Omnidirectional Surveillance System for Digital Home Security

Wai Kit Wong, Joanne Tze Yen Liew, Chu Kiong Loo and Wei Kin Wong.
Faculty of Engineering and Technology, Multimedia University, 75450 Jln Ayer Keroh Lama, Malaysia
e-mail: wkwong@mmu.edu.my, jli_lty@yahoo.com, ckloo@mmu.edu.my, kin2031@yahoo.com.

Abstract— **This paper proposed an efficient omnidirectional surveillance system for digital home security. In this surveillance system, the omnidirectional scenes in a room, kitchen, car porch, dining hall, corridor, garage or backyard within a digital home are first captured using a web camera attached to a hyperbolic optical mirror. The captured scenes are then fed into a laptop computer for image processing and alarm purposes. Log-polar mapping is proposed to map the captured omnidirectional image into panoramic image, hence providing the observer or image processing tools a complete wide angle of view. A trespasser detection algorithm is also designed for trespasser detection purpose. The observed significances of this new proposed surveillance system include: it can cover a wide angle of view (360° omnidirectional), small in size, low cost, lightweight, output images are with higher data compression, and accurate trespasser detection.**

Keywords- Surveillance System, Log-Polar Mapping, Trespasser Detection.

I. Introduction

Visual surveillance had been an active area of research in the field of computer vision for the last few decades. Surveillance has a wide range of applications ranging from our day-to-day life to high security applications such as defense activities. Accordingly, researchers in the field of image and video technology have paid much attention to the study and development of highly sophisticated surveillance system.

One of the earliest real time video surveillance system for use in unattended outdoor environment was proposed by G.L. Foresti in 1998 [1]. Following this, several other advanced surveillance systems were developed, for examples: multiagent surveillance by Remagnino *et al.* [2], remote-cable based surveillance system by Sacchi *et al.* [3], etc. However, their surveillance systems are based on single stationary camera. Such systems are limited by the small field of view (FOV) that can be captured by the camera, so omnidirectional coverage is impossible.

Several surveillance systems using multiple camera/sensors had been proposed to cover omnidirectional view [4-7] but require prior calibration in order to establish correspondence between objects viewed by different cameras. Another approach to wide area surveillance is by using pan-tilt-zoom camera [8], [9] where a single camera effectively moves along the x, y and z axes in the 3D space images so obtained are stitched together to form a panoramic image by employing image mosaicing technique. The main advantage of the two mechanical solutions stated above is the possibility of acquiring very high-resolution images and its major drawback is much time required to mechanically scan the scene to obtain a single omnidirectional image.

Therefore, researchers proposed another option, which is the optical approach. It is by the use of special purpose lenses (such as fish eye lenses [10]) and the use of hyperbolic optical mirror [11]. As we intend to investigate real-time processing and applications, only the optical approach will be pursued. Hyperbolic optical mirror is used in this paper because it is less expensive as compare to fish eye lenses with almost the same reflective quality. The image obtained is not immediately understandable because of the geometric distortion introduced by the optical mirror. Hence, we need to introduce a log-polar mapping which maps the original omnidirectional image into a plane that providing a complete panoramic image. The advantage of such a mapping lays in providing the observer and image processing tools with a complete wide angle of view for the surrounding of the area under surveillance by preserving fine output image quality in a higher data compression manner.

Digital home [12] refers to a residence that is fully automated. The digital home is the result of trends leading to increase communications, entertainment, home control and convenience in a residence. There are five key trends for the design of the digital home which are: digital consumer devices, broadband access, home networking, residential gateways and middleware. While most researches and technology development in the field of digital home have aimed for remote accessing and controlling of various equipments and appliances inside a house, automatic and remote monitoring of a digital home has also been important from safety point of view. Accordingly, in this project, we propose to build up a system for surveillance of digital home. Our research is aimed to develop a new product that fall under digital consumer devices category, which is an embedded surveillance system that includes three main features: 1.) Omnidirection – for wide coverage. 2.) Capture images are unwarp to panoramic view. 3.) Surveillance for digital home with a trespasser detection algorithm.

This paper is organized in the following way: Section II will comment on the omnidirectional surveillance system model. Then in, Section III we will summarize the log polar mapping techniques that use to unwrap omnidirectional images into panoramic images. Section IV will briefly summarize the trespasser detection algorithm. Section V will show some experimental results and finally in Section IV, we will draw some conclusion and envision future developments.

II. Omnidirectional Surveillance Sysyem Model

The omnidirectional surveillance system model for digital home security is shown in Fig. 1.

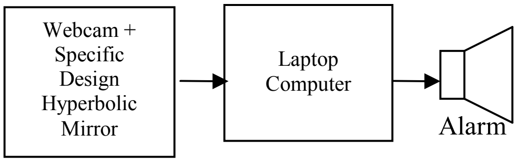

Fig. 1. Omnidirectional surveillance system model.

A. Webcam + Specific Design Hyperbolic Optical Mirror

The surveillance camera set used in this surveillance system consists of a webcam and an attached specific design hyperbolic optical mirror as shown in Fig. 2.

Fig. 2. Surveillance Camera Set (webcam and specific design hyperbolic optical mirror).

The webcam used in the surveillance system is E3500 PLUS QUICKCAM by LOGITECH. The webcam cost around US$25 and can capture high-quality VGA (640 x 480) video and 1.3-megapixel (software-enhanced) images. It is small in size and cheap compare to digital and CCTV camera with fine resolution output. The digital control is also accomplished through the USB port connected to a laptop or PC via Plug-and-play on Windows XP or Vista. It can be interface with MATLAB too. Therefore, it is best outfitted in omnidirectional surveillance system.

The specific design hyperbolic optical mirror used in omnidirectional surveillance system is a small size wide view type, with outer diameter 40mm and angle of view 30° above horizontal plane manufactured by ACCOWLE VISION. The mirror can reflect a 360° view surrounded by itself and as the webcam plug on it, omnidirectional images within a bedroom, kitchen, living room, corridor, garage etc. can be captured and send to a laptop computer in a monitoring room to be processed for surveillance purpose.

A custom-made bracket that is shown in Fig. 3(a) and 3(b) is designed to attach the hyperbolic mirror to the webcam via a socket.

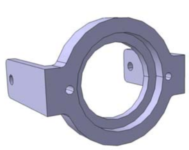

Fig. 3(a) Front view of the custom-made bracket.

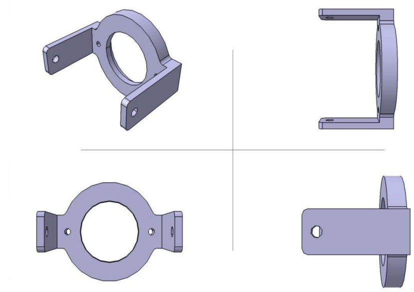

Fig. 3(b) Isometric, back and side views of the custom-made bracket.

B. Laptop Computer

A laptop computer can be used for image processing in a monitoring room. A core 2 Duo laptop computer with specs: 1.83GHz processor, 2 GB DDR2 RAM with MATLAB Ver. 7.0 is choose to be used here. MATLAB has toolbox interface with the webcam and can be used to perform log-polar mapping on the images captured to unwarp them into panoramic form. Then, the trespasser detection algorithm that is programmed in MATLAB can be used to process the panoramic images to detect whether there is any trespasser or not. Alarm will be signalled and suspected images with trespasser will be stored in a database for further police report and insurance claim purposes.

III. LOG-POLAR MAPPING

Log-polar geometry or log-polar transform in short, is an example of foveated or space-variant image representation used in the active vision systems motivated by human visual system [13]. It is a spatially-variant image representation in which pixel separation increases linearly with distance from a central point [14]. It provides a way of concentrating computational resources on regions of interest, whilst retaining low-resolution information from a wider field of view. One advantage of this kind of sampling is data reduction. Foveal image representations like this are most useful in the context of active vision system where the densely sampled central region can be directed to pick up the most salient information. Human eyes are very roughly organized in this way.

In robotics, there has been a trend to design and use true retina-like sensors [15], [16] or simulate the log-polar images by software conversion [17], [18]. In the software conversion of log-polar images, practitioners in pattern recognition usually named it as log-polar mapping. The advantages of log-polar mapping is that it can unwarp an omnidirectional image into panaromic image, hence providing the observer and image processing tools a complete wide angle of view for the surveillance area's surroundings and preserving fine output image quality in a higher data compression manner.

The spatially-variant grid that represents log-polar mapping is formed by i number of concentric circles with N samples over each concentric circle [13]. An example of a spatially-variant sampling grid is shown in Fig. 4.

The log-polar mapping can be summarized as following:

Initially, omnidirectional image is captured using a conventional camera and a hyperbolic mirror. The geometry of the captured omnidirectional image is in Cartesian form (x_1,y_1). Next, the Cartesian omnidirectional image is sampled by the spatially-variant grid into a log-polar form (ρ,θ) omnidirectional image. After that, the log-polar omnidirectional image is unwarped into a panoramic image (x_2,y_2), another Cartesian form. Since the panoramic image is in Cartesian form, subsequent image processing task will become much easier.

The center of pixel for log-polar sampling is described by [13]:

$$\rho(x_1, y_1) = \ln_b \left(\frac{R}{\rho_o} \right) \quad (1)$$

$$\theta(x_1, y_1) = \left(\frac{N}{2\pi} \right) \tan^{-1} \left(\frac{y_1}{x_1} \right) \quad (2)$$

The center of pixel for log-polar mapping is described as:

$$x_2(\rho, \theta) = \rho \cos \left(\frac{2\pi\theta(x_1, y_1)}{N} \right) \quad (3)$$

$$y_2(\rho, \theta) = \rho \sin \left(\frac{2\pi\theta(x_1, y_1)}{N} \right) \quad (4)$$

where R is the distance between given point and the center of mapping = $\sqrt{x_1 + y_1}$.

ρ_o is the scaling factor which will define the size of the circle at $\rho(x_1,y_1) = 0$.

b is the base of the algorithm,

$$b = \frac{N+\pi}{N-\pi} \quad (5)$$

N is the number of angular samples over each concentric circle.

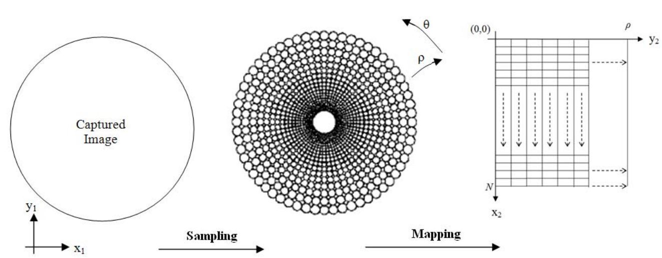

Fig. 4. A graphical view of log-polar mapping.

A graphical view illustrating the log-polar mapping is shown in Fig. 4 [13]. To sample the Cartesian pixels (x_1,y_1) into log-polar pixel (ρ,θ), at each center point calculated using (1) and (2), the corresponding log-polar pixel (ρ_n,θ_n) covers a region of Cartesian pixels with radius:

$$r_n = br_{n-1} \quad (6)$$

where n=1,2,3,…..,N-1. Fig. 5 shows the circle sampling method of log-polar mapping [13], [19].

The region of Cartesian pixels covered by an individual log-polar pixel will have the same color intensity with respect to the original Cartesian center sampling point. During the unwarping process, the (ρ,θ) pixel will map to each corresponding (x_2,y_2) pixel as shown in Fig. 6.

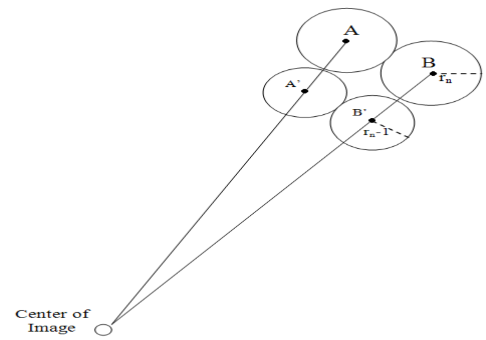

Fig. 5. Circular Sampling Structure in Log-Polar Mapping.

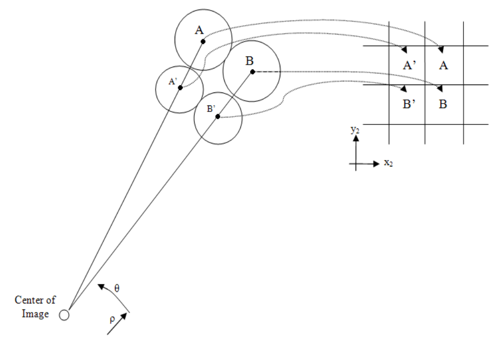

Fig. 6. Unwarping Process.

IV. Trespasser Detection Algorithm

Since the omnidirectional surveillance system is worked for security purpose, a trespasser detection algorithm is designed for tracking trespasser. During initialization, the first captured in image after unwarping into panoramic form will be treated as reference frame, R. After that, the subsequent images will be denoted I. I will be compared to the R iteratively by computing the pixel-wise absolute difference among them:

$$PWAD_t(r,c) = | I_t(r,c) - R_t(r,c) | \quad (7)$$

where r and c represents the 2-dimensional coordinates of row and column of the corresponding image pixels, t is the time instant. The resulting difference image characteristics will be stored in a matrix:

$$M_t(r,c) = \begin{cases} 1, & if \ PWAD_t(r,c) > T \\ 0, & Otherwise \end{cases} \quad (8)$$

where T is the predetermined threshold value. A variable named F_t is used to calculate the total number of 1s in matrix M_t. If F_t is within a predefined number of pixels, it indicates that a trespasser is present:

$$Trespasser = \begin{cases} present, & P < F_t < Q \\ not \ present, & otherwise \end{cases} \quad (9)$$

where P is the lowest limit of pixels number for a possible trespasser and Q is the highest limit of pixels number for a possible trespasser.

The trespasser detection technique will also encounter situations where there are sudden or gradual luminance changes, for instance the movements of shadows cast by static objects [21]. The occurrence of luminance changes can cause the designed algorithm to be deviated in terms of accuracy. Therefore, the reference frame is always updated by the expression [21]

$$R_{t+1} = \alpha I_t \; + \; (1-\alpha) R_t \qquad (10)$$

where α is the learning rate which controls the rate of the reference frame update. When the frequent abrupt luminance changes in which $F_t > Q$, the reference frame is corrected and updated by:

$$R_{t+1} = I_t \qquad (11)$$

V. Experimental Results

In this section, we briefly illustrate the application of the proposed omnidirectional surveillance system for digital home security. For log-polar mapping, we have chosen ρ_o = 1 and N = 79. Fig. 7 shows an omnidirectional image captured by the webcam in a study room within a digital home without any human being. Fig. 8 shows an omnidirectional image captured by the webcam in the same room after half minutes, when there is a human (trespasser) entering that room. Fig. 9 shows the panoramic image unwarps from Fig. 7, while Fig. 10 shows the panoramic image unwarps from Fig. 8. When the system detected the trespasser, it will signal the alarm, and store the suspected image with trespasser in a folder for user to trace back afterwards. From Fig. 7 and Fig. 9, we can observe that Fig. 7 is with resolution 408x408 while the panoramic form in Fig. 9 is with resolution 408x79. This means that the mapping scale is by 166464: 32232, with 5.16 fold of data reduction.

In human detection algorithm, there are 3 parameters to be optimized, which are the predetermined threshold value T, the lowest limit of pixels number for a possible trespasser P and the highest limit of pixels number for a possible trespasser Q. Since the images after log-polar mapping are converted into grayscale and class double, therefore the difference between $I_t(r,c)$ and $R_t(r,c)$ is in between 0 and 1.

For T value, 1000 sample images are used to test for every $|I_t(r,c)-R_t(r,c)|$ point with step size of 0.1. The graph Accuracy vs. $|I_t(r,c)-R_t(r,c)|$ is plotted in Fig. 11. From the plot, the optimum T value is 0.3, with the highest accuracy of 97%.

As for P value, we tried the algorithm with pets (hamster, cat, dog) and human, moving toward and away from the captured region. 1000 sample images are captured. By using the sample images, we repeated the simulation with P = 0.05, 0.1, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0% of number of pixels difference to total pixels in one image ratio. The graph Accuracy vs. Number of Pixel Difference to Total Pixels in One Image Ratio is plotted in Fig. 12. From the plot, the optimum P value is 0.1%, with the highest accuracy of 97%.

As for Q value, we tested the algorithm with human moving toward and away from the captured region and by changing the luminance. 1000 samples are captured, and we repeated the simulations on these sample images with Q = 0.5, 1, 5, 10, 20, 30 40 and 50% of number of pixels difference to total pixels in one image ratio. The graph Accuracy vs. Number of Pixel Difference to Total Pixels in One Image Ratio is plotted in Fig. 13. From the plot, the optimum Q value is 1.0%, with the highest accuracy of 97%.

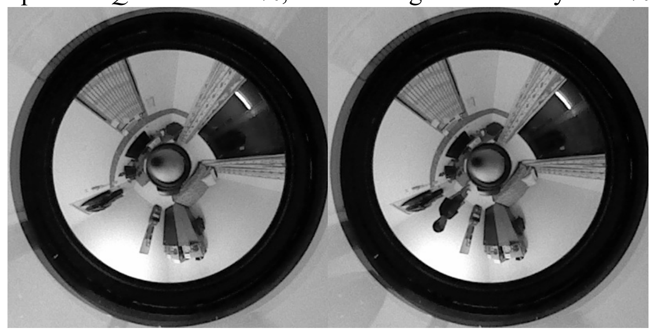

Fig. 7 Fig. 8

Fig. 7. An omnidirectional image captured by the webcam in a study room within a digital home without any human being.

Fig. 8. An omnidirectional image captured by the webcam in a study room within a digital home with a human (trespasser).

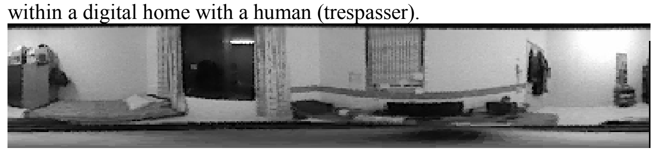

Fig. 9. The panoramic image unwarps from Fig. 6 (without trespasser).

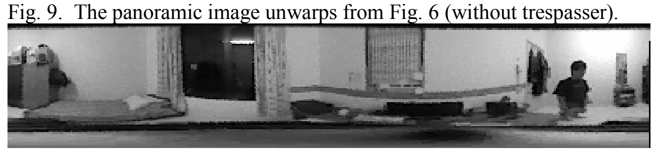

Fig. 10. The panoramic image unwarps from Fig. 7 (with trespasser).

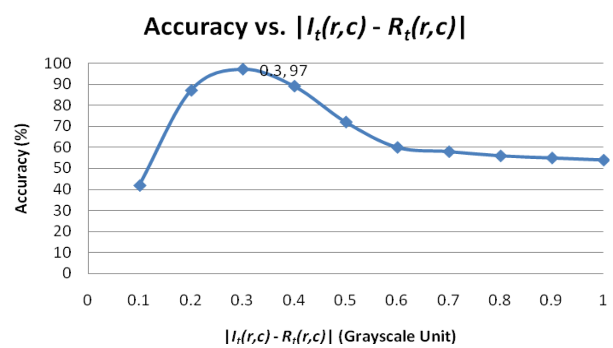

Fig. 11. Accuracy vs. $|I_t(r,c) – R_t(r,c)|$ plot.

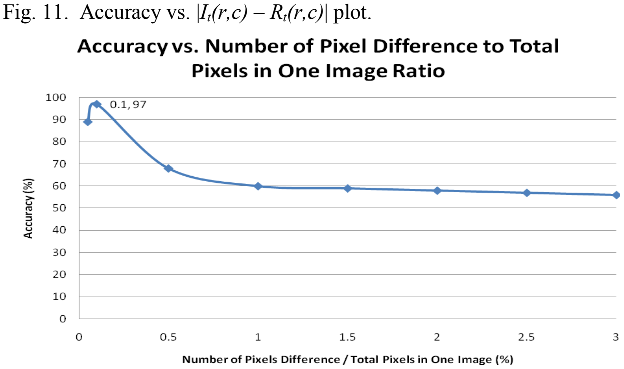

Fig. 12. Accuracy vs. Number of Pixel Difference to Total Pixels in One Image Ratio plot.

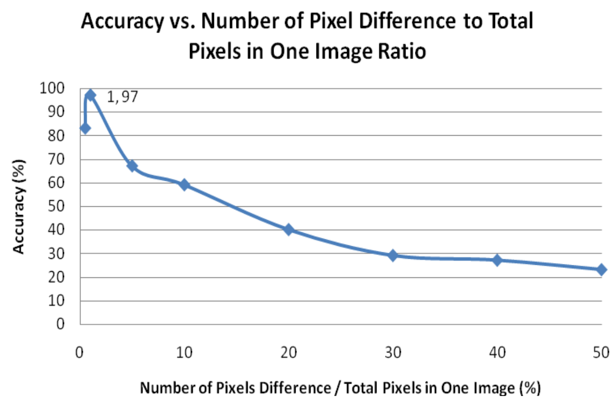

Fig. 13. Accuracy vs. Number of Pixel Difference to Total Pixels in One Image Ratio plot.

The omnidirectional surveillance system was evaluated with respect to the thermal images captured live and displayed on monitor scene as interpreted by a operator (human observer), the overall description of which could be called the "operator perceived activity (OPA)[21]". The operator will comment on the images captured by the webcam and whether there is any trespasser or not. The optimum values of *T*, *P* and *Q* obtained are applied into the system for high accuracy trespasser detection. A total of 10,000 samples images are captured for testing the performance of our surveillance system. This also include without trespasser, more than one trespassers and animals (dogs, birds, etc which are not counted as trespassers). From the total of 10,000 samples, 9741 were detected perfectly (trespasser-or-not condition agree by both observer and surveillance system), i.e. an accuracy of 97.41%. The surveillance system are also function in a fast way whereby the routine time to capture in an image, log polar mapping, trespasser detection till signal alarm is only 1.4 seconds.

VI. Conclusions

In this paper, we have proposed a new omnidirectional surveillance system for digital home security. There are some significant advantages of this new approach observed. This include: the surveillance system covers a wide angle of view (360º omnidirectional), small in size, lightweight, cheap, can cover every angle (omnidirectional) as compare to conventional type (small FOV), output images are in panoramic form and with higher data compression, and the trespasser detection is with high accuracy. In future, a thermal camera will be used to replace webcam for detection in outdoor environment where the lighting condition is naturally varying, and also in night vision too.

References

[1] G. L. Foresti, "A Real-Time System for Video Surveillance of Unattended Outdoor Environments", *IEEE Trans on Circuits and Systems for Video Technology*, Vol. 8, No. 6, p.p. 697-704, OCT. 1998.

[2] P. Remagnino, T.Tan, and K.Baker, "Multi-agent visual surveillance of dynamic scenes", *Image Vision Comput.* 16(8), p.p. 529-532, 1998.

[3] C. Sacchi, and C. S. Regazzoni, "Multimedia Communication Techniques for Remote Cable-Based Video-Surveillance Systems. *ICIAP 1999*, p.p.1100-1103, 1999.

[4] M. Greiffenhagen, V. Ramesh, D. Comaniciu, and H. Niemann, "Statistical Modeling and Performance Characterization of a Real-Time Dual Camera Surveillance System", *CVPR 2000*, p.p. 2335-2342, 2000.

[5] C. S. Regazzoni, and P. K. Varshney, "Multisensor surveillance systems based on image and video data", *ICIP (1) 2002*, p.p. 497-500, 2002.

[6] R. Bowden, A. Gilbert, and K.T.K. Pong, "Tracking Objects Across Uncalibrated Arbitary Topology Camera Networks", *Chapter 6, In Intelligent Distributed Video Surveillance Systems, S A Velastin and P Remagnino (Eds), IEE Professional Applications of Computing* Series 5, pp157-183, 2005

[7] Z. Yang, and A. F. Bobick, "Visual Integration from Multiple Cameras", WACV/MOTION 2005: p.p.488-493, 2005.

[8] G. Thiel, "Automatic CCTV surveillance-towards the VIRTUAL GUARD", *IEEE Aerospace and Electronic Systems Magazine*, Vol. 15, Issue 7, p.p. 3-9, 2000.

[9] S.R. Thirthala, S.N. Sinha, and M. Polleyface, "Calibration of pan-tilt-zoom (PTZ) cameras and omni-directional cameras", IEEE Computer society Conferecnce on Computer Vision and Pattern Recognition, Vol. 2, p.p. 1198, 2005.

[10] J. Kumar, and M. Bauer, "Fisheye lens design and their relative performance", *Proc. SPIE,* Vol. 4093, p.p. 360-369, 2000.

[11] T. Padjla and H. Roth, "Panoramic Imaging with SVAVISCA Camera-Simulations and Reality", *Research Reports of CMP*, CzechTechnical University in Prague, No. 16, 2000.

[12] http://en.wikipedia.org/wiki/Digital_home

[13] H. Araujo, J. M. Dias (1996): An Introduction To The Log-polar Mapping. Proceedings of 2nd Workshop on Cybernetic Vision, p.p. 139-144.

[14] C. F. R. Weiman and G. Chaikin (1979): Logarithmic Spiral Grids For Image Processing And Display. Computer Graphics and Image Processing, Vol. 11, p.p. 197-226, 1979.

[15] LIRA Lab, Document on specification, *Tech. report*, Espirit Project n. 31951 – SVAVISCA- available at http://www.lira.dist.unige.it.

[16] R. Wodnicki, G. W. Roberts, and M. D. Levine, "A foveated image sensor in standard CMOS technology", *Custom Integrated Circuits Conf.* Santa Clara, p.p. 357-360, May 1995.

[17] F. Jurie, "A new log-polar mapping for space variant imaging: Application to face detection and tracking", *Pattern Recognition, Elsevier Science*, 32:55, p.p. 865-875, 1999.

[18] V. J. Traver, "Motion estimation algorithms in log-polar images and application to monocular active tracking", PhD thesis, Dep. Llenguatges.

[19] R. Wodnicki, G. W. Roberts, and M. D. Levine, "A foveated image sensor in standard CMOS technology", *Custom Integrated Circuits Conf.* Santa Clara, p.p. 357-360, May 1995.

[20] Hongzhi Gao and R. Green, Dr., "A Quantitative Comparison Research on Frame Level Background Subtraction Algorithms", Proceedings of Image and Vision Computing New Zealand 2007, pp. 31–34, Hamilton, New Zealand, December 2007.

[21] J. Owens, A. Hunter and E. Fletcher, "A Fast Model-Free Morphology-Based Object Tracking Algorithm", British Machine Vision Conference, p.p. 767-776, 2002.