International Journal of Engineering

Journal Homepage: www.ije.ir

Modelling of Eyeball with Pan/Tilt Mechanism and Intelligent Face Recognition Using Local Binary Pattern Operator

M. Sai Viswanathan*, V. Manoj Kumar

Robotics laboratory, Department of Mechanical Engineering, SRM University, Chennai, India

PAPER INFO

Paper history:
Received 15 March 2017
Received in revised form 05 April 2017
Accepted 21 April 2017

Keywords:
Bio-mimic
Blinking
Camera-orienting Device
Face Recognition Using Local Binary Pattern Histogram
Classifier

ABSTRACT

This paper describes the vision system for a humanoid robot, which includes the mechanism that controls eyeball orientation and blinking process. Along with the mechanism designed, the orientation of the camera, integrated with controlling servomotors. This vision system is a bio-mimic, which is designed to match the size of human eye. This prototype runs face recognition and identifies, matches with a face in the database. Recognition of face leads to capture the facial image and synchronize with the face. As the individual shows any motion, the system also moves according to it.

***doi**: 10.5829/ije.2017.30.07a.20*

1. INTRODUCTION

Study of the vision system of a humanoid robot is very important for the communication between human and humanoid. This paper describes the characteristic feature of the humanoid head through vision system in an attempt to mimic the structure of vision system. A humanoid eye with pan and tilt mechanism is integrated with the vision camera and explore the image of the surroundings with low resolution. It studies the recognized images which are stored in the database to certain samples and with the regularly trained data. Thus, it enables to recognize the object/face in front of the camera, which is more useful in real time interaction with humanoid. This approach is accomplished by obtaining gray scale images which use low memory and even surface [1]. However, the image captured by the camera will be in BGR, which is converted to grayscale to train the data and to detect. This information can be used as the input to the classifier as stored in the database [2]. The trained data will be in its 2D position and orientation.

*Corresponding Author's Email: *saiviswanathan31290@gmail.com* (M. Sai Viswanathan)

2. DESIGN AND MECHANISM

The dimensions of the human head are taken with reference to the average age of human being which includes the pan/tilt of camera and blink mechanism fabricated model with a supporting neck mechanism which acts as a neck to provide output more precisely. The eyeballs are designed to provide independent eyeball movements, so each camera has different degrees of freedom. This prototype has six degrees of freedom. The overall size of this robot is 300mm x 210 mm. The distance between the two Centre's of an eye is 110mm. The dimension includes neck like supporting along with its base. It is shown in Figure 1.

2. 1. Mechanism of Blinking and Pan/Tilt Orientation of Camera Both the eyes of robot contains eyelids which provide blinking mechanism [1]. The blinking process is provided in the view to calibrate the camera. To increase the field of view, multiple degrees of freedom with individual panning and tilting range. This process is synced with the camera to mimic the human eye [3, 4].

Please cite this article as: M. Sai Viswanathan, V. Manoj Kumar, Modelling of Eyeball with Pan/Tilt Mechanism and Intelligent Face Recognition Using Local Binary Pattern Operator, TRANSACTIONS A: Basics Vol. 30, No. 7, (July 2017) 1098-1104

Figure 1. Fabricated model of a humanoid vision system with camera holding structure and connecting plate which support pan and tilt mechanism. The Camera is covered by 180deg eyeball

The mechanism of the panning and tilting range is fixed, and the range for the pan is 35^o and for tilt is 32^o [3]. Optimal configuration is achieved by a selective algorithm. It is shown in Figure 2. Fabricated design of tilting is shown in Figure 3.

Position and orientation are controlled by direct actuators. An alternative approach is provided for future modification by using two-bar mechanism [5] which can provide diagonal trajectory.

3. METHODOLOGY

Supporting base is used to give the neck like rotation along x and y-axis (configured with the idea generated from joint arm articulated robot)

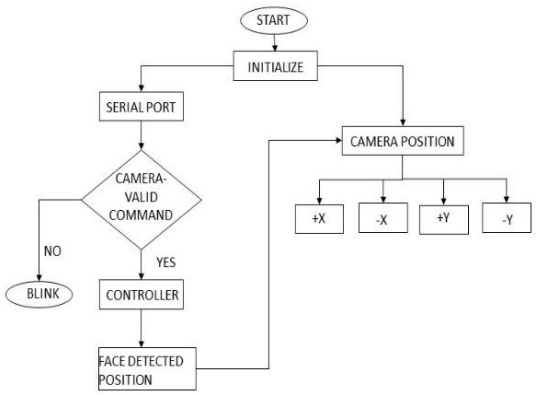

Figure 2. Algorithm for pan/tilt

Figure 3. Fabricated model of tilting mechanism with fixed angle

The main supporting element holds the actual working elements and it is supported by a base. A separate side frame is designed to hold the eyelids. It is independent of pan/tilt mechanism and camera. The lower lid remains to be fixed while the upper lid is used for calibrating by opening and closing the eyelid. It is controlled by direct drive actuators. Camera holder is particularly designed to hold high definition Microsoft camera. This design helps the eyeball to rotate in either of the direction. The working methodology is configured in two processes. Initially, the fabricated prototype is controlled using Arduino (to control direct drive actuators). Active camera is controlled by a separate master controller (i.e. laptop). After both are configured and with respect to the pan/tilt algorithm, Arduino is interfaced in open CV to provide coordinated movement of camera and pan/tilt mechanism which are controlled by motors. A separate algorithm is developed. It is obvious to vary the rpm of servos used [1]. It can be done by adopting a time delay of varying angle. Face recognition is done by using python by detecting the face initially and forming a database. This database helps to store the captures image of stipulated sample sets. An extension of '.yml' is created to train a data by the captured image. Set of images stored in the database is given with a separately used name, so it enables the programmer to assign names of the object/face detected by the camera. In this method, an haarcascade [5, 6] method is used which detects the frontal face. LBPH (Local binary pattern histogram) is used in particular because of structured lighting environment [6]. This methodology is shown in Figure 4. The two used cameras act as a stereo vision, capture the image, store and do matching in the database.

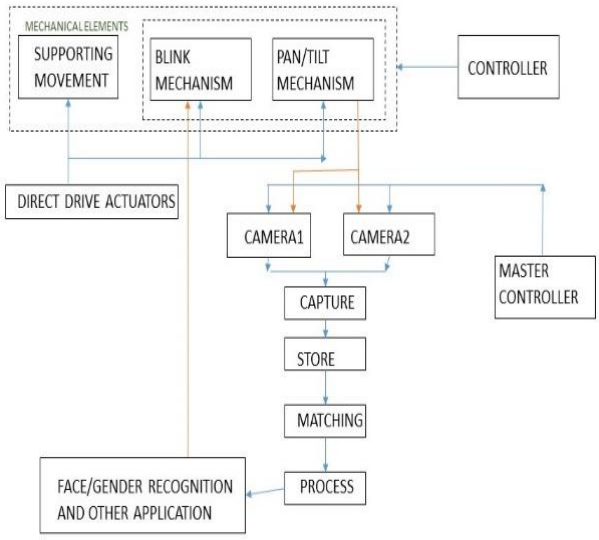

Figure 4. Block diagram

4. HISTOGRAM CLASSIFICATION-LBP

This type of operator is used to describe the color saturation of the image. It relates the pixel value with its neighbor. The threshold value is limited to 3 x 3 matrix, and then it is labeled as descriptor using a histogram. It is shown in Figure 1. This LBP as the advantage of grey-scale and rotating invariance is used to classify the pixels of the image [7, 8]. In this paper, we modify this algorithm by extracting the features which are based on the appearance of cropped section of the recognized spot. Initially, the images are captured by the active camera of different expression and processed in the dataset. At the time of sequential display, algorithm crops to the section of the facial area, eliminating other parts of images [7, 9] and reduced by limiting the samples to be stored in the dataset.

4. 1. Analyzing by Capturing the Face in Different Lighting Condition by using Local Binary Pattern

The above histogram represents the image of faces stored in database captured in different lighting conditions by using local binary pattern histogram. Figure 6 represents the image captured in poor lighting condition, whereas the surface recognition fluctuates in the region surrounded by detected areas. Figure 7 represents the surface of the detected face with evenly distributed recognition rate. Hence, using modified LBP operator ahead of other detecting algorithms gives more precise information in the structured lighting condition.

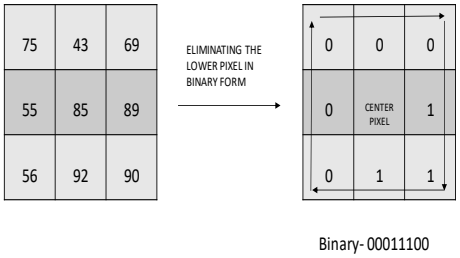

Figure 5. Elimination of lower value pixel with respect to center pixel

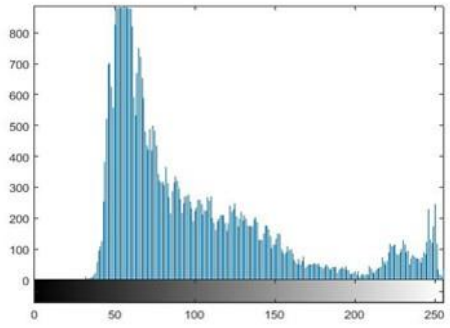

Figure 6. Histogram of image captured in lowlight condition using LBP

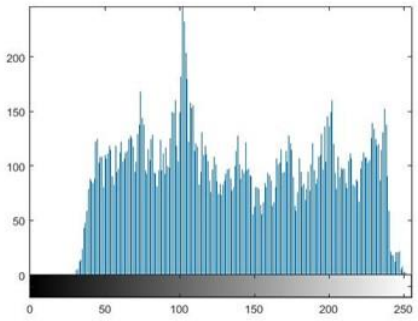

Figure 7. Histogram of image captured in normal lighting condition using LBP

4. 2. Modified LBP Basic LBP based classifier using frontal face detects the face with its most common parameter which includes left and right eye, nose and lips. This LBP is modified by separating the image stored in database into 20 different regions and operating the each region by 3x3 matrix formation, and eliminating the pixel which is less than the center pixel. This results in quicker and more accurate information as described in Figure 7. On detecting the real time face, the system undergoes the matching process with quick intervention. Figure 8 represents the maturity of the cascade classifier about its detection techniques for the frontal face and Figure 9 represents division of the image into 20 regions. The threshold value for the poor lighting condition is explained in section 4.1.

4. 3. Statistical Analysis under Varying Lighting Conditions To carry out the accuracy of face recognition using local binary pattern histogram, it is tested in different lighting condition. Threshold value for recognizing face in day light is high when compared to other detecting algorithm. It is shown in Figure 9. The stability of recognition rate which uses LBP as an operator tends to be more stable without fluctuations. In addition, along with different region of the image, accuracy is increased with the change in brightness of the image.

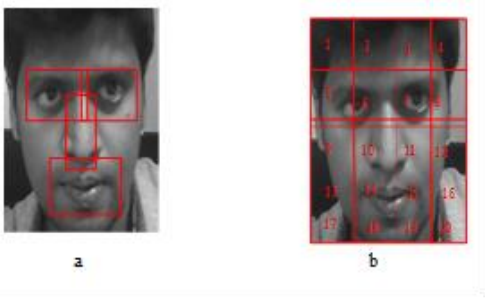

Figure 8. (a) For face recognition, a face image is roughly divided into four semantic patches. (b) For identity recognition, a face image is divided into twenty semantic square regions

TABLE 1. Accuracy of recognizing rate at different lighting conditions

Frames	Dark	Artificial lights	Normal day light	Sunlight With shadow	Sun light
20	0	48	118	142	180
40	2	52	122	160	179
60	3	44	125	163	186
80	3	39	119	158	203
100	5	36	121	155	195

Table 1 shows the intensity value of varying lighting conditions which describes the suitable threshold value to detect the face by using LBP operator. These values are determined by using light dependent resistor and the intensity value is determined by finding the output value of the resistor.

$$intensity = \left(\frac{\frac{2500}{Vout} - 500}{1k\Omega}\right)$$

The intensity level changes with respect to the amount of light falls on it. Figure 9 shows that LBP operator work more accurate in normal day light conditions in a structured environment.

5. VISION PROCESS

Real-time vision system requires two cameras with high definition. The image captured by the camera will be in color mode which is later converted to grayscale. Open CV will work only if the image is converted to grayscale [6, 10, 11].

5. 1. Face Detection Harr cascade classifier is used to detect the face. It is done by configuring the open CV. Then, it detects the faces. It is done by using cascade classifier (opt-front face default). After that, a variable is formed to call the classifier. When the variable is loaded, an image is captured using the external web camera. It is initiated by forming a loop.

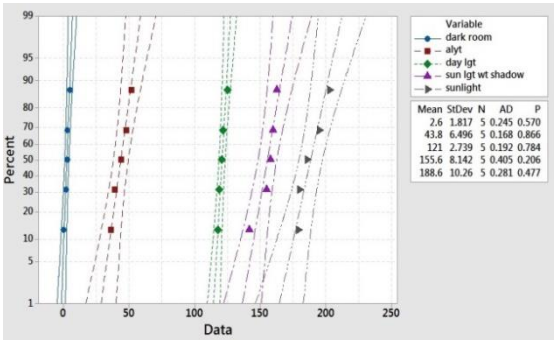

Figure 9. Plot of recognition rate under varying lighting conditions

In this loop, it will be able to capture the picture, but it will be in a colored form (shown in Figure 10). It is then converted to a grayscale image for the cascade classifier to perform. It is converted by using a BGR2gray module [12]. Using a grayscale image, the classifier can detect the faces of multiscale [6]. It forms the rectangular section on the focused part of the image which is captured within 0-255.

5. 2. Forming the Dataset Using face detection algorithm and cascade formation, a data set must be created (shown in Figure 11), face recognition folder is created separately, in which previously done face detection algorithm is used. In this algorithm, when it captures the face, it will be written in a separate folder. Before capturing the face, it is necessary to tell the program whose face it is. It is done by using IDENTIFIER. With this, a separate id is given to a particular face while capturing the image and it can be named later in the coding set after capturing a face, it should be written a file-> a separate folder called dataset. In that folder, name is given by user id+1 (using sample number with respect to rectangular cross section). Since samples are given in incremental order, hence it is needed to break it to required numbers. Only that rectangular section will be stored with respect to height and width.

5. 3. Training the Dataset In training, it is needed to get a sample from the data set folder (which id number is for which face).

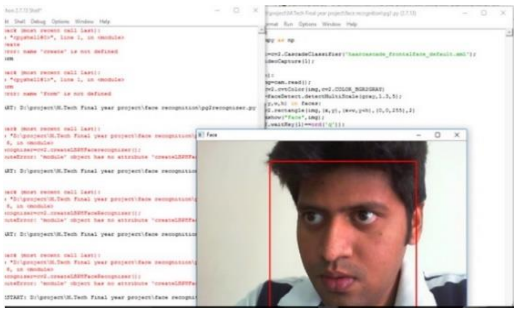

Figure 10. Detecting the face using open cv

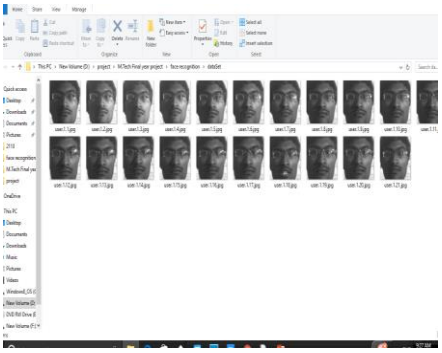

Figure 11. Creating dataset in grayscale

It is necessary to capture all the relative path of the image file. It can be done using python library, import operating system and open CV. As capturing images, PILLOW library is used. Following that, the recognizer is created to train the image and path of the sample is needed (the relative path where samples are stored). A separate method is created to produce corresponding ids and path is created. Lists of the image are created with relative path (os library is used here). It is used to list all of directory which is picture in data set folder and fetching the directory from picture, transferring to file and joint name is appending the file name with its path. Now, all the images can be looped with user id, before that empty list is created for face and user id. After creating an empty path, from the relative path of captured image, it must be converted to numerical value since open CV works only with np. Also, images in python image library are converted to numerical values. Next, we should get user id. This can get from the name of the picture (path splitter and dot splitter is used to get the path). Since it is in string format, it must be converted to an integer. Faces and IDs are appended in to empty a relative path with respect to numbers. Image command is used to relate the camera with numerical values. Similar to previous algorithm, recognizer folder is created and trained data is saved in it. It is stored in '.YML' extension. Each time when it detects, it trains or updates the stored data.

5. 4. Detector Formulation Face detection is recognized using haar cascade classifier. LBPH. Eigen and Fisher face are like holistic towards face recognition. Data is considered as a vector in some high dimensional image. Else-if statement is used to give the name for the sample of dataset images stored, which is of approximately 20 samples. Gray images are again converted to colored image once the data is trained. It detects the object or face by using the cropped part of the image by using its pixel length and width. It is shown in Figure 12. Delay is given, as the open Cv doesn't work without delay.

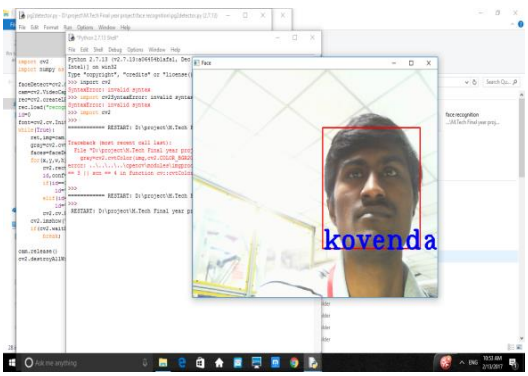

Figure 12. Detecting the face with name by the trained set

5. 5. Local Binary Pattern Histogram This classifier is mainly used for structured lighting environment. It helps in extracting the most useful sections of the image stored in the dataset. This extraction is done by using LBP operator [13, 14]. It separates the image of a pixel into a small region. The algorithm used here is shown in a flowchart in Figure 13. Here, LBP develops the particular structure of section in an image. It compares the pixel intensity from its center, which described in 8-connectivity.

6. NOTATION

6. 1. Pan/ Tilt It is necessary to describe the dimensional view of both independent moving active cameras. It is represented in column vector $[\ x\ y\ z\ 1\ ]^T$. It represents in two dimensions by using single camera. A 4x4 matrix is determined for its homogenous transformation. (as shown in Figure 14).

Each camera is separated by the individual transformation matrix. Therefore, for camera, A= [Ax, Ay, Az]^T and camera B =[Bx, By, Bz]^T.

6. 2. Local binary pattern histogram Formal description for LBP operator is given as:

$$lbp[x,y] = \sum 2\mathrm{p}\, S(ip - ic)$$

where,

$$s(x) = \begin{cases} 1, & if\ x > 0 \\ 0, & otherwise \end{cases}$$

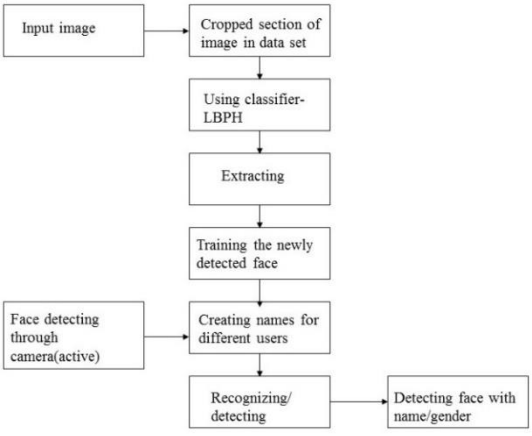

Figure 13. Algorithm used for pattern recognition

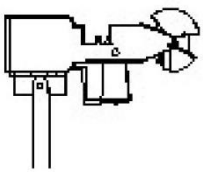

Figure 14. Movement of camera (kinematic)

This formulation helps to get the details of every pixel region of a particular section in an image [13]. The basic idea is done by aligning the neighbors with different radius. By definition, the image can be modified artificially. The histogram is extracted from each image by using LBP to divide the image.

7. CONCLUSION

Design and fabricated prototype of panning and tilting the camera are presented in this paper**.** The result of this paper is obtained by providing a pan/tilt range in and around 35^0 by interfacing with the camera, recognizing the face by a classifier and training the datasets created. Active camera used to present this prototype is Microsoft life cam 3000HD which have been tested using open CV and Python language. Blinking and pan/tilt are controlled by using micro servo motors and a neck like support is provided to make pan/tilt more flexible. This configuration can be converted to binocular vision based on the design. Local binary patterns are the most useful and powerful feature for texture classification. So, it is preferred ahead of other recognizers.

8. ACKNOWLEDGEMENT

We thank the department of robotics for the support and implementation, school of mechanical engineering, Srm university, India.

9. REFERENCES

1. Institute, J.S. and Ude, A., "Active humanoid vision and object classification", in Computer and Information Sciences, ISCIS. 24th International Symposium on, IEEE., (2009), 387-391.
2. Murray, D. and Basu, A., "Motion tracking with an active camera", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, Vol. 16, No. 5, (1994), 449-459.
3. Lee, Y.-C., Lan, C.-C., Chu, C.-Y., Lai, C.-M. and Chen, Y.-J., "A pan–tilt orienting mechanism with parallel axes of flexural actuation", ***IEEE/ASME Transactions on Mechatronics***, Vol. 18, No. 3, (2013), 1100-1112.
4. Villgrattner, T. and Ulbrich, H., "Design and control of a compact high-dynamic camera-orientation system", ***IEEE/ASME Transactions on Mechatronics***, Vol. 16, No. 2, (2011), 221-231.
5. Wang, K. and Zhang, H., "Research on an improved algorithm of face detection based on skin color features and cascaded ada boost", in Communication Technology (ICCT), IEEE 16th International Conference on, IEEE., (2015), 779-782.
6. Goel, P. and Agarwal, S., "Hybrid approach of haar cascade classifiers and geometrical properties of facial features applied to illumination invariant gender classification system", in Computing Sciences (ICCS), International Conference on, IEEE., (2012), 132-136.
7. Xu, Y., Zhao, Y. and Zhang, Y., "Multi-scale local binary pattern histogram for gender classification", in Image and Signal Processing (CISP), 8th International Congress on, IEEE., (2015), 654-658.
8. Phillips, P.J., Moon, H., Rizvi, S.A. and Rauss, P.J., "The feret evaluation methodology for face-recognition algorithms", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, Vol. 22, No. 10, (2000), 1090-1104.
9. Xie, L., Wei, H., Yang, W. and Zhang, K., "Video-based facial expression recognition using histogram sequence of local gabor binary patterns from three orthogonal planes", in Control Conference (CCC), 33rd Chinese, IEEE., (2014), 4772-4776.
10. Korayem, M., Ehtemam, V., Sabzevari, R., Madani, M. and Azimirad, V., "An intelligent vision system on a mobile manipulator", ***International Journal of Engineering-Transaction A: Basics***, Vol. 21, (2008), 279-294.
11. Hajihashemi, V. and Borna, K., "An adaptive hierarchical method based on wavelet and adaptive filtering for mri denoising", ***International Journal of Engineering-Transactions A: Basics***, Vol. 29, No. 1, (2016), 31-40.
12. Hassanpour, H., Azari, F. and Asadi, S., "Improving dark channel prior for single image dehazing", ***International Journal of Engineering-Transactions C: Aspects***, Vol. 28, No. 6, (2015), 880-889.
13. Abdur, R., Shafiul, A. and Rashedul, H.N.I., "Face recognition using local binary patterns (LBP)", ***Global Journal of Computer Science and Technology. Graphics & Vision***, Vol. 13, No. 4, (2013).
14. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D. and Huang, T.S., "Total variation models for variable lighting face recognition", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, Vol. 28, No. 9, (2006), 1519-1524.

Modelling of Eyeball with Pan/Tilt Mechanism and Intelligent Face Recognition using Local Binary Pattern Operator

M. Sai Viswanathan, V. Manoj Kumar

Robotics laboratory, Department of Mechanical Engineering, SRM University, Chennai, India

P A P E R I N F O

Paper history:
Received 15 March 2017
Received in revised form 05 April 2017
Accepted 21 April 2017

Keywords:
Bio-mimic
Blinking
Camera-orienting Device
Face Recognition Using Local Binary Pattern Histogram
Classifier

چکیده

این مقاله، سیستم دید برای یک ربات انسان را توصیف می کند که شامل مکانیسمی است که جهت گیری چشم و فرورفتگی را کنترل می کند. همراه با مکانیزم طراحی شده، جهت دوربین، یکپارچه با سروو موتورهای کنترل شده است. این سیستم بینایی یک زیست شناسی است که برای اندازه گیری چشم انسان طراحی شده است. این نمونه اولیه تشخیص چهره را تشخیص می دهد و شناسایی می کند و با یک صورت در پایگاه داده مطابقت می کند. تشخیص چهره منجر به گرفتن عکس چهره و هماهنگ سازی با صورت می شود. هنگامی که به طور شخصی هر حرکتی را نشان می دهد، سیستم نیز با توجه به آن حرکت می کند.

***doi**: 10.5829/ije.2017.30.07a.20*