

Gesture Recognition using Principle Component Analysis & Viola-Jones Algorithm

Adita Nimbalkar¹, Swati Pawar², Vivek Ugale³

^{1,2,3} Department of E&TC, SITRC, Nasik, Savitribai Phule Pune University, India

Abstract: Gesture recognition pertains to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. It is of utmost importance in designing an intelligent and efficient human-computer interface. The applications of gesture recognition are manifold, ranging from sign language through medical rehabilitation to virtual reality. In this paper, we provide a survey on gesture recognition with particular emphasis on hand gestures and facial expressions. Applications involving wavelet transform and principal component analysis for face and hand gesture recognition on digital images.

Keywords: Wavelet transform, Principle component analysis, Viola-Jones method

I. Introduction

Object recognition is one of the most difficult and urgent tasks of image processing. People can very easily solve this problem, but it is not such really simple for computer science. In this paper, the task of recognition of particular objects as faces and hand gestures are considered. Face recognition is mostly investigated and used in the areas as digital image processing, computer vision, biometric system, videoconference organization, access control system and etc. Hand gesture recognition system is used to identify determined human gestures in order to transform information or control variety equipments. The object recognition process often consists of two phases. The first one is extraction and preservation of known objects features in database. The second phase is comparison of unknown object features with the features locating in the database. Nowadays, wavelet transform is the good method for image representation and features extraction. And principal component analysis is successfully used for objects features comparison in order to solve the object[1][2][3].

The main aim of this paper is creation of a novel algorithm using wavelet transform and principal component analysis in order to recognize faces and hand gestures on digital images, and creation of a novel complex algorithm based on Viola- Jones method, wavelet transform, kNN method and principal component analysis.

II. Wavelet Transform

Wavelet transforms are widely used to solve the class of image processing tasks. Since wavelet transform coefficients contain the information about analysis process and which type of wavelet transform is used, when we choose the type of wavelet transform, we need define which information should be extracted.

The Haar wavelet transform is used to extract plane image features, and then the plane image is classified by multilayer neural network which inputs are wavelet transform coefficients [1,2]. In this paper Haar and Daubechies wavelet transforms are used to extract object image features (face and hand gesture). Examples of the use of Haar wavelet transform to extract face image features and Daubechies to extract hand gesture image features.

Wavelet transform uses a sub-band coder, to produce a pyramid structure where an image is decomposed sequentially by applying power complementary low pass and high pass filters and then decimating the resulting images. These are one- dimensional filters that are applied in cascade (row then column) to an image whereby creating a four-way decomposition: LL (low-pass then another low pass), LH (low pass then high pass), HL (high and low pass) and finally HH (high pass then another high pass). The resulting LL version is again four-way decomposed as shown in Figure 1. Each level has various bands information such as low- low, low-high, high-low, and high-high frequency bands. Furthermore, from these DWT coefficients, the original image can be reconstructed. This reconstruction process is called the inverse DWT (IDWT).

III. Principle Component Analysis

This method is one of the most popular methods of data reduction with the least amount of lost information. This method is a mathematical procedure that uses an orthogonal transformation to convert an input vector P with size N to output vector Q with size M and $M < N$. The output vector Q is set of linearly uncorrelated variables called principal component.

The main idea of principal component analysis is to convert face image to image of principal components called eigenfaces. The principal components are computed for each face image and the number of them is about from 5 to 200 components. The recognition process is a comparison of principal components of known face image with components of unknown face image. In this case face images of one person are grouped to clusters of eigenspace. The candidate images having the least distance from input (unknown) image are chosen from database.

IV. Viola-Jones Method

Viola and Jones introduced a fast algorithm which minimizes the computation time while achieving high detection accuracy. In this algorithm, the concept of “Integral Image” is used to compute a rich set of image features. Compared with other approaches which must operate on multiple image scales, the integral image can achieve true scale invariance by eliminating the need to compute a multi-scale image pyramid, and significantly reduces the initial image processing time which is a must for most object detection algorithms. Another technique used in their approach is the feature selection algorithm based on the AdaBoost (Adaptive Boost) learning algorithm. Boosting is an aggressive feature selection technique in machine learning that can effectively improve the accuracy of a given learning algorithm. The Adaboost learning algorithm is a variation of the regular boosting algorithm, and can adaptively select the best features at each step and combine a series of weak classifiers into a strong classifier. The Viola-Jones algorithm has been primarily used for face detection systems which is approximately 15 times faster than any previous approaches while achieving equivalent accuracy to the best published results. However, limited research has been done to extend the method to hand detection and gesture recognition[2].

V. Feature Extraction

Features are the crucial elements for hand gesture recognition. Large number of features, such as, shape, orientation, textures, contour, motion, distance, centre of gravity etc. can be used for hand gesture recognition. Hand gesture can be recognized using geometric features, like, hand contour, fingertips, finger detections. But these features may neither be always available nor reliable due to occlusions and illuminations [4]. Some non-geometric features (such as color, silhouette, texture) are also available for recognition. But they are inadequate for the purpose. Therefore, the image or the processed image can be fed to the recognizer to select the features automatically and implicitly, rather than using single type of feature alone. Following three approaches [4] are useful for extraction of features.

VI. KNN Method

The k-nearest neighbor algorithm is one of the most common classifier in the literature. The main idea behind this algorithm is to select the k-nearest neighbors of a certain input from a training database and then assign it to the output that cast a majority vote among the ones associated to the selected inputs. In order to obtain always a majority vote, the “k” parameter is usually an odd number since even ones can cause ties in case of two-class classification problem. The main advantage of this algorithm is that it is an universal approximator and can model any many-to-one mapping very well. The drawbacks consist of the lack of robustness for high dimension spaces and the high computational complexity with huge training data-set.

Both for classification and regression, it can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones[5].

The neighbors are taken from a set of objects for which the class (for k -NN classification) or the object property value (for k -NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.

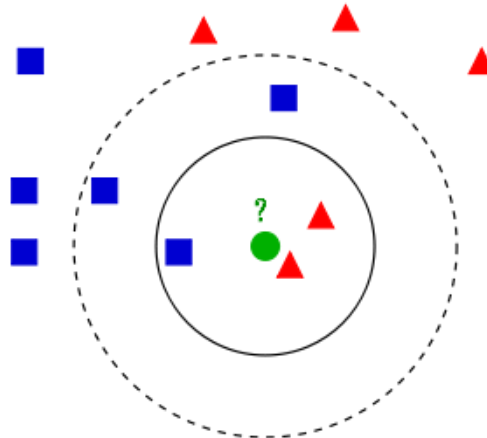
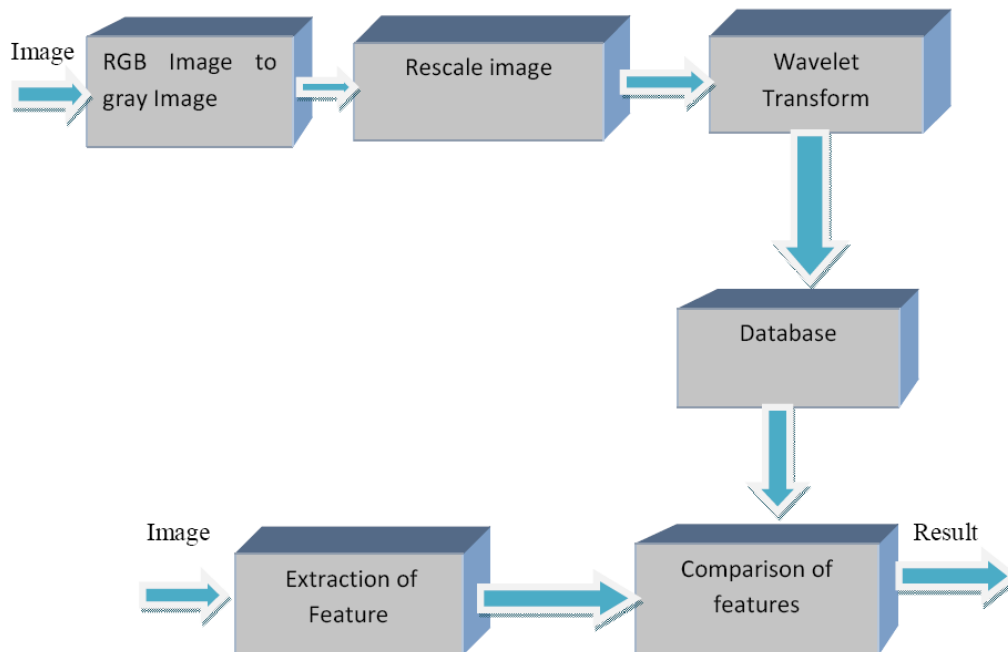


Fig 1 knn Classifier

The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles. If $k = 3$ (solid line circle) it is assigned to the second class because there are 2 triangles and only 1 square inside the inner circle. If $k = 5$ (dashed line circle) it is assigned to the first class (3 squares vs. 2 triangles inside the outer circle)

VII. Proposed Algorithm



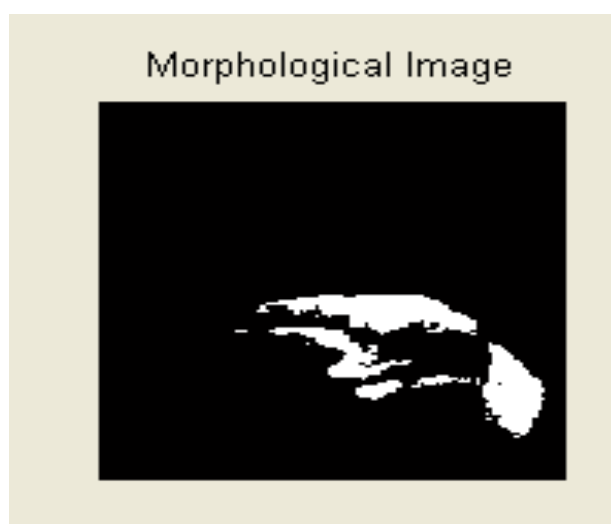
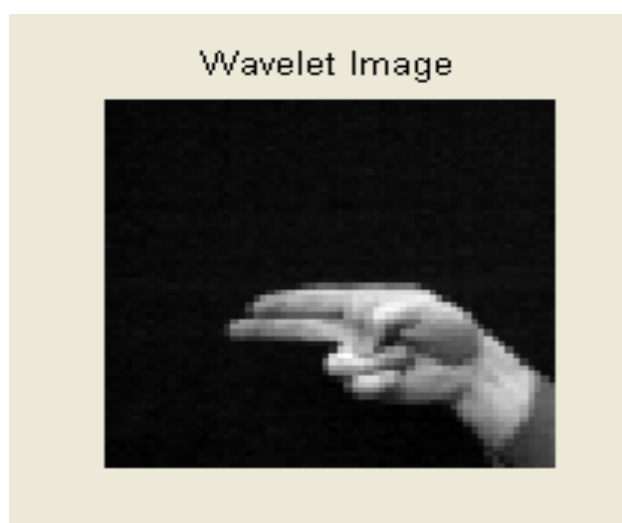
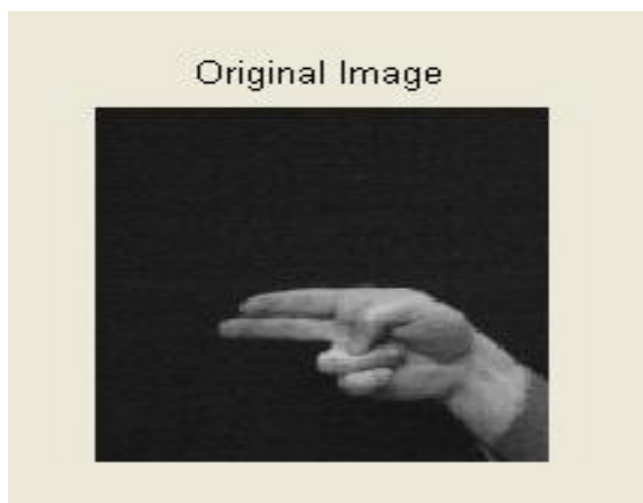
The main vision of this paper is to recognise features of face and hand gesture. The proposed algorithm is based on wavelet transform and principle component analysis. The algorithm has two parts: first is to extract and preserve the features of known images into data base and second is unknown image feature extraction by using data from data base. The flow of algorithm is as follows:

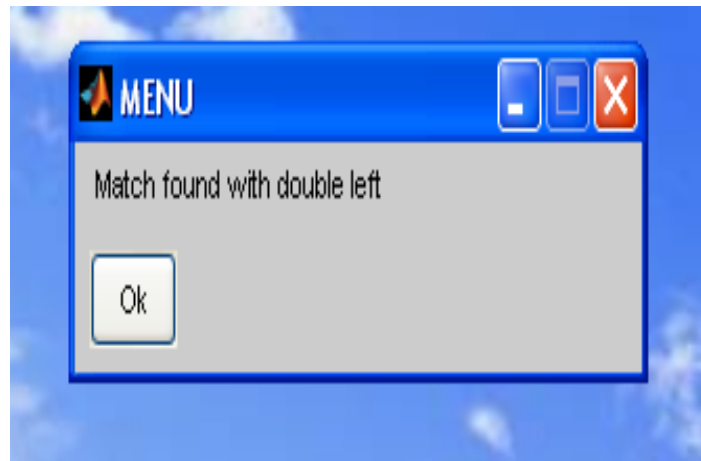
- Step 1: Transform the image into gray scale
- Step 2: Resize the image
- Step 3: Apply wavelet transform for feature extraction
- Step 4: Save this extracted features into data base.

In the second half of algorithm features of unknown image is extracted and it is compared with data base which is already stored.

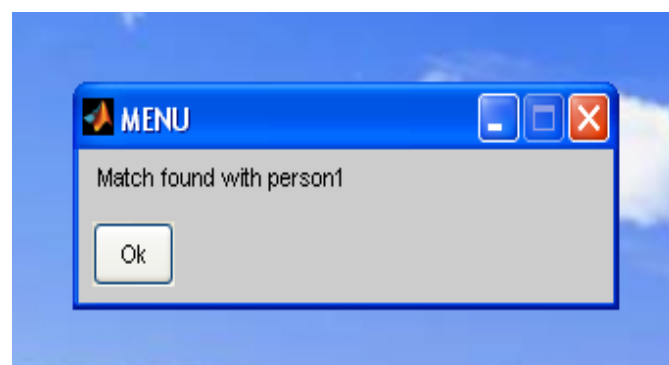
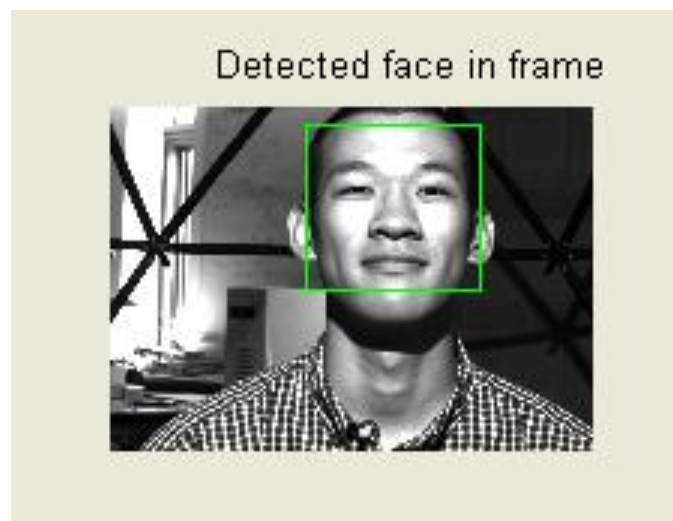
VIII. Results

1) Hand Gesture Recognition





2) Face Recognition



IX. Conclusion

The study in this paper is a complex algorithm based on wavelet transforms, PCA, kNN Classifier & Viola-Jones Algorithm for face & hand gesture recognition on digital image. The results of this system for gesture recognition are 94.80% and for face recognition it is 98.70%. Gesture recognition can be applied in virtual environment, sign language translation, and medical system. But the disadvantage of gesture based system is suitable in controlled lab setting but does not generalize to arbitrary setting. The gesture recognition system uses gestures like hand wave, stop, punch. All these and many more gestures can be useful to turn on light, control TV, music system.

REFERENCES

- [1] Bui.T.T.T, Phan.N.H, Spitsyn V.G, "Face and hand Gesture Recognition Algorithm based on Wavelet Transform & Principle Component Analysis", 978-1-4673-1773-3/12/\$31.00 ©2013 IEEE
- [2] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518, Kauai, Hawaii, USA, 2001.
- [3] P. Viola and M.J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [4] Murthy, G. R. S. and Jadon, R. S. A Review of Vision Based Hand Gestures Recognition. *Int. J. of Information Technology and Knowledge Management*, 2(2) (2009), 405 – 410
- [5] Richard Duda, Peter Hart, David Stork, "Pattern Classification"
- [6] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Reading MA: Addison-Wesley, 2001.
- [7] C. Papageorgiou, M. Oren and T. Poggio, "A general framework for object detection," // *International Conference on Computer Vision*, 1998.
- [8] T.K. Kim, S.F., Wong and R. Cipolla, "Cambridge Hand Gesture Data set," [Online]. Available: http://www.iis.ee.ic.ac.uk/~tkkim/ges_db.htm