

Orientation Invariant Object Recognitions Using Geometric Moments Invariants and Color Histograms

Saleha Samad*, Anam Haq, Shoab A. Khan

Department of Computer Engineering, College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan.

* Corresponding author. Email: salehasamad69@ce.ceme.edu.pk

Manuscript submitted September 9, 2014; accepted March 1, 2015.

doi: 10.17706/ijcee.2015.v7.876

Abstract: Object recognition is a very challenging task in artificial intelligence and robotics. Many approaches have been implemented to achieve this task with greater precision and accuracy. In this paper we have implemented the approach of detecting objects in images undergo with the change in scale, rotation, and orientation. Extracting Geometric moments invariant which are extensively use to extort global features and using color histogram approach we have improved the previously recognition rate to a significant measure. The accuracy of classification is increased by adding the new feature of color Histogram which is also an invariant feature for change in scale rotation, translation, and orientation of objects and using support vector machine learning algorithm for classification.

Key words: Canny edge detection, color histogram, geometric moments, radial basis function, support vector machine.

1. Introduction

Identifying the objects using computer algorithms has become an essential part of today's technology. Many efforts have been put forward to complete this task with maximum accuracy. The object to be recognized can undergo with the change of scale, rotation, and orientation. In such cases the segregation of distinct and efficient features is an important task. Gabriel Takacs and Vijay Chandrasekhar anticipated a method of Orientation Invariant Image Matching by computing and comparing Histogram of Gradient and work on fast descriptor computation [1]. David G. Lowe develops an algorithm for scale invariant feature extraction by using key point localization, which are invariant to rotation translation and scaling [2]. Andres Solis Montero proposed a new algorithm for shape recognition based on line segment detection and extraction of edge contour from colored images [3]. Jan Flusser and Tomáš Suk work for arbitrary binary 2-D shapes and state that geometric moments play vital role in object recognition [4]. Wei Huang extracted the exact geometric moment of gray level images and introduced the Overlapped Rectangle Image Representation (ORIR) and compute geometric moment is on the ORIR basis [5].

In this paper an object recognition system is deliberated which is invariant to the change in scale, rotation, translation, and orientation. The organization of paper is as follows: Section 2 describes the color histogram of the images. Section 3 discusses the canny edge detection. Section 4 states the geometric invariant moment's calculations. In Section 5, support vector machine is described. Section 6 shows the calculated results and Section 7 gives the conclusion of proposed method.

2. Color Histogram

Another important feature that could be invariant to scale, rotation, and orientation of colored object could be their color histogram. Color Histogram is illustration of the spread of colors in an image. A color histogram represents the numbers of occurrence of pixels for each color in the image. It can be draw out for N-dimensions i.e. for each color present in the image but the three main colors usually are taken into account are Red, Green and Blue [6]. In Colored images 24 Bits are required to correspond the intensity value of each pixel, each 8 bit is used to store the intensity value of Red, Green, and Blue pixels respectively. For 8-bits the intensity value is formed as:

$$d = \sum_{i=0}^7 b_i 2^i \quad (1)$$

where i represent the serial number of pixel information bit, and b_i is the value of information bit. For $i=0$ the b_0 comprise the 0th bit plane b_1 the first bit plane and so on. Feature of bit plane images steadily become complex going from high to low bit. The most of the information of intensity of color pixel of image is stored in high bit plane [7]. So taking only high bit plane under consideration we can extract the required information. This practice is known as Intensity slicing. The high bit plane is known as significant bit image plane as it comprises the main information about the image. The histogram of image in a color space C is formed as:

$$H_i = \{N(I, C_i) | i \in [1, 2 \dots N]\} \quad (2)$$

H_i is the histogram which gives number of occurrence of each color pixel in the image. Where i specify the color levels of color space C and $N(I, C_i)$ is the number of the pixel of image I that fall into the cell C_i [8].

3. Canny Edge Detection

In Image Processing techniques, edge detection is the prior step to detect outlines and boundaries of the object in order to separate it from background. Purpose of edge detection is to change a 2-D image into sets of curve. Edges of the objects are inclined to give rapid changes in the intensity of image e.g. color, hue and light intensities. So the intensity change in images depicts the geometric information about the objects in image specially the edges. Canny operator is considering being most optimum edge detector. It works in multi stages. First the Gaussian Convolution is done to smooth the image. Then to emphasize the regions of image first derivative operator is applied. Edges give rise to ridges in the gradient magnitude image. Then non maximal suppression is done to highlight the required edge pixels by setting the other pixels to zero that are not laying on the ridge which gives a thin fine line in the output. In the end the hysteresis is done by using two threshold values on the obtained thinned edge image. The edge pixels below the lower threshold value are neglected while that having pixel values greater than high threshold are connected to labeled as Edge pixels. The three input constraints are needed by the canny edge detector from the user. Standard deviation (Sigma) of the Gaussian filter, define in pixels. The second and third are low and high threshold values mentioned as "low" and "high" respectively. The "low" is stipulate as fraction of calculated high threshold and "high" is specified as a percentage point in the distribution of gradient magnitude values for the entrant edge pixels to use in hysteresis [9].

4. Geometric Moment Invariants

Moment invariants have been used in many approaches of pattern recognition. In 1961 the inspiration of using moments for object or shape recognition came forward. The use of Geometric Moment in image

processing and pattern recognition was first implemented by Hu and Alt. They came up with the idea to utilize the approach of known mathematician Boole, Sylvester and Cayley. These moments remain unaffected if the image endures changes like:

- Scaling
- Translation
- Rotation
- Orientation

Hu declared that if $f(x, y)$ is a continuous hybrid function and has nonzero values in a finite region of x, y plane, then the moment succession $\{m_{pq}\}$ uniquely describes the $f(x, y)$.

$$m_{pq} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} x^p y^q f(x, y) dx dy \quad (3)$$

where $p, q = 0, 1, 2, \dots, \infty$, while the product of x^p and y^q is the source function for this moment and the set of n moments contain all m_{pq} for $p+q \leq n$, i.e there are $\frac{1}{2}(n+1)(n+2)$ entities [10]. As the segment of the image is finite region a moment set can be generated which will have distinct information of the image segment. For two-dimensional images these moment invariants gives the distinct features. They are the peculiarities of connected regions in binary images that are invariant to scale translation and rotation. They characterize a simple calculated region behavior in a set that can be utilized for recognition of shape or any area. η_{pq} are denoted as normalized central moments.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (4)$$

where $\gamma = \frac{p+q}{2} + 1, p+q=2, 3, 4\dots$

From second and third normalize moment Hu derived a set of seven invariant moments and this set of moment from (5) to (11) is invariant to change in size, rotation and translation [11].

$$\phi_1 = \eta_{20} + \eta_{02} \quad (5)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (6)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (7)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (8)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} - \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} - \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]] \quad (9)$$

$$\phi_6 = (\eta_{20} - 3\eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})] \quad (10)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} - \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + (\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]] \quad (11)$$

5. Support Vector Machine

In 1990s the Cortes and Vapnik proposed the theory of Support Vector Machine. It works on the principal of structural risk minimization. The local optimal solution in support vector machine is the globally optimal

solution. The classification for more than two but finite classes is done by creating sets of hyperplanes in a high dimensional space for nonlinear data. The SVM uses two parameters the mean and the standard deviation [12].

$$f(x) = \text{sgn}[\sum_{i=1}^n \lambda_i y_i K(x_i, x) + b] \tag{12}$$

The performance of the SVM is based on the kernel function used for generalization. Where K is the kernel function and for multi class nonlinear classification the most optimal kernels are Gaussian radial basis function. This is given as follows:

$$K(|x - x_i|) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \tag{13}$$

In fact, the kernel function lies in high dimensional space. Providing the kernel function must conglomerate the Mercer Theorem to correspond to the inner product of transformation a space so that the non-linear classes can be separated linearly forming suitable hyperplanes by this dot product of kernel functions. The learning and classification ability of RBF kernel is better and reliable for multi class labeling. RBF is considered being the most studied neural networks among all and hypothetically consists of two layers of neurons. The hidden neuron layer also called basis function is associated with a sample vector while the output neuron corresponds to the predicted class. An RBF network is made up of three layers, input layer, hidden layer and output layer.

The input layer generates the coordination between the input vector and the units in the hidden layer. An activation function is produced based on the associated RBF by the each unit in the hidden layer. In the end, output layer provides a linear combination of each unit in output layer with the units in the hidden layer. The response of RBF network for the given input is completely determined by the activation functions related to the hidden layer units and weights connected between the hidden and output layer [13], [14].

6. Proposed System

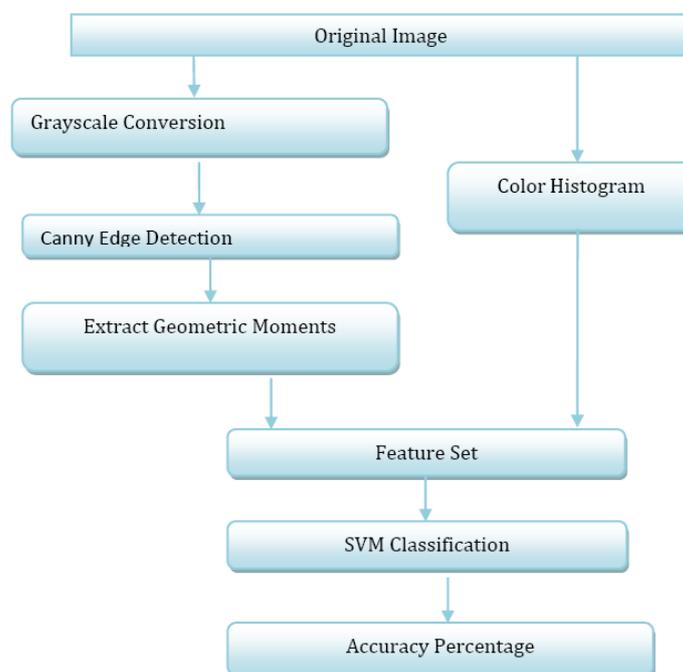


Fig. 1. System block diagram.

In our Proposed System as shown in Fig. 1 the original image first takes into account to obtain the color Histograms of the Objects. On the other hand, for extracting the geometric moments the colored images are converted to grayscale. Then canny edge Detection is applied to get the edge contours of the objects. From these edges detected images the seven geometric moments are calculated. Then with the RGB component values and the seven moment invariants the feature set is formed. The feature set is divided into training and testing data. Using Support Vector machine with Radial Basis Kernel Function classification is done. The accuracy of classification for each accurately classified object is calculated.

7. Experimental Results

The proposed method for feature extraction is experimented on Coil-100 dataset which is publicly available on Columbia University Image Library website with latest version of MATLAB with 15 Objects captured by different poses and rotating with an angle of 5 degrees in each image. Each object has 72 different viewpoints with a noticeable change in scale, angle and viewpoints. During experiment first step was to obtain the color histogram of the images shown in Fig. 2 and then after that convert them to gray images and perform some pre-processing steps to remove noise and then applying canny edge detection in Fig. 3 and these edges detected images are then used to extract the normalized geometric moment invariants.

Table 1. Seven Geometric Moment Invariants

Objects	First	Second	Third	Fourth	Fifth	Sixth	Seventh
Object 1 at 0'	25.35022	390.7508	8.6477e+03	2.1198e+04	2.78e+08	4.1900e+05	-5.93E+08
Object 1 at 5'	25.61637	401.3392	9.050e+03	2.1658e+04	2.82e+08	4.3385e+05	-5.91E+08
Object 1 at 90'	25.76227	354.7333	7.4370e+03	2.0886e+04	2.48e+08	3.93305e+05	-4.96E+08
Object 1 at 110'	26.53294	390.546	9.056e+03	2.3715e+04	3.54e+08	4.6866e+05	-7.45E+08
Object 2 at 80'	31.56873	787.3904	2.3389e+04	3.4879e+04	1.09e+09	9.7845e+05	-1.83E+09
Object 2 at 170'	31.56198	772.8264	2.2769e+04	3.5599e+04	1.1e+09	9.8932e05	-1.86E+09

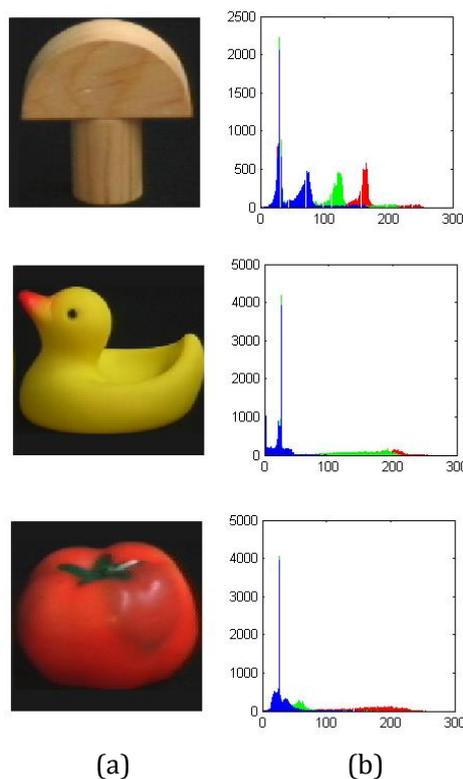


Fig. 2. a) Original image. b) Color histogram.

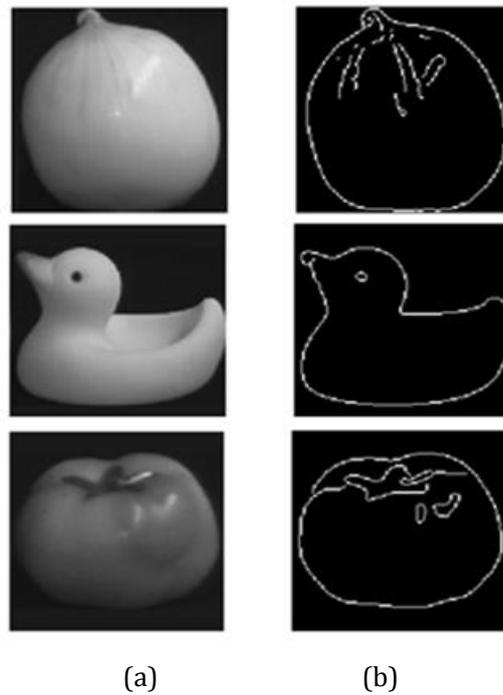


Fig. 3. a) Grayscale image. b) Edge detected image.

Table 1 shows the geometric moments of two objects at different angles and the Fig. 4 displays that the geometric moment invariants of object 1 with different angles are lies so close to each other while that of objects 2 with different angles also lies closely irrespective of the change in angle and viewpoints. Then a feature vector is formed by arranging the color histogram values and these moment invariants in a matrix for testing and training. A support vector machine algorithm for multi classes is then applied to for the classification. SVM using radial basis function achieved accuracy to 95%, which was previously calculated up to 93% [11]. Accuracy for each object has also been calculated individually.

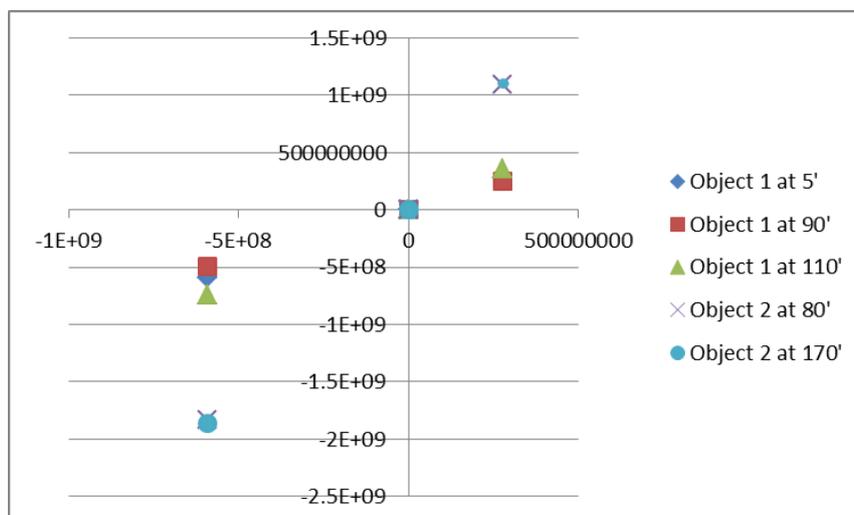


Fig. 4. Scatter graph for Table 1.

8. Conclusion

In this paper the color histogram for the images is obtained and also the seven geometric moment invariants after changing the images to gray scale and detection of edges from these gray scale images.

Table 2. Accuracies Comparison

Object	Previously Calculated Accuracy	Improved Accuracy
Bottle	95%	97%
Car	93%	95%
Can	92%	97%
Cup	90%	92%
Toy	92%	94%
Tomato	95%	99%

A feature vector is formed by these moment invariants and RGB component values. Then support vector machine is applied for classification. From the results shown above it is evident that the geometric moment invariants and color histogram are the features that are invariant to change scale, rotation, translation and orientation. Also the SVM classifier using RBF kernel is the more efficient than previously used classifiers. Using this method recognition rate has been improved to a significant measure. The accuracy for individual class has been measured up to a remarkable improvement (shown in Table 2).

References

- [1] Takacs, G., Chandrasekhar, V., Chen, H. Z., Chen, D., Tsai, S., Grzeszczuk, R., & Girod, B. (2007). Information systems laboratory. *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*.
- [2] Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the International Conference on Computer Vision: Vol. 2*.
- [3] Montero, A. S., Nayak, A., Stojmenovic, M., & Zaguia, N. (2009). Robust line extraction based on repeated segment directions on image contours. *Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications*.
- [4] Flusser, J., & Suk, T. On the calculation of image moments. Institute of Information Theory and Automation, Academy of Sciences of Czech Republic, Czech Republic.
- [5] Huang, W., Chen, C., & Sarem, M. Exact geometric moment computation for gray level images. School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, Hubei, 430073, China.
- [6] Lee, K., & Lee, C. Fast object detection based on color histograms and local binary patterns. Department of Electrical and Electronics Engineering, Yonsei University, Seoul, South Korea.
- [7] Phan, D., Oh, C.-M., Kim, S.-H., Na, I.-S., & Lee, C.-W. (2013). Object recognition by combining binary local invariant features and color histogram. *Proceedings of Second IAPR Asian Conference on Pattern Recognition*.
- [8] Chang, P., & Krumm, J. Object recognition with color cooccurrence histograms. The Robotics Institute Carnegie Mellon University, Pittsburgh, PA, USA.
- [9] Gao, J., Xie, Z., & Wu, X. (2007). Generic object recognition with regional statistical models and layer joint boosting. *Pattern Recognition Letters*, 28(16), 2227-2237.
- [10] Hosny, K. M. Exact and fast computation of geometric moments for gray level images. Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Zagazig, Egypt.
- [11] Muralidharan, R., & Chandrasekar, C. Scale invariant feature extraction for identifying an object in the image using moment invariants. Department of Computer Applications, KSR College of Engineering, India.
- [12] Wang, F., He, K., & Liu, Y. (2013). Research on the selection of kernel function in SVM based facial expression recognition. *Proceedings of IEEE 8th Conference on Industrial Electronics and Applications*.



Saleha Samad was born in Bahawalpur, Pakistan in 1989. She received the BS degree in computer engineering from Islamia University of Bahawalpur (IUB), in 2010. She is currently pursuing the MS degree in computer engineering with the Department of Computer Engineering, NUST, College of Electrical & Mechanical Engineering.



Anam Haq was born in Islamabad, Pakistan in 1988. She received a BS degree in computer engineering from Comsats Institute of Information Technology Islamabad in 2010. She has completed her MS degree in computer engineering with the Department of Computer Engineering, NUST, College of Electrical & Mechanical Engineering, Rawalpindi Pakistan.



Shoab Ahmed Khan was born in Lahore, Pakistan in 1965. He has done his PhD degree in digital signal processing from Georgia Institute of Technology, Atlanta, GA. Dr. Khan has more than 17 years' industrial experience in companies like Scientific Atlanta, Picture Tel, Cisco Systems, and Avaz Networks.