Simultaneous Determination of Dimension and Intensity of an Object by Image Fusion

Umesh C. Pati, *Member, IEEE*
Dept. of Electronics and Communication Engineering
National Institute of Technology
Rourkela, India.

Pranab K. Dutta, *Member, IEEE*
Department of Electrical Engineering
Indian Institute of Technology
Kharagpur, India.

***Abstract*— This paper presents a novel approach for the simultaneous estimation and display of 3-D coordinates as well as intensity information of an object. A data acquisition system has been developed to acquire range images. The 3-D mesh representation of the object is obtained by registration and integration of acquired range images. Structural features from the mesh representation as well as intensity image of the object have been extracted by suitable edge detectors. 3-D edge points have been mapped to 2-D plane using perspective transformation. Fusion of both the edge maps is accomplished by affine transformation followed by iterative closest point algorithm. For the combined display of dimension as well as intensity information, an algorithm has been developed. This provides a high level description of the object with all the relevant information in a single image.**

Keywords - range image; intensity image; structural features; perspective transformation; image fusion

I. INTRODUCTION

Multisensor fusion refers to the synergistic combination of different sources of sensory information into a single representational format. The information to be fused may come from multiple sensory devices monitored over a common period of time or from a single sensory device monitored over an extended time period. Many of the advanced sensors produce images. Because each kind of image sensor can only focus on a given operating range and environmental conditions, it may not receive all the necessary information. A human observer cannot reliably combine visual information by viewing multiple images separately. Further, the integration of information across multiple observers is often unreliable. Thus, a fusion system that can provide one single image with more accurate and reliable information than any source image is of great practical value. In image fusion, two or more images, or some of their features, are combined together to form a single image, which contains the relevant and important information from the inputs so that it is more useful for the purpose of human visual perception. Hence, the purpose of image fusion is to generate a single image, which contains a more accurate description of scene than any of the individual source images. Successful image fusion significantly reduces the amount of data to be viewed or processed without significantly reducing the amount of relevant information [1, 2]. Image fusion is a branch of data fusion where data appear in the form of arrays of numbers representing brightness, color, temperature, distance and other scene properties. The concept of multisensor data fusion is not new. It is naturally performed by humans and animals to achieve more accurate assessment of the surrounding environment by the use of multiple senses.

In a broad sense, image fusion is performed at three different processing levels according to the stage at which the fusion takes place. These are pixel level fusion, feature level fusion and decision level fusion [3]. Feature level fusion is a medium level image fusion. This level can be used as a means of creating additional composite features. At first, relevant features are abstracted from the input images and then combined. A range image and an intensity image bring complementary information of a scene or an object. The range image directly provides precise and accurate 3-D geometric information. In addition, acquisition of range image is independent of ambient illumination. On the other side, an intensity image provides a realistic color texture map. The data acquisition is less costly and easy with a digital camera. Fusion of both the images compensates limitations of one with strengths of the other [4]-[6]. It delivers a richness of description, which is not possible with either source in isolation.

In this work, a novel approach for the simultaneous estimation and display of 3-D coordinates as well as intensity information of an object has been presented. The rest of the paper is organized as follows: 3-D mesh generation of the object from acquired range images has been described in Section II. Extraction of structural features from the 3-D mesh as well as intensity image of the object has been discussed in Section III. Section IV presents mapping of 3-D edge points to 2-D plane by perspective transformation. Fusion of edge maps from both the sources has been explained in Section V. The algorithm for the combined display of 3-D and intensity information has been presented in Section VI. The paper finally concludes in Section VII.

II. 3-D MESH GENERATION

The main steps for 3-D mesh generation are data acquisition, registration and integration [7]. According to the measurement principle, range image acquisition techniques can be divided into two broad categories: passive methods and active methods [8]. Registration is the process in which the multiple views and their associated coordinate frames are aligned into a single global coordinate frame. Existing registration techniques can be mainly categorized into feature matching and surface matching. Successful registration aligns all the range images into a common coordinate system. The integration process eliminates the redundancies and generates a single connected surface model from the range samples.

A. Data Acquisition

The object for the experimentation purpose is made of wood and its external dimensions are 12.2 × 12.5 × 3.5 cm. The surface of the object contains features of different shapes like triangle, notch, rectangle and circle. The object is placed on a horizontal and smooth platform. The method of optical triangulation has been used to acquire 3-D data of the object. A red line diode laser which spreads the laser beam into a sheet of light is placed on a mount. The laser plane scans the object with the help of a manually operated mechanical arrangement on the mount. Part of the laser stripe falls on both sides of the object on the platform. A charge-coupled device (CCD) camera interfaced with Silicon Graphics machine acquires series of range images while scanning is performed. The experimental setup and the object are shown in Fig. 1.

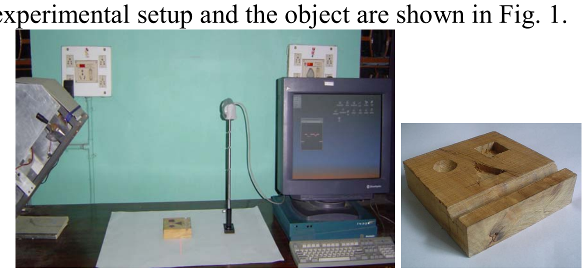


Figure 1. Experimental setup and the object.

The process has been repeated for the rest three different views by rotating the object manually by 90 degrees in anticlockwise direction each time so that all of the surface detail is captured. A laser scanned RGB image of the object is shown in Fig. 2.

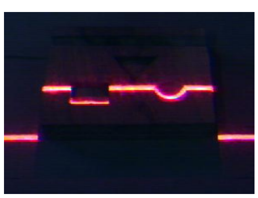


Figure 2. Laser scanned image.

The captured RGB images are subjected to preprocessing and thinning algorithm. With the help of optical triangulation method, the coordinates in 3-D space i.e. (X, Y, Z) values have been calculated for all the illuminated pixels. This process is repeated for all the images of the first view and then for the rest three views.

B. Registration

For successful registration of range images from different views, feature matching technique for coarse registration followed by iterative closest point (ICP) algorithm for fine registration have been applied. We have chosen the left bottom corner point on the surface of the object in the range image as the feature for registration of different views. For fine registration, ICP algorithm performs the process of registration iteratively.

C. Integration

The laser points on the platform have been removed as the points on the surface of the object are the desired points for 3-D reconstruction. The steps of integration algorithm include detection of overlapping points, merger of corresponding points and generation of a single connected surface model. The cloud of 3-D points needs to be connected in such a way that it looks like a smooth surface. Data gridding and surface fitting have been accomplished using the *griddata* function of MATLAB. It uses triangle-based cubic interpolation method based on Delaunay triangulation of data for surface fitting. The mesh representing the surface of the object has been shown in Fig. 3.

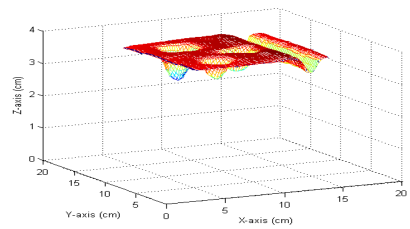


Figure 3. Mesh representation of the surface.

III. EXTRACTION OF STRUCTURAL FEATURES

A. Edge Detection in Range Image

The Laplacian of a Gaussian (LoG) edge detector [9] has been applied on the interpolated depth coordinates of the mesh surface. The edge features are extracted based on the discontinuity in depth. The detected edge points using LoG edge detection algorithm are shown in Fig. 4.

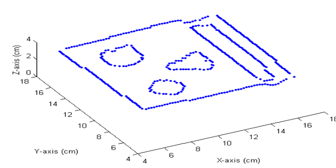
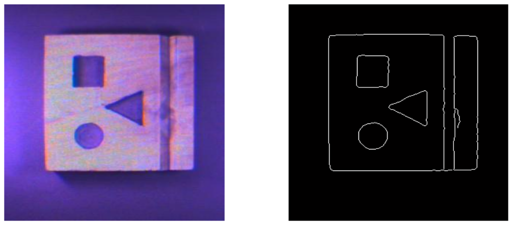


Figure 4. Detected edge points.

B. Edge Detection in Intensity Image

Different shapes on the surface of the object are coated with black color to minimize the effect of shadows. The object is placed on a horizontal and smooth platform. The illumination light is turned on and an intensity image is acquired with the same CCD camera interfaced with Silicon Graphics machine. The captured intensity image of the object is converted into gray scale image and Canny edge detection algorithm [10] is applied to the converted image. The intensity image and its edge map have been shown in Fig. 5..

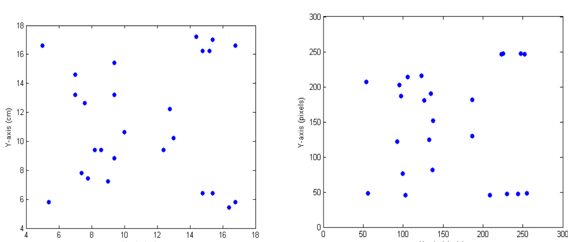


(a) (b)

Figure 5. (a) Intensity image (b) Edge map.

C. Corner Point Detection

In the present case, those points are considered to be corner points which are at a significant distance from the centroid. From the edge image obtained from range and intensity images, the connected shapes on the surface of the object have been further segmented and labeled using 8 connectivity. In this work, shape signatures have been chosen for feature representation. The original signature has been constructed by measuring and plotting the distance from the centroid of the shape to the boundary at all discrete positions along the digitized boundary as a function of angle [11]. Signatures along with the peaks for all the segmented regions of both range and intensity images have been obtained. Corner points have been detected out of the peak values from the shape signature. Prominent peaks are obtained with values above the mean of the peaks. Prominent peaks corresponding to corners of the various shapes have been obtained. We have repeated the above process for all the segmented regions of range and intensity image. Fig. 6 (a) and (b) shows all the detected corner points in range image and intensity image respectively.



(a) (b)

Figure 6. Corner points in (a) range image (b) intensity image.

IV. MAPPING OF 3-D EDGE POINTS TO 2-D PLANE

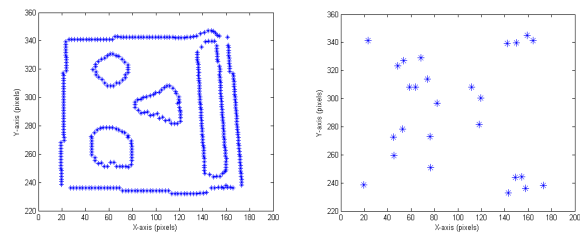
Fusion of extracted features can be performed in two spaces. These are 2-D space and 3-D space. In our approach, we chose 2-D space for fusion of extracted edge maps. For this purpose, 3-D edge points in range image have been projected onto 2-D plane by perspective transformation [12]. With the help of experimentally obtained scene-image coordinate pairs of the control points on the object, the perspective transformation matrix has been obtained as

$$\begin{pmatrix} c_{h1} \\ c_{h2} \\ c_{h3} \\ c_{h4} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -35.0657 \\ 0 & -0.8660 & 0.5 & 308.8160 \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & -0.0004 & 0.0041 & 0.6880 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{1}$$

where $(c_{h1}, c_{h2}, c_{h3}, c_{h4})$ are camera coordinates in homogeneous representation and (X, Y, Z) are world coordinates. The camera coordinates in Cartesian form are

$$x = \frac{c_{h1}}{c_{h4}} \text{ and } y = \frac{c_{h2}}{c_{h4}}. \tag{2}$$

For any world point (X, Y, Z) on the object, the image coordinates (x, y) can now be computed with the help of (1) and (2). The image coordinates are computed for all the points in the 3-D edge map of the object obtained using LoG edge detector and the projected 2-D edge map has been shown in Fig. 7 (a). The corner points detected using shape signatures have also been perspectively projected onto 2-D plane and the projected points have been shown in Fig. 7 (b).



(a) (b)

Figure 7. Projected (a) 2-D edge map (b) corner points.

V. FUSION OF EDGE MAPS

The fusion of edge maps has been accomplished using affine transformation followed by ICP algorithm.

A. Affine Transformation

The perspective projected corner points from the range image edge map have been shown in Fig. 7 (b) where as the detected corner points in the intensity image edge map have been shown in Fig. 6 (b). In order to fuse two data sets by affine transformation, a set of control points must be identified that can be located on both the maps. Since the general affine transformation is defined by six constants, at least three control points must be identified. The correspondence and detection of control points have been performed by finding the geometric orientation of corner points based on length and angle with respect to the centroid. The corner points in the range image for which corresponding points are identified in the intensity image satisfying the angle and length criterion have been saved as the control points.

A least-square technique has been used to determine the six affine transformation parameters from fifteen matched control points. We have computed the affine transformation matrix A using the coordinate values of the control points which is given by

$$A = \begin{pmatrix} 1.4152 & -0.0363 & 31.9664 \\ 0.0046 & 1.6523 & -341.7584 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

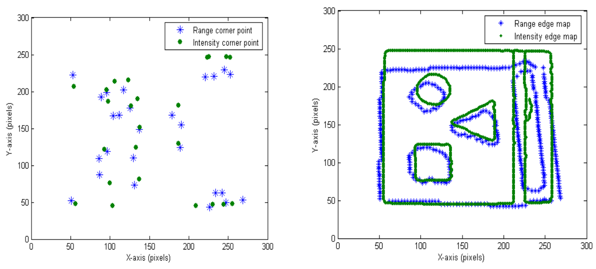
The transformation equations are obtained as follows

$$x' = 1.4152x - 0.0363y + 31.9664$$

and

$$y' = 0.0046x + 1.6523y - 341.7584. \quad (4)$$

Initially, (4) is applied to the projected range corner points. The corner points from both modalities after fusion have been shown in Fig. 8 (a). Then, the transformation is applied to all the points in the range image edge map for fusion with the intensity image edge map. Both the edge maps after fusion have been shown in Fig. 8 (b).



(a) (b)

Figure 8. Fusion of (a) corner points (b) edge maps.

B. ICP Algorithm

The coarsely fused edge points by affine transformation have been subjected to fine fusion using ICP algorithm. The calculation of transformation parameters has been made using the singular-value decomposition (SVD) method [13]. Both the edge maps after fusion are shown in Fig. 9.

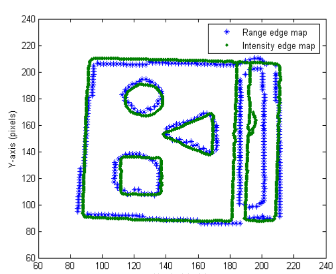


Figure 9. Fusion of edge maps by ICP algorithm.

VI. DISPLAY OF COMBINED INFORMATION

Several situations in image processing simultaneously require high spatial and high spectral information in a single image. In the present work, the coordinate information from range image and the intensity information from camera image have been displayed together. The steps for the display of combined information are as follows:

Step 1: The data matrices containing X, Y, Z coordinate values and intensity for every point obtained from automatic fusion of edge maps are passed to the customized *datacursormode* function in MATLAB.

Step 2: Handle to the figure object to be plotted is obtained.

Step 3: The application data is set as X, Y, Z coordinates and intensity value with the figure handle.

Step 4: The *datacursormode* is enabled and it is passed with the user defined *datacursor* function to display additional data in the figure on clicking the mouse.

Step 5: The structure of figure handle is obtained from the data cursor object in the *data cursor* function.

Step 6: In the user customized *datacursor* function, the structure of position (X, Y, Z position) where the user has clicked is retrieved back. The data index in the array of data is passed for the same position.

Step 7: The application data, which has been set earlier, is obtained from the figure handle.

Step 8: All the information i.e. X, Y, Z coordinate values and intensity value are displayed together in a text box on clicking the mouse. This has been shown in Fig. 10.

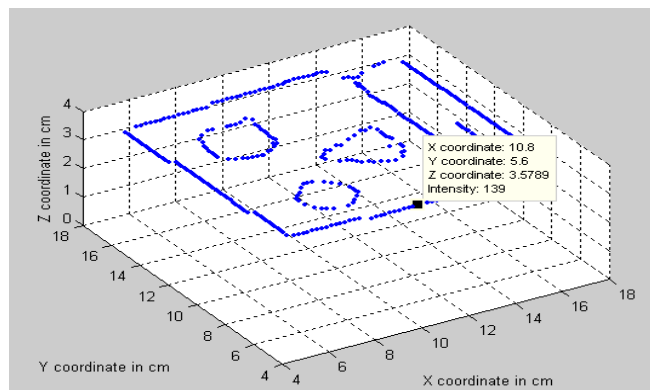


Figure 10. Display of 3-D coordinates and intensity information.

Fig. 10 provides a high-level description of the scene. The 3-D information is obtained from the range image whereas intensity information is obtained from camera image. Complementary information from both the modalities provides all the relevant information in a single image. This method can be augmented with more number of complementary parameters to be displayed in the same figure plot.

VII. CONCLUSION

In this paper, we have presented a novel method for simultaneous display of dimension and intensity information of an object in a single image. The optical triangulation method for range data acquisition provides dense and accurate 3-D data compared to passive sensing methods. The registration of range images takes advantages of both feature and surface matching techniques. The integration stage leads to reduction of redundant data points considerably which minimizes the computational effort.

A technique involving shape signatures has been implemented for the detection of corner points in both the edge maps. The correspondence between the corner points in both the edge maps based on certain criterion has been established. The coarse fusion of edge maps from both the sources has been performed using affine transformation followed by fine fusion using ICP algorithm. *Datacursormode* of MATLAB has been customized to display 3-D and intensity information from both modalities together in the same figure plot. It has the flexibility to accommodate more number of complementary parameters to be displayed together. This approach can find application in situations where multiple complementary parameters are to be displayed in same figure plot. Combined visualization has the potential to provide more information than the individual data sets as it takes advantage of individual strength of each mode.

REFERENCES

[1] R.S. Blum, Z. Xue and Z. Zhang, "An overview of image fusion", in Multi-sensor image fusion and its applications, R. S. Blum and Z. Liu , Eds. Taylor and Francis (CRC Press), 2005, Chapter 1, pp. 1-36.

[2] A.A. Goshtasby and S. Nikolov, "Image fusion: advances in the state of the art", Information Fusion, vol. 8, no. 2, pp. 114-118, April 2007.

[3] H. Wang, "A new multiwavelet-based approach to image fusion", Journal of Mathematical Imaging and Vision, vol. 21, no. 2, pp. 177-191, 2004.

[4] P. Dias, V. Sequeira, F. Vaz and J.G.M. Goncalves, "Registration and fusion of intensity and range data for 3D modeling of real world scenes", in Proc. 4th Int. Conf. 3-D Digital Imaging and Modeling, Banff, Canada, 2003, pp. 418-425.

[5] C. Boucher, J.C. Noyer and M. Benjelloun, "3D structure and motion recovery by fusing range and intensity image sequences", in Proc. 3rd Int. Conf. Information Fusion, 2000, vol. 2, pp. WEB3/17-WEB3/23.

[6] E.K. Forkuo and B. King, "Automatic fusion of photogrammetric imagery and laser scanner point clouds", in Proc. XXth Int. Society for Photogrammetry and Remote Sensing Congress, Istanbul, Turkey, 2004, vol. XXXV, part B4, pp. 921-926.

[7] Z. Zhang, X. Peng, W. Shi and X. Hu, "A survey of surface reconstruction from multiple range images", in Proc. 2nd Asia-Pacific Conf. Systems Integrity and Maintenance, Nanjing, China, 2000, pp. 519-523.

[8] Z. Zhang, X. Peng and D. Zhang, "Transformation image into graphics", in Integrated image and graphics technologies, D. Zhang, M. Kamel and G. Baciu, Eds. USA: Kluwer Academic Publishers, 2004, Chapter 7, pp. 111-129.

[9] R.C. Gonzalez and R.E. Woods, Digital image processing, Pearson Education , Delhi, 2003.

[10] J.F. Canny, "A computational approach to edge detection", IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679-698, 1986.

[11] R.C. Gonzalez, R.E. Woods and S.L. Eddins, Digital image processing using MATLAB, Pearson Education, Delhi, 2004.

[12] R.C. Gonzalez and R.E. Woods, Digital image processing, Addison-Wesley Publishing Company, 1999.

[13] R. Madhavan and E. Messina, "Iterative registration of 3D LADAR data for autonomous navigation", in Proc. IEEE Intelligent Vehicles Symposium, Columbus, OH, 2003, pp. 186-191.