Optical-Ultrasound Data Processing for Hybrid 3D Vision on Air

Jose R. Llata, Carlos Torre-Ferrero, Esther G. Sarabia, Sandra Robla
Electronic Technology and Control Systems Department
University of Cantabria
Avda. De los Castros s/n. 39005, Santander, Cantabria, Spain
SPAIN
llata@teisa.unican.es http://www.teisa.unican.es/~llata

Abstract: - This work explores an approach for obtaining three-dimensional scene representation by using optical and ultrasonic information. It combines a set of three-dimensional data obtained by air-transmission ultrasonic sensors, using the echo-pulse method, with a set of three-dimensional data captured by stereo vision or active triangulation techniques. Both data sets are 3D point clouds which are fused in order to get a hybrid depth map with three-dimensional information about the scene of interest. Because hybrid images contain not only visual information but also ultrasonic information they can be useful for scenarios with poor illumination, with special optical properties or even with transparent elements.

Key-Words: - 3D artificial vision, 3D ultrasonic vision on air, image processing, hybrid images

1 Introduction

At the end of eighties and beginning of nineties, on parallel with the electronics and informatics progress, using ultrasounds in medical applications became very common. At same time, ultrasounds were used in control systems and automatic systems as presence-distance sensors, also in materials testing, quality control and new applications, such as object recognition. Numerous researchers [1], [2], [3], [4], [5], [6] were attracted by the construction simplicity and its low cost and they tried to explore all the possible capacities of this type of sensors for object recognition applications, an fundamental activity in robotics and automation fields. More recently, some authors begin to talk about acoustic vision [7] and three-dimensional ultrasonic images [8], [9] [10], [11], [12], [13] starting to introduce some acoustic image processing with artificial intelligence techniques, especially fuzzy logic [14], [15]. Nowadays, however, it is not usual to fuse the information provided by both technologies (ultrasound and optics) in order to enhance the three-dimensional information, obtaining new 3D visual data.

So, this work explores possibilities for three-dimensional surface reconstruction by using a combination of artificial vision and ultrasonic vision on air. This could be of special interest for applications that require (due to object's transparency, smoke presence, lack of illumination, etc.) to use both technologies to retrieve the scene.

Using ultrasonic arrays and a scan of the scene it is possible to obtain range ultra-sonic images by means of triangulation techniques. In some situations it is also possible to calculate coordinates of the contact point (echo with the scene), obtaining an ultrasonic image with a better approximation (but still with a high uncertainty). On the other hand, information about the scene will be captured by using artificial vision cameras. This could be done by stereoscopic vision or by active triangulation.

2 3D Optical Visual Information

In order to obtain visual information from the scene, a 3D vision system has been developed. It is able to obtain three-dimensional data by using both stereo configuration methods and active triangulation methods. As it is shown in figure 1, two CCD cameras are arranged fixed in opposite sides of the scene while a linear guide shifts a laser projector that sweeps along the objects. Subsequently, the images acquired by both cameras are processed by a computer in order to make a 3D reconstruction.

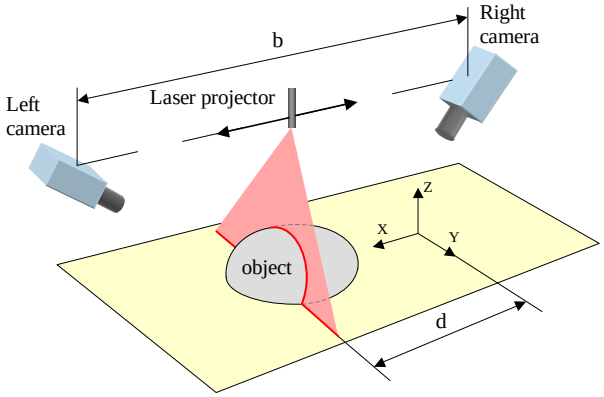

Fig. 1. Three-dimensional artificial vision system arrangement.

As can be seen in figure 1, the cameras are separated by a distance b, in such a way that the disparity produced between their respective images can be used to recover 3D data as a stereo configuration does. In addition, since it is used a laser plane, whose position and orientation are known by calibration, those points that are only seen by either one or other camera can be recovered by active triangulation.

Accordingly, 3D points can be obtained by two reconstruction methods: stereo triangulation (ST) and active triangulation (AT). Hence, there will be three sets of points (ST, AT with left camera and AT with right camera) that will be referred to the same coordinate system, as it is shown in figure 1.

2.1 Calibration

In order to recover 3D data will be necessary to accomplish a complete system calibration. Since the cameras are fixed elements whereas the laser plane is mobile, both will be calibrated apart though taking into account that the results are consistent between them. For this purpose, the same templates are used to calibrate both laser and cameras.
On the one hand, the software we have developed allows performing the camera calibration, also regarding lens distortions, in a semiautomatic way. On the other hand, the laser plane must be parallel to the YZ plane (see Fig.1). With this aim, a cross-line projection head is used to adjust and align the laser projection with the templates.

2.2 Three-Dimensional Reconstruction

When the system calibration has been completed, two projection matrices, Ml (relative to the left camera) and Mr (relative to the right camera), will be available.
In the following subsections, the two methods that have been used to reconstruct the scene are reviewed [16].

2.2.1 Stereo triangulation case

First of all, the corresponding points between a pair of images (stereo pair) have to be obtained. This correspondence is carried out by means of epipolar lines; therefore, for each image point (u_l , v_l) there will be a corresponding point in the other camera (u_r , v_r). Hence, the 3D point searched for will be calculated by equation (1), where m_{kl} and m_{kr} are the k^{th} rows of the matrix M_l (left camera) and M_r (right cam-era) respectively.

$$\begin{bmatrix} u_l \cdot m_{3l}^T - m_{1l}^T \\ v_l \cdot m_{3l}^T - m_{2l}^T \\ u_r \cdot m_{3r}^T - m_{1r}^T \\ v_r \cdot m_{3r}^T - m_{2r}^T \end{bmatrix} \cdot \tilde{P} = A \cdot \tilde{P} = 0 \quad (1)$$

being

$$\tilde{P} = \left[\tilde{X},\ \tilde{Y},\ \tilde{Z},\ \alpha\right]^T = \left[\alpha \cdot X,\ \alpha \cdot Y,\ \alpha \cdot Z,\ \alpha\right]^T \quad (2)$$

The solution to this equation will be the 3D point that minimizes. This is given by the eigenvector of least eigenvalue of $(A^T \cdot A)$.

2.2.2 Stereo triangulation case

As it can be inferred from the figure 1, the laser plane equation has to be x=d. Therefore, the unknowns to be solved are: y, z, α. Accordingly, in order to calculate the 3D point by AT, the following expression, valid for both cameras, has to be used:

$$\begin{bmatrix} -m_{12} & -m_{13} & u \\ -m_{22} & -m_{23} & v \\ -m_{32} & -m_{33} & 1 \end{bmatrix} \cdot \begin{bmatrix} y \\ z \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11} m_{14} \\ m_{21} m_{24} \\ m_{31} m_{34} \end{bmatrix} \cdot \begin{bmatrix} d \\ 1 \end{bmatrix} \Rightarrow A \cdot c = B \cdot D \quad (3)$$

Solving for the coordinate vector *c* results in

$$c = A^{-1} \cdot (B \cdot D) \quad (4)$$

The matrix A is formed from the second and third columns of M, with opposite sign, and from the coordinates of the image point (u,v); whereas the matrix B (3x2) is formed from the first and fourth columns of M.
The image points are obtained by an algorithm that calculates, with sub-pixel precision, the maxima of the laser-stripe intensity profiles. The peak location is achieved by applying the Blais-Rioux filter as is indicated in [17].

3 3D Ultrasound Information

3.1 Ultrasonic Emission and reception

The whole system is based on travel time measurement for ultrasonic signals when they are used with the echo-signal method
From [18] and [19] it can be seen that for a flat and disk-shaped ultrasonic emitter such of figure 2, the pressure field generated for a point b, is given by:

$$\mathbf{p}(r,\theta,t) = j \cdot \frac{\rho_0 \cdot c \cdot k \cdot U_0}{2 \cdot \pi} \cdot \int_s \frac{e^{j(wt - k \cdot r')}}{r'} \cdot ds \quad (5)$$

where c is the wave velocity, k is the wave-number, t is the time in seconds, w is the sensor frequency and ρ_o is the air density at 20 Celsius degrees

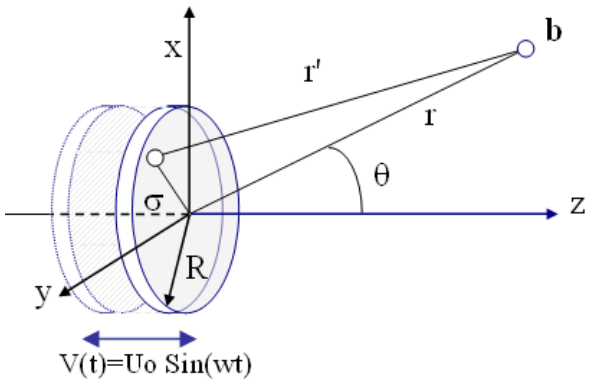

Fig. 2. Disk-Shaped ultrasonic emitter

It can be simplified to obtain the pressure at any point on the axis, in this way:

$$P = 2\rho cUo\left|Sin\left[\frac{kr}{2}\left(\sqrt{1+\left(\frac{R}{r}\right)^2}-1\right)\right]\right| \quad (6)$$

Expression (6) permits the definition of two well-known zones: The near field (Fresnel zone) and the far field (Fraunhofer zone), see (Krautkramer and Krautkramer, 1990). In the far field zone, the wave front tends to take a curved shape inside a cone, defined by the directional term:

$$\left[\frac{J_1(k\cdot R\cdot\sin(\theta))}{k\cdot R\cdot\sin(\theta)}\right] \quad (7)$$

Drawing expression (7) on a polar graph the form of the emitted beam can be seen, see figure 3

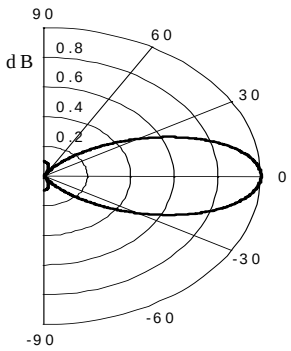

Fig. 3. Beam shape for a circular ultrasonic emitter

Visual information provided by an ultrasonic sensor is function of both sensor characteristics (beam angle, frequency, attenuation, emitter-receiver distance, etc.) and focused-object characteristics (object-sensor distance, shape, rough, etc.).

From several articles [9], [13] and many previous experiments, it has been high-lighted that the main information about the focused object is not in the carrier signal but in its envelope. Therefore, the signal coming from the receiver, see figure 4, has been rectified and digitally filtered with a third order Butterworth low-pass filter. It has been designed for a cut-off frequency of 1 kHz in order to obtain the envelope of the incoming signal. In this way, the number of data to work with is reduced with no significant information losses.

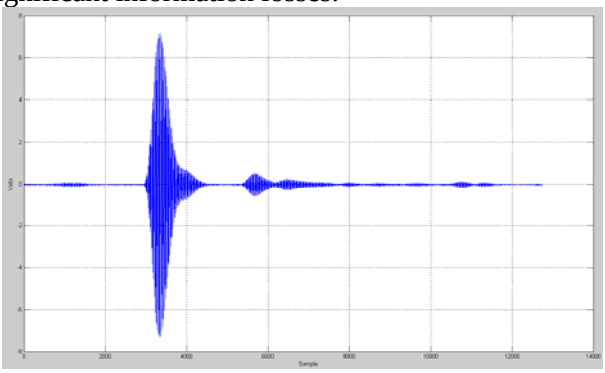

Fig. 4. Echo Signal received.

3.2 Ultrasonic information

Time of flight is the main information provided for the echo signal. It can be used directly as distance measurement without taken into account the aperture of ultrasonic lobe. However, this could produce measurement errors. Also, in order to reduce the errors motivated by the lobe width, a triangulation process could be carried out for obtaining the real object-echo contact point coordinates.

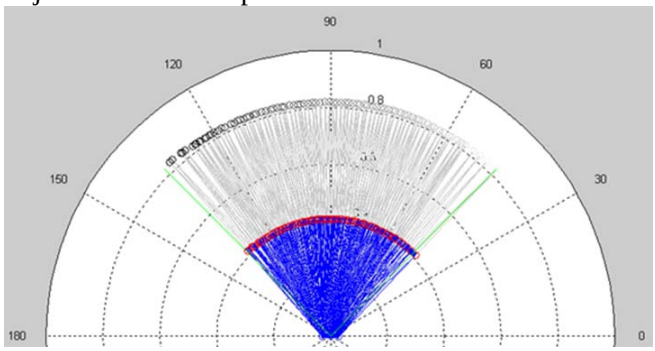

Fig. 5. Echo-object contact points calculated by triangulation for flat surface.

For example, figure 5 shows the real echo-object contact points for a flat surface, with a Receiver-Emitter-Receiver ultrasonic structure, and with an angle from -45º to +45º.

However, there is much more information on the echoes, such as the number of lobes in the envelope, the energy below the envelope and the energy below each lobe could be also important, etc.

For example, defining the energy of the signal as:

$$E = \lim_{t\to\infty}\int_0^T \left|x^2(t)\right|\cdot dt \quad (8)$$

It is possible to obtain the normalized energy below the envelope as figure 6 shows.

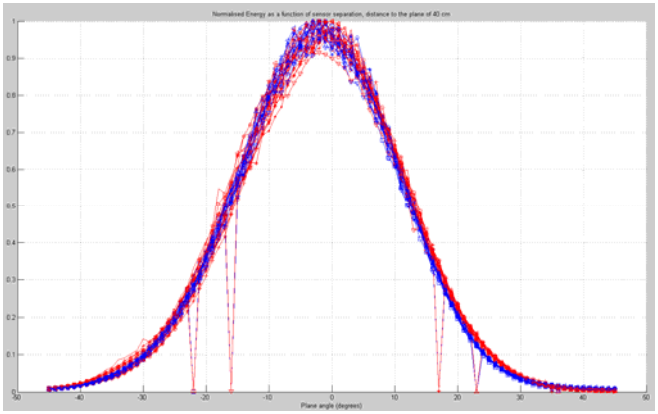

Fig. 6. Normalized energy below the echo signal envelope as function of the sensors separation, with a R-E-R structure, for angles from -45º to +45º.

These curves can be fitted to a Gaussian one:

$$E_{norm} = a \cdot \exp\left(-(\theta - b)^2 / c^2\right) \qquad (9)$$

Table I shows the parameters to fit the normalized energy to a Gaussian function for different distances between sensors and for several gaps between sensors an object. This information can be used to discriminate the angle between the object and the sensors. It is clear that time of flight is not the only information on the echo signal, and it can be useful to get more information about the object's surface.

3.3 Ultrasonic image obtained

By using the time of flight of the echo signal, it is possible to obtain a depth map from the scene. Different configurations of sensors are suitable for obtaining this three-dimensional information. The simplest one is a configuration with only one transducer working as emitter and receiver. Other common configuration is composed by a transducer working as emitter and other transducer working as receiver. Besides, with combinations of emitter-receiver pairs it is possible, under specific conditions, to use triangulation in order to estimate the contact point of the echo signal with the scene, as it was indicated in paragraph 3.2. In this work, an industrial robot has been used for scanning the scene of interest. Besides, a rectangular array structure with an emitter at the centre and eight receivers have been used

TABLE I
FITTING ENERGY TO A GAUSSIAN: PARAMETERS OBTAINED

Distance to the Reflection Plane	Sensors Separation	Gauss function Parameters		
(cm)	(mm)	a	b	c
20	30	1,00	0,56	21,2
	100	1,02	2,18	18,7
	150	0,95	0,39	18,27
40	30	0,99	-2,11	17,81
	100	0,97	-1,86	18,02
	150	0,99	-2.13	18,03
60	30	0,97	3,74	17,84
	100	0,98	2,54	18,15
	150	0,96	3,22	17,61
80	30	0,98	5,05	17,96
	100	0,97	3,61	19,5
	150	0,95	2,28	18,6

Table I. Parameters to fit the normalized energy below the echo signal envelope to a Gaussian Curve for several object-sensors distances and gaps between sensors.

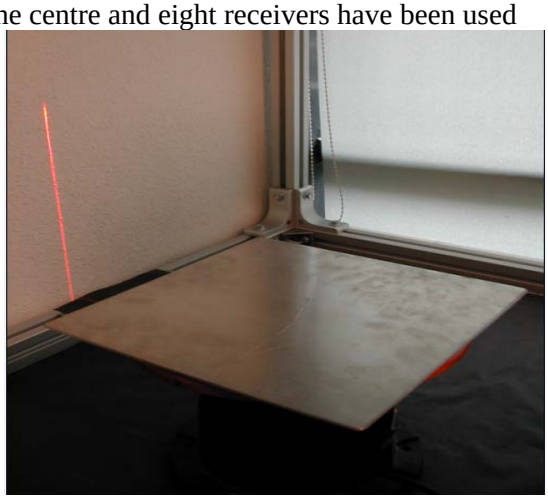

Fig. 7. Flat real specimen

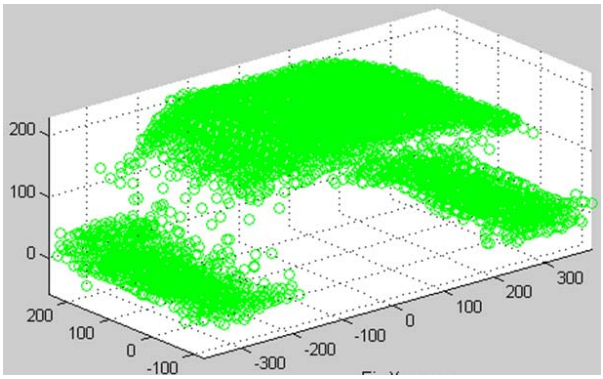

Fig. 8. 3D representation obtained as a cloud of 3D points by using ultrasounds. Axes are in mm

3.3 Ultrasonic image processing

Once the cloud of 3D points captured by ultrasounds is available it is necessary to process this information for improving its quality. It would be a good deal to use the powerful algorithms developed for artificial vision images to process the three-dimensional ultrasound images. Therefore, a solution is to transform the three-dimensional data into a gray-scale image. This can be done by using a linear transformation to map the depth of the 3D points into gray-scale levels. In order to obtain correct results, a convenient resolution should be chosen. Also, it could be convenient to use a three-dimensional interpolation for increasing the definition before the transformation to a gray-scale image. See figure 9.

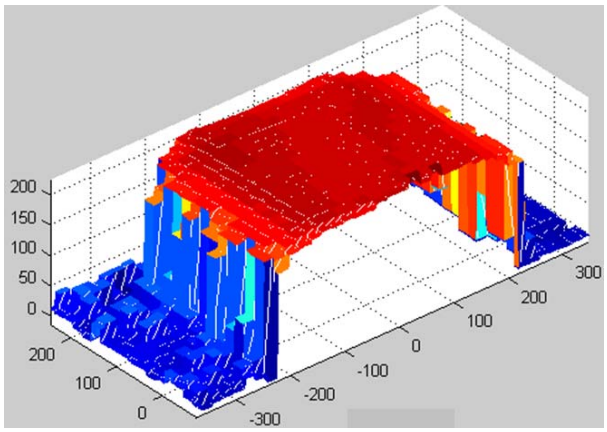

Fig. 9. Results obtained for the real specimen of figure 7 after a three-dimensional interpolation with the nearest method, and a gray-level representation of the previous interpolated three-dimensional image. Axes are in mm

Once the transformation from 3D ultrasound image to gray-scale image is carried out, it is possible to apply powerful algorithms developed for images obtained by artificial vision cameras, such as Gaussian filtering, unsharp filtering, threshold filtering, Hough transform, etc.

So, it is possible to process the image presented in figure 7 and to get different results, such as those presented in figure 10. So it is clear that transforming the three-dimensional ultrasonic data into a gray-level image permits to use powerful algorithms developed initially for artificial vision and, therefore, to increase the possibilities for the ultrasonic image

4 Hybrid Images

Optical-ultrasound hybrid images are obtained by combination of two images from the same scene.

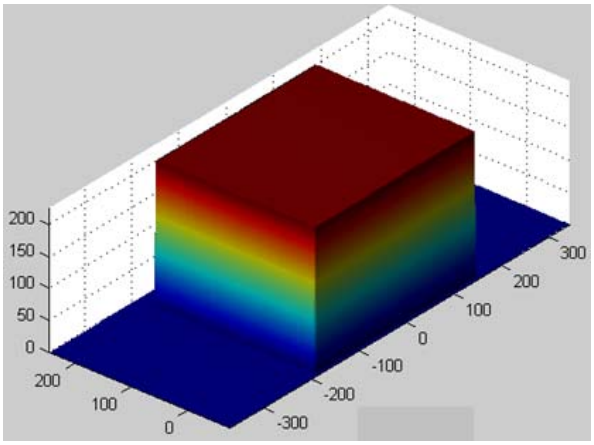

Fig. 10. Results obtained after edge detection by using Hough transform and its 3D representation. Axes in mm.

A 3D image is captured with artificial vision cameras by using three-dimensional techniques (stereo vision or active triangulation) and the other 3D image is obtained with ultrasonic sensors through a scene's scanning process.

In order to fuse images from both technologies, it is necessary to use the same reference frame and for this, to estimate the transformation matrix between the camera frame and the reference frame and the transformation matrix between the ultrasonic sensor frame and the reference frame. These matrices are calculated during the calibration process.

Figures 11, 12 and 13 show the three-dimensional data captured by artificial vision with active triangulation technique, the three-dimensional data captured by ultrasounds after a filtering process, and the fusion of both images, that is, a hybrid image. This image is made of data obtained from both techniques, that is, ultrasounds and artificial vision. Besides, all of these data have been captured for the real specimen from figure 7.

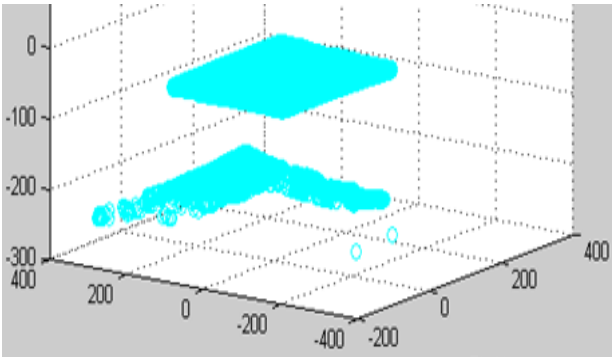

Fig. 11. Results obtained for the real specimen of figure 7. Data captured with three-dimensional artificial vision. Axes are in mm.

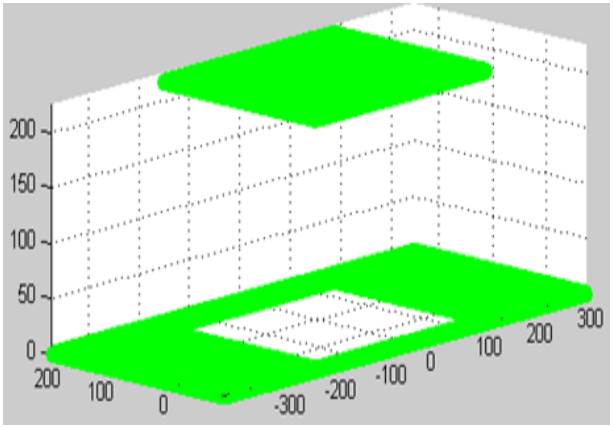

Fig. 12. Results obtained for the real specimen of figure 7. Ultrasonic three-dimensional data after a three-dimensional interpolation with the nearest method. Axes are in mm

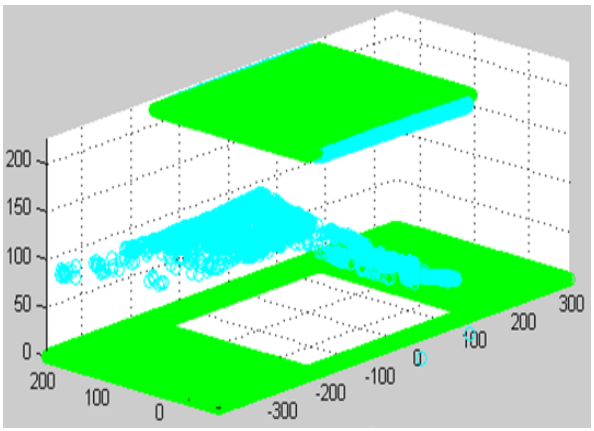

Fig. 13. The hybrid image. Axes are in mm.

4 Conclusions

It has been analyzed the possibility of collaboration between two technologies, that is, ultrasounds and artificial vision, in order to obtain three-dimensional information. It has been proved that three-dimensional ultrasonic information can be transformed on a gray-level image. Besides, under these conditions is possible to use powerful processing algorithms already developed for image processing. Also, three-dimensional information captured by artificial vision and ultrasound techniques can be fused to obtain more information about the scene analyzed. Hybrid images could be interesting for three-dimensional applications that include transparent objects or poor illumination conditions.

Acknowledgment

This work has been carried out under the Spanish Science and Innovation Office sponsorship in the Spanish Project DPI2006-15313

References

[1]. Kuroda,S., Jitsumori,A., Inari, T. Ultrasonic Imaging System for Robot Using Electronic Scanning Method.,. Proc. Int. Con£ Advance Robotics, pp. 187-194, 1983.
[2]. Estochen, E.L., Neuman, C.P., Prinz.,F.B. Application of Acoustic Sensors to Robotic Seam Tracking. IEEE Trans,.Ind. Electronics, vol.IE-3 1, pp. 219-224, 1884,
[3]. Brown, M.K. Feature Extraction Techniques for Recognizing Solid Objects with an Ultra-sonic Range Sensor. IEEE J.Robot Autom., vol. RA-1, pp. 191-205, 1985.
[4]. Abele,E.,Ahrens, U., Drunk,G. Applications of Ultrasonic Sensors for Handling Task with Industrial Robots.", E.Abele, U.Ahrens y G.Drunk. 16th ISIR. Bruselas, 1986.
[5]. Martín, J.M., Ceres,R., Calderón, L., Freire,T. Ultrasonic ranging gets thermal correction., MCB University Press, Sensor Review, Vol Nº 9, pp. 153-155, 1989.
[6]. Parrilla, M., Anaya,J.J., Fritsch,C. Digital Signal Processing Techniques for length Accu-racy Ultrasonic Range Measurements.,. Proc. IEEE Ultrasonic Symposium, 1991.
[7]. Acampora, A. S., Winters, J.H., A Three-Dimensional Ultrasonic Vision for Robotic Appli-cations., IEEE Trans.Pattem Anal. Machine Intell., vol. 11, no. 3, pp. 291-303, 1989.
[8]. Watanabe,S., Yoeyama,M. An ultrasonic visual sensor for three-dimensional object recog-nition using neural networks. IEEE trans. on Robotics and Automation, Vol. 8, Nº 2, pp. 240-249, 1992.
[9.] Llata, J.R., Sarabia,E.G., Arce,J., Oria,JP. Probabilistic expert systems for shape recogni-tion applied to ultrasonic techniques. Proceedings of IEEE International Symposium on Computational Intelligence in Robotics an Automation.
[10]. Llata, J.R., Sarabia,E.G., Oria,JP Short distance ultrasonic vision for mobile robots., Proceedings of intelligent components for vehicles, pp. 327-332, Sevilla, 1998.
[11]. Sarabia,E.G., Llata, J.R., Arce,J., Oria Shape recognition and orientation detection for industrial applications using ultrasonic sensors., Proceedings of IEEE international joint symposia on intelligence and systems, pp. 301-308, Maryland, 1998.
[12]. Llata, J.R., Sarabia,E.G., Arce,J., Oria,JP. Pattern recognition with ultrasonic sensors: A neural networks evaluation., J. R. Llata, E. G. Sarabia, J. P. Oria. Sensor Review. The in-ternational journal of sensing for industry, Volumen 1, pp 45-51, 2001
[13]. Llata, J.R., Sarabia,E.G., Oria Evaluation of Expert Systems for Automatic Shape Recog-nition by Ultrasound. Kluwer, Journal of Intelligent Manufacturing, Volumen 13, pp 177-188, 2002
[14]. Llata, J.R., Sarabia,E.G., Arce,J., Oria, JP. Ultrasonic robotic vision using fuzzy logic for surface reconstruction., Proceedings of the 14th World congress of international federation of automatic control, Vol. Q, pp. 209-214, Beijing, 1999.
[15]. Llata, J.R., Sarabia,E.G., Gil, M., Oria, JP Fuzzy enhancement of three-dimensional ultra-sonic images, Proceedings of 3rd international conference on industrial automation, pp. 14.9-14.12, Montreal, 1999.
[16]. El-Hakim, S. F., A comparative evaluation of the performance of passive and active 3D vision systems, St. Petersburg Conference on Digital Photogrammetry, Volume 2646, pp 14-25. June 25-30, 1995.
[17]. Trucco, E., Calibration, data consistency and model acquisition with a 3-D laser stripe. International Journal of Computer Integrated Manufacturing 11: (4) 293-310, Jul-Aug 1998.
[18]. Rayleigh. Theory of the sound. Dover Publications, New York. 1975.
[19]. Kinsler, L.E., Fundamentals of Acoustics. Wiley, New York, 1982.