ISSN No. 0976-5697

Volume 4, No. 9, July-August 2013

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Performance Comparison of Transform Techniques for Hand Gesture Recognition

Dr. Tanuja K. Sarode
Dept. of Computer Engineering,
ThadomalShahani College of Engg,
Mumbai, India
tanuja_0123@yahoo.com

Vaishali Sakpal
Dept. of Information Technology,
Don Bosco Institute of Technology,
Mumbai, India
vaishali_sakpal@rediffmail.com

Abstract -Human computer interaction is a major issue in research industry. In order to offer a way to enable untrained users to interact with computer more easily and efficiently gesture based interface has been paid more attention. This paper presents a new approach for hand gesture recognition. The approach consists of threemodules: a) Preprocessing of the image b) Feature extraction c)Pattern matching for gesture recognition. Feature extraction is based on feature vector of transformed image using Discrete Cosine Transform, Walsh Transform, Haar transformand Kekre's transform. The various size of feature vector is generated such as 4X4, 8X8, 16X16, 32X32, and 64X64. Results found to be better compare to the existing systems.

Keywords- Hand gesture Recognition, pattern matching, feature vector, HCI, transform.

I. INTRODUCTION

Hand gesture recognition is one of the important aspects in various upcoming fields like human computer interaction, virtual reality, and computer vision. For hearing impaired community, the development of automatic gesture translation based natural languages (e.g. the American Sign Language; ASL) is highly expected to improve their communication means among humans. Compared with the traditional interaction approaches, such as keyboard, mouse, pen etc. vision based hand interaction is more natural and efficient. Therefore hand gesture recognition has received considerable attention.

There are number of techniques and algorithms available in the literature of hand gesture recognition. Jagdish Raheja etal [1] proposed Principle Component Analysis method for hand recognition. In his proposed system, gesture recognition is done with 10 different hand gestures which gives 90% accuracy in proper light arrangement and is further decreasing in poor light arrangement.

Yuehai Wang et al [2] used Morphological Method for hand gesture recognition. It mainly uses skin color extraction, dilation, erosion, and subtraction for hand gesture recognition. A recognition rate is 84%.Yikai Fang, Kongqiao Wanget al [3] proposed Hand gesture Recognition using Blob and Ridge Method; hand detection is done with ad boost to trigger tracking and recognition, Adaptive hand segmentation is executed during detection and tracking with motion and color cues, scale-space features detection is applied to find palm-like and finger-like structures. Deg-Yuan Huang et al [4] used Gabor filter and support vector machine. The principle component method is used to reduce the dimensionality of the feature space; a gesture recognition rate is 95%. Qing Chen et al [5] used Haar-like featuresto describe the hand posture pattern with the computation of integral image. ad boost learning algorithm is used to construct a strong classifier by combining a sequence of weak classifiers. This method gives accuracy about 90%. Rajeshree Rokadeet al [6] in his paper proposed a novel technology for hand gesture recognition which is based on thinning of segmented image. The average recognition rate is 92.13%. Zhong Yang et al[7] Proposed Hidden Markov Model for gesture recognition its recognition rate is 96.67%with 18 different gestures. Karishma Dixit et al [8] used a multi-class Support Vector Machine (MSVM)for training andrecognition; its recognition rate is 96%. Archana Ghotkar et al[9] used HSV color model and Genetic algorithm for gesture recognition. Prateem Chakrabortyet al [10] have given comparative study of Gradient and PCA techniques.

This paper presents hand gesture recognition techniques based on feature vector of transformed image using Discrete Cosine Transform, Walsh, Haar, Kekres transform. The result of these transformed techniques is compared with one of the existing and popular technique that is principle component analysis.

II. SYSTEM DESCRIPTION

This sectionconsists of three major components such as preprocessing, feature extraction and pattern matching. The block diagram is as shown in theFig. 1.

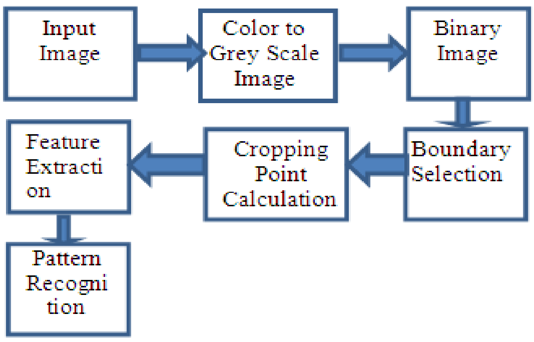

Fig.1: Block Diagram of the System

A. Image Preprocessing

Image Preprocessing is required to remove extra unwanted part for better result in pattern matching.

The steps for Image Preprocessing are as follow:

1. Take the input image from stored database.
2. Convert the input image into Binary image using global thresholding.
3. Extract the object of interest from binary image. This is done by searching white pixel horizontallyandvertically. Occurrence of first and last white pixels are boundary points. The image is cropped according to boundary points as shown in Fig. 2.

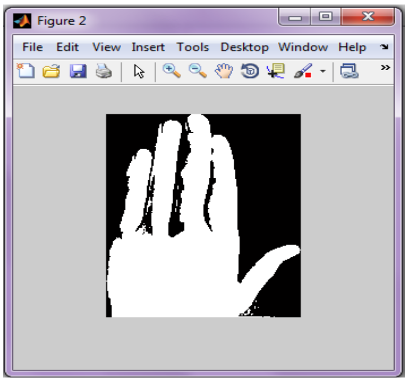

Fig. 2: Boundary Selection

4. Cropping point calculation is to remove part of a hand which is not used in gesture presentation as shown in Fig. 3. Determination of column in which maximum number of white pixels are presents. (Ex. wrist area) cropping points are between the columns of maximum and minimum number of white pixels.

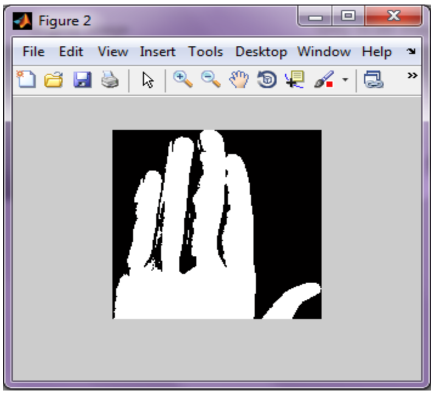

Fig. 3: Cropping Point Calculation

5. Resize the image to some standard size as shown in Fig. 4, in our proposed system after preprocessing the maximum size of the image is 64X64 therefore all the images are resize to 64X64.

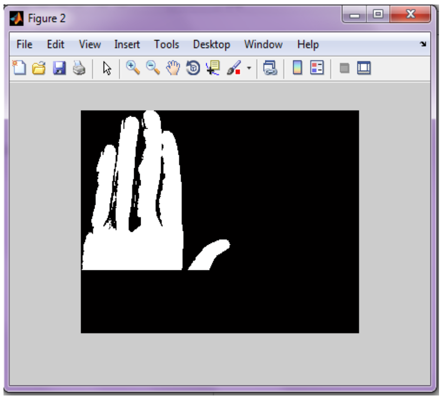

Fig. 4: Resize Image

B. Feature Extraction

Extraction of feature vector from images is done using different techniques. Feature vector of query image is compared with feature vector of the images in the database. Euclidean distance is used to check the closeness of the query image and database image. The techniques for feature extraction is given below,

1.Principle Component Analysis:

Principle Component Analysis [11] is used for feature extraction, algorithm is as follow:

a. Take the input image from test database. The database is stored in Train database.
b. Calculate the mean of each image. Subtract the mean from each of the data dimension. The mean subtracted is the average across each dimension.
c. Calculate the covariance matrix.
d. Calculate the eigenvectors and eigenvalues of the covariance matrix.
e. Choosing component and forming feature vector.
f. Apply step 2 to 5 on input image.
g. Matching the eigenvalues of stored image and captured image.
h. Final image is selected whose Euclidean distance is small between the original image and Captured image

2. Discrete Cosine Transform:

The discrete cosine transform (DCT) [13] is closely related to the discrete Fourier transform. The definition of the two-dimensional DCT can be written as follows, in terms of pixel values f (x, y) for x, y = 0, 1, N-1 and the frequency domain transform coefficients F (u, v),

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left(\frac{(2x+1)u\pi}{2N}\right)\cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (1)$$

For u, v = 0, 1, 2, .N-1 and

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)F(u,v)\cos\left(\frac{(2x+1)u\pi}{2N}\right)\cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (2)$$

The N X N cosine Transform Matrix C={C (u, v)} is define as,

$$C(u,v) = \begin{cases} \sqrt{1}/N & u=0,\ 0\le v\le N-1 \\ \sqrt{2}/N\ \cos\left[\frac{\pi(2v+1)u}{2N}\right] & 1\le u\le N-1,\ 0\le v\le N-1 \end{cases} \quad (3)$$

The Cosine Matrix is real and orthogonal but not symmetric,

$$C = C^* \Longrightarrow C^{-1} = C' \quad (4)$$

$$C.C' = I \quad (5)$$

The 2D DCT of an image can be generated using equation, $F = C\,f\,C'$ (6)

If DCT is applied to real data, the result is also real. The DCT tends to concentrate information, making it useful for image compression applications and also helping in minimizing feature vector size for different applications [12]. For full 2-Dimensional DCT for an NxN image the number of multiplications required are N^2 (2N) and number of additions required are N^2 (2N-1).

3. Walsh Transform:

Walsh transform matrix [13] is defined as a set of N rows, denoted W_j, for j = 0, 1, ... , N - 1, which have the following properties:

- W_j takes on the values +1 and -1.
- Wj[0] = 1 for all j.
- $W_j \times W_k^T=0$, for $j \neq k$ and $W_j \times W_k^T=N$, for j=k.
- Wj has exactly j zero crossings, for j = 0, 1, ..., N-1.
- Each row W_j is even or odd with respect to its midpoint.

Walsh transform matrix is defined using a Hadamard matrix of order N. The row of Walsh transform matrix is considered as the row of Hadamard matrix.It specified by the Walsh code index, which mustbe an integer in the range [0, ..., N - 1]. For the Walsh code index equal to an integer j, the respective Hadamard output code has exactly j zero crossings, for j = 0, 1,..., N - 1.

For the full 2-Dimensional Walsh transform applied to image of size NxN, the number of additions required are $2N^2$ (N-1) and absolutely no multiplications are needed in Walsh transform [13].

4. Haar Transform:

The Haar Transform is deriving from Haar matrix. The Haar transform like most of the other transform is separable and can be expressed in matrix form,

$$F = H\,f\,H' \quad (7)$$

Where f is an NXN image, H is a NXN matrix and F is resultant NXN transform The transform H contains the Haar basis function $h_k(x)$ which are defined over the continuous closed interval x[0,1]

Haar basis functions are,

When, k=0 the Haar function is defined as a constant

$$h_0(t) = 1/\sqrt{N} \quad (8)$$

When, k>0 the Haar function is defined by

$$H_k(t) = 1/\sqrt{N} \begin{cases} 2^{p/2} & (q-1)/2^p \le t < (q-0.5)/2^p \\ -2^{p/2} & (q-0.5)/2^p \le t < q/2^p \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

From the definition, it can be seen that p determines the amplitude and width of the non-zero part of the function, while q determines the position of the non-zero part of the function.

5. Kekre's Transform

Kekre's transform matrix is the generic version of Kekre'sLUV color space matrix [12]. Kekre's transform matrix can be of any size NxN, which need not have to be in powers of 2 (as is the case with most of other transforms). All upper diagonal and diagonal values of Kekre's transform matrix are one, while the lower diagonal part except the values just below diagonal is zero. Generalized NxNKekre's transform matrix can be given as:

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 1 \\ -N+1 & 1 & \dots & 1 & 1 & 1 \\ 0 & -N+2 & \dots\dots & 1 & 1 & 1 \\ . & . & \dots & . & . & . \\ . & . & \dots & . & . & . \\ 0 & 0 & . & 0 & 1 & 1 \\ 0 & 0 & . & 0 & -N+(N-1) & 1 \\ & & \dots\dots & & & \end{bmatrix}$$

The formula for generating the term Kxy of Kekre's transform matrix is:

$$K_{xy} = \begin{cases} 1 & ; x \le y \\ -N+(x-1) & ; x = y+1 \\ 0 & ; x > y+1 \end{cases} \quad (10)$$

For taking Kekre's transform of an NxN image, the number of required multiplications are 2N (N-2) and number of additions required are N (N^2+N-2).

III. PROPOSED ALGORITHM

1. Capture the image, or take the input image from stored database.
2. Perform Preprocessing on input image using boundary point selection and cropping point calculation.
3. Feature vector of the query image is generated as shown in the Fig. 6.
4. Compare feature vector of the query image with the feature vectors of all the images in the database. Feature vector is of different sizes as shown in the Fig.5.
5. Measure Euclidean distance to check the closeness of the query image and thedatabase images. Calculate the closest match with query image.
6. Select the smallest value with respect to the corresponding column to find first closest match with query image.
7. Display the equivalent image from the database.

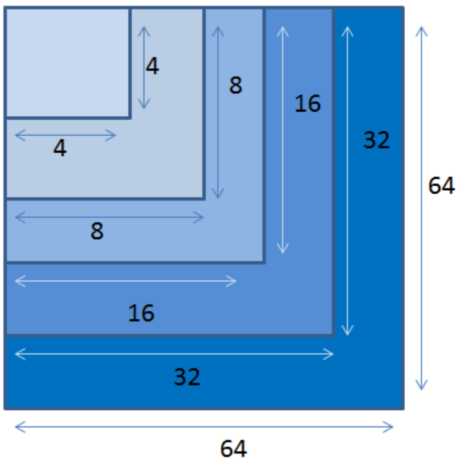

Fig. 5: Feature Extraction of varying size

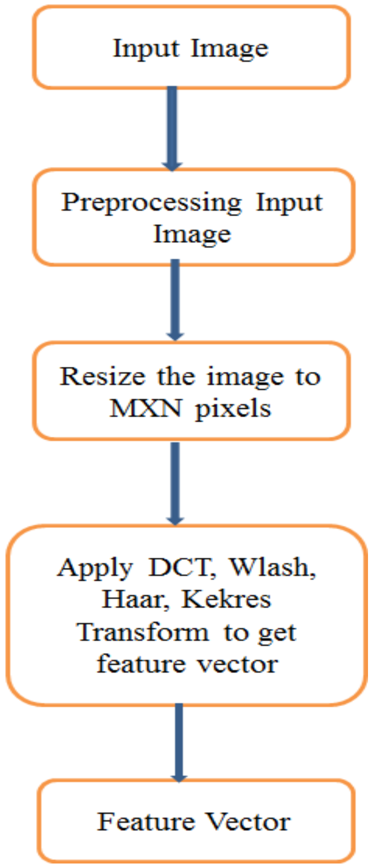

Fig. 6: Flowchart for feature extraction

IV. RESULT AND DISCUSSION

The implementation of the proposed algorithm is done in MATLAB 10 using a computer with Intel Core i3 processor. The proposed algorithm is tested on 48 different gestures representing American Sign Languages [14]. The hand gestures for training were collected with the plane background by same hand gesture 5 times; each time from adifferent angle and position. The dataset contains 120 (=24*5) images of each hand gestures. For testing 48(=24*2) images have been used.

This Database is tested using different feature extraction techniques such as Principle Component Analysisand Transform techniques. The result shows that with Principle component Analysis the Average percentage of accuracy is 100%.Same database is tested with transform techniques with various sizes of feature vectors. Table 1 shows the percentage accuracy for different transform domain method like DCT, Haar, Walsh, Kekres.

Table I.Percentage Accuracywith different sizes oftransformed matrix

Methods	**Accuracy in %**				
Transform	64X64	32X32	16X16	8X8	4X4
DCT	100%	100%	100%	100%	100%
Haar	100%	100%	100%	100%	95%
Walsh	100%	100%	100%	100%	95%
Kekre	100%	98%	63%	70%	58%

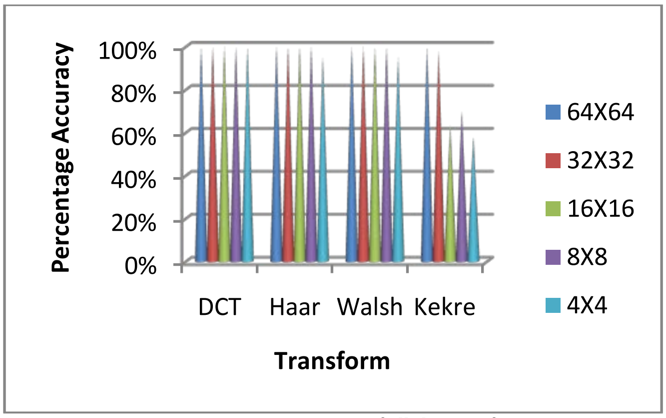

Fig. 7: Percentage Accuracy of all the transforms

The average recognition rate of the principle component analysis method is 9 Sec. whereas the recognition rate of the proposed system 0.92 seconds. Table 2 shows the computational complexity of the transforms for the image of size N X N.

Table II. Computational Complexity for transforms to image of size N X N

Transform	**Number of Additions**	**Number of Multiplication**	**Total Additions for transform of 64 X 64 image**
DCT	$2N^2(N-1)$	$N^2(2N)$	4710400
Haar	$2N^2\log_2(N)$	0	49152
Walsh	$2N^2(N-1)$	0	516096
Kekre	N[N(N+1)-2]	2N(N-2)	329600

V. CONCLUSION AND FUTURE WORK

In today's digitized world, processing speed has increase dramatically, with computers being advanced to the levels where they can assist humans in complex tasks. Yet, input technologies cause a major bottleneck in performing some of the tasks, underutilizing the available resources and restricting the expressiveness of application use. Hand gesture recognition techniques is used to interact with the computer more easily and user friendly.

Our proposed system is simple but effective. The various transforms like DCT, Haar, Walsh andKekres transform are used for gesture recognition. A transforms of different sizes have been used such as 4X4, 8X8, 16X16, 32X32, and 64X64 which gives 100% result and accuracy compare to the existing system such as principle component analysis. The recognition speed is 0.9sec. The complexity comparisonof all the transform shows that the complexity of Haar transform is lesser than the other transforms.

Future Scope for this project will be focused on more complicated gesturesand Gestures with complex background.

VI. ACKNOWLEDGMENTS

We would like to express our sincere thanks to all the Staff of Department of Computer Science, Thadomal Shahani, College of Engineering, for their support and guidance throughout.

VII. REFERENCES

[1]Jagdish La lRaheja, Radhey Shyam, Umesh Kumar P Bhanu Prasad, " Real-Time RoboticHand Control using Hand Gestures" proceeding of International Conference on Machine Learning and Computing (ICMLC) of IEEE, pp.12-16, Feb 2010.

[2]Yuehai Wang, JianfeiLi "Entertainment Robot HandGesture Recognition" proceeding of InternationalConferenceon Database Technology and applications(DBTA), pp.27-28,Nov.2010.

[3] Yikai Fang , Kongqiao Wang ,JianCheng,Hanqing Lu "A Real-Time Hand Gesture Recognition Method" ProceedingofInternational conference on Multimedia and Expo.IEEEpp.995-998, July 2007

[4] Deng-Yuan Huang, Wu-Chih Hu, Sung-HsiangChang "Vision-based Hand Gesture Recognition Using PCA+Gabor Filters and SVM" proceeding of International Conference on Intelligent information hiding and Multimedia Signal Processing, pp 1-4,Sep 12, 2009.

[5] Qing Chen, Nicolas D. Georganas, Emil M. PetriuReal-time Vision-based Hand Gesture Recognition UsingHaar-like Features" Proceeding of International conference on Instrumentation and MeasurementtechnologyIEEE, pp 1-6, May 1-3 2007.

[6]Rajeshree Rokade, Dharmpal Doye, ManeshKokare"Hand Gesture Recognition by Thinning Method" International Conference on Digital Image processing 2009 IEEEpp 284-287, March 7,2009.

[7]Zhong Yang, Yi Li, Weidong Chen,Yang Zheng "Dynamic Hand Gesture Recognition Using HiddenMarkov Models" Proceedingof 7th International Conference onComputer Science and Education. Melbourne, Australia IEEE July 14-17, 2012.

[8] Karishma Dixit, Anand Singh Jalal , "Automatic Indian Sign Language Recognition System" Proceeding of IEEE 3rd international conference on Advance computing , pp. 883-887, Feb 22- 23 , 2013.

[9]Archana S. Ghotkar, Rucha Khatal ,Sanjana Khupase, Surbhi Asatiand Mithila Hadap “Hand Gesture Recognition for Indian Sign Language” Proceeding of 3rd IEEE International Conference on Advance Computing , pp1-4, Jan 10-12, 2012.

[10]Prateem Chakraborty, Prashant Sarawgi, Gaurav Agrawal, “Hand Gesture Recognition: A Comparative Study” proceeding of the International Multi Conference of Engineers and Computer Scientists ISBN, pp.19-21, March 2008.

[11] Nasser H. Dardas and Emil M. Petriu “Hand Gesture Detection and Recognition Using Principalomponent Analysis” proceeding of International Conference onDigital Object Identifier, pp1-6, Sept 2011.

[12] H. B. Kekre,Sudeep D. Thepade and Akshay Maloo “Performance Comparison of Image Retrieval Using Fractional Coefficients of Transformed Image Using DCT, Walsh, Haarand Kekre’stransform” Proceeding of International Journal of Image Processing (IJIP)

[13] H. B. Kekre, TanujaSarode, MeenaUgale,“Performance Comparison of Image Classifier Using Discrete Cosine Transform and Walsh Transform” Proceeding of 2nd International Conference and workshop on Emerging Trendsin Technology (ICWET) 2011

[14]Thomas Moeslund's Gesture Recognition Database, http://www-Prima.inrialpes.fr/FGnet/data/12-Moeslund Gesture/ database.html