Security Guard Robot Detecting Human Using Gaussian Distribution Histogram Method

Mbaitiga Zacharie
Department of Media Information Engineering,
Okinawa National College of Technology, 905 Henoko,
Nago, 905-2192, Okinawa, Japan

Abstract: Problem statement: The purpose of any robotic is to perform tasks that a human would prefer not to do or hopefully do it with precision in order to avoid mistakes or when a human is out of duty due to fatigue or health reasons. The research into human detection into images has paid the way be aware of what is going on around the houses or buildings where a front-line security is needed 24 h a day. In this research a human detection security robot based on Gaussian distribution histogram was proposed. **Approach:** The proposed method consisted of three steps: (1) the RGB color space histogram was created by subdividing a color space into certain number of bins and then counted the number of pixels that each bin contains. (2) The created RGB histogram was converted into HSV color histogram using Gaussian distribution method. (3) The bell-shape curve of the Gaussian distribution was used to calculate the detection probability between the standard deviation. **Results:** Experimental evaluation had been tested on the images sequences where the experimental results revealed that the proposed method was less sensitive to changes in the scene achieving higher performance detection than traditional method of histogram creation and had been found to be robust. **Conclusion:** The results showed that the histogram based human detection resists to any changes in the image scenes.

Key words: Gaussian distribution, human detection, security guard robot

INTRODUCTION

Robotics today are advancing to the point where many tasks that used to be for humans only have been supplements by machines that can do the same tasks faster and safer than their human counterparts (Byeong-Sang *et al*., 2008). Factories are using automated robots that do repetitive tasks all day long leaving the more skill oriented tasks for qualified personnel. Nowadays, due to the recession that hits many security guard companies, security managers are turning to robots to help get the job done in order to reduce personnel costs. This demand has driven the security robots to be one of the most important topics in the robotics research. But for the robot to be used in such domain, it needs to make a difference between objects and humans. Various modern control strategies have been and still widely investigated to deal with this recognition problem. The essence of this problem is the knowledge representation, that is to say, how to teach the robot to recognize human, or its motions and objects. Computer vision based recognition of human activity involves the understanding of human behavior. Understanding human behavior is a complex and challenging task in computer vision due to ambiguity caused by non-rigid body articulation and mutual occlusion. Although the visual sensor gives a lot of information, many difficulties appear during its use. We will not do an exhaustive list of these difficulties but we will raise the main problems encountered during the development of vision application. In order to detect a change in images, computer vision programs store in memory a model which serves as background. Each image is then compared with the background model to detect a difference. However, if a lightning, snow or rain change happens in the environment, the program will be disrupted by this sudden noise and weather condition and will not return good results. The task of understand human behavior can be approached from various levels of details according to the complexity involved in the behavior. The overview article by (Aggrawal and Park, 2004) described the large variety of approaches, ranging from statistical modeling techniques such as Hidden Markov Model to biologically motivated recognition. (Park *et al*., 2008) applied these synergies in their approach for modeling motion. In this study we proposed a new method of human presence detection for security guard robot using Gaussian distribution and statistical probability method which consists of comparing two Hue Saturation Value (HSV) color

histograms: (1) The HSV color histogram model that serves as template. This histogram has been created using Gaussian distribution on the basis of model pictures taken with human presence. (2) The HSV color histogram that is formed when the robot detects objects or humans during its patrol mission.

MATERIALS AND METHODS

Related work: A computer vision-based human detection is a very difficult task especially in complex and unstructured environments. To detect human presence or human motions many physical parameters of human need to be perceived by the sensor module. These parameters can be summarized as follows: Skin color, face, body shape, voice, temperature and motion. But with the continuous development of technology, many methods for solving this complexity problem have started to be available over the internet, books and conference proceedings. Most of the pursued approaches use a multi-modal strategy that combines range sensors and color cameras in order to detect human and its motions. In many projects, additional hardware is used to improve the interaction capabilities, like for example microphones (Soltau *et al.*, 2001), pan-tilt units and unidirectional vision or an artificial face for gesture expression. Braun *et al.* (2005) used the combination of color vision and a laser range for detecting and following human with mobile robots. In their method, they used the Sequential Reduced Support Vector Machine (SRSVM) for classification of image space patches instead of regular one. Prior to the application of the SRSVM, the search space is reduced to image parts containing face candidates. These candidates are determined using skin color filtering and geometrical constraints. The search space is further reduced by fusing range information from the laser scanner with the captured image. This yields the distance information for each face candidate which is then used to restrict the scale at which to look for faces. Gehrid and Schults (2008) focused on human motion recognition as they appear in a kitchen and food preparation scenarios, such as placing objects and pouring fluid into container and discriminate 10 motions sequences. Each motion is described in terms of a sequence of motion units, such as fetching, maneuvering and placing back an object. Human subjects were asked to perform these motions in a controlled setting. To recognize these motions, the features vectors are fed into 3-states from left-to-right Hidden Markov Model which represents a motion unit. Each state of the motion unit from left-to-right HMM has two equally likely transitions; one to the current and

the other to the next state. The emission probabilities are modeled by Gaussian mixtures initialized by the K-Means algorithm based on the mutually segmentation data. Dia *et al.* (2007) proposed an approach for pedestrian detection and tracking from infrared based on joint shape and appearance cues. In their method a layer representation is first introduced and a generalized Expectation-Maximization Algorithm (EMA) is developed to separate infrared images into background (still) and foreground (moving) layers regardless of camera panning. In the two-pass scheme of detecting pedestrians from the foreground layer: shape cue is first used to eliminate non-pedestrians moving objects and then appearance cue helps to locate the exact position of pedestrians. These methods work well but have difficulties in environment changes.

Robot Platform: Robot mechanics: The robot shown in Fig. 1 consists of an iRobot originally developed for Humanoid Robot (HR).

Using this approach, high-level control of the robot is maintained by a remote or local PC/server communicating by a secure wireless link. Low-level functionality is managed by an onboard Digital Signal Processor (DSP) while computationally intensive operations are performed off board. The result is a robot that is lighter, draws less power, runs longer and is dramatically less expensive than fully or self-contained system. Moreover since primary processing resides in a server any hardware upgrades to the central unit are shared by all robots it controls. With its high bandwidth, Wireless Fidelity (Wi-Fi 802.11g) module, the robot can upload all sensors data to a PC or server at rates in excess of 10 Hz. Similarly, streaming audio and video of up to 4 fps either for direct monitoring or for processing by high-level Artificial Intelligent (AI) schemes is a snap. Commands and instructions sent to the robot via the same wireless link also pass at rates exceeding 10 Hz, providing real-time control and access.
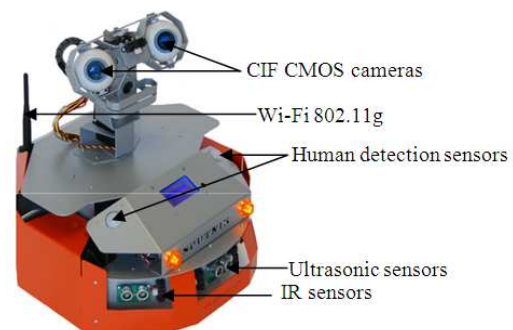

Fig. 1: Integrated cameras and robot sensors

Fig. 2: Samples photos for histogram creation

The wheels-based platform shown in Fig. 1 consists of two 12 V DC motors each supply 300 inches (22 kg cm) of torque to the robot's 18 cm (7 inch) wheels, yielding a top speed in excess of 1 m sec^{-1} (3.3ft sec^{-1}). Two high-resolution (1200 count per wheel cycle) quadrature encoders mounted on each wheel provide high-precision measurement and control of wheel movement. Weighing only 6.1 kg (13.5 lb), the system is light, but it can carry an additional payload of 10 kg (22 lb).

Robot sensors: The sensors consist of i90 that offer full Wi-Fi 802.11g wireless, multimedia and sensing and motion capabilities and comes with a wide range of sensor, camera and audio modules, sufficient to serve in any variety of applications. The robot offers broad expandability as well for projects that may require additional sensors, even specialized modules. Powered by separate RC servo motors, the integrated cameras head can pan and tilt independently.

Human detection strategies:
HSV histogram creation: Human detection in images is one of the most difficult research topics in computer vision but has been proven useful in applications such as, monitoring pedestrian junctions, monitoring students at school, people at hospital and surveillance in various sites where there is a need for advanced frontline security. But human has been proven to be difficult to detect because of the wide variability in appearance due to clothing, articulation and illumination condition that are common in outdoor scenes and the wide range of poses that they can adopt. The first need is a robust feature set that allows the human form to be discriminated clearly, even in cluttered background under difficult illumination. We studied the issues of feature sets for human detection, showing that using the Gaussian Distribution Histogram (GDH) provides excellent performance relative to other existing feature sets including wavelets (Mohan *et al.*, 2001; Viola *et al.*, 2003). In the proposed method, photos with human shown in Fig. 2 have been used to create a histogram in Red, Green and Blue (RGB) by equally subdividing a color space into a number of bins and counted the number of pixels that each bin contains.
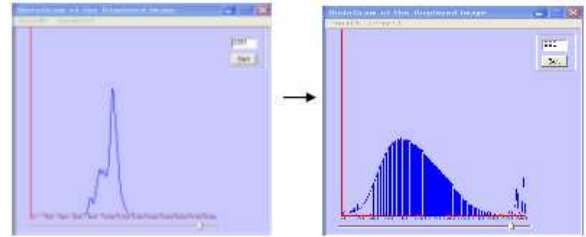
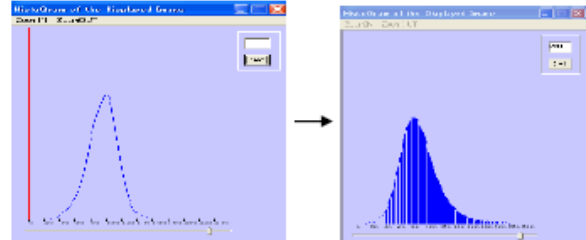

Fig. 3: RGB to HSV color space histogram (Fig. 2a)

Fig. 4: RGB to HSV color space histogram (Fig. 2b)

This method results in a large number of bins which the color represented by adjacent bins would reveal only a trivial difference. So when a change in illumination conditions or the presence of noise is occurred, it will determine a large number of pixels to drift from one bin to another bin. Using this method for human detection may give only a small difference of histogram for two different people with similar skin tones. Consequently a good detection result may not be obtained. To obtain a robust detection system we converted the RGB histogram to Hue Saturation and Value (HSV) color histogram as shown in Fig. 3 and 4 using Gaussian distribution method in order to expand our application to multiple people with different skin tones. The HSV Gaussian color histogram is well suited to the task of human detection due to its variability to implicitly capture complex multi-model patterns of color information. The value of Hue chosen ranges from 0-350° and the corresponding colors vary from red through yellow, green, cyan, blue, magenta and black to red so that there are actually red values in both at 0-350°. Saturation varies from 0-1 and the corresponding colors vary from unsaturated such as shades of gray to fully saturate. Values or Brightness have been chosen to range from 0-1 and let the corresponding colors become increasingly brighter. We described the body color candidate by one-dimension Gaussian model:

$$\lambda(\mu, \delta^2)$$

Where:
μ = Mean of hue
δ^2 = Variance of hue respectively

That is to say:

$$\mu = \frac{1}{\lambda} \sum_{i=1}^{N} hue_i \qquad (1)$$

and

$$\delta^2 = \frac{1}{\lambda} \sum_{i=1}^{N} (hue_i - \mu) \qquad (2)$$

For a pixel i with value hue_i, the distance from μ to hue_i is defined by:

$$D = \frac{(hue_i - \mu)^4}{\delta^2} \qquad (3)$$

The larger D is, the higher the probability that the pixels belong to the different body colors (skin tones) is. In the HSV color histogram the number of pixels with the same hue color is computed between the model histogram Mo and the image histogram for each body color candidate region as follows:

$$CF = \frac{\sum_{i=1}^{\lambda} Min(I_{Ri}), Mo(i)}{\sum_{i=1}^{\lambda} (I_{Ri})} \qquad (4)$$

Where:
CF = The Confidence Factor for the current Region
(I_{Ri}) and Mo(i) = The numbers of pixels in the i-th bin of the histogram
λ = The number of bins

Detection probability between standard deviation: Once the RGB color histogram has been converted to HSV histogram, the bell-curve is used to determine the detection probability between the two standard deviation $\sigma_1 = 2.2$ and $\sigma_2 = 5.8$ shown in Fig. 5.

These standards deviation have been calculated as follows:

- We calculated the square of the difference between each pixel value and the sample mean
- Added those values up
- Divided the sum by N-1(variance)
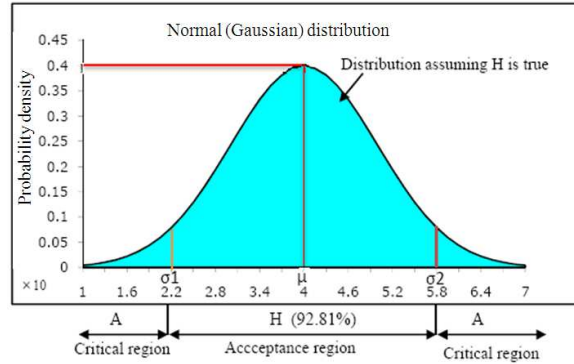- Took the square root to obtain the standard deviation

Fig. 5: Bell-curve and standard deviation

σ_1 denotes the half area under the distribution curve to the right of the center point and σ_2 represents the other half area under the distribution curve from the center to the left. The coefficient of skewness of the Gaussian distribution is set to zero, so that for any number x, P(X = x) = 0and its kurtosis to 4. In mathematics a special notation is used for calculating the probability, but in this research the probability between σ_1 and σ_2 is written as P (2.2<t<5. 8). To find this probability we subtracted the area bounded by t = 2. 2 to the area bounded by t = 5. 8. Using the t-table by taking into consideration the mean value. T-table is used to find the probability that a statistic is observed below, above or between values on the Gaussian distribution. To get P (t<5.8), we scanned down the t column until we located 5. At 5 we read across this row until we intersected the column that contains the place of the t-values. In our application the t-value is 0.8. So in the body of the table, the tabulated probability for t = 5.8 corresponds to the intersection of the row t = 5 with column 0.8. This probability is 0.92814, that is to say, P (t<5.8) = 0.92814. Next to get P (2.2<t) the same method is used and the probability is 0.03593, that is to say, P (2.2<t) = 0.03593.

Hence P = 0.92814-0.03593 = 0.9281 = 92.81%.

Next the statistical inference is used to reject the critical region (Fig. 5) based on the following hypothesis:

$$H = \begin{cases} GH : \mu = 4 \\ \sigma_1 < DP < \sigma_2 \end{cases} \qquad (5)$$

$$A = \begin{cases} GH : \mu \neq 4 \\ DP \leq \sigma_1 \, or DP \geq \sigma_2 \end{cases} \qquad (6)$$

Where:
GH = Gaussian Histogram
A = Alternative of the Gaussian histogram

μ   = The mean value
DP = Detection Probability

Here we are dealing with a human detection in an image with known standard deviation σ_1 and σ_2. We evaluated this hypothesis by supposing that the probability of detecting human presence is 92.81% with an error of ±7.19%. Even if the hypothesis is true, it is unlikely that we would get exactly 92.81% of chances. The decision to accept H or reject H is based on the result of the HSV color histogram obtained during the robot search mission. If the robot search result obtained is improbable H is rejected, that is to say, if the value of detection probability is less or equal to standard deviation sigma 1, greater or equal to standard deviation sigma 2 ($DP \le \sigma_1 or DP \ge \sigma_2$). But if the robot search result obtained is not improbable then, H is accepted. That is to say, if the value of the detection probability obtained is between sigma 1 and sigma 2 ($\sigma_1 < DP < \sigma_2$).

RESULTS

Two different experiments have been performed to evaluate the performance of the proposed method using image sensor camera with lens and digital signal processor LZ0P390M. The resolution of the camera is 353×288 pixels in CIF format and the operation frame rate is up to 15 fps. The sampling time used in the minimum time allowed by the camera frame is T = 0.06 sec. Figure 6a shows the HSV color histogram created during the robot patrol mission and Fig. 6b is the HSV color histogram model created with samples photos. During the robot patrol mission, the data of every detected objects, animals and humans are converted automatically to HSV color histograms which are compared with the model histogram stored in memory prior to make any decision. If the two histograms match, a beep sound is sent to the base station to warm the security manager that a human is detected.
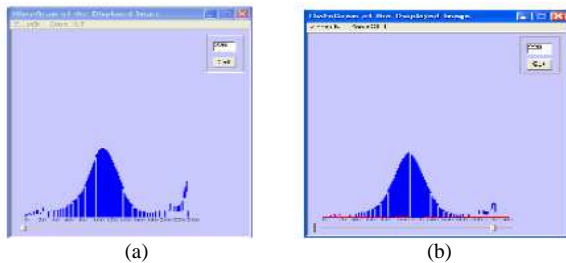

(a)   (b)

Fig. 6: Match of robot HSV color histogram and sample HSV color histogram (a) robot HSV histogram   (b) sample HSV histogram

In the base station the security manager can view a live video sent by the robot and based on the live video information the security guard can take an appropriate action. These images are recorded in both the robot on board computer hard disk and the base station computer hard disk as well and can last up to 15 days.

Experimental evaluation1:
In a room: The first experiment is done in a room with a little girl seating on a chair where two computers display of 15 inches covered almost her face and standing between objects. On the ceiling there are artificial light sources. Figure 7 shows the experimental evaluation result for person seating on a chair and playing around in a room.

In Fig. 7a the head of the girl is well detected in the rectangle box while in the Fig. 7b the entire body of the girl is well appeared in the rectangle box as expected during the algorithm implementation. Table 1 is the detection probability output results of each experiment.

In case 1, the detection probability output of a person seating on a chair is 90.13% and in case 2 it is 91.65% which agree with the Hypothesis H.

Experimental evaluation2:
In a hall: The second experiment has been performed in a hall where the little girl is walking at 20 m, Fig. 8a and b and 35 m, Fig. 8c and d from the robot respectively.

On the ceiling there are an artificial light sources for experiments of Fig. 8a-c while for experiment of Fig. 8d the light have been turned off to evaluate the robot detection capability in a dark environment. From these pictures we can see that the little girl has been well detected by the robot including her bag with an output detection probability ranges from 89.50-91.77% respectively.

Table1: Detection probability results

Case 1	PSCH	DP: 90.13%
Case 2	PSBO	DP: 91.65%

Where: PSCH = Person Seating on a Chair; PSBO = Person Standing Between Objects; DP = Detection Probability

(a)   (b)

Fig. 7: Robot detection results in a room (a) person seating on a chair (b) person standing between objects
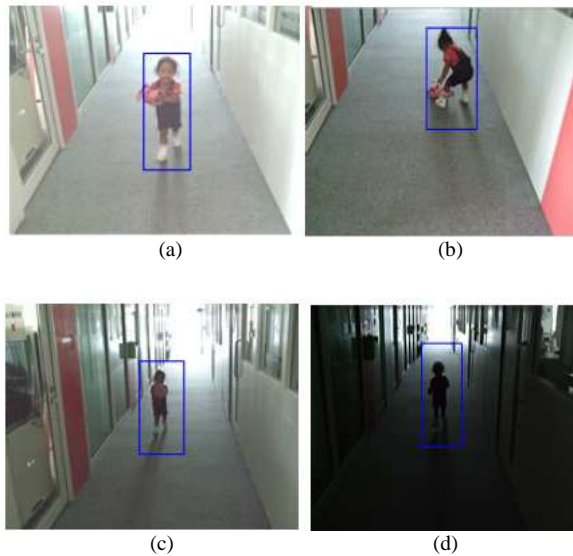
Fig. 8a-d: Evaluation results in a hall

DISCUSSION

The Information Theory approach has proved to be effective in solving many computer vision and pattern recognition problems. Such as: Image matching, clustering and segmentation, extraction of invariant interest points, feature selection, optimal classifier design, human detection or human detection and tracking, face detection (Blanco *et al*., 2007), to mention a few. The computational analysis of images and more abstract patterns is a complex and challenging task, which demands many inter disciplinary efforts. Nowadays, researchers are widely exploiting Information Theory (IT) elements to formulate and solve computer vision and pattern recognition problems. Among these approach (Heisele and Wohler, 1998) have used motion analysis for detecting walking pedestrians in sequences of color images. This approach is based on custom models and study with a fixed camera. A camera mounted on an autonomous robot will face difficulties separating human motion and other moving objects.

Laser based approaches use scanning laser range finders to detect and track human (Fod *et al*., 2002). In this approach, the model consists of two parts: the Background Model and the Foreground Model. The Background Model is used to filter background information that is not immediate relevance to tracking. The Foreground Model is used to predict range measurements that are not explained by the background model. The foreground maintains an estimate of the velocity of each object it tracks. The calculation in the Foreground Model involves the prediction of range readings using the velocity and information from the previous scan as well as re-estimation of the velocity. By repeating the above described process over a sequence of scans the trajectories of the moving objects is estimated. Panangadan *et al*. (2004) proposed a laser range finders-based system for tracking people in an outdoor environment and detecting interactions between them. In this approach to indentify anomalous behavior in the environment they build a model of frequency at which similar activities and interactions are observed. During the monitoring phase, they counted the number of times different activities that rare observed in a small interval. If the probability of observing the detected number of activities in that interval falls below a certain threshold, an anomalous behavior is detected. These techniques are simple to implement and require only a little processing effort and detecting performance is limited to a very plane. Also detecting many people with similar or a little different skin tones is difficult and may not resist to changing in environment.

In our proposed system, the cameras that are mounted into the robot can turn for about 90° left and 90° right covering a wide range of detection area. The robot has been assigned to detect every objects it encounters when patrolling the off-limit area to any suspicious human and report in real-time the detection situation to the base station when moving. The system can make a very clear difference between human, rock, trees including moving object and resist to any environment changes such as: illumination and even in dark environment with a detection performance of up to 90%. When a human is detected it stops and scans the all area around him before proceeding. This system is very useful for construction sites security where a human guard cannot study 24 h a day for the fear of losing life.

CONCLUSION

In this study we proposed a new human detection system for security guard robot using Gaussian distribution. The performance evaluation of the proposed method have been performed in a real environment whose the robot after detecting human presence sent a sound beep to warn the sentinel in the base station so that the security guard can take an appropriate action. The robot navigation has been achieved by Map Based Route Planning method (Zacharie, 2009) where its sensors readings have been used to decide the robot turning angle for obstacles avoidance. Our future study is to continue to evaluate the performance of the proposed method outside the building for improvement.

REFERENCES

Aggrawal, J.K. and S. Park, 2004. Human motion: Modeling and recognition of actions and interactions. Proceeding of the International Symposium on 3D Data Processing, Visualization and Transmission, Sept. 6-9, IEEE Computer Society, Washington DC., USA., pp: 640-647. DOI: 10.1109/3DPVT.2004.75

Blanco, J., W. Burgard, R. Sanz and J.L. Fernandez, 2007. Fast face detection for mobile robots by integrating laser range data with vision. Proceedings of the International Conference on Advanced Robotics, (ICAR'07), Universität Freiburg, Germany, pp: 953-958. http://www.informatik.uni-freiburg.de/~burgard/postscripts/icar2003blanco.pdf

Braun, T., T. Szentpetery and K. Berns, 2005. Detecting and following human with a mobile robot. Proceedings of the EOS Conference on Industrial Imaging and Machine Vision, (EOS'05), Universität Freiburg, Germany, pp: 1-4. http://agrosy.informatik.uni-kl.de/fileadmin/Literatur/Braun05a.pdf

Byeong-Sang, K., K.S. Kim, D.I. Min and J.B. Song, 2008. Development of jumping mechanism for a portable guard robot. Proceedings of the 39nd International Symposium on Robotics, pp: 55-59.

Dia, C., Y. Zheng and X. Li, 2007. Pedestrian detection and tracking in infrared imagery using shape and appearance. Comput. Vis. Image Understand., 106: 288-299. DOI: 10.1016/j.cviu.2006.08.009

Fod, A., A. Howard and M.A.J. Mataric, 2002. A laser-based people tracker. Proceeding of the IEEE International Conference on Robotics and Automation, May 2002, IEEE Press, Washington DC., USA., pp: 3024-3029. http://cres.usc.edu/pubdb_html/files_upload/70.pdf

Gehrid, D. and T. Schults, 2008. Selecting relevant features for human motion recognition. Proceedings of the 19th International Conference on Pattern Recognition, (ICPR'08), IEEE Press, USA., pp: 1-4. http://figment.cse.usf.edu/~sfefilat/data/papers/ThBCT8.47.pdf

Heisele, B. and C. Wohler, 1998. Motion-based recognition of pedestrians. Proceedings of the 14th International Conference on Pattern Recognition, Aug. 16-20, IEEE Computer Society, Washington DC., USA., pp: 1325-1330. DOI: 10.1109/ICPR.1998.711946

Mohan, A., C. Papageorgious and T. Poggio, 2001. Example based object detection in images by components. Patt. Anal. Mach. Intell., 23: 349-361.

Park, A.N., A. Mukovskiy, L. Omlor and M.A. Giese, 2008. Self organized character animation based on learned synergies from full-body motion capture data. University of Tübingen. http://www.compsens.uni-tuebingen.de/pub/pdf/CogSys08.pdf

Panangadan, A. M. Mataric and G. Sukhatme, 2004. Detecting anomalous human interactions using laser range-finders. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Sept. 28-Oct. 2, IEEE Xplore Press, USA., pp: 2136-2141. DOI: 10.1109/IROS.2004.1389725

Soltau, H., F. Metze, C. Fugen and A. Waibel, 2001. A one-pass decoder based on polymorphic linguistic context assignment. Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding, Dec. 2001, The Robotics Institute, USA., pp: 214-217. http://www.ri.cmu.edu/publication_view.html?pub_id=3929

Viola, P., M.J. Jones and D. Snow, 2003. Detecting pedestrians using patterns of motions and appearance. Proceedings of the 9th IEEE International Conference on Computer Vision, Oct. 13-16, IEEE Computer Society, Washington DC., USA., pp: 734-742.

Zacharie, M., 2009. Intelligent okiKoSenPBX1 security patrol robot via network and map-based route planning. J. Comput. Sci., 5: 79-85.