THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

An Efficient Algorithm for Face and Expression Recognition

T. Jeyashree
UG Student, Department of CSE, SNS College of Technology, Coimbatore, India
S. Balasubramaniam
UG Student, Department of CSE, SNS College of Technology, Coimbatore, India
S. Arunkumar
UG Student, Department of CSE, SNS College of Technology, Coimbatore, India
M. Sumathi
Assistant Professor, Department of CSE, SNS College of Technology, Coimbatore, India

Abstract:
Facial expressions convey non-verbal cues. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces. The face is divide into the several regions, and the distribution of the Scale Invariant Feature Transformation (SIFT) features are extracted from them. Facial feature vectors are generated from key point descriptors using Speeded-Up Robust Features. The descriptor can be performed under illumination, noise, expression, and time lapse variations. SIFT detects and uses a much larger number of features from the images, which reduces the contribution of the errors caused by these local variations. When compared with the local Directional Number Pattern, SIFT can increase the efficiency and accuracy in the recognition rate of the face and expression rate.

1. Introduction

Image Processing is the form of signal processing in which the input is a image, photograph and the output is an image or characteristics of an image .In an image processing the image are measured with the pixel and resolution. Most Image Processing techniques involve treating the image as a two-dimensional signal and applying standard signal processing techniques to it. Closely related to Image Processing are Computer Graphics and Computer Vision. The techniques involved in image processing are Image Enhancement is increasing the quality of the image. Image Segmentation is partitioning an image into group of pixels. Image Restoration minimization of known degradations. An image transformation is moving from one domain to another domain. An Image compression is reducing the image size using algorithms. Image Analysis is making quantitative measurements from an image. Scale-invariant feature transform (or SIFT) is an algorithm which useful to identify local features in images.

For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object. The feature, extracted from the input image. To perform efficient recognition the features extracted from the input image are detectable even under changes in image scale, noise & illumination. Such points usually lie on high Contrast region of the image, such as object edges.

2. Related Works

In the Automatic Facial Expression Recognition for Intelligent Tutoring Systems, It explores the idea of facial expression for automated feedback in teaching. It shows how automatic real time facial expression recognition can be effectively used to estimate the difficulty level, as perceived by an individual student, of a delivered lecture. On a video lecture viewing task, training on less than two minutes of recorded facial expression data and testing on a separate validation set, our system predicted the subjects' self-reported difficulty scores with mean accuracy of 0:42 (Pearson R) and their preferred viewing speeds with mean accuracy of 0:29. Our techniques are fully automatic and have potential applications for both intelligent tutoring systems (ITS) and standard classroom environments [3]. In Component-Based Representation in Automated Face Recognition presents a framework for component based face alignment and representation that demonstrates improvements in matching performance over the more common holistic approach to face alignment and representation. The component-based framework consists of the following major steps: (I) landmark extraction using Active Shape Models (ASM), (II) alignment and cropping of components using Pro Analysis, (III) representation of components with Multi-Scale Local Binary Patterns (MLBP), (IV) per-component measurement of facial similarity, and (V) fusion of per-component similarities. We demonstrate on three public datasets and an operational dataset consisting of face images of 8,000 subjects, that the proposed component based representation provides higher recognition accuracies over holistic-based representations. Additionally, we show that the proposed this component - based representations: (I) are more robust to changes in facial pose and (II) improve recognition accuracy on occluded face images in forensic scenarios[4]. In Robust Facial Expression Recognition via Sparse Representation and Multiple Gabor filters, Facial expressions recognition plays important role in human communication. It has become one of the most challenging tasks in the pattern recognition field. It has many applications such as: human computer interaction, video surveillance, forensic applications, criminal investigations, and in many other fields[6]. In this paper we propose a method for facial expression recognition (FER). This method provides new insights into two issues in FER: feature extraction and robustness. For feature extraction we are using sparse representation approach after applying multiple Gabor filter and then using support vector machine (SVM) as classifier. We conduct extensive experiments on standard facial expressions database to verify the performance of proposed method and we compare the result with other approach [3].

3. Existing System

3.1. Local Directional Number Pattern

The Local Directional Number Pattern (LDN) is a six bit binary code assigned to each pixel of an input image that represents the structure of the texture and its intensity transitions. The edge magnitudes are largely insensitive to lighting changes. Consequently, we create our pattern by computing the edge response of the neighborhood using a compass mask, and by taking the top directional numbers, that is, the most positive and negative directions of those edge responses. We illustrate this coding scheme in Fig.1[1]. The positive and negative responses provide valuable information of the structure of the neighborhood, as they reveal the gradient direction of bright and dark areas in the neighborhood. Thereby, this distinction, between dark and bright responses, allows LDN to differentiate between blocks with the positive and the negative direction swapped (which is equivalent to swap the bright and the dark areas of the neighborhood, as shown in the middle of Fig. 1) by generating a different code for each instance, while other methods may mistake the swapped regions as one. Furthermore, these transitions occur often in the face, for example, the top and bottom edges of the eyebrows and mouth have different intensity transitions. Thus, it is important to differentiate among them; LDN can accomplish this task as it assigns a specific code to each of them. Local Directional Number Pattern (LDN), for robust face recognition that encodes the structural information and the intensity variations of the face's texture. LDN encodes the structure of a local neighborhood by analyzing its directional information. Consequently, we compute the edge responses in the neighborhood, in eight different directions with a compass mask. Then, from all the directions, we choose the top positive and negative directions to produce a meaningful descriptor for different textures with similar structural patterns. This approach allows us to distinguish intensity changes (e.g., from bright to dark and vice versa) in the texture, that otherwise will be missed see Fig. 1. Furthermore, our descriptor uses the information of the entire neighborhood, instead of using sparse points for its computation like LBP. Hence, our approach conveys more information into the code, yet it is more compact as it is six bit long. Moreover, we experiment with different masks and resolutions of the mask to acquire characteristics that may be neglected by just one, and combine them to extend the encoded information. We found that the `inclusion of multiple encoding levels produces an improvement in the detection process.

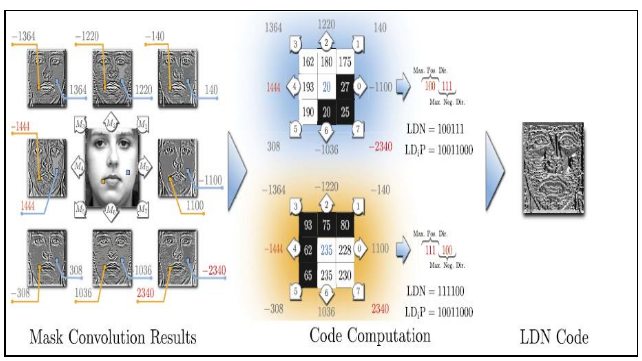

Figure 1

Fig.1. LDN code computation. The (Kirsch) compass masks are convoluted with the original image to extract the edge response images (shown in the left).From these images, we choose the prominent directional numbers (positive and negative directions) to encode the texture in the neighborhood. We show an example of a neighborhood in the middle - top, that corresponds to the colored marks on the edge response images. It shows the different response values, the top directional numbers (in red and orange), and the final LDN code (shown in the right). Moreover, LDN can detect changes in the intensity regions by producing a different code (as shown in the middle-bottom) while other directional patterns cannot (like LDIP), as they produce the same code for different textures.

4. Proposed System

- Scale Invariant Feature Tansformation

  Scale-invariant feature transform (or SIFT) is an algorithm which useful to identify local features in images. For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object. The feature, extracted from a training image, can then be used to identify the object of an input image. To perform efficient recognition the features extracted from the training image are detectable even under changes in image scale, noise & illumination. Such points usually lie on high contrast region of the image, such as object edges. SIFT detects and uses a much larger number of features from the images, which reduces the contribution of the errors caused by these local variations. SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. The SIFT consists of 1. Key localization – DoG (Difference of Gaussian), 2. Scale space extrema detection, 3. Rotation - orientation assignment, 4.Geometric distortion-re sampling of local image orientation plane, 5. Index match - nearest neighbor as shown in fig.2.[7]

4.1. Scale space extrema detection

This is the stage where keypoints detected. For this, the image is convolved with Gaussian filters at different scales, and then the difference of Gaussian - blurred images are taken. Once DoG images have been obtained, keypoints are identified as local minima/maxima of the DoG images across scales. This is done by comparing each pixel in the DoG images to its eight neighbors at the same scale and nine corresponding neighboring pixels in each of the neighboring scales. If the pixel value is the maximum or minimum among all compared pixels, it is selected as a candidate key point.

4.2. Key point localization

Scale-space extrema detection produces too many keypoints candidates, some of which are unstable. The next step in the algorithm is to perform a detailed fit to the nearby data for accurate location, scale, and ratio. This information allows points to be rejected that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge.

4.2.1. Interpolation of nearby data for accurate position

First, for each candidate keypoints, interpolation of nearby data is used to accurately determine its position. The initial approach was to just locate each keypoints at the location and scale of the candidate keypoints. The approach calculates the interpolated location of the extreme, which improves matching and stability.

4.2.2. Discarding low – contrast keypoints

To discard the keypoints with low contrast, the value of the second-order Taylor expansion is computed. If this value is less than 0.03, the candidate keypoints is discarded.

4.2.3. Eliminating edge responses

The DoG function will have strong responses along edges, even if the candidate keypoints is not robust to small amounts of noise. Therefore, in order to increase stability, we need to eliminate the keypoints that have poorly determined location the principal curvature across the edge would be much larger than the principal curvature along it. Finding these principal curvatures amounts to solving for the eigen values Therefore the higher the absolute difference between the two eigenvalues, which is equivalent to a higher absolute difference between the two principal curvatures of D, the higher the value of R. Harris corner detection.

4.3 Orientation assignment

Each keypoints is assigned one or more orientations based on local image gradient directions. This is the key step in achieving invariance to rotation as the keypoints descriptor can be represented relative to this orientation for gaussian smoothed image and scale sigma, gradient magnitude and orientation. An orientation histogram with 36 bins is formed, with each been covering 10 degrees. Peak histogram is the dominant orientation**.**

4.4 NN Classifier

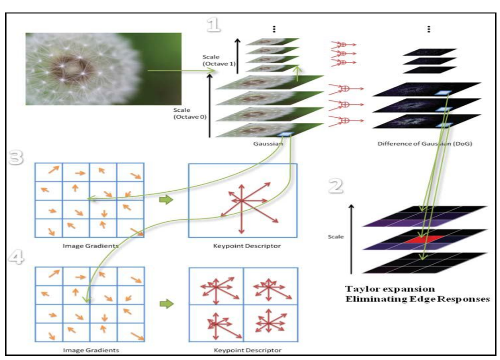

Figure 2: SIFT algorithm: (1) scale-space extrema detection, (2) accurate Key point localization, (3) orientation assignment, and (4) the local image descriptor

Given a set of training samples and test point and find k training point closest vector collect the labels and classify which has the greatest number of representatives. In short classification is performed by taking the majority vote among k-nearest neighbor. We need to be able to compute the euclidean distance two points to the test and training points. D (a, b) =sqrt (sum (a-b) ^2) where a and b are two points and D is the euclidean distance between them.

5. Conclusion

This paper presents a real time SIFT based feature extraction engine that is capable to compute 2000 feature points for HD1080p30 at 100 MHz The proposed design adopts the layer parallel restructured box kernel to replace iterated Gaussian blur operations for simple and parallel computation.

This also reduces the latency to a few image lines instead of several frames. The final flow uses the on-the-fly feature extraction flow so that only partial intermediate results have to be stored. With these techniques, the presented design easily achieves the real time demand with significantly lower cost, which saves 56% gate count and 90.4% memory cost when compared to the previous design.

6. References

1. Adin Ramirez Rivera, Student Member, IEEE, Jorge Rojas Castillo, Student Member, IEEE, and Oksam Chae, Member, IEEE " Local Directional Number Pattern for Face Analysis: Face and Expression Recognition" , IEEE transactions on image processing, vol.22.no.5,may 2013.
2. H. Feng, E. Li, Y. Chen, and Y. Zhang, "Parallelization and characterization of SIFT on multi-core systems," in Proc. IEEE Int. Symp. Workload Characterized Sep. 2008, pp. 14–23.
3. Jacob White hill, Marian Bartlett, and Javier Movellan, "Automatic Facial Expression Recognition for Intelligent Tutoring Systems", 2008 IEEE Computer Society conference.
4. Kathryn Bonnen, Brendan F. Klare, Member, IEEE, and Anil K. Jain, Fellow, IEEE, " Component- Based Representation in Automated Face Recognition" , IEEE transactions on information forensics and security, vol. 8, no. 1, Jan 2013
5. Liang-Chi Chiu, Tian-Sheuan Chang, Sen, Jiun-Yen Chen, and Nelson Yen-Chung Chang," Fast SIFT Design for Real-Time Visual Feature Extraction" IEEE Transactions on Image Processing, VOL. 22, NO. 8, Aug 2013
6. Prof.Rania Salah El-Sayed, Prof.Dr. Ahmed El Kholy "Robust Facial Expression Recognition via Sparse Representation and Multiple Gabor filters", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.3, 2013
7. Zhang, Y. Chen, Y. Zhang, and Y. Xu, "SIFT implementation and optimization for multi-core systems," in Proc. IEEE Int. Parallel Distribution Process., Apr. 2008, pp. 1–8