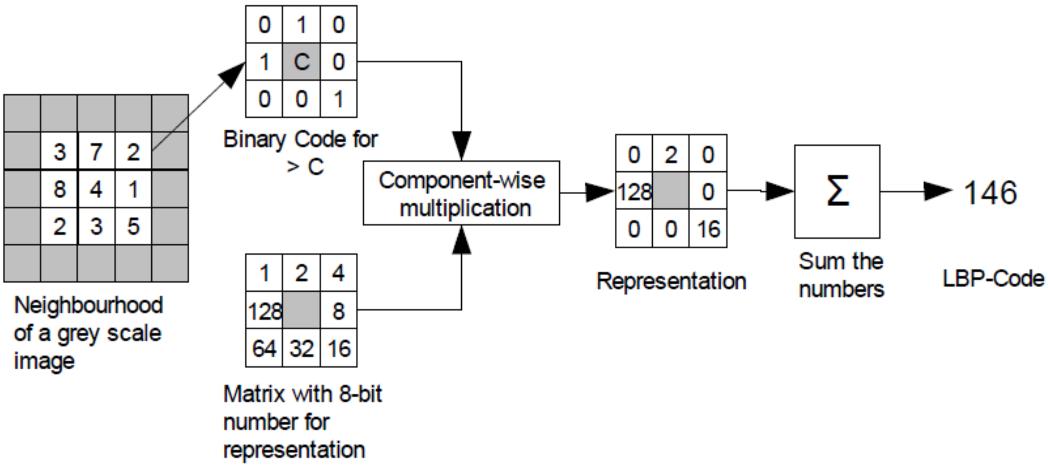

Figure 4. Example of how LBP-operator works [34].

The $\mathbf{LBP_{P,R}}$ operator is defined as

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} s(g_p - g_c)2^P,$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{4.1}$$

where g_c is the gray value of the central pixel, g_p is the value of its neighbors, P is the total number of involved neighbors, and R is the radius of the neighborhood.

In practice, this equation means that the signs of the differences in a neighborhood are interpreted as a P-bit binary number, resulting in 2^P distinct values for the LBP code. The local gray-scale distribution, i.e. texture, can thus be approximately described with a 2^P-bin discrete distribution of LBP codes:

$$T \approx t(LBP_{P,R}(x_c, y_c)) \tag{4.2}$$

In calculating the $LBP_{P,R}$ distribution (feature vector) for a given N × M image sample ($x_c \in \{0, \ldots, N-1\}, y_c \in \{0, \ldots, M-1\}$), the central part is only considered because a sufficiently large neighborhood cannot be used on the borders. The LBP code is calculated for each pixel in the cropped portion of the image, and the distribution of the codes is used as a feature vector, denoted by S [33]:

$$S = t\left(LBP_{P,R}(x, y)\right), x \in \{\lceil R \rceil, \ldots, N-1-\lceil R \rceil\}, y \in \{\lceil R \rceil, \ldots, M-1-\lceil R \rceil\} \tag{4.3}$$

4.2 Uniform Local Binary Patterns

Uniform Local Binary Patterns are an extension to the original LBP and is used to reduce the length of the feature vector and a simple rotation-invariant descriptor. A local binary pattern is called uniform if the binary patterns contain at most two bitwise transitions (0/1 or 1/0) when considered as a circular binary string. For example, the patterns 00000000 (0 transitions), 11110001111 (2 transitions) and 000111000000 (2 transitions) are uniform as long as the patterns 111110001001 (4 transitions) and 000010100100 (6 transitions) are not. In uniform LBP mapping, there is a separate output for each uniform pattern and all the non-uniform patterns are labeled with a single label. An example of uniform pattern can be explained in a simple way as demonstrated in Table 1.

Table 1. An example of how Uniform LBP works

Pattern	Circular transition	Uniform	Label
00000000	0 transitions	Yes	Label 0
00011110	2 transitions	Yes	Label 1
11001111	2 transitions	Yes	Label 3
11001001	4 transitions	No	Label 4
01010010	6 transitions	No	Label 4

The number of different output labels for mapping for patterns of p bits is calculated as follows:

$$\text{No. of bits} = p(p-1) + 3 \tag{4.4}$$

where p is the number of bits. For example, the uniform mapping generates 59 labels for neighborhoods of 8 sampling points and 243 labels for neighborhoods of 16 sampling points. Figure 4 indicates the uniform pattern LBP in (8, R) neighborhood.

The reasons of omitting the non-uniform patterns are twofold. First, most of the local binary patterns in natural images are uniform and the next reason is the statistical robustness. Uniform patterns have produced better recognition instead of all the possible patterns. Uniform patterns are more stable, i.e. less prone to noise. On the other hand, considering only uniform patterns makes the number of possible LBP labels significantly lower and reliable, estimation of their distribution requires few samples [34]. Figure 5 shows an example of the 58 different uniform patterns in (8,R) neighborhoods.

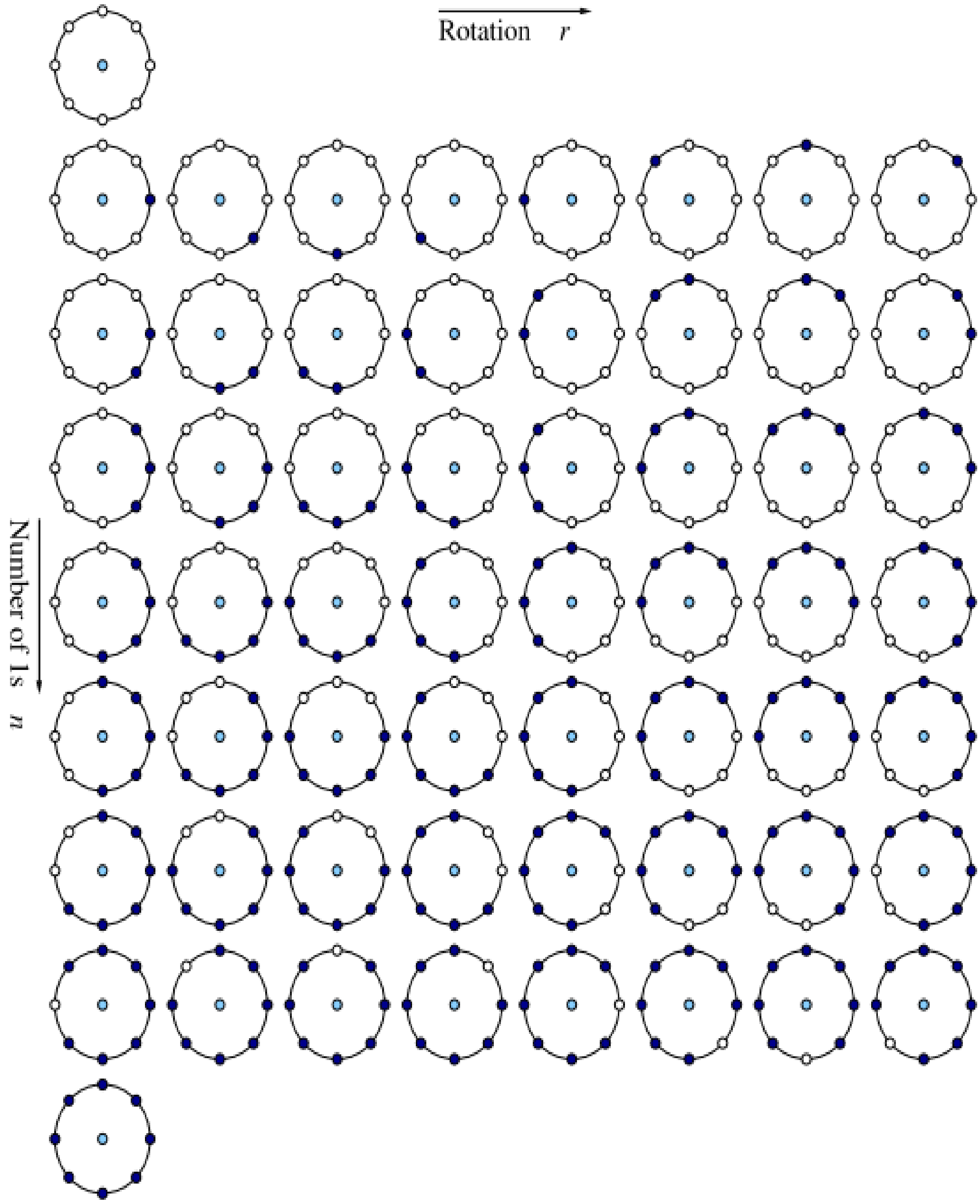

Figure 5. The 58 different uniform patterns in (8, R) neighborhood [34].

4.3 Completed Local Binary Patterns

Completed Local Binary Pattern (CLBP) is a generalized version of LBP which is proposed by Z. Guo et al. [35] and it has been proved to be effective on texture analysis [34] and is one of the best performers.

CLBP adds contrast information to the final feature histogram (LBP does it too). Apart, instead of using sign [-1, 0, 1], it uses magnitude information. In CLBP, a local region is represented by Center pixel and the difference between the values with local Center pixel with magnitude that is called as Local Difference Sign-Magnitude Transform (LDSMT). CLBP has three different components namely CLBP_S, CLBP_M and CLBP_C. CLBP-S indicates the sign which can be positive or negative, of difference between the local pixel and center pixel, CLBP-M illustrates the magnitude of the difference between the Center pixel and local pixel and CLBP-C shows the difference between local pixel value and average central pixel value.

4.3.1 Local Difference Sign-Magnitude Transform

Given a central pixel g_c and its P circularly and evenly spaced neighbors $g_p, p = 0,1,\dots,P-1$, we can simply calculate the difference between g_c and g_p as $d_p = g_p - g_c$. The local difference vector $[d_0,\dots,d_{P-1}]$ characterizes the image local structure at g_c. Because the central gray level g_c is removed, $[d_0,\dots,d_{P-1}]$ is robust to illumination changes and they are more efficient than the original image in pattern matching. d_p can be further decomposed into two components [34] as follows:

$$d_p = s_p * m_p \text{ and } \begin{cases} s_p = \text{sign}(d_p) \\ m_p = |d_p| \end{cases}$$

$$S_p = \begin{cases} 1, & d_p \geq 0 \\ -1, & d_p < 0 \end{cases} \quad (4.5)$$

where S_p is the sign of d_p and m_p is the magnitude of d_p. $[d_0, \dots, d_{P-1}]$ is transformed into a sign vector $[s_0, \dots, s_{P-1}]$ and a magnitude vector $[m_0, \dots, m_{P-1}]$. Obviously, $[s_0, \dots, s_{P-1}]$ and $[m_0, \dots, m_{P-1}]$ are complementary and the original difference vector $[d_0, \dots, d_{P-1}]$ can be perfectly reconstructed from them. Figure 6 shows an example. Figure 7(a) is the original 3 × 3 local structure with central pixel being 25. Figure 7(b) is The difference vector which is [3, 9, _13, _16, _15, 74, 39, 31]. After Local Differences Sign-Magnitude Transform, the sign vector which is shown in Figure 7(c) is [1, 1, _1, _1, _1, 1, 1, 1] and the magnitude vector which is shown in Figure 7(d) is [3, 9, 13, 16, 15, 74, 39, 31]. It is clearly seen that the original LBP uses only the sign vector to code the local pattern as an 8-bit string "11000111" ("-1" is coded as "0").

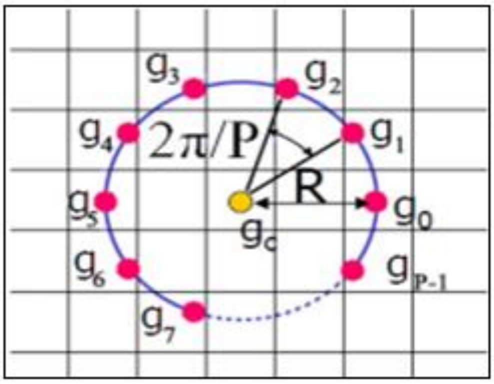

Figure 6. Central pixel and its P circularly and evenly spaced neighbors with radius R [34].

9	12	34
10	**25**	28
99	64	56

(a)

-16	-13	9
-15		3
74	39	31

(b)

-1	-1	1
-1		1
1	1	1

(c)

16	13	9
15		3
74	39	31

(d)

Figure 7. (a) 3 × 3 sample block; (b) the local differences; (c) the sign; and (d) Magnitude components [34].

4.3.2 CLBP_S, CLBP_M and CLBP_C operators

The sign component preserves the information of local difference. This explains why the simple LBP technique can reasonably represent the image local features. Meanwhile, the magnitude component may contribute additional discriminant information if it is properly used. In addition, the intensity value of the center pixel itself can also have useful information [36], [37]. In this subsection, a completed LBP (CLBP) framework to explore all the three types of features is presented. The CLBP framework is illustrated in Figure 6. The original image as its center gray level (C) and the local difference are represented. The local difference is then decomposed into the sign (S) and magnitude (M) components by the LDSMT. Consequently, three operators, namely CLBP_C, CLBP_S and CLBP_M, are used to code the C, S, and M features, respectively. Then, the CLBP_C, CLBP_S, and CLBP_M codes are combined to form the CLBP feature map of the original image. Finally, a CLBP histogram can be built, and some classifier, such as the nearest neighborhood classifier, can be used for texture classification. The CLBP_S operator is the same as the original LBP operator defined. Since the M components are of continuous values instead of the binary "1" and "-1" values, they cannot be directly coded as that of S. Inspired by the coding strategy of CLBP_S (i.e., LBP) and in order to code M in a consistent format with that of S, CLBP_M operator is defined as follows:

$$\mathrm{CLBP\_M}_{P,R} = \sum_{p=0}^{P-1} t(m_p, c)2^p,$$

$$t(x, c) = \begin{cases} 1, x \geq c \\ 0, x < c \end{cases} \tag{4.6}$$

where c is a threshold to be determined adaptively. Both CLBP_S and CLBP_M produce binary strings so that they can be conveniently used together for pattern classification. There are two ways to combine the CLBP_S and CLBP_M codes: in concatenation or jointly. In the first way, the histograms of the CLBP_S and CLBP_M codes are calculated separately, and the two histograms are concatenated together. This CLBP scheme can be represented as "CLBP_S_M". In the second way, a joint 2-D histogram of the CLBP_S and CLBP_M codes is calculated. This CLBP scheme is represented as "CLBP_S/M". The center pixel, which expresses the image local gray level, also has discriminant information. In order to make it consistent with CLBP_S and CLBP_M, it is defined as

$$\mathrm{CLBP\_S}_{P,R} = t(g_c, c_t) \tag{4.7}$$

where t is defined in the previous equation and the threshold c is set as the average gray level of the whole image. The three operators, CLBP_S, CLBP_M, and CLBP_C, could be combined in two ways, jointly or hybridly. In the first way, similar to the 2-D joint histogram, a 3-D joint histogram can be built, denoted by "CLBP_S/M/C". In the second way, a 2-D joint histogram, "CLBP_S/C" or "CLBP_M/C" is built first, and then the histogram is converted to a 1-D histogram, which is then concatenated with CLBP_M or CLBP_S to generate a joint histogram, denoted by "CLBP_M_S/C" or "CLBP_S_M/C".

4.4 Rotational Invariant Local Binary Patterns

The rotational invariant Local Binary Patterns (RLBP) is calculated by circular bitwise rotation of the local LBP to find the minimum binary value. The minimum value LBP is used as rotation invariant signature and is recorded in the histogram bins. The RLBP is computationally very efficient.

Rotations of a textured input image causes the LBP patterns to translate into a different location and to rotate about their origin.

Computing the histogram of LBP codes normalizes for translation, and normalization for rotation is achieved by rotation invariant mapping. In this mapping, each LBP binary code is circularly rotated into its minimum value

$$\mathrm{LBP}_{\mathrm{P,R}}^{\mathrm{ri}} = \min_{\mathrm{i}} \mathrm{ROR}(\mathrm{LBP}_{\mathrm{P,R}}, \mathrm{i}) \quad (4.7)$$

where ROR (x, i) denotes the circular bitwise right rotation of bit sequence x by i steps. For instance, 8-bit LBP codes 10000010b, 00101000b, and 00000101b all map to the minimum code 00000101b. Omitting sampling artifacts, the histogram of $\mathrm{LBP}_{\mathrm{P,R}}^{\mathrm{ri}}$ codes is invariant only to rotations of input image by angles $\frac{360}{\mathrm{P}}$ such that a = 0, 1, . . . , P − 1. However, classification experiments show that this descriptor is very robust to in-plane rotations of images by any angle [38].

In order to illustrate the method, Figure 8 shows a pattern of three consecutive LBP bits, in order to make this descriptor rotation invariant, the value is left-shifted until minimum value is reached.

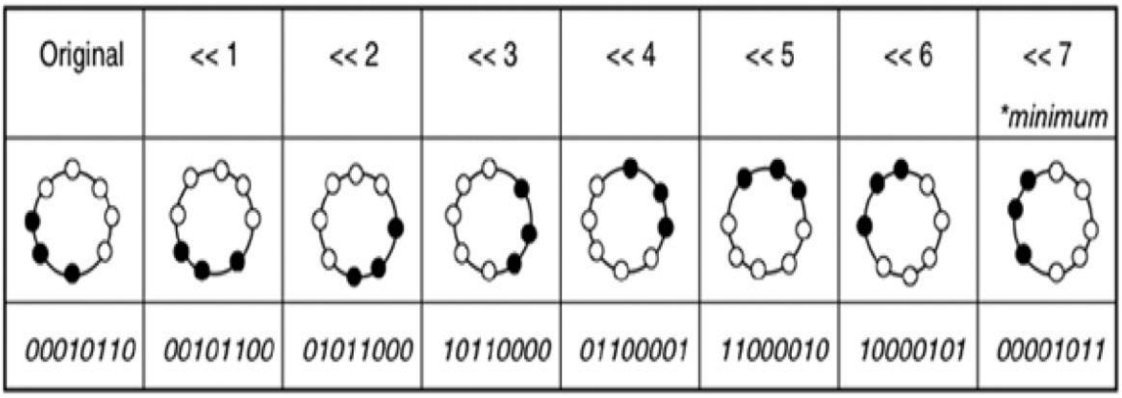

Figure 8. Calculation of the minimum LBP by using circular bit shifting of the binary value to find the minimum value [39].

Chapter 5

DATASETS

There is a great number of face databases available such as FERET, LFW, ORL, AR, BioID, FRGC, SCface, PIE and etc. Each of these databases has a role in the problems of face recognition or face detection and range in size, scope and purpose. The photographs in many of these databases were acquired by small teams of researchers specifically for the purpose of studying face recognition. Acquisition of a face database over a short time and particular location has advantages for certain areas of research, giving the experimenter direct control over the parameters of variability in the database.

Most face databases have been created under controlled conditions to facilitate the study of specific parameters on the face recognition problem. These parameters include such variables as position, pose, lighting, background, camera quality, and gender. While there are many applications for face recognition technology in which one can control the parameters of image acquisition, there are also many applications in which the practitioner has little or no control over such parameters.

In this chapter, we summarized several publicly available databases that have been used for estimating gender. The number of images and the number of subjects in each database are illustrated in Table 2. The controlled variations during collection of the data are also shown. These databases were collected for evaluating face recognition or detection systems; hence gender labels may not be available. Therefore, researchers need to label ground truth (original images) using visual inspection by hand.

For training and estimating their gender, some researchers take only a subset of the databases (e.g. only frontal images, without background clutter), or, to use a large amount of images, combine several datasets. Also, in some datasets, researchers need to detect faces by using some face detectors such as Viola and Jones face detector.

Table 2. Publicly available face datasets (P: pose or view, L: lightening or illumination, X: expression, O: occlusion)

Dataset	Number of images	Number of subjects	Gender labels	Controlled variations
FERET	14126	1199	No	P, L, X
LFW	13233	5749	No	Uncontrolled
AR	<4000	126	Yes	X, L, O
ORL	400	40	No	P, L, X, O
Standard FERET subset	3541	1196	No	P, L, X

The face datasets from different databases are used in this thesis to demonstrate the performance of different approaches for gender recognition. These facial databases and datasets are described below.

5.1 FERET Database

FERET [40] has been established and widely used for estimation of face recognition systems, and also has been used by many researchers for gender classification. The aim of the FERET program is to develop algorithms on a common database and to report results in the literature using this database. Results reported in the literature did not provide a direct comparison among algorithms because each researcher reported results using different assumptions, scoring methods, and images. The independently administered FERET evaluations allowed for a direct quantitative assessment of the relative strengths and weaknesses of different approaches.

More importantly, the FERET database and evaluations clarified the state-of-the-art in face recognition and pointed out general directions for future research. The FERET evaluations allowed the computer vision community to assess overall strengths and weaknesses in the field, not only on the basis of the performance of an individual algorithm, but in addition on the aggregate performance of all algorithms tested. Through this type of assessment, the community learned in an unbiased and open manner of the important technical problems that needed to be addressed.

FERET database holds 14,126 images of 1199 subjects and there is no gender annotation, though for a subset of images with 199 females and 212 males, this has been made available by Makinen and Raisamo [3]. The faces have a variety pose, and some variation in expression and illumination. The images are noise-free, without background clutter and have consistent lighting. Figure 9 illustrates (5 subjects with 7 sample face images with under illumination and expression) cropped and resized images, after face detection, from the FERET database.

Figure 9. Sample face images from the FERET Database

5.2 Labeled Faces in the Wild (LFW) Database

Labeled Faces in the Wild (LFW) [41] has been collected to help the study of unconstrained face recognition. The database contains faces captured in uncontrolled conditions, showing a large range of variation typically encountered in everyday life, exhibiting natural variability in factors such as lighting, occlusion, accessories, race, and background. LFW has more than 1300 images (10256 females, 29771 males), with the number of males outnumbering females significantly by roughly 3 times, and with many subjects appearing more than once. Furthermore, it contains mostly of public figures such as politicians and celebrities. Figure 10 shows sample face images before face detection from LFW database.

The primary contribution of LFW is providing a large set of relatively unconstrained face images. By unconstrained, we mean faces that show a large range of the variation seen in everyday life. This includes variation in pose, lighting, expression, background, race, ethnicity, age, gender, clothing, hairstyles, camera quality, color saturation, and other parameters. The reason we are interested in natural variation is that for many tasks, face recognition must operate in real-world situations where we have little control over the composition, or the images are pre-existing. For example, there is a wealth of unconstrained face images on the Internet, and developing recognition algorithms capable of handling such data would be extremely beneficial for information retrieval and data mining. In contrast to LFW, existing face databases contain more limited and carefully controlled variation. LFW fills an important gap for the problem of unconstrained face recognition. Figure 10 shows some sample face images from LFW database.

Figure 10. Some captured face images from LFW database

5.3 AR Database

AR database [42] contains over 4,000 color images corresponding to 116 people's faces (53 females and 63 males). All faces correspond to frontal view faces with different facial expression, different illumination conditions and with different characteristic changes (people wearing sun-glasses or scarf). The images were captured at the Computer Vision Center (CVC) at University of Barcelona in Barcelona, under strictly controlled conditions in two sessions per person. There were no restrictions to wear (clothes, glasses), makeup, hair style and etc.

All images are acquired using the same system. The camera parameters, the illumination conditions and the distance from the subject to the camera are strictly controlled during the whole acquisition process. In order to guarantee such a strict, set up, they calibrated the system twice a day: one time in the morning and another time in the afternoon. Calibration is done referring to distance, pose and illumination conditions. Each person participated in two sessions, separated by two week times. The same images were captured in both sessions. Figure 11 shows sample face images from AR database.

Figure 11. Sample face images from the AR Database

5.4 ORL Database

ORL database [43] contains a set of face images taken during two years (April 1992 and April 1994) at the laboratories of Computer Engineering Department at Cambridge University. The database was established to use in the context of a face recognition project carried out in collaboration with the Vision and Robotics Group of the Cambridge University Engineering Department. There are 400 images corresponding to 40 individual's face. For some subjects, the images were taken at different times, varying the facial expressions (open / closed eyes, smiling / not smiling) and occlusion (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position. Figure 12 shows sample face images from ORL database.

Figure 12. Part of the face image from the ORL database

5.5 Standard FERET Subsets

The FERET program ran from 1993 through 1997. Sponsored by the Department of Defense's Counterdrug Technology Development Program through the Defense Advanced Research Products Agency (DARPA), its primary mission was to develop automatic face recognition capabilities that could be employed to assist security, intelligence and law enforcement personnel in the performance of their duties [40].

The FERET image corpus was assembled to support government monitored testing and evaluation of face recognition algorithms using standardized tests and procedures. The final corpus, presented here, consists of 14051 eight-bit grayscale images of human heads with views ranging from frontal to left and right profiles.

Standard FERET subset contains many images which is divided into some specific parts as shown in Table 3.

Table 3. Standard FERET Subsets

Subsets	Duplicate 1	Duplicate 2	fb	fc	Gallery
Number of images	239	73	1195	194	1196
Number of samples for female	2	2	1	1	1
Number of samples for male	2	2	1	1	1
Number of Train subjects	1196	1196	1196	1196	1196
Number of Test subjects	478	146	1198	194	-
Number of Female subjects	75	21	475	88	477
Number of Male subjects	164	52	723	106	719

Duplicate I set holds 239 images whose matches were taken between 0 and 1031 days after the match. The median is 72 days and the mean is 251 days. The Duplicate II set contains 73 images from subjects whose gallery match was taken between 540 and 1031 days beforehand. The median is 569 and the mean is 627 days. Fb contains 194 alternative facial expressions and fc contains 194 face images under illumination. All these four subsets are used as test sets and the Gallery dataset that contains 1196 images is used for training set.

Chapter 6

EXPERIMENTS AND RESULTS

In this chapter, we will provide information about our implementation and comparison of the performance of the Local Binary Patterns and its variants on different databases and variations.

6.1 Experimental Setup

We conduct experiments first by using part of FERET images in the training set. We chose 1000 images corresponding to 200 individuals (92 females and 108 males).

Later, we carried out the experiments by using other databases as follows:

- LFW database contains 13233 images. We selected 80 individual's face images (40 females and 40 males), 4 sample images for each subject.
- AR database contains 1700 color face photographs. We used all the images (50 females and 50 males) with 17 sample images for each individual.
- ORL dataset contains 400 gray scale face images (40 females and 40 males) that we used all of them.

A summary on the organization of experiment sets are shown in Figure 13. There are mainly two set of experiments, each comparing LBP and its variants' performance on either FERET, LFW, ORL, AR datasets or on the standard FERET subsets.

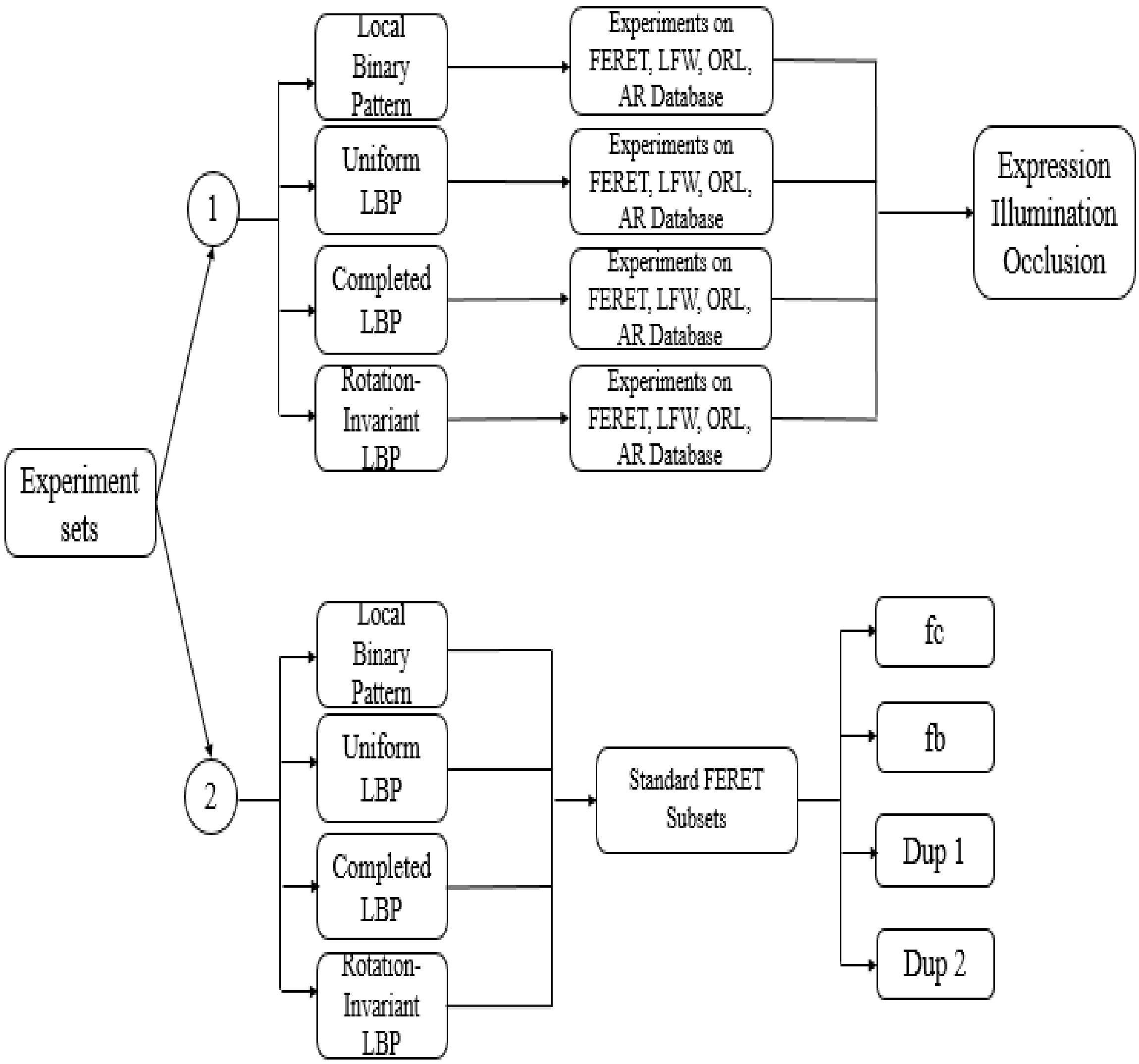

Figure 13. Organization of Experiment sets

For all the databases we put 80% of images for training set and 20% of the images for testing set. All information about databases are shown in Table 4.

We manually labeled the ground truth (original images) regarding gender for each face for all databases except AR database. Images were detected by Viola and Jones face detector for LFW and FERET database. Therefore, they only include the head of individuals.

Table 4. Databases used in Experiments

Database	FERET	LFW	AR	ORL
Number of subjects	200	80	100	40
Number of samples for female	5	4	17	10
Number of samples for male	5	4	17	10
Number of Train subjects	600	240	1020	330
Number of Test subjects	175	80	340	60
Number of Female subjects	80	20	40	4
Number of Male subjects	120	20	40	36

Generally, we have two main steps that exist in order to classify facial images. Training phase is the first step in which first of all a dataset of resized and cropped images are given to the feature extraction algorithm, then the output of train images is obtained. Similarly, in test phase a dataset of resized and cropped images are given to the feature extraction algorithm, then the output of test images is available to be compared with the output of train images.

First, as it is explained above, we divided the dataset into 2 subsets (training set and testing set) of similar size, keeping the same ratio between male and female. The images of a particular individual appear only once in one subset.

Each image that is given to the algorithm as an input may have different size, therefore, sometimes it is necessary to unify size of all images. In order to divide images to equal size with proper values, the images are resized to 78×63, 80×65, 78×60 and 77×63 pixels for 9, 25, 36 and 49 partitions respectively. In our work, we decided to scale down the images to 65×80 pixels from the original sizes and divide each image into 25 sub-region (blocks).

Later, we applied LBP, ULBP, CLBP and RLBP in different experiments to each block to extract the features first for train dataset and then for the test dataset. First we applied LBP descriptor to extract the features for each block. Thus each face image was described by LBP histogram of the 6400 (25 × 256) bins. So, Local Binary Pattern histograms are extracted and concatenated into one feature histogram to represent the whole image. As a next feature extractor we used Uniform LBP which is used to reduce the length of feature vector. Therefore, gender classification using (1,8) and (2,8) neighborhoods with uniform pattern is performed for female and male facial images. Completed LBP is a generalized version of LBP which is more effective on texture analysis which is applied as the third feature extractor.

Rotation invariant LBP is also used to increase the discriminative power of the LBP operator. The results are in (%) and are shown in Table 5 for the classification rates.
In the last step, after calculating LBP and its variants for each block, we concatenated them into a single vector. Then we used Manhattan Distance Measure to find the minimum distance between the test set and training set and compare them for gender recognition. We calculated the classification accuracy as follows:

$$\text{Accuracy} = (\text{number of correct images} / \text{number of all images}) \times 100 \qquad (6.1)$$

6.2 Gender Classification Results on Facial Images

The first set of experiments is done with Local Binary Patterns and its variants such as Uniform LBP, Completed LBP and Rotation-Invariant LBP on 4 large databases namely FERET, LFW, AR and ORL database which contain face images under controlled conditions and uncontrolled conditions. The facial images from all subsets of the databases are divided into 25 blocks. Local Binary Patterns and its variants' histograms are extracted and concatenated into one feature histogram to represent the whole image. Gender classification using (2, 8) neighborhoods with uniform pattern is performed for female and male facial images. The overall gender classification performances for 4 different datasets of facial images are presented in Table 3. We used Manhattan Distance Measure for LBP and its variants for gender classification experiments in this study.

We measured accuracies of the systems in each test setup separately and in this section we report the results of our work. The average classification accuracy on female and male images are calculated and shown in Table 5.

Table 5. Results of the Local Binary Patterns, Uniform LBP, Completed LBP, Rotation-invariant LBP

Method	Accuracy (%) on the dataset of			
	FERET	LFW	AR	ORL
LBP	94.64	83.37	78.82	98.01
CLBP	**96.60**	**85.83**	79.12	97.23
ULBP	89.10	62.68	77.05	**98.50**
RLBP	84.99	83.33	**80.00**	93.17

According to these experiments, the classification rates between approaches are slightly different on each dataset. However, CLBP has better performance compared to the other approaches on both FERET and LFW datasets.

LBP, Completed LBP, Uniform LBP and Rotation-Invariant LBP gender classification results are different on FERET, LFW, AR and ORL database. As it is shown in the above table, CLBP has the highest accuracy for FERET and LFW database, achieving the classification rate of 96.60% and 85.83%. For AR and ORL database the best accuracies belong to RLBP and ULBP respectively, achieving classification rate of 80.00% and 98.50%. Therefore, CLBP provides better performance compared to the other the approaches.

6.3 Gender Classification Results under Illumination, Expression and Occlusions

Over the last decade, computer vision researchers have improved the performance of recognition systems on a variety of challenging face recognition tasks. The real challenge in the face detection and recognition technologies is the ability to handle all those scenarios where subjects are non-cooperative and acquisition phase is unconstrained. There are many factors that cause the appearance of the face to vary. These sources of variation in the facial appearance can be categorized into two groups namely intrinsic factors and extrinsic factors. Intrinsic factors are due purely to the physical nature of the face and are independent of the observers. Extrinsic ones cause the appearance of the face to alter via the interaction of light with the face and the observer.

The common problems and challenges that a recognition system can have while detecting and recognizing are pose, illumination, facial expression and occlusion. Pose (pose of a face changes with viewing and relation in the head position) and illumination (light variation) correspond to extrinsic factors while facial expression (emotions) and occlusion (blockage, when the whole face is not available) correspond to intrinsic factors.

In this section, we illustrate gender classification results under illumination, occlusion and expression for each database separately. We have conducted the experiments for this work in the same way as described in the previous sections. The gender classification performances for 4 different databases on variations of facial images are presented in Tables 6 to 9.

Table 6. Recognition rate under expression, illumination and occlusion on FERET Database

FERET - Database	Expression	Illumination	Occlusion
Number of Samples for female	4	4	4
Number of Samples for male	4	4	4
Number of female subjects	29	7	18
Number of male subjects	30	50	36
Number of Train	160	136	152
Number of Test	76	92	64
CLBP	79.34	**63.07**	67.41
LBP	78.8	58.81	57.6
ULBP	**88.4**	54.6	**80.4**
Rotation - invariant LBP	63.54	54.6	49.43

Table 6 shows results of gender recognition with FERET dataset under illumination, occlusion and expression. According to that table, ULBP achieves better performance compared to the other approaches for the experiment set with facial expressions, achieving classification rate of 88.4%. In the experiment set with facial images having illumination, CLBP has the best performance with 63.07% accuracy. RLBP shows better performance for facial images with occlusion, achieving classification rate of 80.4%.

Table 7. Recognition rate under expression, illumination and occlusion on LFW Database

LFW - Database	Expression	Illumination	Occlusion
Number of Samples for female	4	4	4
Number of Samples for male	4	4	4
Number of female subjects	16	15	2
Number of male subjects	18	33	20
Number of Train	80	140	64
Number of Test	56	52	24
CLBP	**85.61**	67.10	79.17
LBP	62.9	52.63	77.15
ULBP	56.63	65.78	66.7
Rotation - invariant LBP	63.3	**71.05**	**83.33**

Table 7 indicates results of gender recognition with Labeled Faces in the Wild (LFW) dataset under illumination, occlusion and expression. We summarized the results of LBP and its variants on variations in that table. It is observed that CLBP achieved the classification rate of 85.61% for facial images under expression which is better than other approaches. On the other hand, for the experiments corresponding to illumination and occlusion, RLBP has the best performance, achieving classification accuracy of 71.05% and 83.33%. In this uncontrolled database, RLBP shows better performance for facial images with occlusion and illumination changes.

Table 8. Recognition rate under expression, illumination and occlusion on AR Database

AR - Database	Expression	Illumination	Occlusion
Number of Samples for female	6	6	4
Number of Samples for male	6	6	4
Number of female subjects	50	50	50
Number of male subjects	50	50	50
Number of Train	360	360	360
Number of Test	240	240	240
CLBP	85.00	**83.75**	**89.4**
LBP	74.58	78.33	57.62
ULBP	**87.91**	80.83	83.75
Rotation - invariant LBP	80.00	79.6	80.62

Table 8 indicates the classification rate of different conditions with AR dataset. We summarized the results of LBP and its variants on variations such as expression, illumination and occlusion in that table. It is observed that CLBP achieved the classification rate of 83.75% and 89.4 for facial images with illumination and occlusion, respectively which is better than the other approaches. On the other hand, ULBP has the best performance, achieving classification accuracy of 87.91% for facial images under expression. In that case, CLBP shows better performance for facial images with occlusion and illumination changes.

Table 9. Recognition rate under expression, illumination and occlusion on ORL Database

ORL - Database	**Expression**	**Illumination**	**Occlusion**
Number of Samples for female	3	3	3
Number of Samples for male	3	3	10
Number of female subjects	4	4	2
Number of male subjects	27	15	15
Number of Train	60	39	27
Number of Test	33	18	12
CLBP	72.6	**91.31**	76.7
LBP	66.7	65.51	80.00
ULBP	**73.60**	76.83	83.33
Rotation - invariant LBP	70.36	76.83	**91.7**

Table 9 shows the results of gender recognition with ORL dataset under illumination, occlusion and expression. For the images with facial expression, ULBP achieved better accuracy compared to the other approaches, achieving classification rate of 73.60%. In the experiment using facial images with illumination, CLBP has the highest accuracy. Besides, for facial images with occlusion, RLBP achieved the best performance for gender classification.

In this study, we presented the performance of Local Binary Patterns and its variants as feature extractors and to classify genders we used Manhattan Distance Measure. The presented results are based on testing on FERET, LFW, AR and ORL database (80% training set and 20% testing set). We also evaluated LBP approach and its variants on these databases that include illumination, expression and occlusion for gender recognition. On the average, the experimental results indicate that recognition rate for female is consistently about 5% higher than the recognition rate of males.

The experimental results demonstrate that there were no significant differences in the classification rates between approaches. There were small differences between methods when calculated from the several test setups with different databases and under illumination, expression and occlusion. LBP features are used and based on the results reported, good performance has been achieved. The first presented results were gained when the FERET images were used for training and we got a 96.60% accuracy that belongs to CLBP. The results indicate that the best accuracy for LFW database also belongs to CLBP. Therefore, CLBP worked better in these two databases because FERET and LFW databases hold more images under illumination. Also when Standard FERET database was used, CLBP has the highest accuracy in fc subset which holds images corresponding to the facial images under illumination. We achieved classification rate of 80.00% in AR database with RLBP which was the highest one. RLBP has better performance compared to the other approaches for images with occlusion, it had also better accuracy based on the results presented in the previous subsections for databases which contains more images with occlusion. As presented in the previous subsections, ULBP achieved better results on ORL database and also on images that were captured under occlusion.

6.4 Gender Classification Results on Standard Subsets

We have conducted a set of experiments on Standard FERET subset in the same way as described in the previous sections. For Duplicate 1 and Duplicate 2 which contain the images which are captured during 18 months, the CLBP method achieved better classification rate compared to the other methods. ULBP has the highest accuracy for fb subset which holds facial expression images. In fc subset, we have images corresponding to the facial images under illumination. CLBP has the highest accuracy for fc subset. LBP gave the highest accuracy for fc subset. Table 9 illustrates the results of Standard FERET subsets.

Table 10. Gender Classification Accuracy on Standard FERET subset

Standard FERET subsets	Accuracy (%) on			
	LBP	CLBP	ULBP	RLBP
Dup 1	63.63	**84.67**	78.91	73.97
Dup 2	60.28	**72.73**	69.85	67.79
Fb	82.36	93.18	**97.21**	92.36
Fc	74.97	**75.23**	61.14	59.33

Gender classification results on the standard FERET subsets demonstrate that CLBP method achieves the best accuracy on the images having illumination and on the images taken in different time periods.

Chapter 7

CONCLUSION

As humans, we are able to recognize people's gender from their face and facial images and are often able to be quite precise in this estimation but computers cannot detect it. Gender recognition is a fundamental task for human beings, as many social functions critically depend on the correct gender perception. Automatic gender classification has many important applications, for example, intelligent user interface, visual surveillance, collecting demographic statistics for marketing, etc. Human faces provide important visual information for gender perception. Gender classification from face images has received much research interest in the last two decades.

In this thesis, we carried out experiments on several state-of-the-art gender classification methods. We compared the performance of Local Binary Patterns approach and its variants with different databases and variations in illumination, expression and occlusion. We used different databases for face images under controlled conditions and uncontrolled conditions namely FERET, LFW, AR and ORL database to compare the approaches. We divided each database into three subsets including illumination, expression and occlusion and applied LBP and its variants. We also used Standard FERET subsets to find which method or approach has the best performance on several variations.

First of all, we found that the database used affected the classification accuracy a lot and this should be taken into account when doing experiments. Second, we found that

CLBP has better performance compared to the other approaches and it is achieving the best performance for images with variations in illumination. For the other variations such as expression and occlusion, in general, ULBP and RLBP have better accuracy for expression and occlusion, respectively.

As a future work, the experiments can be extended on the other uncontrolled datasets. Gender classification may be tested on both controlled and uncontrolled large datasets that are publicly available. Additionally, a more powerful classifier such as SVM can be used for further experiments.

REFERENCES

[1] N. ChoonBoong,Y. Tay & B. Goi.(2015) A review of facial gender recognition: A survey. Volume abs/139 - 755.

[2] N. Choon Boong, Y. Tay & B. Goi. (2012) Vision - based human gender Recognition, A survey. Volume abs/1204-1611.

[3] E. Makinen & R. Raisamo. (2008) An experimental comparison of gender classification methods. volume 29, p.p. 1544 - 1556.

[4] W. Garrison Cottrell & J. Empath. (1990) Face, emotion, and gender recognition using holons. p.p. 564 - 571.

[5] B. A. Golomb, D. T. Lawrence & T. J. Sejnowski. Sexnet. (1990) A neural network Identifies sex from human faces. *In Proceedings of the 1990 Conference on Advances in Neural Information Processing Systems 3,* NIPS-3, p.p. 572 – 577.

[6] R. Brunelli & T. Poggio. (1993) Face recognition: features versus templates pattern Analysis and Machine Intelligence, *IEEE Transactions* on, 15(10):1042 – 1052.

[7] H. Abdi, D. Valentin, B. Edelman & J. O'Toole. (1995) More about the difference between men and women: Evidence from linear neural networks and the principal - component approach.

[8] L. Wiskott, J. Fellous Z, & Y. Norbert Kruger. (1995) Face recognition and gender Determination.

[9] M. J. Lyons, J. Budynek, A. Plante & S. Akamatsu. (2000) Classifying facial attributes using a 2-d gabor wavelet representation and discriminant analysis. p.p. 202 – 207.

[10] Z. Sun, G. Bebis, Xiaojing Yuan & S. J. Louis. (2002) Genetic feature subset selection for Gender classification: a comparison study. p.p 165-170.

[11] B. Moghaddam & Ming-Hsuan Yang. (2002) Learning gender with support faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. p.p. 707 – 711.

[12] A. Jain & J. Huang. (2004) Integrating independent components and linear discriminant Analysis for gender classification. p.p. 159 – 163.

[13] N. P. Costen, M. Brown & S. Akamatsu. (2004) Sparse models for gender Classification. p.p. 201 – 206.

[14] N. Sun, W. Zheng, C. Sun, C. Zou & L. Zhao. (2006) Gender classification based on boosting local binary pattern. p.p.194 – 201.

[15] R. Rojas. (2009) Adaboost and the super bowl of classifiers a tutorial introduction to adaptive boosting.

[16] Y. Saatci & C. Town. (2006) Cascaded classification of gender and facial expression using active appearance models. p.p. 393 – 398.

[17] L. Bourdev, J. Malik & S. Maji. (2011) Describing people: A poselet - based approach to attribute classification. *IEEE International Conference on Computer vision*, (ICCV), p.p. 1543 – 1550.

[18] X. Fu, G. Dai, C. Wang & L. Zhang. (2010) Centralized Gabor gradient histogram for Facial gender recognition. p.p. 2070 - 2074.

[19] T. C. I. Lin & Yi. Zhao.(2011) A feature - based gender recognition method based on color information. p.p. 40 – 43.

[20] D. Chen and P. C. Hsieh. (2012) Face-based gender recognition using Compressive sensing. p.p. 157 – 161.

[21] I. Ullah, M. Hussain, H. Aboalsamh, G. Muhammad, A. Mirza & G. bebis. (2012) Gender recognition from face images with dyadic wavelet transform and local binary pattern. p.p. 409 – 419.

[22] J.E. Tapia & C.A. Perez. (2013) Gender classification based on fusion of different spatial scale features selected by mutual information from histogram of LBP, intensity and shape. *IEEE Transaction on Information Forensics and Security*, p.p. 488 – 499.

[23] C. Ding & H. Peng. (2003) Minimum redundancy feature selection from microarray gene expression data. p.p. 523 – 528.

[24] H. Cheng, Z. Qin, W. Qian & W. (2008) Liu. Conditional mutual information based feature selection. p.p. 103 – 107.

[25] P. A. Estevez, M. Tesmer, C. A. Perez, & J. M. Zurada. (2009) Normalized mutual, information feature selection. *IEEE Transactions on Neural Networks*, p.p. 189 – 202.

[26] J. G. Wang C. Y. Lee, J. Li & W. Y. Yau. (2010) Dense sift and gabor descriptors-Based face representation with applications to gender recognition. *International Conference on Control Automation Robotics and Vision*, p.p. 1860 – 1864.

[27] A. Lanitis, C.J. Taylor & T.F. Cootes. (2002) Toward automatic simulation of aging effect on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p.p. 442 – 455.

[28] A. Lanitis. (2002) On the significance of different facial parts for automatic age Estimation. p.p. 1027 - 1030 vol.2.

[29] L.L. Yu, D. Gonzalez-Jimenez, J. L. Alba-Castro & P. Dago-Casas. (2001) Single and Cross database benchmarks for gender classification under unconstrained settings. *IEEE International Conference on Computer Vision Workshops*, p.p. 2152 – 2159.

[30] T. Ojala, M. Pietikäinen & D. Harwood (1994), "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", *Proceedings of the 12th IAPR International Conference on Pattern Recognition,* (ICPR 1994), vol. 1, pp. 582 - 585.

[31] J. Huang, B. Heisele & V. Blanz. (2003) Component-based Face Recognition with 3D Morphable Models, *AVBPA.*

[32] Z. Gengtao, Z. Yongzhao & Z. Jianming. (2006)"Facial Expression Recognition based on Selective Feature Extraction". *Sixth International Conference on intelligent Systems Design and Applications*, 2006. ISDA '06 . Volume 2, p.p. 412 - 417.

[33] T. Ojala, M. Pietikainen & D. Harwood. (1996) "A Comparative Study of Texture Measure with Classification Based on Feature Distribution," pattern Recognition, vol. 29, No. 1, pp. 51-59.

[34] L. Nanni & A. Lumini. (2010) A Local Binary approach based on local binary Patterns and its Variants texture descriptor. Expert syst. Appl. 37, p.p.7888 - 7897

[35] Z. Guo, L. Zhang & D, Zhang. (2010) A Completed Modeling of Local Binary Pattern Operator for Texture Classification, "*IEEE TRANSACTIONS ON IMAGE PROCESSING"*, vol. 19, no. 6, pp. 1657 – 1663.

[36] M. Varma & A. Zisserman. (2005) “ Texture Classification: Are filter banks necessary?, ” in *Proc. Int. Conf. Comput. Vision Pattern Recognition,* pp. 691 – 698.

[37] M.Varma & A. Zisserman. (2009) A statistical approach to materiall classification Classification using image patch examplars , *IEEE Trans. Pattern Anal. Match. Intell.*, vol. 31, no. 11, pp. 2032 – 2047.

[38] H. Zhou & R. Wang. (2008) A novel extended local binary pattern operator for texture Analysis. *Inf. Sci.* 178, pp. 4314 – 4325.

[39] G.Zhao & M. Pietikäinen. (2007) Dynamic texture recognition using local binary Patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach*. pp. 915 – 928.

[40] P. Phillips, H. Moon, S. Rizvi & P. Rauss. (2000) The FERET evaluation methodology for face-recognition algorithms. *Pattern Anal Mach Intell IEEE Trans,* pp. 1090 - 1104.

[41] G. Huang, M. Mattar & T. Berg. (2008) Learned-Miller E, Labeled faces in the wild: a database for studying face recognition in unconstrained environments. In: workshop on faces in ’ Real Life’ Images: detection, alignment, and recognition.

[42] AR Database. (2016, March 16). Retrieved from http://www2.ece.ohio-state.edu/

[43] ORL Database. (2016, March 16). Retrieved from http://www.cl.cam.ac.uk/