Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 93 (2016) 503 – 512

Procedia Computer Science

6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

Robust Face Recognition System in Video using Hybrid Scale Invariant Feature Transform

Mohanraj. V[a*], Vimalkumar. M[b], Mithila. M[c], Vaidehi. V[d]

[b,c]*Department of Information Technology, Madras Institute of Technology, Anna University, Chennai-44.*
[a,d]*Department of Electronics Enigneering, Madras Institute of Technology, Anna University, Chennai-44.*

Abstract

Face recognition plays a significant role in the research field of biometric and computer vision. The important goal of an efficient Face Recognition (FR) system is to have negligible misclassification rate. In video-based face recognition system, the illumination and pose variation problems are predominant. Most of the efficient FR systems are developed for controlled or indoor environment, hence they fail to give accurate recognition in outdoor environment of different illumination variation. Other challenges include occlusion and facial expression. The illumination problem is handled by Histogram Equalization in existing methods. The original Scale Invariant Feature Transform (SIFT) also works well only for pose variation and fails to produce satisfactory results under varying illumination. Hence Hybrid Scale Invariant Feature Transform (HSIFT) with Weighting Factor in feature matching is proposed in this paper which uses a fixed facial landmark localization technique and orientation assignment of SIFT to extract illumination and pose invariant features. The extracted features are then matched using Fast Library for Approximation of Nearest Neighbor (FLANN). The proposed method has been implemented in OpenCV to give a recognition rate of 98% and 95.5% in YouTube celebrity and Extended Yale B dataset respectively.

Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: Face recognition; SIFT; Ensemble of regression trees; Weighting factor;

1. Introduction

Face recognition is one of the fastest growing research areas, owing to its significance in security surveillance, building/store access control and several other applications. It is an application used to detect, identify

* Corresponding author. Tel.: +91-9944257261.
E-mail address: mohanraj4072@gmail.com

Peer-review under responsibility of the Organizing Committee of ICACC 2016
doi:10.1016/j.procs.2016.07.240

and recognize human faces from video or images. Face is the most distinctive region of human and is very hard to forge, hence it plays a vital role in security applications. Some of the challenges in face recognition are illumination, pose, occlusion, ageing, expressions and low-resolution. There are several works that deal with each of this challenge, in the proposed work the focus is illumination and pose. Most of the applications of face recognition are in uncontrolled environments and in such cases illumination and pose are the most challenging factors to overcome. Face detection is the first step in recognition as only after the face region is detected, features can be extracted and recognized. Face detection techniques based on Haar and HoG[1, 2] have several advantages and disadvantages. HoG is used in the proposed work as it outperforms the other methods in terms of detection accuracy. It has high detection and low false positive rates which makes it robust. In order to overcome illumination variations and shadowing it performs contrast normalizations.

There are two major approaches for feature extraction, typically holistic feature and local feature. In holistic feature based approach like Eigen face the features are extracted from the face as a whole which may sometimes be affected by occlusion and expression changes. Whereas in local features based approach, these issues are overcome as only patches of the image are considered, also they are scale and rotation invariant. There are several local feature extractors like Gabor, LBP, SIFT, SURF etc that are found to be effective, yet fail in certain conditions. For instance, SIFT is scale and rotation invariant as it is based on local features, but it suffers from varying illumination conditions. Similarly each feature extraction method fails in handling more than one challenge; hence there is a need for a hybrid feature extractor. In the proposed work, SIFT is adopted and modified by including fixed landmark points and light adaptation filter based on retina modelling to overcome illumination changes, giving better recognition rates when tested in real-time video datasets as well as standard datasets such as YouTube celebrity. The paper is organized as follows; section 2 consists of related works, section 3 explains video based Face Recognition system, section 4 shows the Implementation and Results and section 5 gives the conclusion and future work.

2. Related Works

CemilTosik et al.[1] proposed an Illumination Invariant Face Recognition system. In their system, Histogram equalization, Discrete Cosine Transforms (DCT) and steerable Gaussian filters are applied to face images as a pre-processing technique. It is found that Histogram equalization with Steerable Gaussian filters gives the best performance under varying illuminations. Ngoc-Son Vu et al.[2] proposed an Illumination-robust face recognition. Illumination normalization in their work is based on retina modelling, which combines two adaptive nonlinear functions and a Difference of Gaussians filter. It achieves very high recognition rates even for the most challenging illumination conditions. Paul Viola and Michael Jones[3] in their work have described a machine learning algorithm for object detection called “Rapid Object Detection using a Boosted Cascade of Simple Features” which is capable of rapid processing and high detection rates. This detector used in real-time applications, runs at 15 frames per second without getting affected to skin color detection. However a few misclassifications are found in certain lighting conditions. Navneet Dalal and Bill Triggs[4] proposed Histograms of Oriented Gradients for Human Detection. Their work is based on the idea that the objects appearance and shape can be characterized by the local intensity gradients and edge directions. Using this type of locally normalized Histogram of Oriented Gradient features gives very good results and is unlike the well-known Haar object detection algorithm which has many false positive rates.

Shaoqing Ren et al.[5] proposed a method of Face Alignment via Regressing Local Binary Features in which two methods are introduced such as local binary features and a method for learning those features. It is observed that their method is computationally less complex and achieves over 3,000 fps for localizing the landmarks. Vahid Kazemi et a1.[6] proposed an one millisecond face alignment with Ensemble of Regression Trees that localize facial landmark positions directly from pixel intensities, achieving good performance and quality predictions in real time. David G. Lowe[7] has proposed Distinctive Image Features from Scale-Invariant Key points that are invariant to rotation, scale and change in viewpoint as it is based on the local features rather than the holistic image. These features can generally be used for object recognition and in the proposed work it is used for face recognition. Renliang Weng et al.[8] proposed a Robust Point Set Matching for Partial Face Recognition. Generally holistic facial