

Object Tracking System using Stereo Vision and LabVIEW Algorithms

Rodica HOLONEC, Romul COPÎNDEAN, Florin DRAGAN, Valentin Dan ZAHARIA

Faculty of Electrical Engineering, Technical University of Cluj-Napoca

Abstract – An artificial system of tracking must acquire information about the environment exterior to it and to use them in order to establish a relationship between it and the objects in the respective area. So, in the surveillance system where the use of video cameras for tracking is needed for the objects in movement, the stereo vision gives the advantage of obtaining some important pieces of information regarding the area of interest. This work presents a low cost system of tracking 2D, in real time, an object by using stereo vision techniques. The application is made in the LabVIEW programming language (National Instruments) and proposes algorithms for the detection, tracking and 3D depth estimation of one object in movement within the surveillance area.

Key words – object tracking, stereo vision, LabVIEW, depth estimation,

1. INTRODUCTION

The world we live in is a tri-dimensional one. That is why it is extremely important that artificial vision system to perceive the world in three dimensions. Ordinary sensors used in video cams are only bi-dimensional. Although, by combining two or more cameras in a stereoscopic configuration, the tri-dimensional aspects of the world may be quantified. The movement represents on its turn an important source of information in a sequence of images. Combining the data extracted from tri-dimensional space using the stereo vision with those regarding movement in bi-dimensional plan allows obtaining some performing system of tracking objects.

A common method for extracting depth information from 2D images is to acquire a pair of such images using two cameras displaced from each other by a known distance. The result will be a density stereo map.

Object tracking is an operation by which the object is maintained in the field of view of video camera. There are different tracking methods that can use a stationary or fixed camera or a moving camera [1]. Pan and tilt type cameras differ from a fixed camera because they can move left and right (pan) or up and down (tilt) and in consequence they are able to perform the tracking operation.

A tracking system may use one or more objects and it may contain one or more video cameras. As consequence, a stereo vision system can be successfully used in a computer vision application for tracking and depth information extraction.

By analogy with the human visual system, the binocular stereo vision uses a pair of images of the same scene, but it is captured in different positions. The aim of binocular stereo vision algorithms is to calculate the position of the object within the frame of the scene and in particular the depth of the elements contained in the two images. So, for that at least four stages are needed:

- Acquisition of the images – the two images of the same scene are captures in the same time by two different video cameras
- The calibration of the image capture system – this stage consists in determining the internal and external parameters of the geometrical model of the image capture system. In most of the cases it is adopted the model of pinhole camera for which the geometric shaping of the images consists in a projection in perspective.
- Establishing the correspondence among pixels – it refers to re-finding in the two images of the pairs of pixels corresponding to the projection of the same element in the scene and compute the disparity (distance in pixels) of these points. In this stage the correspondence problem occurs. As two cameras are used it is necessary that both of them have to work at once and both images should contain same matching points.
- 3D reconstruction– consists in calculating for each pixel the position in space of the point projected on that pixel. This stage, named triangulation, requires knowing the correspondences resulted in the previous stage.

The principle of binocular stereo vision [2] is illustrated in Fig. 1, where b is the baseline, f is the focal length of both cameras, X_A is the X-axis of a camera, Z_A is the optical axis of a camera, $P(X,Y,Z)$ is a real-world point, u_L is the projection of the real-world point P in an image acquired by the left camera and u_R is the projection of the real-world point P in an image acquired by the right camera.

Since the two cameras are separated by distance “ b ”, both cameras view the same real-world point P in a different location on the 2-dimensional images acquired. The X-coordinates of points u_L and u_R are given by relations (1) and (2):

$$u_L = f * \frac{x}{z} \quad (1)$$

$$u_R = f * \frac{x-b}{z} \quad (2)$$

The distance between the two projected points is known as “disparity” and it is calculated using relation (3).

$$disparity = u_L - u_R = f * \frac{b}{z} \quad (3)$$

The depth information, which is the distance between real-world point “ P ” and the stereo vision system, is given by the relation (4).

$$depth = f * \frac{b}{disparity} \quad (4)$$

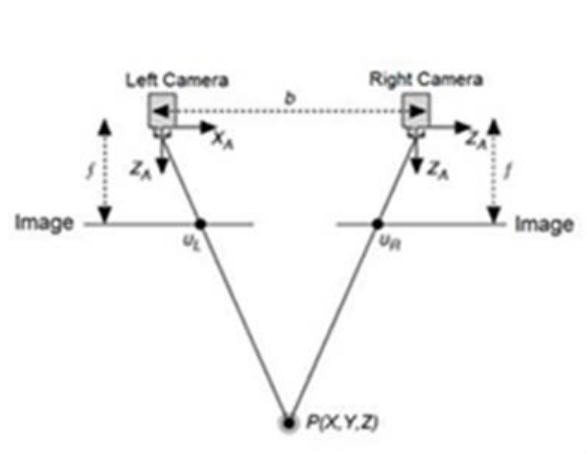


Fig. 1 Stereo Vision System principle [2]

The stereo vision systems are used in a lot of tracking applications, but the algorithms assume that the two images are taken at exactly at the same time [3]. That means that expensive stereo vision systems must be used. In the case of low cost stereo vision systems there are many algorithms for motion detection and object tracking [4]. The most simple is the one in which object moves against the uniform background [5]. The moving object, which is the blob to be detected, is isolated easily by the threshold technique. In noisy environments, the threshold level can be detected adaptively based on the stochastic

analysis. The color of the object may be a determining feature in obtaining the threshold level.

Once the object detected by techniques of image processing, one may determine also the coordinates of the mass center of it in the two images. Based on those pieces of information the tracking algorithm may lay on the computation of average of object's coordinates and keeping the target on the middle point between the two cameras [6].

Graphical programming language LabVIEW allows the realization of some complex, performing applications, in a simple and elegant manner, having in the same time a high flexibility and a relatively low cost. Programs in LabVIEW are based on the modularization and arborescent hierarchy concepts. A virtual tool created may be used as a self-existing application or may be used as a virtual sub-tool in order to realize another virtual instrument. Practically sub-IV has the role of a sub-routine in the main program.

The graphic interface of such an automation system must be simple and easy to use. From this point of view also, LabVIEW is an excellent programming language, offering by its libraries all the components needed by such an application. The stereo vision features in the LabVIEW Vision Development Module helps programmers to implement 3D vision systems in complex tracking applications.

In this paper, the LabVIEW-based proposed system is based on the fusion of stereo data with and visual motion data in order to realize a reliable object tracking system.

2. THE PROPOSED SYSTEM STRUCURE

The implemented system uses hardware and software components, observing the system flow chart from Fig. 2.

2.1. Hardware description

From the hardware point of view in the configuration of the system (Fig.3) are contained the following elements:

- Two web cameras with automate focusing, VX-800-Microsoft, CMOS sensor, VGA, 640x430, USB
- The pan/tilt camera support containing two servomotors: MCN-SEV-03 for the movement of the cameras in horizontal plan and MNC-SEV-06 for the movement of the cameras in vertical plan. The cameras are positioned in parallel on a 6.5 cm distance on the pan/tilt support.
- The Mydaq data acquisition board (National Instruments) for the control of the servomotors on the signal generation AO1 and AO2 channels.
- The voltage source used for feeding the two servomotors
- The visualization area and the object to be tracked

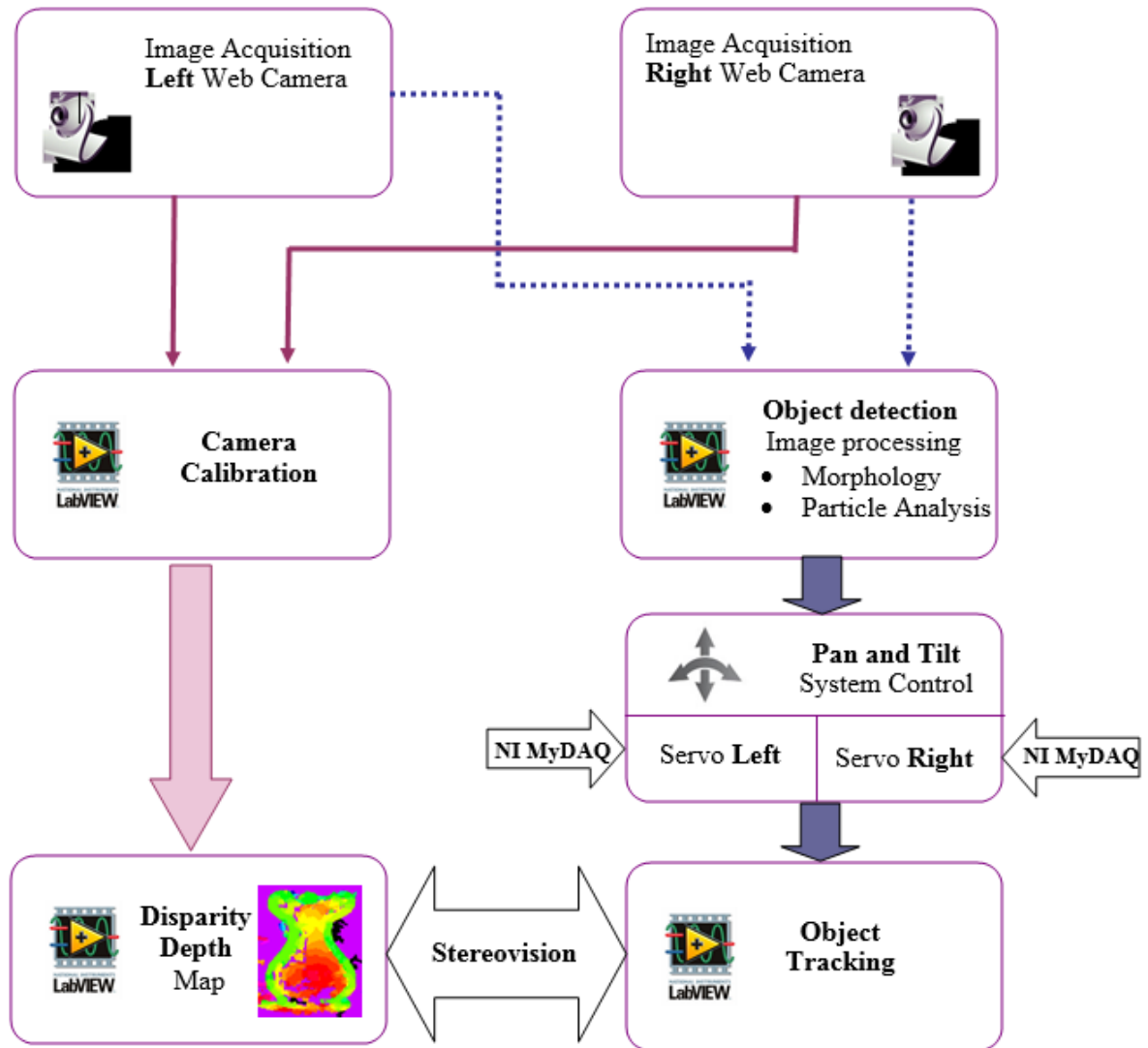


Fig. 2 The system flow chart

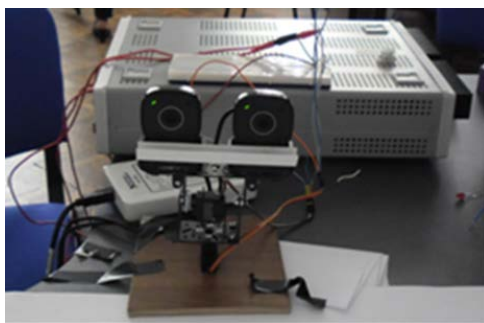


Fig. 3 The hardware setup

2.2. Testing the system for the depth map computation

The two USB web cameras (left, right) have been configured for a continuous acquisition of the image.

The hypothesis made for the ideal stereo vision system may not be done for the applications of stereo vision in the real world. Even the best cameras and lenses will introduce a certain distortion level of the acquired image, so in order to compensate these

distortions the stereo vision system needs to be calibrated. The calibration process supposes the use of a grid of calibration by the help of which pairs of images are acquired from different angles, in order to calculate the distortion of the image as well as the exact spatial relationship between the two cameras. The calibration grill is included in the Vision development module. The obtained calibration images are displayed in Fig. 4.

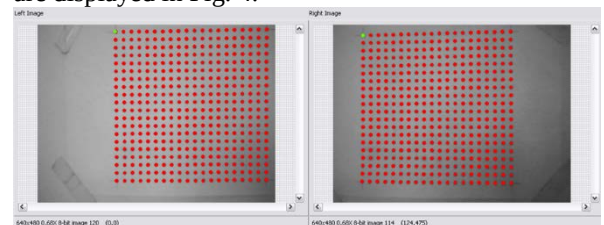


Fig. 4 Stereo Vision System Calibration

Disparity information provides relative depth information. The disparity computation is made by using the semi-block-matching algorithm from NI Vision [7] because it provides more details and works

in regions with little or no apparent texture which is the case of the proposed system.

The depth images from Fig. 5 were obtained during the experiments.

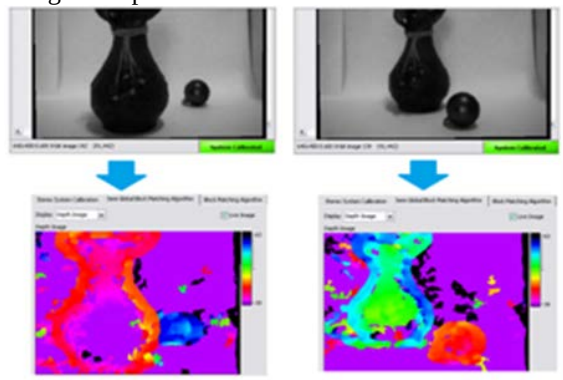


Fig. 5 The depth images

The obtained results are acceptable in the conditions of a slow movement of the object in the visual field of the cameras.

3. THE APPLICATION

The application is meant to track only one object in movement in the visual field of the cameras. The chosen object has the following features: it is round, with diameter $D \approx 5\text{cm}$ and it is dark red colored. The region of interest has a low noise level background. The images coming from the system of stereo cameras are simultaneously used in the tracking process and in that of displaying the depth map.

3.1. The sequence of detecting the object

For detecting the object, the two cameras images are processed by:

3.1.1. Applying a color threshold to the three plans of RGB image

By the use of LabVIEW National Instruments Vision Assistant, there were experimentally established the RGB optimum thresholds levels for the studied object: $R=[2, 255]$; $G=[0, 255]$; $B=[0, 30]$. Consequently to this processing the binary image in Fig. 6 has been obtained.

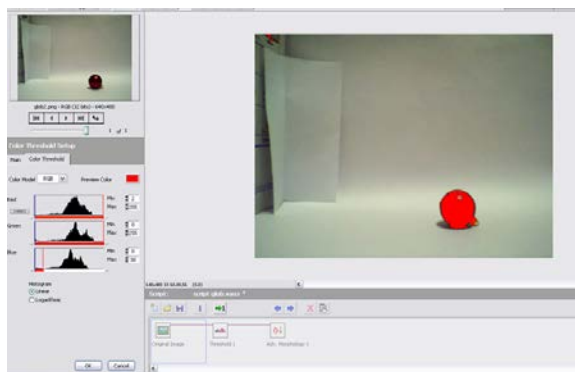


Fig. 6 Application of RGB threshold

3.1.2. Applying morphological operations

Due to the luminosity difference in the images acquired, it was necessary to morphological process them. So, in a first step the operator “Fill holes.vi” was used, followed by that named “Remove small object.vi” in 5 iterations. Consequently to this processing the binary image obtained (Fig. 7) contains only the object of interest.

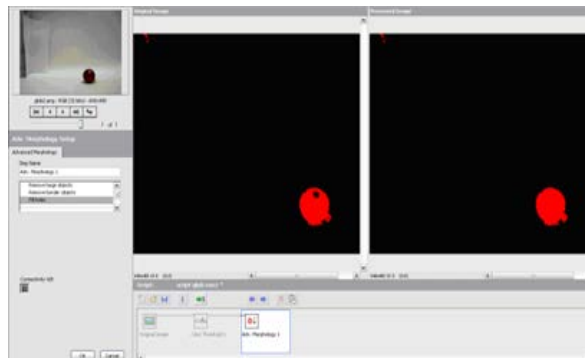


Fig.7 Image obtained consequently to morphological processing

3.1.3. Detecting of the object and its center of mass

An operator type “Particle Analysis.vi” was applied in order to determine the mass center of the object (x,y) as in Fig. 8. These pieces of information will be used in the tracking sequence.

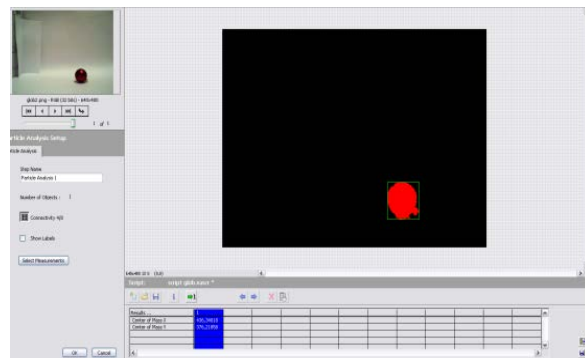


Fig.8 The effect of „Particle Analysis” operator

3.2. Tracking sequence

The coordinates of the centre of the object (x_1, y_1) and respectively (x_2, y_2) obtained from the two cameras (left, right respectively) are transformed relatively to the centre of the image of sizes 640×480 according to the relations (5)-(8):

$$x' = \frac{x_1 + x_2}{2} \quad (5)$$

$$y' = \frac{y_1 + y_2}{2} \quad (6)$$

$$x'' = \frac{x' - 320}{320} \quad (7)$$

$$y'' = \frac{y - 240}{240} \quad (8)$$

These computations are made because the tracking algorithm is based on the principle of keeping the centre of the object opposite to the middle point between the cameras [6].

So, the coordinates (x'' , y'') will take values in the interval $[-1,1]$; values that may be easily used in the process of tracking by controlling the servomotors. These will be directed left-right, up-down, by PWM ($f=50$ Hz) signals with a duty cycle situated in the domain of $[3,11]\%$. (Fig. 9)

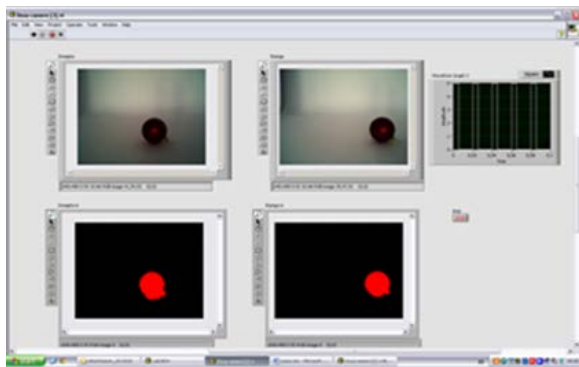


Fig. 9 The servomotors control signal

So, any movement of the object in the visual field of the cameras is tracked by the pan/tilt system (Fig. 10).



Fig. 10. Tracking the object by the pan/tilt camera system

In parallel with the tracking application it runs also the application for obtaining the depth image by means of which there are obtained pieces of information regarding the distance from the object to the cameras within the stereo vision system (Fig. 11).

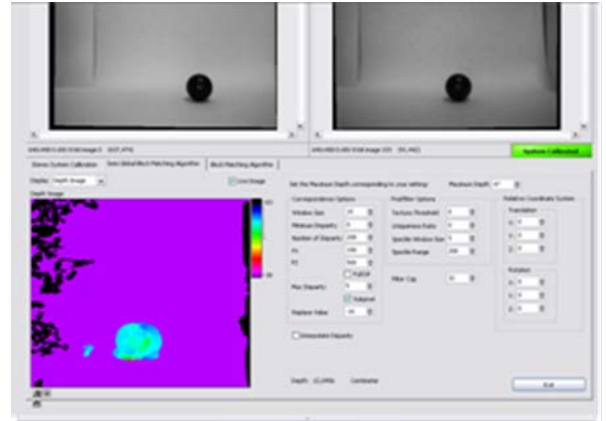


Fig. 11. Depth map of the tracked object.

3.3. Results

The performance of our tracking method was evaluated by using real time video recording of the system behavior.

Experimental results presented in Fig. 12 demonstrate the efficiency of our method.

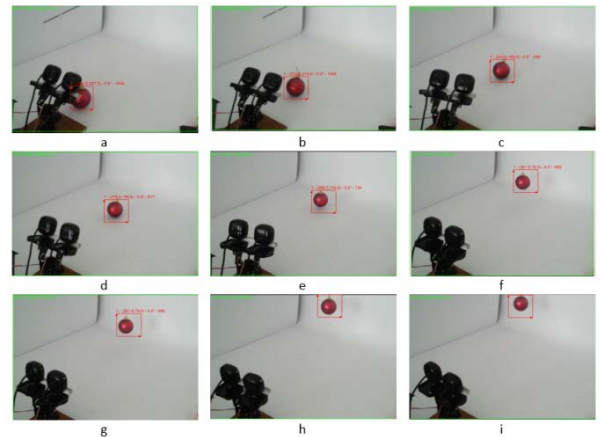


Fig.12. Experimental results from real time tracking process of the moving object

4. CONCLUSION

The main objective of the paper was to implement a low cost stereo vision tracking system using LabVIEW. The object tracking algorithm is made by first, detecting the object in both web-cameras images and then by maintaining the middle point between the cameras in the centre of the image. In the same time the acquired images are used to estimate the distance between the object and stereo vision system. Accurate results were obtained in a range of 39 to 60 cm. The proposed system can be improved in the next directions:

- object detection can be made by using other different features than color like shape, size, texture and according to this, different tracking algorithms can be implemented

-by using a Field-Programmable Gate Arrays (FPGA) system configuration based on LabVIEW FPGA Module that is very well suitable to stereo vision applications.

REFERENCES

1. Lee Chong Wan Patrick Sebastian, Yap Vooi Voon, "Stereo Vision Tracking System", International Conference on Future Computer and Communication, 2009
2. D. Nair, "A guide to stereovision and 3D imaging: NASA tech briefs." <http://www.techbriefs.com/component/content/article/14925?start=1>, 2012.
3. K Nickels, C Divin, J Frederick, J Graham, L Powell : "Design of a low-power motion tracking system", Proceedings of ICAR 2003, The 11th International Conference on Advanced Robotics Coimbra, Portugal, June 30 - July 3, 2003
4. Sahil S.Thakare, Rupesh P. Arbal, Makarand R. Shahade: "Artificial Intelligence with Stereo Vision Algorithms and its Methods" International Conference on Recent Trends in Information Technology and Computer Science (IRCTITCS) 2011, Proceedings published in International Journal of Computer Applications® (IJCA)1
5. Sang Wook Lee, K. Wohn: Tracking moving by a mobile camera", Technical Reports, 1988
6. Andrew Kirillov, "Making a step to stereo vision", 2009 [online], http://www.aforgenet.com/articles/step_to_stereo_vision/
7. S Birchfield, C Tomasi: "Depth discontinuities by pixel-to-pixel stereo", International Journal of Computer Vision, 1999 - Springer
8. RYD Xu, JG Allen, JS Jin: "Robust real- time tracking of non-rigid objects", Proceedings of the Pan-Sydney area workshop on Visual information processing, p.95-98, June 01, 2004
9. Karan Gupta, Anjali V. Kulkarni: "Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching", Advances in Computer and Information Sciences and Engineering, 2008, pp 245-250

Rodica HOLONEC
Romul COPÎNDEAN
Florin DRAGAN
Valentin Dan ZAHARIA

Electrical Engineering and Measurements Department
Faculty of Electrical Engineering
Technical University of Cluj-Napoca
Str. Memorandumului nr. 28, 400114 Cluj-Napoca