

Computational model for video surveillance, dual detection and processing of relevant events

Pamela Araceli Rangel Tirado, Luis Pastor Sánchez Fernández, Oleksiy Pogrebnyak

Centro de Investigación en Computación, Instituto Politécnico Nacional,
Av. Juan de Dios Bátiz s/n esq. Miguel Othón de Mendizábal,
Col. Nueva Industrial Vallejo, CP 07738, Mexico City, Mexico
pamela.rangel.t@gmail.com, {lsanchez, olek}@cic.ipn.mx

Abstract. This paper proposes a computational model for automated obtaining images in greater detail, through the combination of multiple fixed cameras and a mobile camera, working a couple at a time. It detects an event of interest and its location to make object pan, tilt, zoom. To achieve this objective, computer vision techniques, image processing, pattern matching and calibration are used and inter-camera Communication Protocol Pelco-D is employed. The application was programmed in LabVIEW 2010. The tests are performed in a restricted parking, resulting in a car zooming in.

Keywords: Surveillance, dual system, LabVIEW, virtual instruments, fixed camera, PTZ, pattern matching, computer vision, image processing.

1 Introduction

Nowadays, the detection and recognition of relevant events is one of the most studied areas of research. Images have a vital importance in the analysis of crime, as evidence of abuse or simply to track events. One of the outstanding problems is the fact that fixed cameras have a wide field of view (FOV), but cannot be used to view details of an object, such as a car plate or a person's face. One possible solution is to leverage commercial video surveillance systems of low and medium performance with advanced algorithms and a combination of fixed and mobile cameras, to perform more complex activities to maximize their utility.

This paper proposes an automated system that employs X_i fixed cameras to detect the presence of an object in a large FOV, processing images to segment them and find valuable information for pattern recognition of the object location. This is done positioning the mobile camera in the center of FOV and making zoom for obtaining an image with more details. The methods can be the basis for achieving more complex activities, such as tracking an object in real time.

The paper is structured as follows. Section II provides an overview of the state of the art. Section III describes the system architecture. Section IV presents the method-

ology to develop this system, with all the algorithms for image processing and control of the dome. Section V provides some results in images. Section VI concludes the paper.

2 Related works

On the social side, the increase in crime is one of the main factors that have influenced the demand for remote video surveillance systems [1], as it seeks to stop, prevent and combat illegal. Technologies such as video surveillance can observe, respond, intervene and mitigate risks that may arise. [2] There are many works in this area of research. Boyd et al. [3] describe a video surveillance system that assimilates information from multiple cameras in a single stage model. In a real time application with separate processors to perform low level operations (motion segmentation and recognition of objects), and then shared through a local area network, information with other parts of the monitoring system.

In such systems, there are some that will track objects of interest, either with a single chamber or more advanced systems. Regarding the first case, and Bilodeau Varcheie [4] conducted a project that seeks to identify and recognize objects and detect and track them using a moving pan, tilt, zoom (PTZ) camera with communication via Internet. This method is based on the comparison of elliptical samples to the objective evaluating the similarity probability when estimating the location of the object, through a study of the diffuse areas movement without employing the optical flow algorithm Lucas-Kanade. Likewise prediction of the object position is made to help track. This system has the peculiarity that the sample must be taken manually, which makes it useful for identifying previously chosen persons.

Another innovative system is proposed by Chu-Sing et. al [5], which in addition to tracking algorithms employed internally also adds geographical location of people. Another system proposed by Hsien-Chou and Wei-Yi [6], which proposes a dual camera system with PTZ using motion sensors to resemble the identification and tracking of an object as eagles do, using the first camera see the wide FOV and the other for the detail in the same area.

3 System Architecture

This paper presents the development of virtual instruments for video surveillance and security system (Figure 1), with getting video through X_i CCTV cameras, which may or may not require conversion to digital video (depending on the type of sensor employed), where $i = 1, 2, \dots, n$ and n depends on the FOV of the dome camera. The obtained frames are used in the algorithms to change the position of the dome camera and get detailed pictures. Once the process is complete, the user reports are generated and sent to a client computer. The method developed for image analysis and PTZ movement will be explained in Section IV.

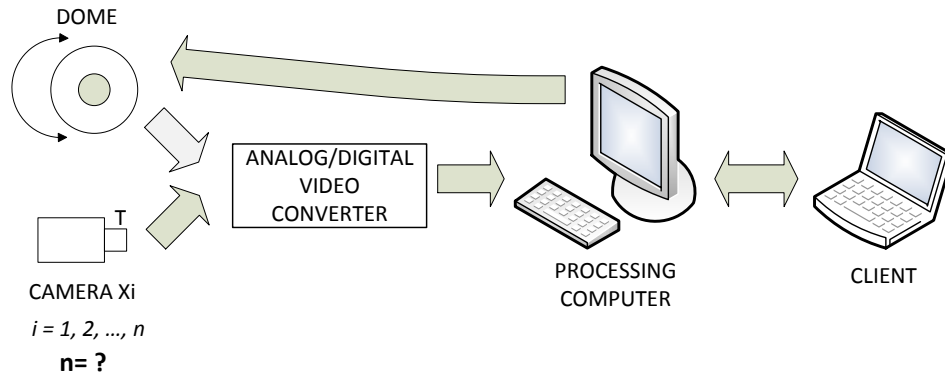


Fig. 1. Method of operation of the proposed system

4 Methodology

The application starts with the scanning of X_i fixed cameras to determine if an event happens in the regions of interest (ROI) that user previously determined, Fig. 2.

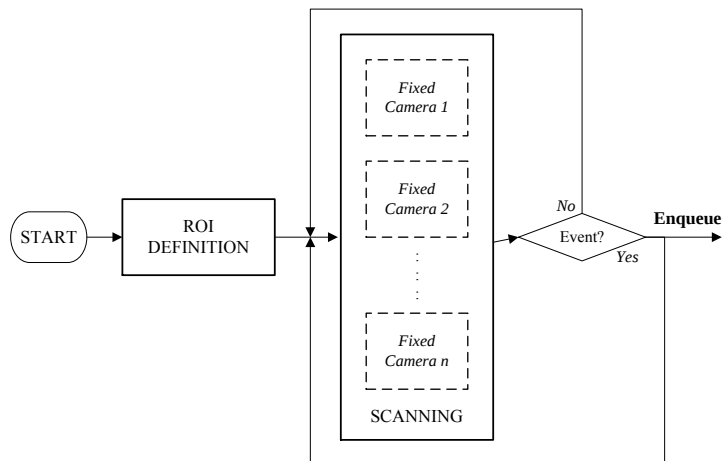


Fig. 2. General block diagram of the system, part 1

When the presence of an object has been detected, the identification number of the camera that generated the event and the images clean and with the object, are placed in a queue and sent to the next part of the application. The feedback line means that even when the data is sent, inspection continues to check for other event.

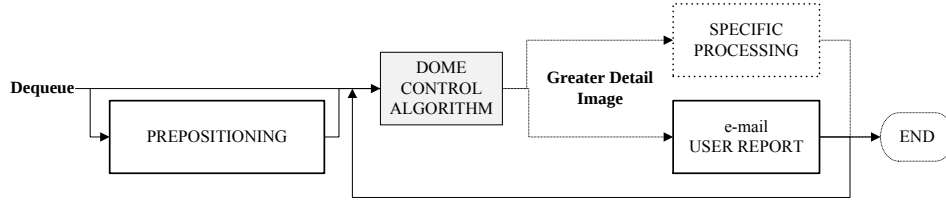


Fig. 3. General block diagram of the system, part 2

Fig. 3 shows that when the queue has elements, system starts running the prepositioning that uses preconfigured routines directly in the memory of the dome. This block is used to move the dome to a known area to resemble its FOV with the fixed camera X_i that generated the event, thus to find the object by pattern recognition.

Moreover, the dome control algorithm is executed as shown in Fig. 4; it will be explained in the following subsections.

Finally, when a focused image is obtained, the user report are generated and sent them to an email account. The specific processing block refers to the utility that can be the final image for other analyzes not attacked in this work, such as the determination of license plates or traits as textures, color and even model.

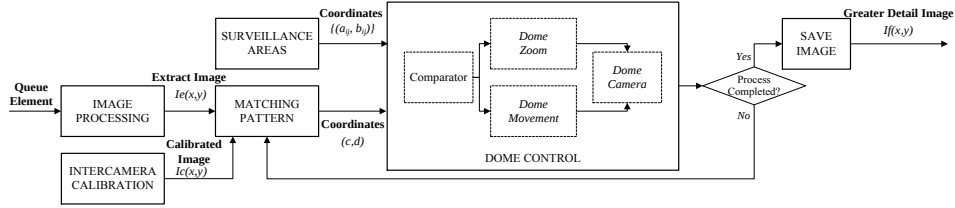


Fig. 4. Block diagram for dome control algorithm

4.1 Image Processing

The following explains the image processing components, which after the object discrimination, extract the object $I_{ext}(x, y)$ that triggered the event. The queue element consists of the identifier of the camera that generated the event, a clean image as a reference and the other with the object of interest.

Edge detection.

This algorithm is used to determine that an event has occurred, through the detection of changes in illumination along the ROI that the user determined before the inspection. Once done, a clean original image is saved; this will be the benchmark for image processing explained in the following subsection. When change lighting is detected, the coordinates of the points are stored in an array. This is very important for correct particle discrimination.

Image subtraction.

This is done taking two images: the reference image $I_A(x, y)$ after selecting the ROI, and the other image $I_B(x, y)$ containing the object, which has entered into the scene. At this step, the subtraction of images pixel by pixel, Eq. (1) is made [7]:

$$I_{Dest}(x, y) = I_A(x, y) - I_B(x, y). \quad (1)$$

This basic mathematical operation has the purpose of indicating in grayscale which pixels are not in the clean image, turning every other to zero (black), having an output image with high contrast. This is the reason that the image is binarized using only the average value of its histogram as the threshold value.

Particle analysis.

Firstly, a dilation Eq. (2) is calculated [7], with a structure element B of 3x3 size to fill internal voids in all non-zero elements:

$$X \oplus \tilde{B} = \{x \in Z^2, X \cap B_x \neq \emptyset\}. \quad (2)$$

Then, 5 erosions with the same structure element B of dilation according to Eq. (3) [7] are performed:

$$X \ominus \tilde{B} = \{x \in Z^2, B_x \in X\}. \quad (3)$$

This ensures that noise decreases its size. Then, particles are filtered per area. Finally, the coordinates of a circumscribed rectangle of their location are stored in an array.

Particle discrimination.

Ideally, only one object should be detected in the image, but it in practice the system observes various objects. To prevent the system might be confused, particle are discriminated by their proximity to the coordinates where the object presence was detected in the ROI. Suppose two coordinates (X_1, Y_1) and (X_2, Y_2) correspond to particles obtained after analysis that eliminated spurious particles. Also, consider that the point where they shot the event in the ROI is (X_a, Y_a) , but it has a margin of 100 pixels to define if the particle corresponds to the object (see Fig. 5). Mathematical logic operation is as follows:

$$P_i = O_i \leftrightarrow [X_a - 100 \leq X_i < X_a + 100] \wedge [Y_a - 100 \leq Y_i < Y_a + 100]. \quad (4)$$

That is, the particle P_i (with coordinates are equal to its subscript) is the object of interest O_i , if and only if it is into the tolerance limits set by ± 100 pixels in both directions.

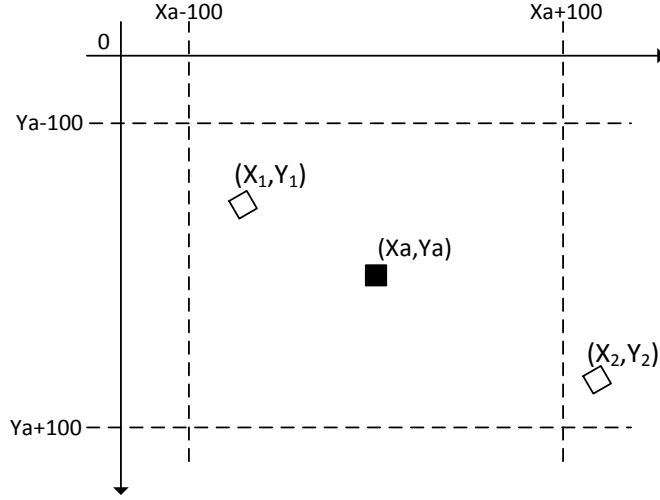


Fig. 5. Particle discrimination

4.2 Inter-camera calibration

The initial image after processing and extraction from a fixed camera is searched only once in the dome FOV, so it is necessary to compensate the difference in lighting levels before pattern matching. This is done by measuring the light intensity of the image of the dome to take some parameters and modified the extracted image histogram of the stationary camera. [8] That is, if the equalization expression, Eq. (5), is taken as a basis, the similarly with the other image is calculated according to Eq. (6):

$$s = \int_0^r P_r(w)dw = T(r) . \quad (5)$$

$$v = G(z) = \int_0^z P_z(t)dt . \quad (6)$$

To equalize the histograms, it is necessary to take the form $T(r) = G(z)$, and then z must satisfy the condition of Eq. (7):

$$z = G^{-1}(s) = G^{-1}[T(r)] . \quad (7)$$

4.3 Pattern matching

Pattern recognition has many uses. In this case it is used to find the object $H(x, y)$ that has entered the scene, giving its location and can thus continuously positioning in the dome FOV until it is centered. [7]

Learning.

At this stage, information is extracted from the objects of interest contained in the template through non-uniform sampling (see Fig. 6), where sections with the same intensities of pixel neighbors are taken to assure that the descriptive features have better quality and to reduce computational consumption. It is also important to define if the particles are rotated, scaled and/or moved.

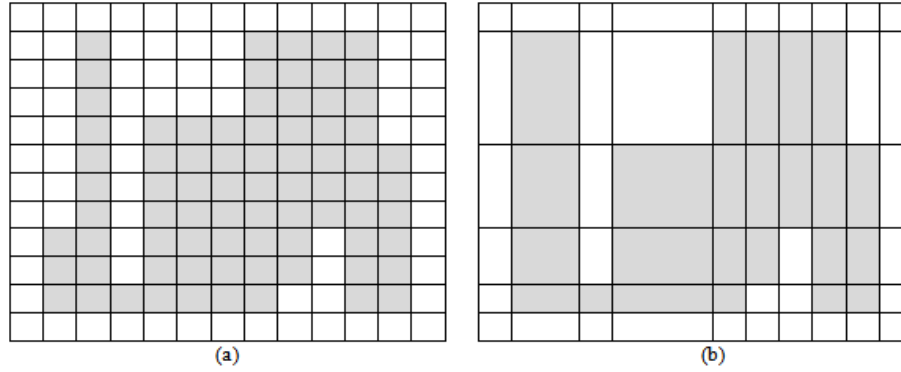


Fig. 6. Non-uniform sampling

Search.

During this phase, the descriptor is used to find the object in the inspected image. As it is shown in Fig. 7, the template is moved pixel by pixel to calculate a cross-correlation matrix according to Eq. (8) [9]:

$$C(i, j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x, y) f(x + i, y + j), \quad (8)$$

where $f(x, y)$ is the image, $w(x, y)$ is the template, x is a horizontal offset, y is a vertical offset, K is a horizontal template size, L is a vertical template size; j is a horizontal coordinate $j = 0, 1, \dots, M - 1$, i is a vertical coordinate $i = 0, 1, \dots, N - 1$, and $M \times N$ is the image size.

Once the cross-correlation matrix C is calculated, its highest value indicates that with a greater probability in the neighborhood of the corresponding pixel is the desired object. Therefore, the output will have the coordinates (c, d) of the template center in the inspected image.

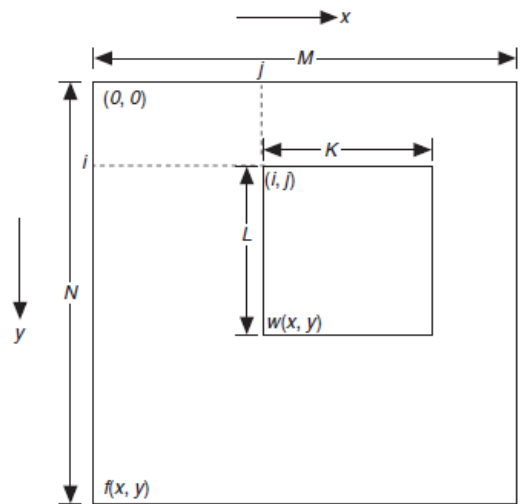


Fig. 7. Cross-correlation calculation

4.4 Surveillance areas

In order to correlate each of the data portions with control actions, the picture block of 2 dimensional (2D) video splits into 81 surveillance areas. Fig. 8 shows the FOV where in the central square there are areas of each color for a fine movement and a better focus on the object. Once located in the zone 40, the dome movement is stopped to zoom in the object.

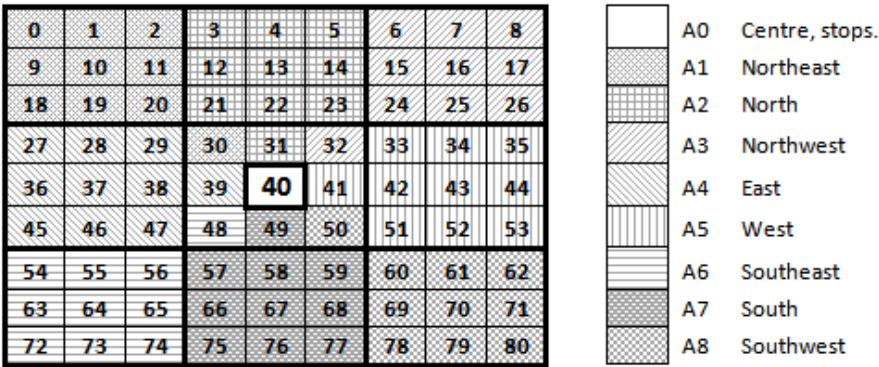


Fig. 8. Division FOV in surveillance areas

4.5 Dome Control

The movement of the dome is one of the most important parts of this application because thanks to this camera that moved to the object of interest location with a

zooming, a better quality image is provided. This is accomplished through communication protocol Pelco-D and RS-485, by sending of a command string in hexadecimal.

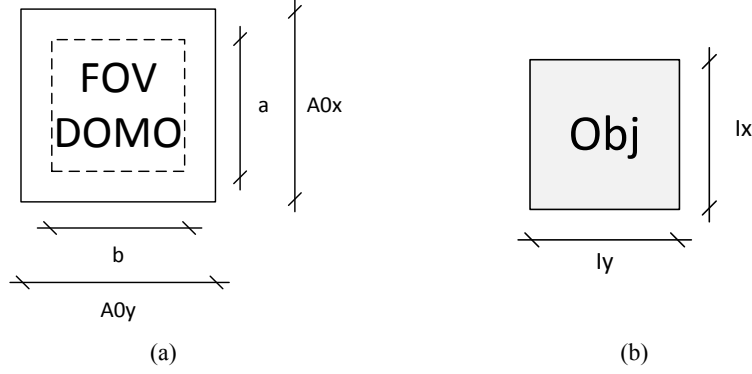


Fig. 9. Logic for zoom in

For zooming, when the surveillance area selector provides the identifier 40, it means that the image is ready to begin zooming. The dimensions of the FOV dome shown in Fig. 9 (a), and for the zoom in, the first step is to calculate the tolerances according to Eq. (9):

$$Tolerances \begin{cases} a = AO_x - 10 \\ b = AO_y - 10 \end{cases} \quad (9)$$

Next, the extracted image is increased by a factor 1.1 preserving the aspect ratio. This value, 1.1, was found experimentally is suitable for resizing after various adjustment tests. This parameter is used only to make increases for zoom in, but not affect completely the image.

The increasing is repeated until the object size is less than or equal to the resolution in either side of the dome FOV, as shown mathematically in Eq. (10):

$$(I_x \geq a) \vee (I_y \geq b), \quad (10)$$

where I_x, I_y are the object size, a, b are the object size in the scale of the dome FOV.

The corresponding pseudocode is shown below:

```
While ((Ix ≥ a) ∧ (Iy ≥ b))
{
    INx = (1.1) (Ix) ;
    INy = (1.1) (Iy) ;
    Ip (x, y) = resize (INx, INy) ;
    Ix = size.x (Ip (x, y)) ;
    Iy = size.y (Ip (x, y)) ;
}
```

5 Experiments and analysis

5.1 Results

The equipment employed for the tests is shown in Table 1. The event was presented for the IP camera.

Table 1. Equipment used

CAMERAS	Minitrax Dome Camera, Syscom	3DNR Super Night Vision Outdoor Camera, Epcom	IP Camera Vi-votek
CONVERTER	4-Channel Digital Video Recorder (DVR), Syscom		

Fig. 10 shows the detection axes of an automobile because of two of the four ROI lines. Red dots indicate where lighting change was detected.



Fig. 10. Edge Detection

When the event happen, the image processing (see Fig. 8) is responsible for searching for the global coordinates of the ROI, which indicates where the subtraction is performed. Fig. 11 illustrates that after thresholding, particle analysis and discrimination, the object of interest is delimited, showing its centroid.



Fig. 11. Presence of objects

Then, when the dome has a FOV similar to having the camera that detected the event (see Fig. 12), the extracted image is compared with the FOV to calibrate its light intensity to find the object.

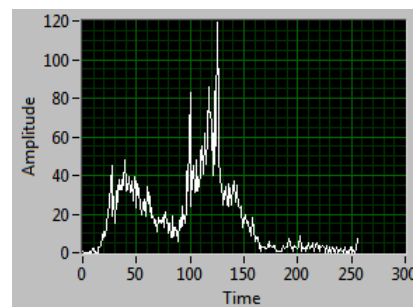


Fig. 12. Dome FOV

Fig. 13 (a) shows the extracted image and Fig. 13 (b) shows its histogram. Furthermore, after the measurement of illumination, the parameters brightness and gamma are modified, maintaining the contrast in the neutral value (45) (see Fig. 14 (a)).



(a)



(b)

Fig. 13. Image search

As shown in Fig. 14 (b), the image is very dark, because it did not change the contrast value. However, this is not a limitation, because the pattern matching algorithm succeeds in finding the object. After transformations, the histogram is modified (see Fig. 14 (c)).

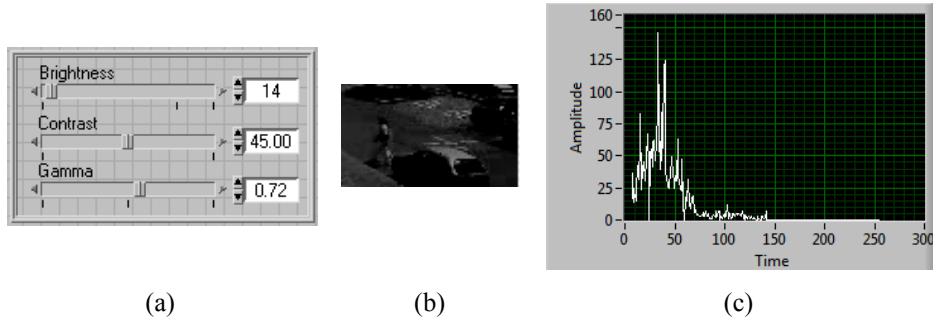


Fig. 14. Modified image search

Pattern matching continually monitors the target location and gives the image coordinates, which are compared by the dome control to locate the relevant surveillance area. Fig. 15 shows how pattern is found. This is repeated until the target is in the center of the FOV of the dome.



Fig. 15. Pattern Matching

After positioning in the center and then making zoom in as necessary, this steps will stop when some of the lengths of the pattern (vertical or horizontal) is greater than or equal to the stated tolerance.

Finally, the result is shown in Fig. 16. This image, $I_f(x, y)$, according to the block diagram of Fig. 1, can be used to perform other advanced processing or for display it in a report that is sent to to the user email account.



Fig. 16. Result

5.2 Discussion

The design of the system corresponds to a dual system that employs a pair formed by a fixed camera and a mobile one that work together, as proposed in the paper of Hsien-Chou y Wei-Yi [6].

The test was conducted in a restricted parking that complicates the algorithm having uncontrolled lighting, unlike Chu-Sing et. al work [5] which is done outside.

One advantage of the proposed system over other developed by Varcheie y Bilodeau [4], is that the present work employs RGB color to grayscale transformations, which can be a benefit for the low-performance commercial systems that provide monochromatic images. Thus, matching pattern is not dependent on color, the object can be found through inter-camera calibration.

6 Conclusion

In this paper a computational model is proposed and applied for the security of buildings, in this case parkings in restricted areas. The presence of an object is detected by a fixed camera then the command is send to the dome camera to move it to that point and zoom in to get a picture with a greater detail. This is achieved using advanced algorithms of image processing, pattern matching and Pelco-D communication protocol. The application was programmed in LabVIEW 2010. The final image can be used for other purposes that computer vision allows, such as feature analysis and optical character recognition (OCR). In this paper, user reports are generated and sent via email. Future work will be related to sharing information in a local network, and integrating other image processing techniques with the purpose of object tracking in real time for a more robust and complete system.

7 References

1. Imai, Y.; Sugiue, Y.; Hori, Y.; Masuda, S.; , "Application of A Remote Surveillance System with Mobile Phone-Enhanced User Interface," Mobile Business, 2006. ICMB '06. International Conference on , vol., no., pp.27, 26-27 June 2006.
2. Jenkins, W., Top Benefits of video surveillance, Articlesbase, Free Online Articles Directory, <http://www.articlesbase.com/customer-service-articles/top-benefits-of-video-surveillance-1233905.html>
3. Boyd, J.E.; Hunter, E.; Kelly, P.H.; Li-Cheng Tai; Phillips, C.B.; Jain, R.C.; , "MPI-Video infrastructure for dynamic environments," Multimedia Computing and Systems, 1998. Proceedings. IEEE International Conference on , vol., no., pp.249-254, 28 Jun-1 Jul 1998.
4. Varcheie, P.D.Z.; Bilodeau, G.-A.; , "Active people tracking by a PTZ camera in IP surveillance system," Robotic and Sensors Environments, 2009. ROSE 2009. IEEE International Workshop on , vol., no., pp.98-103, 6-7 Nov. 2009.
5. Chu-Sing Y.; Ren-Hao C.; Chao-Yang L.; Shou-Jen L.; , "PTZ camera based position tracking in IP-surveillance system," Sensing Technology, 2008. ICST 2008. 3rd International Conference on , vol., no., pp.142-146, Nov. 30 2008-Dec. 3 2008.
6. Hsien-Chou L.; Wei-Yi C.; , "A dual-PTZ-camera system for visual tracking of a moving target in an open area", Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on , vol.01, no., pp.440-443, 15-18 Feb. 2009.
7. Relf, C., Image Acquisition and Processing with LabVIEW, Editorial CRC Press, Estados Unidos de América, 2004.
8. Escalante, B., Apuntes de curso de Procesamiento Digital de Imágenes, 2006. <http://verona.fi-p.unam.mx/boris/teachingnotes/Capitulo4.pdf>.
9. National Instruments Corporation, NI Vision Concepts Help, National Instruments documentation. August, 2011.