A SYSTEM FOR TRIDIMENSIONAL IMAGES FROM TWO DIMENSIONAL ONES USING A FOCUSING AND DEFOCUSING VISION SYSTEM

Modesto G. Medina-Meléndrez
Departamento de Ing. Eléctrica-Electrónica, Instituto Tecnológico de Culiacán, Sinaloa

David Báez-López, Liliana Díaz-Olavarrieta, J. Rodríguez-Asomoza and L. Guerrero-Ojeda
Departamento de Ingeniería Electrónica, Universidad de las Américas Puebla, Cholula, Puebla
México

Keywords: Image processing, Depth information, Focusing and defocusing.

Abstract: Machine vision has been made easier by the development of computer systems capable of processing information at high speeds and by inexpensive camera-computer systems. A Camera-Computer system called SIVEDI was developed based on the shape from focusing (SFF) and shape from defocusing (SFD) techniques. The SIVEDI system has as entries the images captured by the camera, the number of steps of the focusing mechanism, and user specifications. The focusing mechanism used is the internal one of the camera and a computer-camera interface was designed to control it. An equation is introduced to obtain the measure of relative defocusing among many images. A 3D image is produced as the result of the system. SIVEDI consist of a number of modules, each one implementing a step in the system model. The modules are independent of each other and could be easily modified to improve the system. As SIVEDI was developed in a MATLAB environment it can be used in any computer with this software installed on it. These characteristics allow users to access intermediate results and to control the system internal parameters. Results show that the 3D image reconstruction has an acceptable quality.

1 INTRODUCTION

The 3D recovery information of a scene using bidimensional images is a fundamental problem in machine vision systems. Researchers in this area have created a variety of sensors and algorithms that can be classified into two categories: active and passive. Active techniques, as point triangulation and laser radar, produce very precise depth maps and have been used in many industrial applications. But, when the scene can not be controlled, as it is the case of distant objects in big scale scenes, active techniques have proved to be unpractical. As a consequence passive techniques are used because they offer better results. Passive techniques, as stereo and movement structures, depend upon algorithms that establish correspondence among two or more images. The process to determine correspondence is computationally expensive. Other passive techniques are based in the focus analysis of many images. Shape From Focus (SFF) uses a sequence of images captured with small changes in the internal parameters of the camera optics. For each pixel the parameters of the focus that maximizes the images contrast are determined. These focusing parameters can be used to calculate the depth of the analyzed point in the scene and to produce very precise measurements.

In contrast, Shape From Defocus (SFD) uses only a few images with different optics parameters. The relative defocusing of the images can be used to determine depths. The focus level of the images can be varied changing the optics parameters of the lens, moving the camera image sensor with respect of the lenses, or changing the aperture size. SFD does not present correspondence problem and missing parts problem in the analysis. These advantages make SFD an attractive method to determine depths. The objective of this work is to implement a camera-computer system to estimate object depths in a 3D scene using focus analysis in bidimensional imagines. In machine vision systems a typical camera, used in passive techniques, carries many noise sources, as optics noise and electronic

G. Medina-Meléndrez M., Báez-López D., Díaz-Olavarrieta L., Rodríguez-Asomoza J. and Guerrero-Ojeda L. (2005).
A SYSTEM FOR TRIDIMENSIONAL IMAGES FROM TWO DIMENSIONAL ONES USING A FOCUSING AND DEFOCUSING VISION SYSTEM.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Robotics and Automation*, pages 475-478
DOI: 10.5220/0001190004750478

noise. As a consequence, errors in the estimation are inevitable, but they are minimized by establishing two noise thresholds. An important objective is to implement a set of subroutines to execute the SFF and SFD techniques. Also a user interface has been developed to implement these techniques. The final system was verified with a series of experiments. The obtained results present a good precision in the estimation of depth when the objects contain high contrast.

2 SIVEDI

A system called SIVEDI (acronym for the Spanish name Sistema de Visión por Enfoque y Desenfoque de Imágenes) was developed. SIVEDI can be used to verify the SFF and SFD theories. The SIVEDI system is divided into a set of modules, in such a way that it can be easily modified to implement different estimation depth techniques which use image focusing analysis.

The camera-computer system and the set of subroutines that constitute SIVEDI are described in the next sections.

2.1 Camera computer system

SIVEDI was built in the Electronics Laboratory of the Universidad de las Américas, Puebla, México. The image sensor is a camera Sony DCR-TRV120, the video signal is digitized by a Video Frame Grabber Board HRT-218-5, the algorithms are implemented in an Intel Pentium PC. The camera focusing is executed using the LANC protocol of the video-cameras SONY. An interface between the serial port of the computer and the LANC plug of the camera was implemented to execute the different focusing commands. The Video Frame Grabber captures monochromatic images with 640X80 pixels and 8 bits per pixel. A windows interface was implemented to execute the analysis in a direct way. The implemented subroutines execute: the control commands of the camera, the control of the Video Frame Grabber Board, and provide a set of modules to perform the SFF and SFD analysis.

2.1.1 ANC Interface

The computer-camera interface, which controls the camera commands, was implemented using a PIC16F84A.

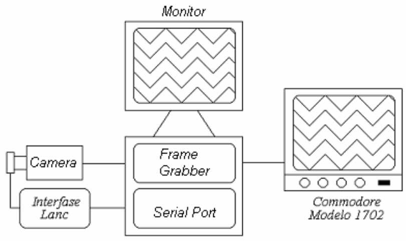


Figure 1: Implemented System

The PIC was programmed such a way that when the computer sends a command (corresponding to the LANC protocol) it synchronizes itself with the camera and resends the command. The use of an interface was needed because it is not possible to synchronize the serial communication port LANC of the camera with the serial communication port of the computer. In addition, the computer indicates the camera how many times it must execute the command.

2.2 Implemented Software

For the SFF and SFD applications, it was needed to design a set of subroutines to find error value thresholds. Also, calibration was needed the system to establish an unequivocal relation between the focus step and the object depths for SFF, as a relation between the relative defocus and the object depths for SFD. In the following subsections are described the implemented algorithms.

2.2.1 Noise Thresholds T1 and T2

In regions of low contrast in an image the methods SFF and SFD do not work properly, so it is needed to define a minimum value of focus measurement or minimum value of contrast. In the case that measurements are not over this value any obtained estimation is rejected. The obtained thresholds are T1 for SFF and T2 for SFD. T1 is the minimum value a focus measure must have to be accepted. This value was obtained with the following procedure: A low contrast object was placed in a known position from the camera and the focus measure was obtained in a 16X16 pixels region. The used focus measure was the Laplacian Energy. The Laplacian is a high-pass filter, the opposite effect of defocusing, that is, a low-pass filter. The measurements were performed over many regions in the image and many positions of the object. T1 is defined as the media of the measurements plus three times the standard deviation. T2 is a threshold which is obtained in a similar way as T1, but in place of using the focus measure of a unique image, it is used the sum of the focus measurements of the captured images for each region and each positions of the object. T2 is the media of the measurements plus three times the standard deviation.

2.2.2 SFF Implementation

A set of subroutines was developed to implement SFF: an autofocus algorithm that find the focus step where the image is in focus, a method to deduce the equation that relates an obtained focus step with the object depth, and the SFF algorithm to estimate depths of every objects in a 3D scene.

The autofocus algorithm performs a search of the best focused image among a sequence of images (to reduce noise, each image is an average of 4 captured images with the same optics). Once the best focused image is obtained among the image's sequence a quadratic interpolation is performed over the focus steps of the rear image, the focused image and the previous image. The interpolation maximum is found and the focus step which corresponds to this maximum is defined as the focus step where the image is in focus.

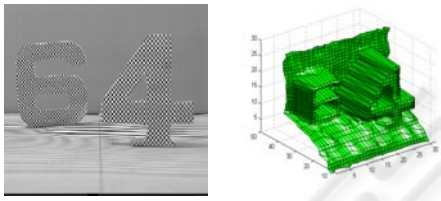
The procedure to find the relation between the focus step and the object depth is described next. First, the object is placed in a known position and the autofocus is executed, this produces the focus step where the best focus measure is obtained, and the match focus step-object depth is recorded. This procedure is repeated for many object positions keeping the step increment uniform. Once every match is obtained, a polynomial equation is adjusted to these points. The obtained polynomial is used in the equation that must be used to calculate the searched correspondence. The final polynomial, with a 300 step zoom from the ZoomWide is

$$dist = -2.82\times10^{-8}X^{5} + 1.34\times10^{-5}X^{4} - 0.0025X^{3} + 0.23X^{2} - 10.88X + 234.73 \quad (1)$$

The SFF algorithm produces estimations in 16x16 pixel regions. Each 16x16 pixel region in an image corresponds to the same 16x16 pixel region in each captured image. This technique consists in finding the focus step in which the maximum focus measure is obtained for each 16x16 pixel region. After capturing the complete image sequence the focus measure is performed over every 16x16 pixel region in every image. Then, the maximum among focus measurements in the image sequence for each region is searched.

To improve the depth definition a quadratic interpolation is performed over the focus measurements of the focused image, the following image, and the previous one. This interpolation is executed over each vision region, obtaining an estimation of focus steps that corresponds to different depths.

To calculate the depth of all the regions in the scene equation 1 is used, where X is the Focus Step. An experimental test using SFF is shown in figure 2.



(a) (b)

Figure 2: a) Analyzed scene and b) Obtained estimation using SFF

2.2.3 SFD Implementation

A few images are needed to be captured in order to determine a 3D scene using SFD technique. In previous works just two images of the scene have been used. As a consequence, the range of detected depths has been limited. In this work a method is proposed to determine the quantity of images to be captured as a consequence of the range of depth to be estimated. First, for a specific object placement, a focus measure curve is analyzed looking for the main lobule. The step increment between captures is established as the half wide of the main lobule of the focus measure curve. As a result of knowing the upper and lower limits of the depth measurements, the quantity of images to be captured can be established. This capture interval is needed because the main lobule must be sampled at least at two points.

The following equation is proposed to be used in order to get the relative defocus of the captured images.

$$\frac{M}{P} = \frac{\sum_{j=1}^{k} Gj[k-2j+1]}{\sum_{j=1}^{k} Gj} \quad (2)$$

where k is the quantity of captured images, Gj is the focus measure (Laplacian) over each one of the k images. M/P is the relative defocusing that observes a linear behaviour in the analysis region, so for each obtained M/P value there exists just one corresponding depth. To obtain the correspondence between M/P and Depth it is needed to use an equation in which the known value is M/P and the calculated value is the distance of the object. The procedure to obtain such an equation is described next. First, an object is placed in a known position (depth) from the camera, M/P is calculated and the match M/P-Depth is recorded. This is repeated for a variety of object positions, getting a table of M/P-Depth matches. Then, a curve adjustment is performed getting a polynomial which is the searched equation. In this work a 300 steps zoom is used from the ZoomWide. The obtained equation is

$$dist = 1.609\times10^{-4}Y^4 + 6.9\times10^{-3}Y^3 + 4.52\times10^{-2}Y^2 + 2.621\times10^{-1}Y + 2.5881 \quad (3)$$

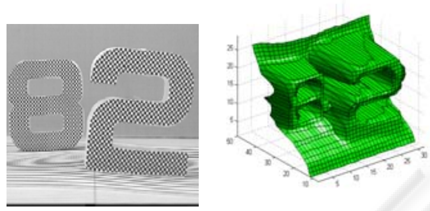
The depth estimation is realized over 16x16 pixel regions. First, k images are captured, M/P is calculated for each region in the scene and equation 3 is used to estimate the depth of each region where Y is M/P value. Figure 3 shows the result of applying this method over a scene.

2.2.4 User Interfase

SIVEDI has a user interface developed with MATLAB subroutines. The user has complete access and control over the system to specify the analysis to perform. When the user interface is executed the main window appears and shows if the system is calibrated or not. In order to be calibrated the system must contain the equation Focus Step-Depth to be applied in SFF and the equation M/P-Depth to be applied in SFD. If the user wishes to calibrate the system, a window appears that guides the user in this task. Once the system is calibrated, the main window shows the options to perform the analysis SFF or SFD. For any analysis it is possible to record the obtained estimation. The default parameters for the analysis are: the region size of analysis (16x16 pixels), the quantity of captured images for SFF (25 images), the increment in focus steps at which the images are captured for SFF, the quantity of captured imaged for SFD (7 images), the focus steps at which the images are captured for SFD, the focus measure (Laplacian), and the estimation range (7 to 50 centimetres). If any of these parameters should be modified, it is done directly in the implemented subroutines in m files.

3 CONCLUSIONS

The execution time of SFF is reduced, because the quantity of captured images is reduced and an interpolation is used among three images to obtain the maximum measure. The noise is minimized using the average of four images for the SFF analysis. The implemented platform prevents the use of two images in the SFD analysis, as a consequence a method to determine the needed captured images to perform the SFD analysis is proposed.



(a) (b)

Figure 3: a) Analyzed scene and b) obtained estimation using SFD

Also, the use of an equation (M/P) to calculate the relative defocus among the captured images is proposed. This measure is related with the depth of the object. The noise is reduced using the average of eight images for the SFD analysis.

The obtained results show that SFD is less sensitive to noise than SFF, but is more sensitive to spectral content over the analyzed regions. It is harder to calibrate SFD than SFF. The execution time is lower for SFD than for SFF. Images must contain high contrast in order to both techniques work properly.

REFERENCES

Hecht E., 2000. “Optica”, Addison Wesley, Madrid, Third Edition.

Jenn K. T1997. “Analysis and Application of Autofocusing and Three-Dimensional Shape Recovery Techniques based on Image Focus and Defocus”, PhD Thesis, SUNY in Stony Brook,.

Nayar S.K., M. Watanabe and M. Noguchi 1995. “Real-time focus range sensor”, Intl. Conference on Computer Vision, pp. 995-1001, June.

Subbarao M., “Spatial-Domain Convolution Deconvolution Transform”, Technique Report No. 91.07.03, Computer Vision Laboratory, State University of New York, Stony Brook, NY 11794-2350.

Subbarao M, T. Choi, y A. Nikzad 1992. “Focusing Techniques”, SPIE Conference, OE/TECHNOLOGY `92, Vol. 1823, pp 163-174, Boston, MA.

Subbarao M. y M.C. Lu, 1994. “Computer Modeling and Simulation of Camera Defocus”, Machine Vision and Applications, pp 277-289.

Subbarao M., y J.K. Tyan, “The Optimal Focus Measure for Passive Autofocusing and Depth-from-Focus”, SPIE Conference on Videometrics IV, Vol. 2598, Philadelphia PA, pp. 89-99, October 1999.