Motion Detection Surveillance System for Human Head

Mohamed Ibrahim Shujaa   **Ahmed bahaa aldeen**

Technical middle university
Technical College Of Management ,Baghdad
Information Technology Department

dralnedawy@yahoo.co.uk   ahmedmeg29@yahoo.com

Abstract

Motion detection for human head and its coordinates is one of the important areas with multiple applications at the present time, such as control of devices for people with disabilities and other modern control devices estimation model to detect human head and follow up coordinates. This system uses frame subtracting to detect motion with development the background frame is to eliminate the problem of wrong alarm for motion when the background frame still without change. Gaussian smoothing filter also has been used to reduce noise caused by bad illumination or web camera low quality, so there is no need for kalman filter, then using Laplacian filter which is dependent on second derivation in order to detect edges of the frames(background and current) , at last detecting the human head using the sum of gradient magnitude around the perimeter of ellipses equation. The system has been tested on several samples indoor and out door , two of these samples for females it head hooded (with Islamic hijab) and the results were very good in detecting heads and tracking their coordinates. The system has been built using visual C# which is considered faster than interpreted MATLAB that have been used in other researches in the same area.
Keywords:-computer vision, image processing, motion detection, human head recognition**.**

<div dir="rtl">

الخلاصة

ان مجال كشف حركة رأس الانسان وتقدير احداثياته من المجالات المهمة ذات تطبيقات متعددة في وقتنا الحالي مثل اجهزة التحكم للاشخاص المعاقين و اجهزة السيطرة الحديثة وغيرها. اعتمد عمل النظام لكشف الحركة على طرح الاطارات مع تطوير الاطار الاولي (الخلفية) كل مرة لتطوير المشهد و التخلص من مشكلة كشف خاطئ للحركة في حالة ثبات الاطار الاولي، كما تم استخدام مرشح كاوس على الاطارات الفديويه لتقليل الضوضاء الناتجة عن سوء الاضاءة او قلة جوده كاميرات الويب و بذلك تنتفي الحاجة لاستخدام مرشح كالمان ، ومن ثم استخدام مرشح لابلاس المعتمد على المشتقة الثانية لاستخلاص حواف صورة الاطار الفديوي وصولا الى عملية تحديد الرأس البشري بأستخدام معادلة تقدير الشكل البيضوي التي تعمل على جمع حجوم الانحدار للنقاط المحيطة بالشكل البيضوي . تمت تجربة النظام على عينات عدة منها عينتان لرأسين بالحجاب الاسلامي وهذا قد جرب لاول مرة في هذا البحث و كانت نتائج كشف رأس الانسان و متابعه احداثيات حركته جيده جدا و في مختلف الظروف داخل بناء او خارجه، واخيرا تم بناء هذا النظام بلغة سي شارب و التي تعتبر اسرع من ماتلاب المفسرة و التي استخدمت سابقا في بحوث اخرى بنفس المجال.

الكلمات المفتاحية :رؤيه الحاسوب، معالجة الصور ، كشف الحركة ، تمييز رأس الانسان

</div>

1- Introduction

Motion detection is identifying the image changes in some areas, detecting the presence of object motion avoiding interference by light changes [zhengie 2010]. Motion detection for surveillance system is detecting region of interest present in a region of awareness or portion of environment being monitored [Yan 2010]. Head detection is currently one of the most active research topics in the domain of computer vision; this interest is driven by a wide spectrum of applications in many areas such as anti terrorism, video surveillance, virtual reality, perceptual user interface and so on. [Li 2009] this research aims to build a system that can detect motion of the human head and track its position, this work overcomes motion detection problems of noisy frame caused by blurring or tree movement using frame updating and Gaussian filter with detecting human head by its shape and tracking the motion coordinates
 Frame differencing is a technique widely used for change detection in dynamic image, it compares each incoming frame with previous and classifies those pixels of significant variation into

foreground, let $f(x,y,Ti), f(x,y,Tj)$ taken at times Ti, Tj is compared with two frames pixel by pixel. One procedure for doing this is to form a difference image. Comparing this image against subsequent image of the same scene but including moving object result in the difference of the two images [Awni 2010].

$$dij = \{1 \, if \, |f(x,y,ti) - f(x,y,tj)| \geq T$$
$$0$$

Otherwise ---------------------------- (1)

This method has a problem if there was a person on the first frame (background frame) , but then he is gone , so the system will always have a motion detected on the place , where that person was , to overcome this obstacle we can use developing the scene by moving the background frame to the current frame on the specified amount (1 level per frame) , by moving the background frame slightly in the direction of the current frame[Amina 2011]. Also there are problems of noise caused by web camera and its effects on the accuracy of detecting the elliptical shape of the head.

2-Related researches

Yan zhoo and Jiao-Min[2010] improved the method of subtracting frames in their paper by blocking and subtracting consequence frames and background, this new method separates the image into different blocks first, then determines the object movement by calculate the difference between consequence frames inn one series. This blocking and frame difference method separate one image taken by camera to several blocks and detect change in every block, once moving object is detected the video recording equip will turn on[Yan 2010].

Li fang and Zhang Meng [2009] proposed a method using both temporal different and optical flow methods together with morphological filter for the purpose of motion detection. Temporal difference is used to obtain initial coarse image so as to reduce the number of the pixels that downstream tracking algorithms have to process, no knowledge of background is required as in background subtraction where training prides in the absence of foreground are required for bootstrapping. The optical flow method is then used to further analyze the detected motion area to reduce noises such as movement of trees and the capturing devices, and then morphological filter applied which is used to suppress noises while preserving the main object characteristics. It consists of ways for digital image processing based on mathematical morphology, which is able to decompose complex shapes into meaningful parts and separate from background, but this method has inability o detect the entire shape of the objects of interest, considering the fact that the system is for simple office and home uses. [Li 2009]

Ammar Awni[2009] in his paper used frame differencing using matlab which is interpreted without any mention about object recognition or treating noise caused by blurring or other cases like tree movement, the disadvantage in his approach is if the object is moving smoothly will receive small change from frame to frame, so it is impossible to get the whole moving object, so when the object is moving so slowly, the algorithm will not give any result at all, and this system is also for indoor usage. [Awni 2010]

Amina Dawood and Balasim hussien [2011] in their paper improve the work frame subtracting by updating the background frame by moving the background frame to the current frame slightly but they did not try for object detection or illumination of lighting problems using any filtering system [Amina 2011].

3-proposed system methods

The proposed system is built to detect human head motion, with tracking the coordinates for this movement, firstly, a moving object detection algorithm is used to find whether there is a target in the view field of the camera or not, after that a head segmentation and detection algorithm which is based on head shape model is active to detect the persons head in the image, the following flowchart of the proposed system Fig 1:
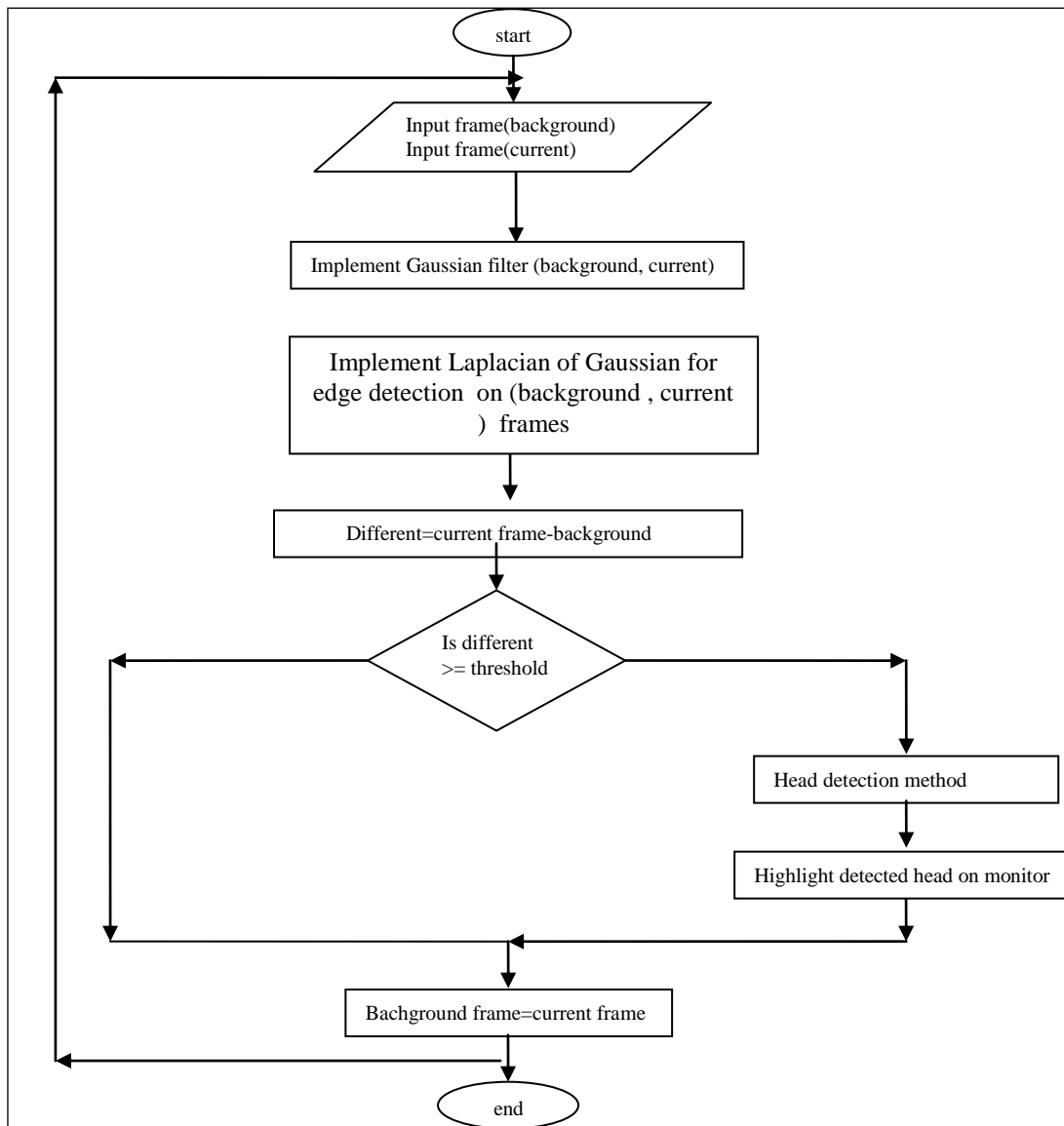
Figure 1 : flowchart explain the procedure of the proposed system

The system takes a current frame (image), a gray scale copy of the current frame and background frame and gray scale of it. At the beginning, the system deploys Gaussian filter on these two frames (current and background) frames in order to illuminate noises caused by blurring or bad web camera adjustment. Gaussian is commonly used with edge detection. Most edge detection algorithms are sensitive to noise; using Gaussian before edge detection aims to reduce level of noise in the image, which improves the result of any edge detection algorithm. This approach is commonly referred to as Laplacian of Gaussian or (LOG filtering).

In 2-D, Gaussian has the form:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2 + y^2}{2\sigma^2}}$$ ---(2)

The Gaussian output's a weighted average of each pixel's neighborhood, with the average weighted more towards the value of the center pixels. This is in contrast to the mean filter's uniformly weighted average. Because of this, a Gaussian provides gentler smoothing and preserves edges better than a similarly sized mean filter. By choosing an appropriately sized Gaussian filter we can be fairly

1738

confident about what range of special frequencies are still present in the image after filtering [Awni 2010][Amina 2011].

(a)   (b)

figure 2 : picture(a) an image with noise and (b) the image after Gaussian filter smoothing

Then we use the Laplacian of an image which highlights regions of rapid intensity change and is therefore after used for edge detection, the kernels of laplacian are approximately second derivative measurement on the image, they are very sensitive to noise. To counter this, the image is often Gaussian smoothed before applying the Laplacian filter. The operator simply takes a single grayscale image as input and produces another grayscale image as output. The laplacian L(x,y) of an image with pixel intensity values I(x,y) is given by[Gonzales 2007][Fisher 2003]:

$$L(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{I \partial^2}{\partial y^2} \text{--(3)}$$

Since the input is represented as a set of discrete pixels, we have to find a discrete convolution kernel that can approximate the second derivatives, second derivation must be zero inconstant intensity, nonzero at onset and end of an intensity step or ramp and zero along ramp of constant slop , in the definition of the laplacian, two commonly used kernels are shown in the figure below[9]:

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

LOG response will be positive on the darker side, and negative on the lighter side. This means that a reasonably sharp edge between two regions of the uniform but different intensities, the Log response will be:

- Zero at long distance from edge.
- Positive just one side from edge.
- Negative just to other side of the edge.
- Zero at some point in between on the edge itself.[Fisher 2003]
  Second derivation produces a double edge on pixel that is separated by zero , second derivation enhances fine detail much better than first derivation.[Gonzales 2007]

The system gets first the frame of the video sequences as background frame, moves the background frame into current frame [Awni 2010]

Background frame = current frame ------------------------ (1)

So the system moves the background frame slightly in the direction of the current frame in the background frame the human head must be detected [Amina 2011], then an elliptic model, is used to find persons head. The basic idea of the ellipse model is to calculate the mean gradient of the ellipse outline, because the head is modeled by an ellipse with a fixed vertical orientation and fixed ration of $.1.2^2$ each frame the image becomes available the ellipse state $s = (x, y, \sigma)$, where x, y is the position and σ is the size (length of the minor axis) is maintained by performing local search to maximize the normalized sum of the gradient magnitude around the perimeter of the ellipse, we move the center of the ellipse in the searching space Is to find the S max

$$Sj = 1/N \sum_{i=1}^{n} |gi|$$

$$S\max = \max(Sj) \qquad \text{------------------------------- (5)}$$

Here N is the number of the pixels on the ellipse's outline, gi is the gradient on the *ith* pixels[Birchfield 1997][Liu 2001]
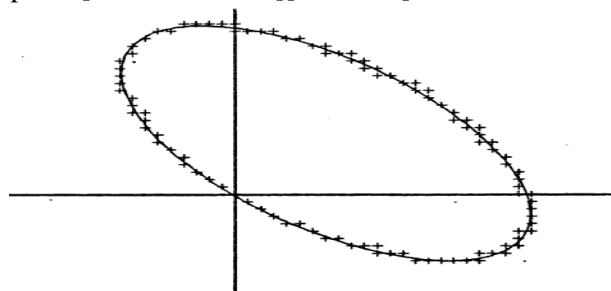

Figure 3: explain the calculate the mean gradient of the ellipse outline

The origin of this equation is that , first we use the center of the ellipse as a reference point and assume that it is centered at X0 and Y0 with major and minor diameters a and b. for moment, we will assume that ellipse is oriented with its major axis parallel to the x –axis . later we will relax this requirement by introducing an additional parameter for arbitrary orientation . for moment assume a and b are fixed then the equation of the ellipse is :-

$$\frac{(x-x0)^2}{a^2} + \frac{(y-y0)^2}{b^2} = 1 \text{ ----------------------------- (6)}$$
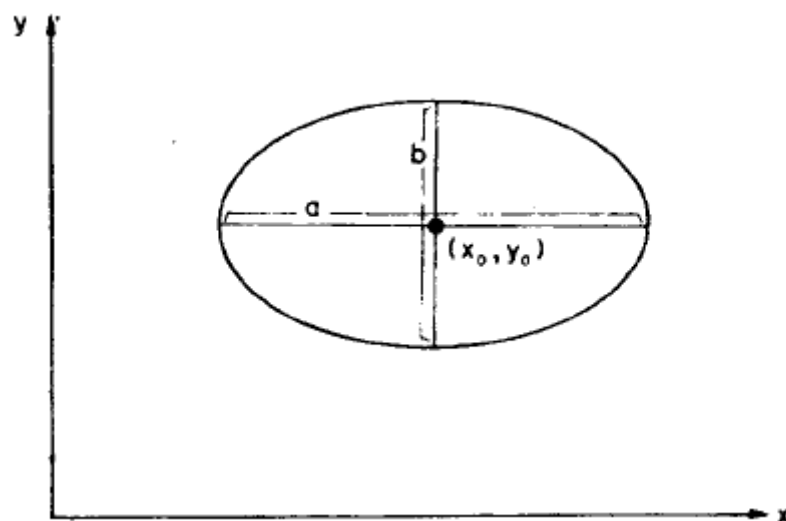
See figure below

Figure 4: parameterization of an ellipse with major axis parallel tp x- axis

let $x = x - x0$ and
$y = y - y0$ then

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ --- (7)}$$

Differentiating with respect to x

$$\frac{2x}{a^2} + \frac{2y}{b^2}\frac{dy}{dx} = 0 \text{ ---(8)}$$

But dy/dx is known from edge pixel information . let $dy/dx = z$ then from (7)

$$x^2 = \left(\frac{a^2}{b^2}\right)y^2 \text{ ---(9)}$$

Substituting in (7)

$$\frac{y^2}{b^2}\left(1 + \frac{a^2}{b^2}z^2\right) = 1 \text{ ---(10)}$$

So x and y will be

$$y = \frac{+}{-}\frac{b^2}{\sqrt{1 + \frac{a^2}{b^2}}} \qquad\qquad x = \frac{+}{-}\frac{a^2}{\sqrt{1 + \frac{a^2}{b^2}z^2}} \text{ -----------------(11)}$$

And finally given a ,b, x, y and dy/dx

$$x0 = x\frac{+}{-}\frac{a^2}{\sqrt{1 + \frac{b^2}{a^2 z^2}}} \text{ -----------------------(12)}$$

$$y0 = x\frac{+}{-}\frac{b^2}{\sqrt{1 + \frac{a^2 z^2}{b^2}}} \text{ ---------------------------------(13)}$$

Final step is to handle rotation by introducing a fifth parameter θ , for an arbitrary θ we can calculate (x,y) using :

Z=tan $(\phi - \theta - \pi/2)$------------------------------------(14)

And rotate this (x,y) by θ to obtain the correct x0 and y0 notice the figure below
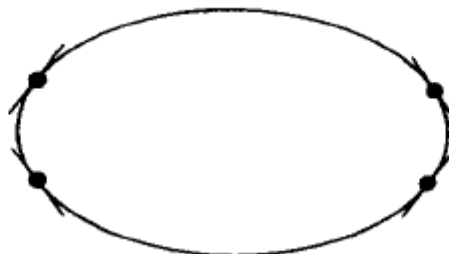

Figure 5: four reference point solutions resolvable with gradient quadrant information

4-Experiments and Results

The proposed system was implemented by using standard PC hardware Intel CORE i5 2.6 GHz To evaluate the system stability and performance in human head motion detection and tracking, the program has been tested on 4 different people, the sample of 4 people is different in sex, skin color and amount of hair and head shape, two of them are covered women (a women with veil or the Islamic hijab), the proposed system is tested indoor and outdoor areas, the tested persons are asked to move their heads indoor to make the system track their head coordinates, the program was implemented using visual C# 2008. and the program interface consists of two screens , the first on left showing the real time picture with implementing Laplacian edge detector filter, and one text box to write the coordinates of human head while the head is moving, with one button for starting and stopping the camera. see figures below:
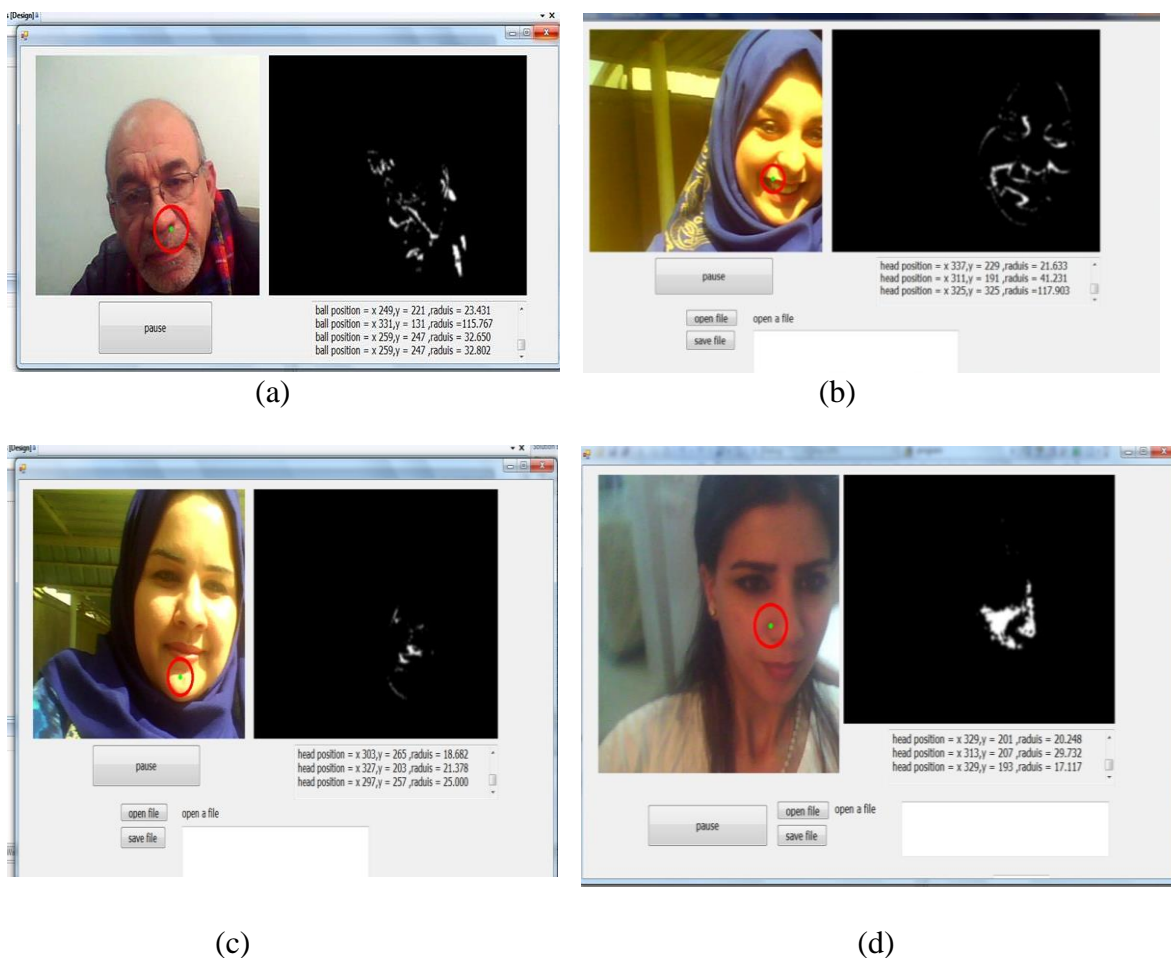

(a)   (b)

(c)   (d)

Figure 4:- a sample for a man, b and c two girls with hijab , and (d) sample for a girl with hair

Picture (a) for a man without hair that has been taken in an indoor office, the system succeeded in detecting his head motion and track coordinates for his movements then we can see pictures (b) and (c) for two girls with a cover on their hair (Islamic hijab) with a brown skin , the system also succeeded in tracking their head movement while each one of them was standing about 1 meter from web camera outside building. The last sample picture(d) is for a girl with a black girl inside room, the system also succeeded to track her head movements standing about half meter from the camera.

The system gained a good result for head between 1 to 3 meters with a good executing time between 5 to 10 seconds to catch a head.

In this work we used an index area matching $\Omega(k)$ to evaluate shape precision by using the equation

$$\Omega(\kappa) = \omega\varepsilon(\kappa) / \omega\eta(\kappa) \text{---(15)}$$

Here $\omega\varepsilon(\kappa)$ is the area of the ellipse, $\omega\eta(\kappa)$ is the area of the human head , if the $\Omega(k)$ is less than 0.2 we take it as false detection see the table of performance

picture	Area matching	Computer time	$\Omega(k)$
A	0.9	10ms	0.98
B	0.8	20 ms	0.81
C	0.7	22 ms	0.75
D	0.9	10 ms	0.91

We can notice that the picture with veil needs more processing time in sample (b and c) than samples (a and d) area matching was good for unveiled heads with its $\Omega(k)$ versus the veiled heads, but the system succeeded in catching all samples.

5- Conclusion

In this paper we have presented a method for tacking head motion, the reader can notice the following conclusion:-

1-enhancing the frame subtraction method by developing scene to overcome the problem of existing human head in the first frame , when this head is gone in the next frame the system will always inform head detection , so developing scene by making next frame is background frame is a solution to overcome obvious problem.

2- The use of Gaussian filters on the frames in order to illuminate noises caused by bad web camera adjustment or illumination problem is effective before edge detecting filter.

3-Laplacian filter is if efficient in edge detection in order to calculate the mean gradient of the ellipse to detect the human head.

4- Ellipse searching space equation succeeded in detecting human heads in good time even with head for women covered with veil (Islamic hijab) which is tested in this work, this method is better than depending on skin color approach because the skin color may be near the intensity of the background or something else in the frame.

5- The system has to work with head in a front picture in distance between 1 to 5 meter otherwise its accuracy decreased with increasing the distance.

6-Suggestion for future work

The researchers can develop this paper for head detecting using other biometrics like eyes, nose, mouth distances to recognize the person not only detect the head, the researchers can also improve in the motion detection area, by building systems that detect strange actions like fighting or other strange behavior for surveillance system for antiterrorism.

References

Amina, A.; Balasim, A. *"motion detection in real time video streams using moved frame background*", journal of Babylon university,2011, issue 18 p (386-395)

Awni. Ammar "motion *detection for surveillance camera*",Journal of Baghdad for Science,2010 , issue 7 ,1070-1075

Birchfield, S.   *"An elliptical head tracker*", Signals, Systems & Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on   (Volume:2)

Fang,Li ; Meng Zhang , "*Smart motion detection surveillance system*" (2009),IEEE computer society ; International Conference on Education Technology and Computer; IEEE DOI: 10.1109/ICETC.2009.10

Fisher, R. Perkins, S.; Walker A.; and E. Wolfart. 2"image processing learning resources" 2003, http://homepages.inf.ed.ac.uk/rbf/HIPR2/log.htm

Gonzales R. , Woods ,"digital image processing" 3rd edition ,prentice hall , 2007

Liu,qingshan, Songad MA , LU Hanqing, "*head tracking using shapes and adapting color histogram*" , the Chinese academey of sciences, Beijing, 100080, P.R. china ,2001

Wang,Zhengie; Zhao,Yuan(2010), " *Research on motion detection of video surveillance system* ",(2010),3rd international congress on image and signal processing (193)(194)

Zhao, YAN & LU,Jiao "*Improved method for human motion detection and application*" , 3rd international congress on image and signal processing(CISP2010) , IEEE, 2010.