The 18th IEEE International Symposium on
Robot and Human Interactive Communication
Toyama, Japan, Sept. 27-Oct. 2, 2009

Real Time Detection and Evaluation of Region Of Interest by Mobile Robot using Vision

Md. Rokunuzzaman, *Student Member, IEEE*, K. Sekiyama and T. Fukuda, *Fellow, IEEE*

***Abstract*— This paper describes an approach for real time detection and evaluation of Region Of Interest (ROI) based on saliency and relevance. In our approach, we have considered moving background updating technique to adapt with real environment with camera equipped with mobile robot. We have introduced a new decision making criterion which selects the relevant objects based on human psychology of relevance. We have provided some experimental results which prove to be promising and realistic with this approach.**

I. Introduction

How do we evaluate a scene in real time? How can we select a best region to describe a scene or video when it is changing rapidly? Actually Region of Interest (ROI) bears important information of a scene and can be used to describe a scene or interpret a scene. It is sometime very hard to fix the proper size of Region of Interest. Why? This is because there is no general rule to fix a region of interest. The selection of ROI depends on user's interest or intention and differs with individual's choices. Moreover, we can expand or contract the size of our ROI. However, the expansion of ROI can give many and irrelevant information of the scene. The consequences of this are memory overflow and large computational time. On the other hand if we contract our ROI unwisely, it may give a partial information and ambiguity. Therefore we need a balance between Information level and ambiguity. Human visual system is very sensitive to motion objects than static objects. In observing motion object, human attains objects which have faster motion than other object. That's why in that situation attention overrides the decision of choice or individuals interest. Our method is based on human psychology of attention and relevance, therefore the chance of failure of ROI detection is comparatively less.

The main problem of ROI selection is to select the cognitive boundary of a region. Which objects should be inside the boundary and which objects are to be excluded from the boundary? To solve this problem we introduce selection criteria which help to make this decision. The derivation of this criterion is based on human psychology of relevance. Relevance theory claims that humans do have an automatic tendency to maximize relevance, not because we have a choice in the matter – we rarely do – but because of the way our cognitive systems have evolved [1].

Evaluation of ROI is important for justification of the selection based on important information in terms of attentional and relevance value.

Automatic detection of the ROI requires human like attention which enables robots to behave like Human Visual System. For automatic detection of ROI we have formalize Visual Saliency which is first introduced by Itti et al [2] in a different way for real time application.

Different methods have been proposed to detect regions of interest in an image or video. Stark et al. [3] defined visual regions of interest in relation with eye fixations. Automatic video region-of-interest determination based on user attention model is developed by Cheng et al. [4], locating regions of interests for Mars rover expedition is reported by Privitera in [5], Liu et al. [6] proposed a ROI detection and ranking algorithm based on attention model by using segmentation of perceptive regions. Tarkan et al. [7] propose a tool for the automatic detection and tracking of salient objects, and derivation of spatio-temporal relations between them in video. Fast approximation to a Bayesian model of visual saliency is proposed by Nicholas et al. [8]. Salient Region detection and tracking in video is developed by Y. Li et al. [9] .In [10] real time surveillance of people and their activities are presented. Martin Clauss et al. [11] present an evaluation of ROI based on attention algorithm using probabilistic measure which handles situations of unordered ROIs. Paulo estimated video object's relevance using segmentation and evaluated with human observer in [12]. The other part of the title 'Evaluation of ROI', the literature is not sufficient and well developed. There are few literatures which addresses this issue.

The existing methods are not sufficient as they do not relate human psychology of vision with ROI evaluation. They only quantify the region based on information entropy which is just a mathematical measure not with human behavior. Moreover, in real time there is no such evaluation to judge the region for attention. The current detection methods of ROI are based on interest objects detection which attracts human attention. However, only selection of interesting objects has little meaning than the relation with other objects which are relevant to the most salient object. In [11], the Evaluation of ROI is demonstrated with simple test scenarios, not using real environment. In this paper, we have presented a clear formalization of the ROI detection and Evaluation problem.

The rest of the paper is organized as follows. Section 2 gives an overview of our mobile robotic vision architecture on which we implement our algorithm. We describe the System Overview in Section 3. In Section 4, we give an overview of the proposed framework and explain our

Md. Rokunuzzaman is currently a doctoral student in the Department of Micro-Nano Systems Engineering at Nagoya University, Japan (corresponding author to provide phone: 052-789-4481; fax: 052-789-3115; e-mail: rzaman@robo.mein.nagoya-u.ac.jp).

K. Sekiyama is currently as an Associate Professor in the Department of Micro-Nano Systems Engineering at Nagoya University, Japan (e-mail: sekiyama@mein.nagoya-u.ac.jp)

T. Fukuda is currently as the Professor in the Department of Micro-Nano Systems Engineering at Nagoya University, Japan (e-mail: tofukuda@attglobal.net).

approach for automatic salient object detection, ROI determination and Evaluation. We present experimental results by using video movie and evaluate the performance of our system in Section 5. We conclude with possible future improvements to the system in Section 6.

II. MOBILE ROBOTIC VISION ARCHITECTURE

A. *Vision System Configuration*

Table I shows the Vision System Configuration for automatic detection of ROI and its evaluation.

TABLE I
VISION SYSTEM CONFIGURATION

Vision Processor	Intel Core 2 Duo, 2.20 GHz, 2.0 GB RAM
Vision Sensor	Canon Pan-Tilt-Zoom (PTZ) Camera
Development Platform	Microsoft Visual Studio 2005
Language Used	C++, Visual C++
Code Development	Intel's Open Computer Vision Library

We have used DirectShow to process the video frames after initialization of the vision system.

B. *Robot System*

Fig.1 shows the Robot system which controls the PTZ Camera and moves to observe the ROI with the Decision Module (PC).The Decision module is a small SONY Viao Model PC operating with Windows. The Robot is a Pioneer 3AT type which is equipped with Zigbee communication module, Camera and other sensors as modalities. The output of the video can be observed through user interface using OpenCV functions. The user interface also provides information about the Robot's positions with some reference coordinate that we can choose. Zigbee module uses Peer-to-Peer Communication through Teleoperation.

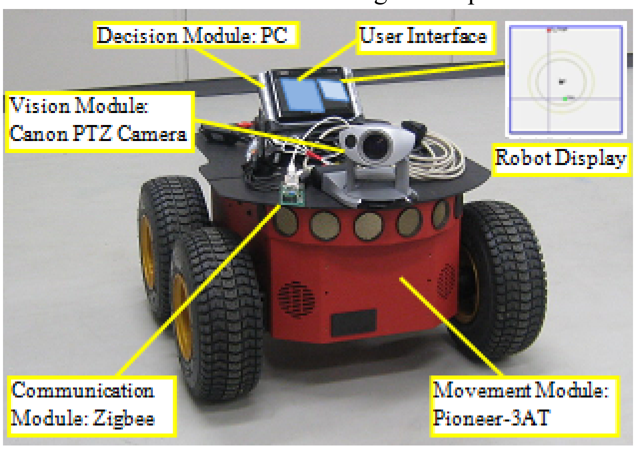

Fig.1 Robot system

III. REAL TIME ROI DETECTION AND EVALUATION

The method of ROI detection and Evaluation in real time is illustrated in the Fig.2. We first initialize our vision module and capture video frames using OpenCV [13] functions based on theories mentioned in the book [14]. In our approach we consider moving background model as it is very crucial in visual surveillance. We initialize background parameters and update the background at each frame.

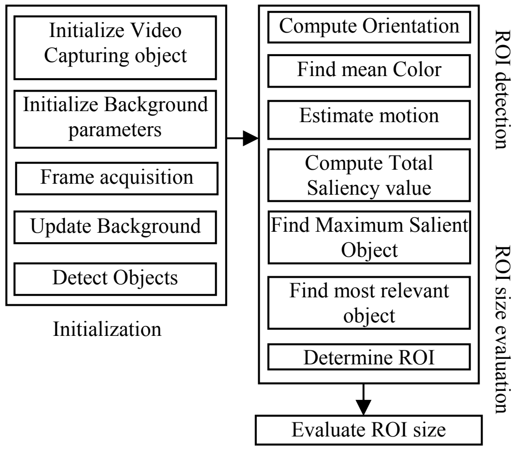

Fig.2 ROI detection and evaluation process

We detect objects in the foreground using blob detection module and measure each objects property. Then we calculate each objects saliency values and sort out according to saliency values. The object with maximum saliency has been selected which will be in the ROI first. Then we sort all the relevant objects with respect to this maximum salient object. Most relevant object is then determined. Two objects are then combined by fitting with a minimum bounding rectangle and make an ROI. Finally, the ROI size is evaluated by computing its salient information, maximum relevant information and compactness ratios based on areas covered by the ROI. The detail of each process is described in Section 4.

IV. FORMULATION OF THE PROBLEM

In this section, we formalize the Real time ROI detection and Evaluation for Mobile Robot. The steps are as follows:

A. *Image Processing*

In Real time detection there need to develop various image processing algorithms. The purpose of this algorithm is to model the background, detect object and estimate motion.

1. Background Modeling

Since the background is moving it is essential to model the moving background. Therefore, we have used Mixture of Gaussian Background Model which is based on Expectation Maximization (EM) Algorithm as proposed by Stauffer [15].

In this model background is assumed as the distribution which has higher evidence (ω) and lower variance (σ). Therefore, Gaussians are ordered by the value of ω/σ. The highest value ordered at the top and lower valued distributions are gravitate toward bottom and replaced by new distributions. The first distribution is chosen as background model, where

$$B = \arg\min_b \left(\sum_{k=1}^{b} \omega_k > T \right) \qquad (1)$$

Where T is a measure of the minimum portion of the data that should be accounted for by the background and ω_k is the weight for k th distribution. The weights are adjusted at time t as $\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t})$ (2)

Where α is the learning rate and $M_{k,t}$ is 1 for the model which matched and 0 for the remaining models. A match is defined as a pixel value within 2.5 standard deviations of a distribution.

Due to slow adaptation problems of the Stauffer-Grimson Model, an improved version is developed in [16]. In the improved version Gaussian mixture model is estimated by expected sufficient statistics update equations and then switch to *L*-recent window version when the first *L* samples are processed. The weight is updated by equation 3 as

$$\omega_{k,t} = \omega_{k,t-1} + \frac{1}{t}(M_{k,t} - \omega_{k,t-1}) \quad \text{(Sufficient Statistics)}$$
$$\omega_{k,t} = \omega_{k,t-1} + \frac{1}{L}(M_{k,t} - \omega_{k,t-1}) \quad \text{(L-recent window)} \tag{3}$$

The *L*-recent window update equations gives priority over recent data therefore the tracker can adapt to changes in the environment.

2. Object Detection

We have updated background, thus get foreground objects for each frame. We have clustered each objects in the foreground and taken as a blob by using a blob filter. Then we measure different properties of each blob in real time at each frame. The process is shown schematically in Fig.3.

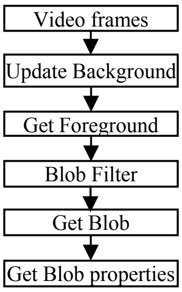

Fig.3 Object Detection process

3. Motion Estimation

We estimate the motion of each blob by their center to center euclidean distance in two consequtive frames.

Let the center position of i_{th} blob at frame $_n$ be Cx_n^i and Cy_n^i. The center position of i_{th} blob at frame $_{n+1}$ be Cx_{n+1}^i and Cy_{n+1}^i. Therefore the motion of the i_{th} blob can be computed as the absolute difference between the centers at frame $_{n+1}$.

For *x* coordinate,

$$dx_n^{n+1}(i) = \left|\left(Cx_{n+1}^i - Cx_n^i\right)\right| \tag{4}$$

and for *y* coordinate

$$dy_n^{n+1}(i) = \left|\left(Cy_{n+1}^i - Cy_n^i\right)\right| \tag{5}$$

The resultant motion of the i_{th} blob can be found by the Euclidean distance as

$$M = \sqrt{(dx_n^{n+1}(i))^2 + (dy_n^{n+1}(i))^2} \tag{6}$$

B. Saliency Determination

For real time ROI detection we used the concept of Visual saliency. We define saliency as the normalized addition of Orientation, Color Intensity and Motion of each blob at each frame. Different from the other approaches, we compute center surround as the difference each object's property from their global minimum.

C. Total Saliency Value Determination

Real time ROI detection needs attentional mechanism, which can be formulated by Saliency. In this method, we simply define it as a combination of Orientation, Color Intensity and Motion cues weighted by information density. Pixels convey information about an object. Another aspect of human vision system is that it attains objects with larger area as it covers most of the portion of the retina. Based on this concept we define a term information density with eq.7 as

$$I_D = A_O / A_{OBB}, 0 \le I_D \le 1 \tag{7}$$

Where, A_O is the area measured by number of pixels inside the object and A_{OBB} is the rectangular area of the box which fits the periphery of the object.

We compute the saliency of each object by total of each cue with eq.8 as

$$T_{sv} = (O'_{sv} + C'_{sv} + M'_{sv}).I_D \tag{8}$$

Where O'_{sv}, C'_{sv} and M'_{sv} is normalized value of orientation, color and motion. We have further normalized to obtain the total saliency value between 0 to 1. Finally Eq. 8 becomes

$$T'_{sv} = T_{sv} / \max(T_{sv}) \tag{9}$$

1. Orientation Saliency Value:

Orientation is calculated by fitting an ellipse in each object and calculating its angle between horizontal axis and the first side (i.e. length) in degrees. Then the Orientation Saliency Value is computed as:

$$O_{sv} = (O_i - O_{min}) \tag{10}$$

Where O_i is the orientation in terms of angle of i_{th} object, O_{min} is the minimum among n objects in the frame at that instant. Finally after normalization we get,

$$O'_{sv} = O_{sv} / \max(O_{sv}) \tag{11}$$

2. Color Saliency Value:

The color R, G, B components are extracted by a convolution mask of ROI from the blob coordinate with the current frame. Then mean color of each component is computed for each object. We then calculate Average Intensity of the values by the following equation:

$$\bar{I} = (\bar{\mu}_R + \bar{\mu}_G + \bar{\mu}_B)/3 \tag{12}$$

Where $\bar{\mu}_R$, $\bar{\mu}_G$ and $\bar{\mu}_B$ are the mean values of R, G and B components of the object respectively. The Color saliency is therefore

$$C_{sv} = (\bar{I} - \bar{I}_{min}) \tag{13}$$

Finally after normalization,

$$C'_{sv} = C_{sv} / \max(C_{sv}) \tag{14}$$

3. Motion Saliency Value:

Motion Saliency value of an object is the difference of motion of object and minimum of the surrounded object

$$M_{sv} = (M - M_{min}) \tag{15}$$

After normalization we obtain

$$M'_{sv} = M_{sv} / \max(M_{sv}) \tag{16}$$

Using equation (11), (14) and (16), we obtain Total saliency value of each object.

D. Relevancy determination:

In our approach we consider relevancy as the nearest object of the most salient object. Moreover, any object which has more area has more information and can be a candidate to include in a ROI. Since there are many objects that compete to attract human attention and the winner stands out among the candidate, therefore we need to select the winner. To select the winner, we introduce a new term Chance Factor which is defined as

$$CF = I_D / D_R \tag{17}$$

Where I_D is the information density defined in eq.7 and D_R is the Relative distance calculated as the Euclidean distance of the surround objects from the maximum salient object and formulated as

$$D_R = \sqrt{(Cx_{MaxSalObj} - Cx_i)^2 + (Cy_{MaxSalObj} - Cy_i)^2} \tag{18}$$

Where, $Cx_{MaxSalObj}$ is the center x coordinate of the maximum salient object and Cx_i is the ith object's center x coordinate except maximum salient object. The second term applies for y coordinate.

E. ROI determination

We assume ROI as a rectangular component consist one most salient rectangle R_1 and most relevant rectangle R_2. According to minimum bounding rectangle theory, the ROI size is illustrated by Fig. 4 and calculated as follows

$$W_{ROI} = \{(\text{Max } x \mid x \in R_1, R_2) - (\text{Min } x \mid x \in R_1, R_2)\} \tag{19}$$

$$H_{ROI} = \{(\text{Max } y \mid y \in R_1, R_2) - (\text{Min } y \mid y \in R_1, R_2)\} \tag{20}$$

$$S_{ROI} = W_{ROI} \times H_{ROI} \tag{21}$$

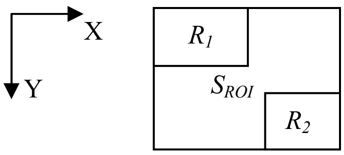

Fig. 4 ROI size determination

F. ROI Size Evaluation

To evaluate ROI size lets define some variables as Image size S_I, most important information, I_{MI} and size ratio, S_r. The definitions are: $S_I = Im_w \times Im_H$ (22)

Where Im_w, Im_H are image width and image height in pixels respectively. The maximum relevancy value of object is determined by its saliency value corresponding to maximum CF denoted by $T_{sv}\,|_{max(CF)}$. Making the summation we obtain,

$$I_{MI} = \max(T_{sv}) + T_{sv}\,|_{max(CF)} \tag{23}$$

$$S_r = S_I / S_{ROI} \tag{24}$$

Where S_{ROI} is the size of ROI defined in eq. (21). Evaluation of ROI is quantified by EF_{ROI} which can be expressed as

$$EF_{ROI} = I_{MI} / S_r = (\max(T_{sv}) + T_{sv}\,|_{max(CF)}) \times S_{ROI} / S_I \tag{25}$$

V. Experimental Results

In Experiment we have test our algorithm with different real time video sequences at various lighting and contexts. The results are summarized as follows:

A. Real Time ROI Detection Results

1. Indoor Video: Video of Amigobot's Motion

Real time ROI detection is applied in indoor as the Amigobot moves in the Laboratory and its relevant and salient objects are captured in the ROI as shown in Fig.5.

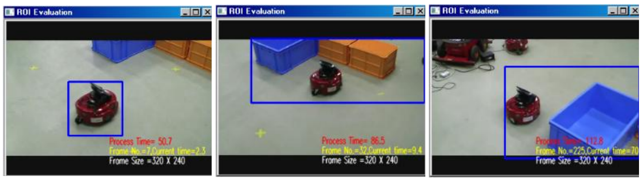

Fig.5 ROI detection in indoor video

2. Outdoor Video1: Video of people, busy road and cars

The Real time ROI detection in a busy road with cars is depicted in Fig.6. From this figure it is seen that salient objects and its relevant objects are in the same ROI.

Fig.6 ROI detection in Busy Road video

3. Outdoor Video2: video of Peoples crossing the Road

In the figure 7 it is seen that the most salient objects are detected first and its relevant salient objects are included in the ROI in real time. Moreover it adapts itself quickly with moving background.

Fig.7 ROI detection in crossing road video

4. Outdoor Video3: Video of peoples around campus

Another outdoor video in which a car is standing and people is approaching or moving around is used for test. In Fig.8 the ROI includes interesting objects with adaptation of the moving background as surveillance system.

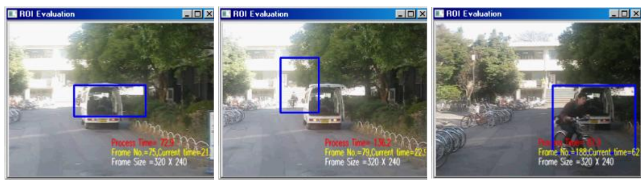

Fig.8 ROI detection in Around Campus video

B. Time Evaluation Results

The ROI detection is evaluated through total processing time needed which includes Background updating, process looping up to viewing. In our system we use Intel Core2 Duo processor with 2.20 GHz and 2.00 GB physical RAM. The processing time is calculated based on the internal clock counter on x86 types PC. Time based analyses are shown in Table II and Table III.

TABLE II
BACKGROUND UPDATE, ROI DETECTION AND TOTAL TIME

Video Title	Time (ms)			
	Background Update	ROI detection		Total Process
	Avg.	Max.	Min.	Max.
Indoor video	23.5	113.1	6.0	138.9
Outdoor video1	25.1	131.1	4.6	158.8
Outdoor video2	25.8	363.5	10.1	393.6
Outdoor video3	24.6	361.4	6.5	391.0

TABLE III
TOTAL PROCESSING TIME FOR DIFFERENT IMAGE SIZE

Image size (pixels)	80×60	160×120	320×240	640×480
Time (ms)	2.98	9.10	43.72	144.83

C. ROI Detection Time for Different Videos

In Fig. 9 we have plotted ROI detection time for various test videos w.r.t 400 frame sequences taken at speed of 30 frames/sec.

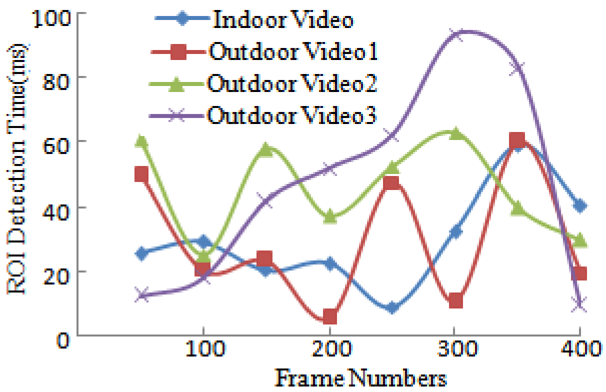

Fig. 9 ROI detection time for 4 sample test video sequences

D. Evaluation of ROI size Results

1. Evaluation Factor Vs Total(Average) Process Time

In Fig.10 we have plotted Evaluation Factor and corresponding Process time for various types of video sequences. The video frame has a size of 320×240 for all types of video. From this figure we can see the average process time is between 40~50 ms when the EF reach its maximum. The results in Fig.10 are shown as (a) for indoor video, (b), (c) and (d) for Outdoor video 1, Outdoor video 2 and Outdoor video 3 respectively.

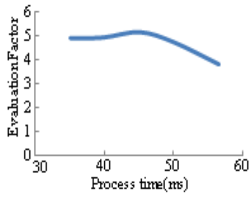

Fig.10 (a)

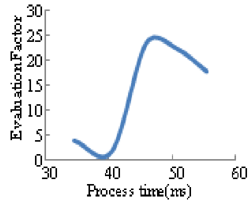

Fig.10 (b)

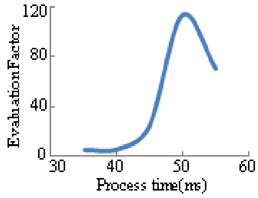

Fig.10(c)

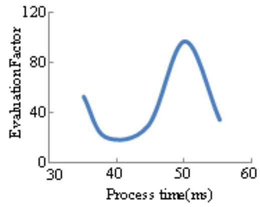

Fig.10 (d)

Fig. 10 Evaluation Factor Vs Total (Average) Process time for 4 sample test video sequences

2. ROI Size Vs Total Process Time(TPT)

In Fig.11 we show some characteristic graphs for different types of video sequences. From these graphs we can see that ROI size is at a minimum value in the interval between 40-50 ms process time. This means that the Evaluation of ROI is maximum at its minimum ROI size. Therefore ROI has its richness of useful information at its minimum size for memory allocation. The results in Fig.11 are shown as (a) for indoor video, (b), (c) and (d) for Outdoor video 1, Outdoor video 2 and Outdoor video 3 respectively.

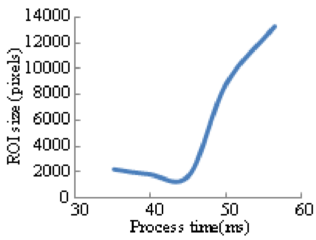

Fig.11 (a)

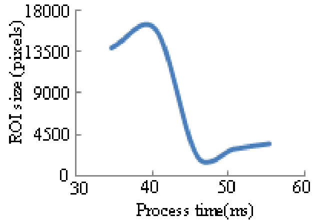

Fig.11 (b)

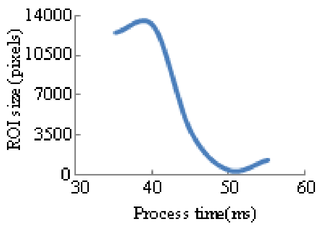

Fig.11 (c)

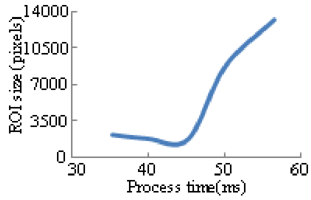

Fig.11 (d)

Fig. 11 ROI Size Vs Total (Average) Process time for 4 sample video sequences

E. Comparative Detection Accuracy

We have compared our ROI detection with existing methods of detection of salient objects. First we detect ROI in real time and save the video in an AVI (Audio Video Interleave) file. Then we run the video and pause it according to frame sequence. In this way we obtain some static images. To compare results we take 12 sample frames of each video and determine the average detection rate (D) and false alarm rate (F). The True Positives (TP) and False Positives (FP) are set by human eye fixation data. If total number of frames is N then D and F are defined by the eq.25 and 26 respectively as

$$D = TP / N \tag{25}$$

$$F = FP / N \tag{26}$$

Then we draw a ROC curve with only one point (D, F) of each video results in Fig.12.

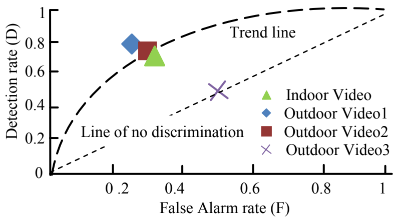

Fig. 12 ROI detection performance

Since we have only one point, we compute the area under ROC curve as an index of detection accuracy with a method described in [17]. The area under curve (AUC) is calculated as follows:

$AUC_{indoor\ video}$ = 0.75+0.25× (D-F)-F× (1-D) = 0.75+0.25× (0.67-.33) +0.33× (1-0.67) = 0.735= 73.5%

With similar calculations: $AUC_{OutdoorVideo1}$ = 81.25%, $AUC_{OutdooVideo2}$=0.76% and $AUC_{OutdoorVideo3}$=62.5%.

The detection accuracy is compared and shown by Table IV. From this table it is evident that our method performs better than other's method in ROI detection. The existing methods detect only salient objects whereas our method detects relevant objects as well.

TABLE IV
COMPARATIVE DETECTION ACCURACIES

Model	Our method	Neuro-Vision Tool (NVT)	Itti et.al	Informax	Saliency Tool Box (STB)
Ref.		*IEEE Trans. on PAMI*, 20(11), 1254-1259, 1998	*Vision Research* 40:1489-1506, 2000	*NIPS*, 155–162, 2005	*Neural Networks* 19, 1395-1407 , 2006.
ROC area	**0.81**	0.75	0.69	0.72	0.74

VI. Discussion and Conclusions

We have successfully applied our algorithm in Real time Applications by using vision. We compare our method with method described in recent paper [8]. We only compare the saliency computation time for both systems for various size of images. We run the code of method [8] in our system and compare our method. Table V shows these comparative results. From this table, we can see that for 320×240 and 640×480 our method is nearly 2.0 times faster than Fast SUN method in saliency computation time.

TABLE V
SALIENCY COMPUTATION TIME (AVERAGE) IN MILISEC

Image size (pixels)	80×60	160×120	320×240	640×480
Method [8] : Fast SUN	4.5	19.5	92.0	580.5
Our Method	3.2	19.0	45.0	350.0
Time save	1.3	0.5	47.0	230.5

The main contribution of our algorithm is we have developed a new decision making criteria which selects the relevant object without any training data or learning by the robot. We evaluate the scene instantly by which we solve the ROI selection problem by introducing new decision making criteria which is derived from human psychology of relevance. We successfully detect ROI in moving background relative to other moving objects which is very difficult for visual surveillance. In future we will extend our algorithm to select ROI based on interaction in real time.

References

[1] D. Sperber, F. Cara and V. Girotto, "Relevance theory explains the selection task," *Cognition*, vol. 57, pp. 31-95, 1995.

[2] L. Itti, C. Koch, et al., "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.20-11, pp.1254–1259, 1998.

[3] M. Privitera and W. Stark, "Algorithms for Defining Visual Regions of Interest: Comparison with Eye Fixations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.22-9, pp.970-982, 2000.

[4] H. Cheng, W. Chu et al., "Automatic video region-of-interest determination based on user attention model," *IEEE Int. Symposium on Circuits and Systems*, vol. 4, pp. 3219 – 3222, 2005.

[5] M. Privitera, M. Azzariti et al., "Locating regions-of-interest for the Mars Rover expedition," *Int. J. Remote sensing*, vol.21-17, pp. 3327-3347, 2000.

[6] H. Liu, S. Jiang et al., "Region-Based Visual Attention Analysis with Its Application in Image Browsing on Small Displays," in *Proc. of the 15th Int. Conf. on Multimedia*, 2007, pp.305 – 308.

[7] T. Sevilmis, M. Bastan et al., "Automatic detection of salient objects and spatial relations in videos for a video database system," *Int. J. of Image and Vision Computing*, vol. 26, pp. 1384–1396, 2008.

[8] N. Butko, L. Zhang et al., "Visual Saliency Model for Robot Cameras," *in Proc. of the Int. Conf. of Robotics and Automation, ICRA*, 2008, pp. 2398-2403.

[9] Y. Li, H. Zhang et al., "Salient Region Detection and Tracking in Video," *Int. Conf. on Multimedia and Expo*, vol.2, pp.269-272, 2003.

[10] I. Haritaoglu, D. Harwood and L.S. David, " Real-Time Surveillance of People and Their Activities, " *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol.22-8, pp.809-830, 2000.

[11] M. Clauss, P. Bayerl and H. Neumann, "A Statistical Measure for Evaluating Regions-of-Interest Based Attention Algorithms," *Pattern Recognition*, vol.3175, pp.383-390, 2004.

[12] P. Correia and F. Pereira, "Estimation of Video Object's Relevance," *in Proc. of European Signal processing conference (EUPSICO)*, Finland, 2000, pp. 925-928.

[13] Intel's Open Computer Vision (OpenCV) Library, Available : http://sourceforge.net/projects/opencvlibrary

[14] G. Bradsky, A.Kaehler, *Learning OpenCV : Computer Vision with the Open CV Library*, 1st ed. O'Reilly Media, Inc., 2008, ch 4, 5,9,10.

[15] C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," in *Proc. of the IEEE Int. Conf. of Computer Vision and Pattern Recognition,* pp. 246-252,1999.

[16] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection," *Proc. of the 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01*. pp.135-144, 2001.

[17] S.T. Mueller and J. Zhang, "Upper and lower bounds of area under ROC curves and index of discriminability of classifier performance." *Proc. of the ICML 2006 Workshop on ROC Analysis in Machine Learning,* pp.41–46, 2006