Local Octal Pattern: A Proficient Feature Extraction for Face Recognition

Nithya J¹, Suchitra S², Dr. S Chitrakala³  
¹²Computer Science and Engineering, Easwari Engineering College, India  
³Computer Science and Engineering, Anna University, India  
¹ nithya.jagan90@gamil.com  
² suchitra.s@srmeaswari.ac.in  
³ ckgops@gmail.com

Abstract—This paper presents a novel and efficient face image representation based on Local Octal Pattern (LOP) texture features. The standard methods viz., the local binary pattern (LBP), the local ternary pattern (LTP) and the local tetra pattern (LTrP) are able to encode with a maximum of four distinct values about the relationship between the referenced pixel and its corresponding neighbors. The proposed method calculates the diagonal, horizontal and vertical directions of the pixels using first-order derivatives. Thereby, it encodes eight distinct values about the relationship between the referenced pixel and its neighbors. The performance of the proposed method is compared with the LBP, the LTP and the LTrP based on the results obtained in terms of average precision and average recall on PubFig image database.

Keywords— face image retrieval; local binary pattern (LBP); local ternary pattern (LTP); local tetra pattern (LTrP); CBIR

I. INTRODUCTION

A. Motivation

The prodigious growth of camera devices, enable people to freely take photos to capture moments of life. By estimation more than 60% of those photos cover human faces. The importance and the sheer amount of human face photos make manipulations (e.g., search and mining) of large-scale human face images a really important research problem and aid many real world applications. The designing aspects of a face retrieval system take account of noise, illumination, poses and expressions, low cost for processing and eliminating misclassified face images. Thus, an efficient Face Image Retrieval System (FIRS) development is required to automatically search the more relevant image from the large database.

Facial recognition system automatically identifies or verifies a person from a digital image. One of the methods is by comparing selected facial features from a query image and a facial database. Automated FIRSs [1] typically involve finding facial landmarks (such as the center of the eyes) for aligning, normalizing the appearance of face, selecting a suitable feature representation, learning perfect feature combinations, and developing precise and scalable matching approaches.
Fig. 1. Major steps in automatic face recognition as in [1]

The prominent step in FIRS is feature extraction whose efficiency depends upon the methods being implemented for extracting features for a given face image. Color, texture, shape and spatial properties are the major features extracted from the face image. There is difficulty in achieving a general representation for face images from the extracted feature, due to the fact that the photographs are taken in different conditions such as illumination, orientation, pose, expression and aging. A comprehensive and extensive literature survey on face detection, extraction and recognition is presented [2] – [5].

To leverage the representation of the image, the local octal pattern is calculated based on the direction of pixels using horizontal, vertical and diagonal derivatives. By incorporating this method, we build a large-scale content based face image retrieval system by taking advantage of the direction of the pixels. In order to evaluate the performance of the proposed system, we conduct extensive experiments on public datasets named Pubfig [6].

B. Related Works

The local binary pattern (LBP) has been a powerful feature for texture classification and retrieval. The LBP as described in [7] and [8] divides the face image into multiple regions for extracting the LBP feature distribution. Finally, the extracted feature is converted into feature vector. Unlike other LBP feature extraction methods, genetic based LBP [9], attempts to reduce the size of the feature extraction based on genetic and evolutionary computing (GEC). GEC evolves LBP extraction that has been randomly distributed.

The forms of the LBP cannot passably deal with the presence of variations in images due to facial expression, pose, illumination, etc. This problem has been addressed with introduction, the local ternary pattern (LTP) [10]. The proposed method tackles the different lighting conditions by combining local feature extraction, distance transform based matching, kernel-based feature extraction, illumination normalization and multiple feature fusion. LTP solves the sensitivity to noise to some extent by introducing a third state. But such minimal pixel difference can be overcome by noise. To address this issue, relaxed LTP [11] introduces uncertain state to encode the minimal pixel value.

The LBP and the LTP are able to encode images with only two and three different values. However, the LTrP [12] has been able to encode four different values. LTrP supports varied applications like CBIR, fingerprint recognition due to its integration of direction of pixels with respect to center pixel.

The LBP, the LTP, and the LTrP translate the relationship between the center pixel and the referenced pixels based on the distribution of edges. The maximum directions which can be coded is only four. Thus, the performance of these methods can be improved by separating the edges in more than two directions. This has led to proposal of a novel method, referred to as local octal pattern (LOP) for content based face image retrieval.
II. FEATURE DESCRIPTORS

A. LBPs

The LBP operator thresholds the each pixel with the center pixel \( g_c \) value in 3 X 3-neighborhood. It encodes 1 for gray value greater than \( g_c \) and 0 for gray value less than \( g_c \). Fig. 2 illustrates the basic LBP operator based on [13].

\[
LBP_{p,r} = \sum_{p=1}^{P} 2^{(p-1)} \times f_1(g_p - g_c) \tag{1}
\]

\[
f_1(x) = \begin{cases} 
1, & x \geq 0 \\
0, & \text{else}
\end{cases} \tag{2}
\]

where \( g_c \) is the center pixel gray values, \( g_p \) is the gray value of its neighbors, \( P \) is the number of neighbors, and \( R \) is the radius of the neighborhood.

B. LTPs

As introduced in Tan and Triggs [10], LTP is an extension of LBP. Unlike LBP, it uses a threshold constant to threshold pixels into three values rather threshold the pixels into 0 and 1. The gray values when equal to the threshold are quantized to zero, those above the threshold are quantized as +1, and those below the threshold are quantized as -1. Fig. 2 shows the calculation of LTP.

\[
f_1(x, g_c, t) = \begin{cases} 
+1, & x \geq g_c + t \\
0, & |x - g_c| < t \\
-1, & x \leq g_c - t
\end{cases} \tag{3}
\]

C. LTrPs

Murala et al. [12] proposed LTrP that encodes the relationship between the center pixel and the referenced pixel based on the direction of the center pixel \( g_c \). The horizontal and vertical directions are calculated using first-order derivatives.

The center pixel’s first-order derivatives \( (g_c) \), along vertical \( (g_v) \) and horizontal \( (g_h) \) direction can be formulated as:

\[
l_{\psi}^{(v)}(g_c) = I(g_h) - I(g_c) \tag{4}
\]

\[
l_{\psi}^{(h)}(g_c) = I(g_v) - I(g_c) \tag{5}
\]

© 2014, IJCSMC All Rights Reserved
and the direction of the center pixel can be formulated as

\[
I_{\text{DIR}}(g_c) = \begin{cases} 
1, & I_{60}^0(g_c) \geq 0 \text{ and } I_{90}^0(g_c) \geq 0 \\
2, & I_{60}^0(g_c) < 0 \text{ and } I_{90}^0(g_c) \geq 0 \\
3, & I_{60}^0(g_c) < 0 \text{ and } I_{90}^0(g_c) < 0 \\
4, & I_{60}^0(g_c) \geq 0 \text{ and } I_{90}^0(g_c) < 0
\end{cases}
\] (6)

From (6), the possible direction is 1, 2, 3, or 4, for each center pixel and finally the image is converted into four values.

Fig.3 and Fig.4 shows the calculation of LTrP. More details about LTrP can be found in [8].
III. PROPOSED METHOD

In this section, we first outline our scalable content-based face image retrieval system, and then we explain the proposed method: Local Octal Pattern (LOP).

A. System Overview

Fig. 7 shows the overall system representation. First, system detects the location of the faces using Viola-Jones face detector [14] for all image in the database. For each detected face image, we will extract the facial components of face viz., two eyes, nose tip, and two mouth corners. A 5X7 grid is defined at each detected component, where each grid is a square patch [15]. Thereby a total of 175 grids are extracted from five components. A LOP feature descriptor is computed for each patch. Every descriptor is quantized into codewords using sparse code as described in [16], after extracting local feature. Many techniques in information retrieval can be applied in the storage of the inverted index [17]. When a query image arrives, it will go through the same procedure to obtain sparse codewords and use these codewords to retrieve images in the index system.

B. Local Octal Pattern

LOP encodes the relationship between the center pixel and the referenced pixel based on the direction of the center pixel $g_c$. The diagonal, the horizontal and vertical directions are calculated using first-order derivatives. An 8-bit tetra pattern for each center pixel is obtained. From the calculated directions, patterns are divided into four parts. Finally, three binary patterns are generated for each direction. Thus, a total of 12(4X3) binary patterns are obtained.

As described by Guo et. al. [18], using sign and magnitude components will extract more useful information. Thus, the 13th binary pattern by using the magnitudes of diagonal, vertical and horizontal first order derivatives has been included.

Fig. 5 illustrates the flowchart of the proposed image retrieval system and Fig. 6 shows the algorithm of LOP feature descriptor.

---

Fig.4. Calculation of tetra pattern for the center-pixel direction “1”

Fig.5. LOP System Framework
IV. EXPERIMENTAL RESULTS

For experiment, images from PubFig database [6] have been used. This database consists of facial images of people taken in different conditions. The performance of the proposed method is measured in terms of average precision and average recall.

The precision is defined as follows:

\[ \text{Precision} = \frac{A}{A + C} \times 100\% \]

The recall is defined as follows:

\[ \text{Recall} = \frac{A}{A + B} \times 100\% \]

where

- \( A \) = Number of relevant images retrieved
- \( B \) = Number of relevant images not retrieved
- \( C \) = Number of irrelevant images retrieved
Based on this figure, it is evident that the LOP significantly improves the retrieved results.

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel approach named as LOP for face recognition. The LOP calculates the direction of pixels using diagonal, vertical and horizontal derivatives. The overall performance of LOP has been compared with the LBP, the LTP and the LTrP on facial images. The result has improved significantly with respect to the average precision and the average recall. Results can be further leveraged by incorporating human attributes (e.g., gender, race, and hairstyle) in a scalable framework.

REFERENCES


