Scientific Research and Essays Vol. 5(15), pp. 2044-2050, 4 August, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 ©2010 Academic Journals

Full Length Research Paper

Real-time motion-sensitive image recognition system

Musa Peker and Ahmet Zengin*

Department of Computer Science Education, Faculty of Technical Education, Sakarya University, 54187 Esentepe Campus, Sakarya, Turkey.

Accepted 12 July, 2010

This study aims at realizing an effective motion analysis for real time security application by using image processing techniques. After obtaining from the environment, image frames are used to sense and perceive a mobile object, as well as its size and motion degree. In case estimated values exceed its threshold, it is decided whether this image belongs to a human being or a destructive object. Images are stored to the hard disk and then a vocal warning system takes over. The camera is driven by a step motor by which it is possible to track mobile object continuously. The system is enough sensitive to perceive slightest motion and allows user flexibility to select security level. System identifies the person inside by detecting the skin color algorithms and records the image into the computer. The application software has been developed by using C# programming language of .Net technologies.

Key words: Image processing, image recognition, camera security system, face detection.

INTRODUCTION

Image processing deals with retrieving the images from the real environment using a camera, transferring them into computing environment and processing the data after digitalization. Real time image processing applications require hardware such as computers, cameras and image capture cards. The detected image is turned into a matrix in the computer environment and undergoes the required applications. Some of these applications are image enhancing, finding edges and detecting and identifying the position of an object in the image (Köker, 2002). Tracing the mobile objects in image patterns is one of the most important challenges in image processing applicaions (Cucchiara et al., 2000).

The main interest in tracking mobility started increasing after the late 1970s (Kefeli et al., 2005). There are various application areas in the field of motion tracking and analysis such as military, security and detection, traffic, road and parking control (Koller et al., 1993) (Stiller and Konrad, 1999). For example, tracking and destroying a mobile object in a military action is very important. Along the same lines, it is only possible to automatically interpret the human activities in highly sensitive environments through the use of a developed, strong and dependable system that has ability to detect and track human being Image processing interfaces that necessitate human-computer interactions to develop multi-interactive environments require the tracking of human beings/objects (Foresti, 1999).

The goal of this study is to realize highly sensitive motion analysis in real-time and manage face recognition application. The study puts forth and exemplifies a security system in the context of its goals. The system automatically detects tracks and records the image of the mobile object in the database.

The study undertakes a face recognition study along with motion tracking. The basic purpose for face recognition is to correctly and credibly realize security and track objects which is the most important step in the automatic labeling of photos or videos that require assembling/ montage. The article consists of four sections. Following the introduction as first section, the second section provides basic information related to techniques used in the development of the system and the third section presents the phases of the application. The fourth section provides the reader with the results and evaluation of the study and gives suggestions for the future.

*Corresponding author. E-mail: azengin@sakarya.edu.tr.

METHODS

Motion detection algorithms

The visual-based motion analysis starts with the detection of the motion. Motion detection depends on the differentiation of the mobile area in the image from the rest of the areas. Motion detection algorithms are usually examined in three groups (Kefeli et al., 2005):

1. Background difference methods:

a. Difference method.
b. Background modeling methods using median filtering.
c. Weighted sum method.
d. Double Background method.

2. Statistical methods.
3. Visual flow methods.

Background difference method is a very popular approach since it provides ease of use in applications. In this method, the mobile areas are determined by comparing the differences in the pixels of the observed image and the reference image in the same coordinates with those of a pre-determined threshold value. Threshold value is an experimentally pre-determined value that identifies whether the differences between two pixels are based on movement. The study uses the method of differences which is least affected from the elements in the environment in order to identify the change between two consecutive/sequential images. According to the method of measuring differences, assessment of the absolute difference in the sequential frameworks is the basic method in order to detect the possible motion in images with fixed backgrounds (Kefeli et al., 2005). When there is no motion, it is expected that the absolute differences among the frameworks would be rather low (theoretically zero). However, because of the effects of thermal noise and flicker in the sensors, this difference is never found to be zero. Because of this reason, sensitivity and accuracy adjustments need to be set in the best possible degree (Oral and Deniz, 2005).

In this method, sequential color image frameworks that are captured by camera are turned into gray images in order to reduce computational burden in processing. Equation (1) is used in this process (Çetin, 2010):

$$I = (R + G + B) /3 \quad (1)$$

R: red component in the pixel in RGB image.
G: green component in the pixel in RGB image.
B: blue component in the pixel in RGB image.
I: gray_level component.

The next step is to identify whether the ratio of absolute difference of the sequential images to the total pixel number in the image goes over the specified threshold. The detection of the change can be calculated mathematically with Equation (2) (Coşar et al., 2004):

$$OÇF(T) = 1 \frac{1}{w \times h} \sum_{x=1}^{w} \sum_{y=1}^{h} | I_t(x,y) - I_{t-1}(x,y) | > E_1 \quad (2)$$

OÇF: Average framework difference,
I_t: The captured framework of the moment,
I_{t-1}: Previous framework,
w: Horizontal pixel dimension of the captured image,
h: Vertical pixel dimension,
(x, y): The position of the related pixel,
E_1: Threshold value.

According to equation, if the absolute value of the difference between the brightness values in I_t and I_{t-1} pictures which belong to an image component with any (x,y) position, is higher than the specified threshold value, this image component is considered to have motion (Oral and Deniz, 2005). According to this consideration, the areas that show motion are obtained from the difference between the background and the present image. The threshold value is a value that corresponds to a statistically important brightness change. This value is determined experimentally and corresponds to a certain number. The change in the value provides the increase or decrease in the sensitivity of motion detection.

Skin color identification algorithms

The most important step in face recognition methods is the accurate identification of faces that appear in the image. The number and quality of the faces in the captured image affects success of the face recognition method directly. Although there are various methods in the literature that are used for face recognition, many of them are inadequate because of the noise and low resolution of system video archived (Coşar et al., 2004). Skin color based algorithms create successful results in the accurate identification of the facial area. Used skin color algorithm is a method in which facial area is detected by identifying the areas in the image that have similar characteristics by employing the knowledge related to the average facial color of the human face (Yang et al., 2002).

This study utilizes two different skin color algorithms to best identify the human facial areas. These are RGB code technique and YCbCr code technique algorithms (Lyon and Vincent, 2009).

According to RGB code technique, Equations (3) are used since the human face carries red and green color tones mostly:

$$G=G/(G+R+B) \quad (3)$$

Different skin colors were examined by using RGB code technique and the most appropriate value range for the detection of skin color was targeted. Therefore, $R > 0.40$ and $0.25 < G < 0.33$ ranges were found as most appropriated values in the detection of skin color areas.

According to YCbCr code technique, Y component stands for information on brightness, Cb and Cr components present the color information in the YCbCr color space. So information on brightness can easily be gathered. RGB color space can be turned into YCbCr color space by using Equation (4) (Url1, 2010):

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ Cb &= - 0.169R - 0.332G + 0.500B \\ Cr &= 0.500R - 0.419G - 0.081B \end{aligned} \quad (4)$$

In the process of locating the skin color area, Cb and Cr values are used. Maximum and minimum values for Cb and Cr components are calculated and pixels between these values are marked as skin color areas (Kim et al., 2009). Maximum and minimum values for Cb and Cr components are calculated as seen in Equation 4. The minimum and maximum values of Cb and Cr components are determined according to the mean average (Ort_Cb, Ort_Cr) and standard deviations (Std_Cb, Std_Cr) of these components (Nabiyev and Kurt, 2007):

$$\begin{aligned} Cbmin &= Ort\_Cb - Std\_Cb * f \\ Cbmax &= Ort\_Cb + Std\_Cb * f \\ Crmin &= Ort\_Cr - Std\_Cr * f \\ Crmax &= Ort\_Cr + Std\_Cr * f \end{aligned} \quad (5)$$

In this paper, Ort_Cb = -11.1051, Std_Cb = 4.3568, Ort_Cr = 22.9265, Std_ Cr = 3.9479 values were employed and f = 3 was selected. After identifying the skin color areas, these areas are

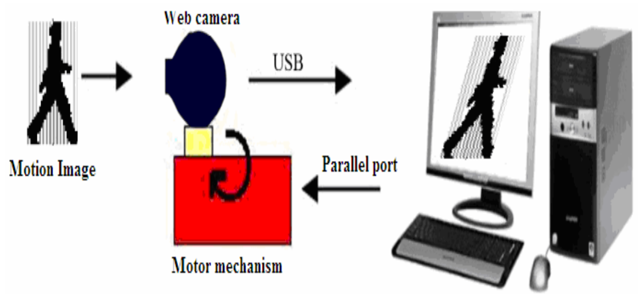

Figure 1. The general structure of the system.

(a) (b)

Figure 2. Web camera and motor mechanism: (a) Inner structure of the system; (b) Outer structure of the system.

classified according to their density and the weighted centers in the dense areas are found. The skin color area is framed in a box and identified by using the weighted center (**Figure 7**).

A SECURITY APPLICATION WITH REAL-TIME MOTION DETECTION AND FACE RECOGNITION

The developed system consists of two parts: Hardware and a user interface. The hardware includes the camera mounted on the step motor mechanism which detects the moving object and transfers the image into the computer via parallel port. On the other hand, camera is connected to computer via USB interface. The general structure of the system that tracks the object and recognizes faces can be seen in Figure 1.

Mechanical part

The web camera on a platform mounted to step motor is connected to the USB port of a PC. An electronic circuit was developed and connection was provided between the hardware and software in order for the step motor to track the mobile object. This circuit was prepared by using 6 phase step motor, ULN 2003 which is integrated via parallel port. The inner structure of the system and the camera mechanism can be seen in Figure 2(a).

Software

C# programming language from.Net Technologies has been used in creating software for managing the system. Program provides great convenience to the user since it allows many functions to be used at the same time. Figure 3 shows the user interface of the program. Numbered parts in the Figure 3 are given as follows:

1. Control panel window has the tools that allow the system to work and for the adjustments to be made.
2. Web camera window shows the images taken from the web camera.
3. Face sensitive window detects the facial area with skin color algorithms and geometric methods and marks it with a red box.
4. Motion-sensitive screen window detects the object in motion and tracks it by calculating the geometric center of the said object.
5. Object of interest screen focuses on the facial area of the individual who is identified according to the motion detection.
6. Motor spin direction screen shows the direction the motor spins to make it easier for the step motor to track the object in motion. The color-marking of the boxes here takes place according to the value of motion pixels in x coordinates of the geometric center.
7. In the tracking information screen, RGB and YCbCr color values; color coding techniques used in the location of the skin color; are given.
8. Motion detection adjustments screen carries the adjustment tools used in the detection of the object in motion. These adjustments are alarm level values that help identifying the capture moment for the motion, threshold value and the size of the object in motion.
9. The ratio of motion screen provides the percentage value of the pixels in motion in the image to the whole image.
10. Face identification adjustment screen shows whether the facial area has been recognized or not. Activate human face option makes it possible to color the face white by filtering and keeping the rest of the area black. Also, which one of the RGB and YCbCr coding techniques to be used is also adjusted in this screen.
11. Motor control screen allows the activation of the step motor mounted under the web camera to allow the tracking of the object in motion.
12. Image processing screen shows the realization time of a given process and the total processing time.
13. Warning system screen activates the vocal warning system, used as a deterrent, when an object that can threaten security enters the environment.

Application of motion detection

When video image of the secured environment is

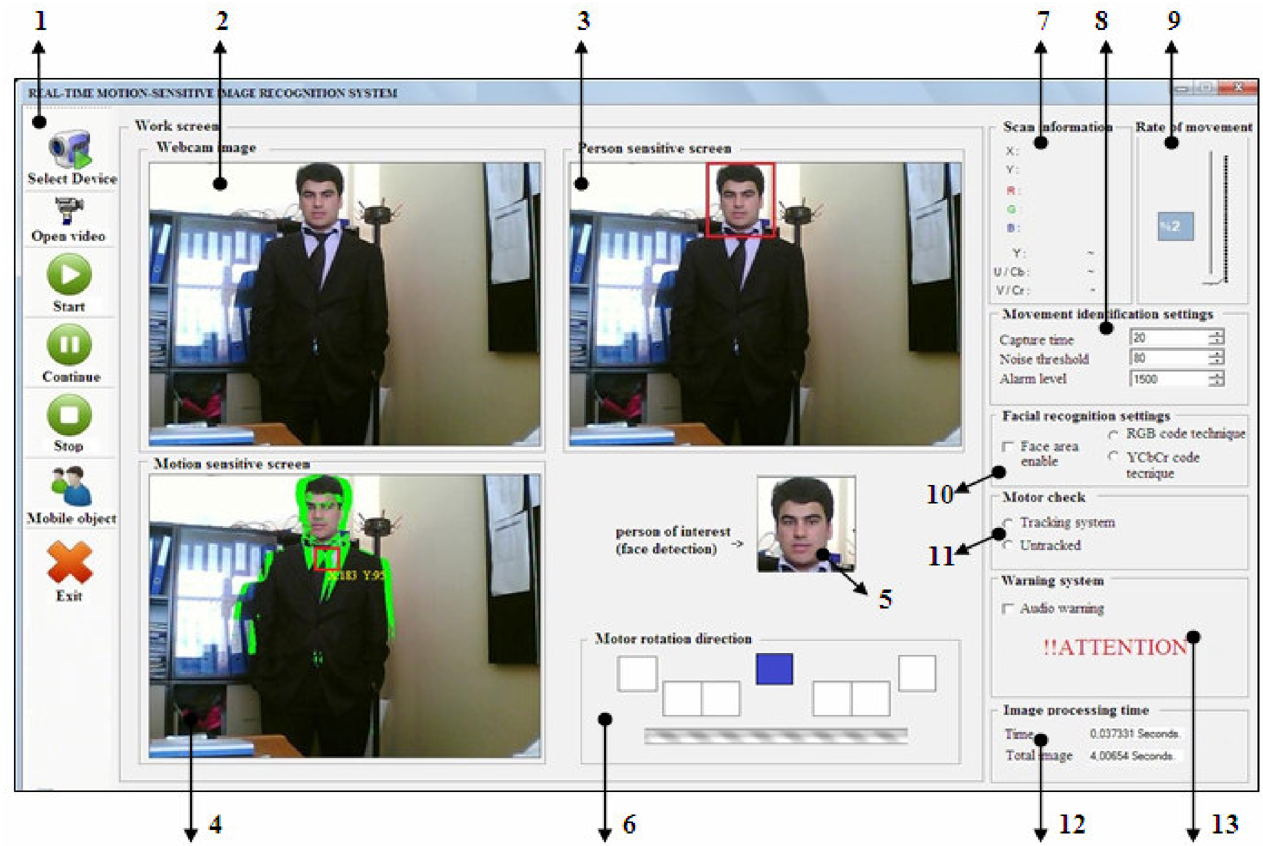

Figure 3. User interface of the security system program.

captured, the image is divided into sequential image frameworks. The object in motion is detected according to block phase differences of the sequential image frameworks of the video. The block phase differences of the sequential image frameworks are calculated with the method used in Equation (2) and compared to the threshold values explained in Section 2. The difference pixels exceeding specific threshold values are represented in color green as seen in Figure 4. When the number of difference pixels exceeds the value specified via user interface for identification of sensitivity and also exceeds the adjusted alarm level, the system detects a large object in the environment and captures and records its image to the computer environment in a file named with date.

Face recognition and identification

The face of the human object in the environment is detected by using skin color algorithms, geometric techniques and filtering techniques. In this study, skin color algorithms given in Equations (3) and (4) are used for face recognition.

The biggest problem encountered in face recognition is the existence of objects that are close to skin color. In this case, the system recognizes them as parts of the human face. In order to solve the problem, image processing techniques are used. To be able to work in an accurate and speedy manner on the color image, image is first converted to gray, then to black and white images. Equation (1) is used to convert the image to gray level images. In order to obtain black and white images, the pixels in the skin color area identified by skin color algorithms are converted to white and the rest of the area is converted to black. That way, the image is converted to binary level using threshold technique. In the binary system, zero represents black colors and one represents white colors. The binary form of the image is shown in Figure 6(a).

After it has been converted to binary form, the image is filtered and protected from unnecessary noise. The median filter which aims to soften the image is used as a 3 × 3 matrix. Median filter is a non-linear process that protects the edges while eliminating random noise (Danisman and Alpkoçak, 2003). It is based on eliminating the anomalous values in the finite pattern by replacing the median value in the pattern (Danisman and Alpkoçak, 2003). Figure 5 provides an example that explains how to obtain the median value and Figure 6 shows the image before and after the median filtering is applied.

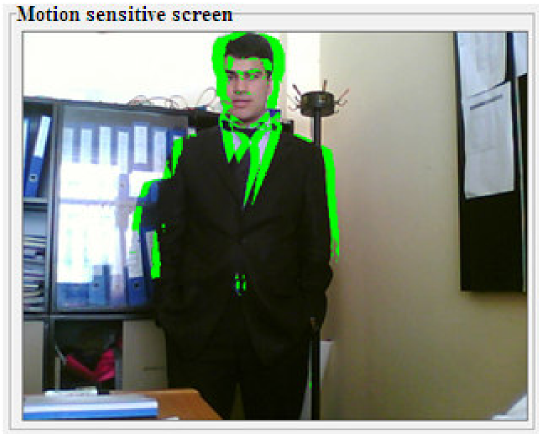

Figure 4. Displaying the different pixels that exceed the noise threshold in color green.

4 15 17
5 40 16 ⟹ 4 5 5 15 **16** 16 17 17 40
5 16 17

Figure 5. Calculation of the median value.

(a) (b)

Figure 6. Application of the median filter: (a) Image with photo noise; (b) Image free of noise.

Figure 7. The identification of the facial area by skin color algorithms and geometric techniques.

(a) (b)

Figure 8. Comparison of skin color identification algorithms: (a) RGB coding technique; (b) YCbCr coding technique.

As can be seen from Figure 6, although the image is filtered, it is still not free from noise. Because of this reason, image is subjected to a filter with a 3 × 3 matrix. This filter combs the image screen and detects the areas with the most skin color. After the detection of the area, it is marked with a box to identify the facial area (Figure 7). The area with the most tan color determined through skin color algorithms and geometric methods are marked as a rectangular box and the areas outside the box is still screened in order to track another facial area. When a facial field is not recognized in the field of vision, the red frame is cleared.

As already mentioned earlier, this study utilizes RGB and YCbCr coding techniques which are both skin color based algorithms. The results of algorithms using the same coding techniques were compared in this study and it was seen that algorithm utilizing YCbCr coding techniques provide better results. The reason may be due to the fact that algorithm utilizing YCbCr coding techniques are less affected by the elements in the environment such as brightness and dust. Figure 8 displays the difference between these two algorithms more clearly.

Tracking mobile objects

The video image is divided into seven equal areas in the horizontal axis in order to track the mobile object easily. By the way, only related areas are controlled instead of controlling all screens. Approach diminishes the processing burden and allows higher control on the system regarding tracking. The performed experiments related to the video image, reflected the fact that division of the image into seven equal areas in the horizontal axis provides good results. In the case of dividing the image into higher number of areas, a better determined tracking would be possible; however, that may create difficulties in controlling more areas and may result in detecting the object in the wrong area due to the transfer of size of the object to other areas.

It is necessary to determine which direction and how many degrees the camera needs to be directed after the blocks with mobility is identified. In order to overcome this

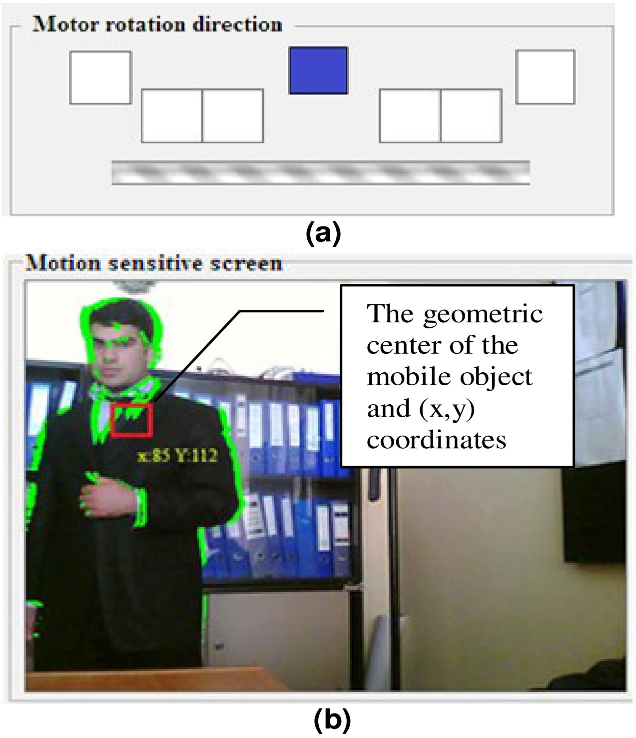

Figure 9. Identification of the mobile object: (a) The identification of the position of the object in video image that is divided into 7 equal parts. (b) Calculating the geometric center of the area which is dense with mobile pixels and identification of the target.

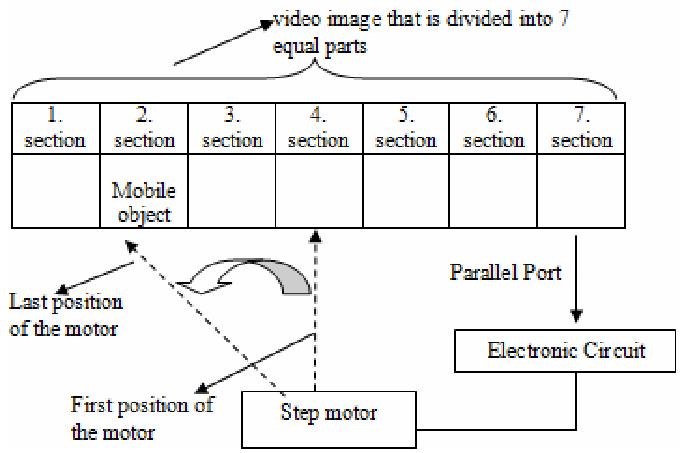

Figure 10. Realization of motion tracking.

issue easily, the geometric center of the image of the mobile object is calculated. The geometric center of an image can be defined as the minimum point, that is, the square sum of its distances to all other points (Köker et al., 2001). Equation (6) is utilized in obtaining the geometric center. Here, n represents the total number of mobile pixels; x represents the x coordinate value of the mobile pixels and y represents the y coordinate value of the mobile pixels. In Figure 9b, geometric centers of pixels with more mobility is detected and marked with a small box:

$$\text{Geometric center} = \left(\frac{1}{n}\sum_{x} x, \frac{1}{n}\sum_{y} y \right) \qquad (6)$$

Figure 9 visually displays the video image that is divided into seven areas and / or blocks with the weighted center. Mobile object in the far left block of the screen is understood from the Motor Stir-Direction screen as seen in the example given in Figure 9b. The (x) coordinate value of the geometric center of the mobile pixels in this block is used in identifying the direction of the motor.

Figure 10 displays the process of directing the camera to the area with more mobility by the rotating of the Mobile blocks according to the (x) coordinates of the geometric center. The location of the mobile objects using geometric center values can be transferred to the electronic circuit over the parallel port with the help of the developed software. The values transferred to the electronic circuit are used in the provision of direction to the step motor. If the mobile object is on the right of the motor, the motor rotates to the right and vice versa. If the position of the mobile object and the motor is the same, the motor does not engage in rotation and the mobile object is centered in the viewfinder all the time.

As seen in the example provided in Figure 10, step motor rotates from the fourth area to the second area since movement is detected in the second area. Since the rotation of the motor is directed to the mobile object, the camera will engage in centering the object according to its geometric center and the object will still be in the fourth area for the motor. Therefore, the motor continuously rotates by centering the object in motion and makes possible to record the threat.

The next step allows for face recognition in real-time, basis for face recognition studies. Skin color algorithms, geometric techniques and filtering techniques are used in this process. Subsequent to the identification of the face, the image is recorded.

The last step involves the design of a mechanical system that will allow the camera to track the object in motion. The mechanism designed, uses the step motor that turns towards the direction of the weighted centers of the mobile pixels of the object and continuously tracks the object. By the way, the camera mounted on motor mechanism connected to the PC with parallel port can be directed to track mobile objects or objects in a continuous manner. The movement angle of the camera is 350 degrees. The camera system that can turn with a wide angle and with the help of the step motor, focused on the area of movement in order to scan all the live images that are proportionally reflected in the field of vision. That is, when the camera is in a fixed position, it turns towards the mobile object as soon as object appears in the environment and is situated so that it can track the field of movement of the object. In that way, it is possible to provide the operations of multiple cameras with a single one.

Conclusions and future remarks

This study provides motion detection in real time. The system is both sensitive enough to perceive the slightest motion in the environment and differentiate among mobile objects.

The application was done by C# .Net; an object oriented language. A computer with an Intel® Pentium® IV 2.4 MHz processor and 1GB memory space was used in the experiments. The results showed that 95 out of 100 mobile images were correct in identifying the facial area and the rate of accuracy was found to be 95%. The average time of the process in the identification of the mobile object and tracking of the facial area in a human face, is 0.03 s.

With developed user interface, the images captured by the camera are processed with image processing techniques and unauthorized breaches to the environment can be detected. The images regarding the breach are stored in the computer to be used later. Since the system allows sensitive adjustments, the factors such as dust and light do not affect the system.

Developed user interface has a simple design with easy adjustments, it allows for an effective motion analysis in the processing of face recognition.

Using step motor in the mechanical system, values such as direction, revs per minute and rotation speed can be controlled by the assistance of computer which makes it possible to know the direction, speed and status of these motors. With the help of these characteristics, very sensitive status control is feasible by the use of step motors. Step motors can be stop and restart as many times without causing any damage. There is no cumulative status damage in each new step. Since their mechanical structures are simple, they do not require maintenance.

The study utilizes a cost-efficient web camera. With the mounting of a web camera to the computer, it will be possible to secure many businesses and homes in a cost-efficient manner.

Since the system only records the mobile objects in the database as opposed to the present security systems that record everything, the system allows for the efficient use of memory capacity.

One of the reasons for high operational speed of the system, there is no requirement in learning the use complicated machine in real time for identifying the face on tan colored areas and the existence of rather simple mathematical operations.

Based on this study, it is possible to track and destroy a mobile object and develop smart weapons for national security in the framework of military applications. Detecting and tracking of objects on motion can be used to develop traffic control. Elements that need to be under control in traffic such as overtaking, following up closely, accidents, and excessive speed can be taken under control by real time identification and accidents can be curbed. Also IP cameras can be utilized instead of web cameras and security can be established without the need for computer intervention. A vocal warning system can be added to the mechanical system and a more developed and a highly deterrent system can be developed in order to prevent and decrease theft.

REFERENCES

Coşar S, Urhan O, Güllü MK (2004). CCD Kamera Kullanarak Sayısal Görüntü İşleme Yoluyla Gerçek Zamanlı Güvenlik Uygulaması,Otomasyon Dergisi sf. pp. 82-84.

Cucchiara R, Mello P, Piccardi M (2000). Image Analysis and Rule-Based Reasoning for a Traffic Monitoring, IEEE Transactions on Intelligent Transportation Systems, IEEE Press, 1(2): 119-130.

Çetin Ö (2008). Hareketli görüntü uygulamaları için sırörtme yaklaşımı ile veri gömme algoritması tasarımı, Doktora Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü.

Danışman T, Alpkoçak A (2003). AVIR: Ten Rengi Tabanlı Yüz Bulma Yöntemi ile Otomatik Video İndexleme ve Sorgulama, 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı Bildiriler Kitabi, İstanbul, Türkiye.

Douglas L, Nishanth V (2009). Interactıve Embedded Face Recognition, J. Object. Technol., 8 (1): 23-53.

Foresti GL (1999). Object Recognition and Tracking for Remote Video Surveillance, IEEE Transactions on Circuits and Systems for Video. Tech., 9(7): 1045-1061.

Inseong K, Joon HS, Jinkyu Y (2010). Face Detection,Stanford University. http://www.stanford.edu/class/ ee368/Project_03/Project/reports/ee368group02.pdf.

Kefeli HC, Urhan O, Ertürk S (2005). Double stage video object segmentation by means of background registration using adaptive thresholding, Proceedings of the IEEE 13: 80-83.

Kim Shin-Hyoung, Woo-Sung Kil, Jeong Hwan Kim (2009) Snake-based Object Contour Tracking using SURF and Mean-Shift. IPCV pp. 825-832.

Koller D, Danilidis K, Nagel HH (1993). Model-Based Object Tracking in Monocular Sequences of Road Traffic Scenes, Int. J. Comp. Vision, 10(3): 257-281.

Köker R, Öz C, Sarı Y (2001). Hareketli cisimlerin bilgisayar görmesine dayalı hareket analizi,Elektrik-Elektronik-Bilgisayar Mühendisliği 9. Ulusal Kongresi, pp. 497-500.

Köker R (2002). Üç Eklemli Bir Robot Kolunun Görmeye Dayalı Olarak Model Tabanlı Zeki Kontrolü, Doktora, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü.

Lyon D, Vincent N (2009). Interactive Embedded Face Recognition. J. Object. Technol. 8: 1.

Nabiyev VV, Kurt B (2007). Facial Expression Recognition, II. Uluslararası Bilim ve eğitimde Bilgi ve İletişim Teknolojileri Uygulamaları Konferansı, Bakü, Azerbaycan, pp. 779-791,

Oral M, Deniz U (2005), Motion detection in moving pictures, Proceedings of the IEEE 13: 187-190.

Stiller C, Konrad J (1999), Estimating Motion in Image Sequence: A Tutorial on Modeling and Computation of 2D Motion, IEEE Signal Process. Mag. 16: 70-91.

Url1 (2010). http://academic.mu.edu/phys/matthysd/ web226/L0221.htm/.

Yang MH, Kriegman D, Ahuj N (2002). Detecting faces in images: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 24(1): 34-58.