Journal of Computational Information Systems 10: 17 (2014) 7619–7628
Available at http://www.Jofcis.com

Obstacle Detection of 3D Imaging Depth Images by Supervised Laplacian Eigenmap Dimension Reduction

Zhenchang ZHANG [1,2,*], Shuang WANG [3]

[1] *Shenzhen Research Institute of Xiamen University, Shenzhen 518057, China*

[2] *College of Computer and Information Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

[3] *Department of Computer Science, Xiamen University, Xiamen 361005, China*

Abstract

In this paper, we propose an obstacle detection method for 3D imaging sensors by supervised Laplacian eigenmap manifold learning. The paper analyses the depth ambiguity problem of 3D depth images firstly, then the ambiguity boundary line and intensity images are used to eliminate ambiguity and extract the non-ambiguity regions of depth image. 3D information without ambiguity is applied to the manifold learning stage directly, and we use a biased distance in supervised Laplacian eigenmap to realize a non-linear dimensionality reduction of the depth data. In the experiment, 3D coordinate information of obstacles and non-obstacles is used as training data of manifold learning respectively. Experiment results show that our model can effectively eliminate the depth ambiguity of 3D imaging images and realize obstacle detection and identification, the method also shows good stability to 3D imaging noise.

Keywords: Manifold Learning; Obstacle Detection; Laplacian Eigenmap

1 Introduction

As a new 3D data measuring technique, the application of time-of-flight 3D imaging sensors in the obstacle detection problem is becoming a hot research area, and TOF has showed great potential in the computer vision field. Obstacle detection is crucial for Automated Guided Vehicle navigation (AGV), intelligent robot moving and motion tracker [1, 2, 3, 4]. Compared to common 2D imaging sensors, TOF sensors have many superiorities. Depending on illumination conditions, surface materials, and object orientations, the appearance of objects in a 2D image varies greatly, but the 3D information of objects is invariant to those factors [5].

In recent years, TOF 3D imaging sensors have been widely used in the obstacle detection problem. In order to adapt to the obstacle detection and segmentation problem, the Method

*Supported by Shenzhen science and technology plan project JC200903180638A.
**Supported by Fujian Agriculture and Forestry University Project 2012xjj06.
*Corresponding author.
Email address: stdin@fafu.edu.cn (Zhenchang ZHANG).

1553–9105 / Copyright © 2014 Binary Information Press
DOI: 10.12733/jcis11613
September 1, 2014

introduced in [6] converted the TOF depth values to 3D point clouds in their AGV coordinate frame. By integrating different pedestrian samples in the scene, a Pedestrian-Scene Map (PS Map) is learnt to measure ground/non-ground regions and other scene information [7]. An effective obstacle extraction mechanism is proposed to capture obstacles by various object proprieties revealing in the depth map [8]. A practical framework is proposed to estimate road surface and obstacle [9], which has a very high accuracy. An obstacle detection system presented in [10] can handle non-flat road surface detection problem and is robust to illumination condition. Combining range and intensity images enables an fast and accurate object segmentation, many useful information such as distances, obstacle types can also be obtained [11].

However, there are still many problems in the actual use of TOF data, TOF imaging sensors derive the depth images via computing the phase offset between the emission signal and the reflected signal. The data itself has many critical problems, such as optical noise, low resolution, depth ambiguity problem and so on [12]. Many researches have been done to enhance the quality of depth images. Traditional multi-frame superresolution method is used to enhance depth image resolution and reduce noise [13, 14]. A method that combines TOF ranger sensor and passive stereo performs well for high accuracy depth maps [15]. Combing luminance information with depth images can efficiently solve the non-stationary noise estimating problem within depth images [16]. There are other ways which combined color images together with depth images [17, 18]. As for the depth ambiguity problem, there is a method which integrates intensity images with depth images after segmentation [19] to eliminate depth ambiguity. Analyzing depth images with two different modulation frequencies within a single capture also proved to be useful to avoid the ambiguity problem [20].

For such problems as AGV, intelligent robot moving and so on, improving the quality of depth images elaborately will cost a of time while it has no obvious effect on the detection results. And the methods mentioned above are all very complex. In our experiments, we focus on the most crucial defect of depth images on the obstacle detection problem: depth ambiguity. We try to find the first ambiguity boundary line of depth cycles on the depth image, then eliminate the information with ambiguity, and extract the non-ambiguity regions. Then we use supervised Laplacian eigenmap to realize nonlinear dimensionality reduction of depth images and obstacle detection.

All the methods forementioned about obstacle detection via TOF 3D imaging sensors did not reckon for the intrinsic topological relation and underlying manifold structure of TOF 3D data. In fact, many high-dimensional data is embedded in a low-dimensional nonlinear manifold structure, and this low-dimensional manifold might be the same as the manifold in human's recognition. Compared to other linear dimensionality reduction methods, such as principal component analysis (PCA), independent component analysis (ICA) and Fourier Transform, Laplacian eigenmap is better in preserving the topological relation of original 3D data and making use of the underlying manifold structure of TOF data. TOF and manifold learning have cooperated very well for some applications, such as human body trackingactivity recognition and posture estimation [21, 22, 23]. Using Laplacian eigenmap to find a low-dimensional representation that best preserves the intrinsic topological relation and local structure properties of TOF data can reduce the complexity of processing. At the same time as a method of machine learning, manifold learning is insensitive to noise, which makes this approach robust and stable to process TOF data.

In our experiment, obstacle and non-obstacle samples are marked on the depth images after ambiguity elimination as training samples, then we use new depth images with median filtered to

perform obstacle detection. Experiment results show that after manifold learning, we can detect the positions and shapes of the obstacles with a high accuracy rate, and the method also shows its good stability and flexibility.

2 Depth Ambiguity Elimination

The 3D imaging system computes distances by measuring the phases of cosine wave, so the measured distances are periodic. When the modulation frequency is 20 MHz, the interval of the distance periodicity is 7.5m. So we must determine which distance periodicity every pixel locates in, thus the correct distances can be obtained. The range image obtained by the system has an obvious characteristic: when the actual distance changes from near to far, the distance is periodic. From near to far, if the distance periodicity is labeled as C_n $(n = 1, 2, 3,)$, then at the boundary of the two distance periodicities C_n and C_{n+1}, the distance changes suddenly from large to small. According to this characteristic, for the scanning laser ladar data, the range ambiguity can be eliminated by range gradient [24]. The algorithm is: for a range image, scan from bottom up, find the points where the distance changes suddenly from big to small of every column. Then the distance periodicity of every point can be determined. The advantage of this method is that it only need single frame of image to eliminate the distance ambiguity, and the calculation amount is small. In this paper, we use a pseudo color image to represent a depth image, the color changes from blue to red, which means the distance is getting farther. Fig. 1 shows the ambiguity phenomena of outdoor scenes.

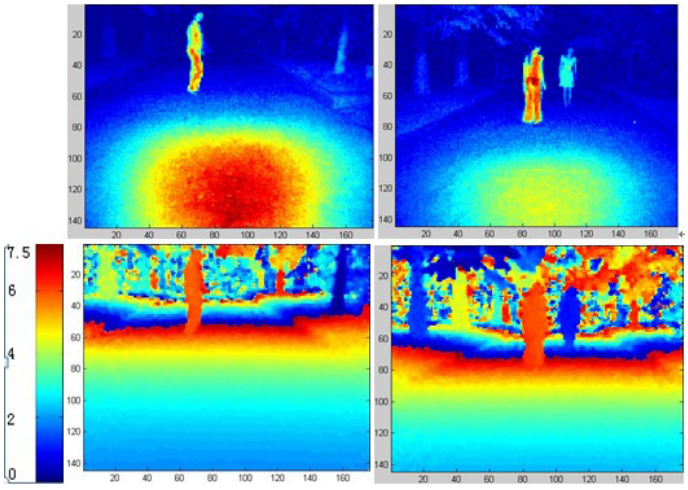

Fig. 1: Ambiguity phenomena of outdoor scenes

3 Laplacian Eigenmap Dimension Reduction

In depth images produced by time-of-flight imaging sensors, the information of each pixel contains (X, Y, Z) coordinates. Thus if process the data directly, we will get a high-dimensional data space. Moreover, considering that many high-dimensional data is embedded in a low-dimensional nonlinear manifold structure, and this low-dimensional manifold might be the same as the man-

ifold in human's recognition. So we attempt to apply manifold learning to obstacle detection problems by TOF depth images.

In this paper, we use k-nearest neighborhood method and a biased distance [25] to reconstruct the weight graph, and supervised Laplacian eigenmap is applied to realize graph embedding.

3.1 Graph reconstruction

Given a database $\mathbf{X}[(x_1, c_1), (x_2, c_2), ...(x_n, c_n)]$, where $x_i \in R^D$, $i = 1...n$, and c_i is sample class label, here c_i is the label of depth image sample (obstacle or non-obstacle). $\mathbf{Y}[y_1, y_2...y_n]$ is the corresponding low-dimensional of $\mathbf{X}$, where $y_i \in R^d$, $i = 1...n$, $d \ll D$. K-nearest neighborhood principle is an effective method to reconstruct the weight graph, and we use Gaussian kernel to define weight graph.

$$\mathbf{W}_{ij} = \begin{cases} \exp(-\frac{d^*(i,j)^2}{2\sigma^2}), & x_i \in N_k(x_j) \vee x_j \in N_k(x_i) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where, $\mathbf{W}(i, j)$ is the weight neighborhood matrix, d^* is biased distance [25] different from the traditional Euclidean distance, which is defined as:

$$d^*(i, j) = d(i, j)f(i, j), \tag{2}$$

where, $d(i, j) = \|x_i - x_j\|$is the Euclidean distance. $f(i, j) = \frac{P(i,j)+\nu}{\max_{ij}P(i,j)-P(i,j)+\nu}$, $P(i, j) = \tau|c_i - c_j|$, ν is noise and τ is control parameter. We use biased distance d^* to choose the k-nearest neighborhood of sample x_i, which can effectively reduced the errors brought by using Euclidean distance $d(i, j)$.

3.2 Supervised Laplacian eigenmap

If the samples are evenly sampled from the low-dimensional manifold which is embedded in high-dimensional space, then the Laplacian Beltrami operator on the manifold can be approximated by the Laplacian matrix of graph. The eigenvector corresponding to the smallest eigenvalue of Laplacian matrix is the discrete approximation of the Laplacian operator on manifold. The weight neighborhood graph encodes the local geometric features and the underlying manifold properties of the 3D information, and the key point of manifold embedding is to keep the underlying manifold structure shown from the embedding process. The form of Laplacian embedding is:

$$\min_{\mathbf{y}} \sum_{i,j}^{n} \|\mathbf{y}_i - \mathbf{y}_j\|^2 \mathbf{W}_{ij} = \mathbf{Y}^T \mathbf{L} \mathbf{Y}, \tag{3}$$

where $\mathbf{L}$ is the graph Laplacian matrix of weight matrix $\mathbf{W}$ and $\mathbf{L} = \mathbf{D} - \mathbf{W}$, $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$, $\mathbf{D}$ is a diagonal matrix. We use $\mathbf{W}_{i,j}$ to punish the neighbor points which are far away from each other after the mapping, which ensures that the neighborhood points are still next to each other after the mapping. Then we get the optimal low-dimensional embedding by minimizing the eigenvalue of the graph Laplacian matrix. Combine the sample classification information of the 3D depth image as supervised information, the supervised Laplacian eigenmap method [26] is:

$$\min_{\mathbf{y}} \sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 \mathbf{W}_{ij} + \lambda \sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 \Omega_{ij} \tag{4}$$

where, λ is a scalar constant and Ω contains the supervised information of 3D sample labels, which is defined as follows:

$$\Omega_{ij} = \begin{cases} \exp(-\frac{|c_i - c_j|^2}{2\sigma^2}), & x_i \in N_k(x_j) \vee x_j \in N_k(x_i) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Then we get the optimal embedding by minimizing the eigenvalue of $L+\lambda\widetilde{L}$, and $\widetilde{L}$ is the Laplacian matrix of Ω.

4 Experiments

In our experiment, we use 3-dimensional non-scanning imaging sensor Swiss Ranger SR-3000 to capture 3D information. The system uses Time-of-flight principle to measure distance, with 0.6 μm CMOS/CCD image sensor, and the space resolution is 144×176. For the modulation frequency of 20 MHz, the non-ambiguity range equals to 7.5m. The experimental results presented in this paper are all based on a sinusoidal modulation of the emitted optical signal with a frequency of 20MHz.

First, we use the method proposed in Part 2 to realize ambiguity elimination of depth images. Examples of this process are shown in Fig. 2.

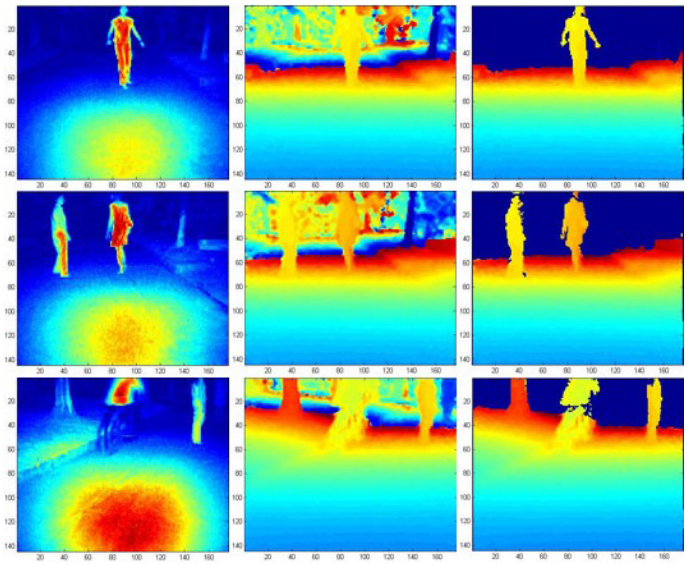

Fig. 2: Depth images without ambiguity

From Fig. 2 we can see that our method can effectively eliminate the ambiguity regions in the depth image and keep the non-ambiguity parts completely. Take the third scene in Fig. 2 as an example, we show its three-dimensional display of distance information Z in Fig. 3. From the figure we can see that before ambiguity elimination, the distance values within the ambiguity regions are confusion and irregular, while after ambiguity elimination, only regular distance values left. This can also illustrate the effectiveness of our method from another view.

After depth ambiguity elimination, we use supervised Laplacian eigenmap to realize dimensionality reduction of 3D data. In the training stage, we choose 800 obstacle samples and non-obstacle

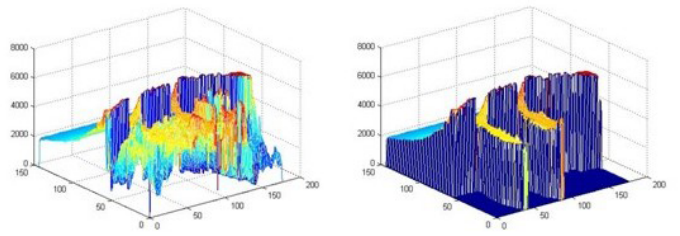

Fig. 3: Three-dimensional display of Z. (a) Before ambiguity elimination, (b) After ambiguity elimination

samples with the size 8×8 respectively as training samples. Every sample directly contains the whole 3D information captured by the TOF imaging sensor: X, Y and Z, where Z is the distance between the target and the sensor. In Fig. 4, the black square shows the training samples.

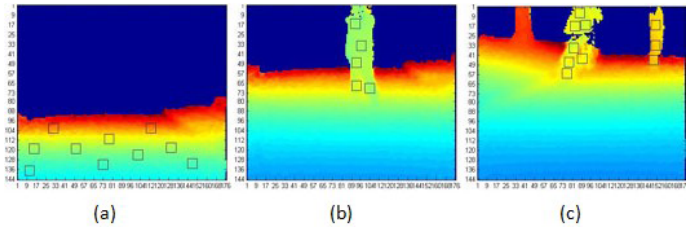

Fig. 4: (a) The non-obstacle samples, (b) and (c) the obstacle samples

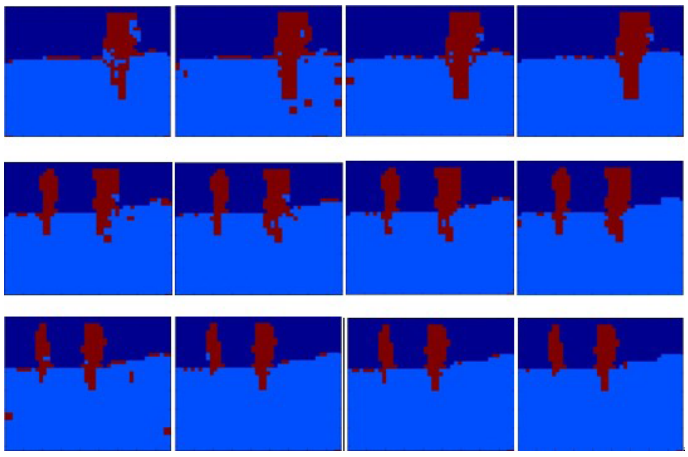

Fig. 5: Detection results corresponding to different numbers of training samples (200, 400, 600, 800)

In the testing stage of manifold learning, the algorithm automatically segments depth images without ambiguity into data blocks with the same size 8×8 as training samples. For the regions around the ambiguity boundary line, the scan space between adjacent blocks is 4, which may cause a little data overlapping, but the precision of the algorithm is increased. To the blocks, the number of whose valid pixels is less than 64, we choose 50 as a threshold and preserve the blocks with more than 50 valid pixels in them.

In order to determine the optimal number of training samples, we set the training number to be 200, 400, 600 and 800 respectively in the experiment. The detection results under different circumstances are shown in Fig. 5.

Table 1: Accuracy rates corresponding to different numbers of training samples

sample number	200	400	600	800
accuracy rate	95.95%	97.68%	98.81%	99.34%

In Fig. 5, from left to right, the number of training samples is 200, 400, 600, 800. From the detection results we can see, as the training number increases, the detection results become more complete, and the number of missing and error points falls. During the experiment, when the training number is 200 or 400, the detection result varies greatly with different training samples. When the training number reaches 600, the algorithm begins to show its good stability. From the accuracy table Table 1, we can see that the algorithm always keeps a high classification accuracy rate. Considering the stability and error rate, we set the final training number to be 800 and do not increase the number.

In the form shown in Fig. 6, we overlap the experiment result with the depth image to check the correspondences of the detected obstacles and the actual obstacles in position and shape.

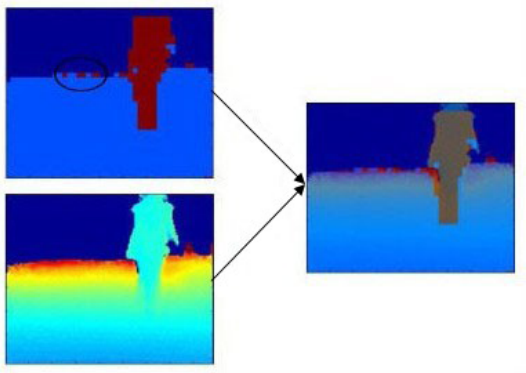

Fig. 6: The overlapping of the experiment result and the depth image

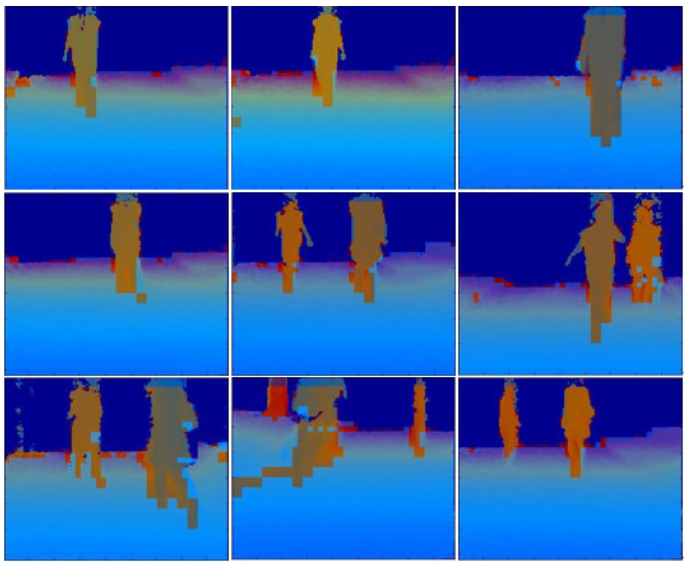

Fig. 7: Results of obstacle detection

In Fig. 6, the image at the top left corner is the experiment detection result, and the lower left

one is the corresponding depth image and the right image is the overlapping result. From the overlapping image we can see that the experiment result can show the position of the obstacles rightly and represent the obstacle's shape in general. In this overlapping way, the results of some scenes are shown in Fig. 7.

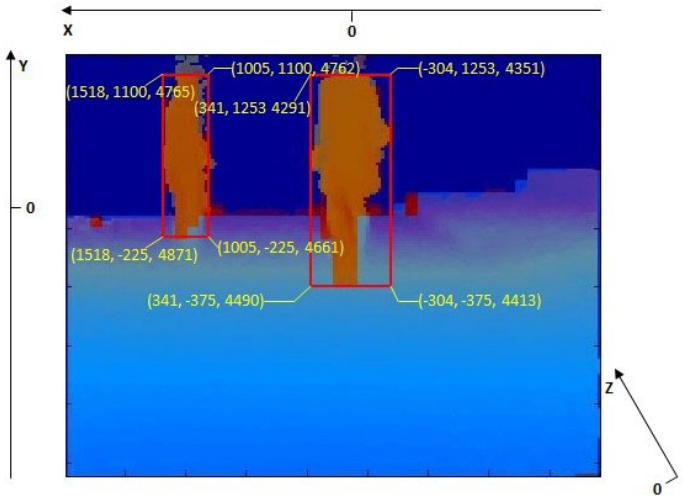

Fig. 8: 3D coordinate values of obstacles

Finally, we can obtain the coordinate values (X, Y, Z) of the obstacles in the 3D coordinate frame, where Z represents the distance between obstacles and sensors. The distance is measured in millimeters. As is shown in Fig. 8, the location of obstacles is complete. From the coordinate information we can see that the widths of the two obstacles are about 500mm and 640mm, the heights of the two obstacles are about 1300mm and 1600mm, and their distances from our observation point are about 4700mm and 4400mm. These 3D values are all important information in the obstacle detection and avoidance problem.

Experiment result shows that the model we introduced in this paper can effectively eliminate the depth ambiguity of time-of-flight depth images and realize obstacle detection and identification. During the whole process, we only did an easy median filter to depth images and the results show that manifold learning method has an ideal noise stability. Moreover, from the experiment detection results shown in Fig. 6, we can see that there are some "pseudo obstacles points" (the points in black circles) in the image. But these points are located around the ambiguity boundary line, which is because the measured depth values are not stable around the line, so they are not real "error points". Moreover, from Table 1 we can see that even when the training number is very low, the algorithm always keeps a high accuracy rate, which shows the stability and superiority of the algorithm.

5 Conclusion

In this paper, we applied the 3D information produced by time-of-flight camera into outdoor obstacle detection problems. After the depth ambiguity elimination stage, we use supervised Laplacian eigenmap manifold learning to realize data dimensionality reduction and obstacle identification. The experiment results are satisfying and can show positions and shapes of the obstacles generally. From the 3D coordinate values of the obstacles we can tell the widths, heights and

distances of the obstacles, these are very important information to the obstacle detection and avoidance problem. As a new and promising technique, time of flight principle is playing an more and more important role in the field of computer vision. Our method can also provide a reference for further study of obstacle detection problem via time-of-flight 3D imaging sensors.

References

[1] S. May, B. Werner, H. Surmann, and K. Pervolz, 3D time of flight cameras for mobile robotics, in IEEE/RSJ International Conference on Intelligent Robots and Systems, (Academic, Beijing, China, 2006), pp. 790-795.

[2] Jens T. Thielemann, Gøril M. Breivik and Asbjørn Berge, Robot navigation and obstacle detection in pipelines using time-of-flight imagery, in Three-Dimensional Image Processing (3DIP) and Applications, Proc. SPIE 7526, (2010).

[3] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert, Robust real-time 3D respiratory motion detection using time-of-flight cameras, International Journal of Computer Assisted Radiology and Surgery 3, 427-431 (2008).

[4] T. Schamm, M. Strand, T. Gumpp, R. Kohlhaas, J. Marius Zöllner, and R. Dillmann, Vision and ToF-based driving assistance for a personal transporter, Advanced Robotics (2009).

[5] S. Hsu, S. Acharya, A. Rafii, R. New, and Canesta Inc, Performance of a Time-of-Flight Range Camera for Intelligent Vehicle Safety Applications, in Advanced Microsystems for Automotive Applications 2006, V.Jürgen, G. Wolfgang, eds. (Academic, Heidelberg, Berlin, 2006) pp. 205-219.

[6] R. Bostelman, T. Hong, and R. Madhavan, Obstacle detection using a time-of-flight range camera for automated guided vehicle safety and navigation, Integrated Computer-Aided Engineering 12, 237-249 (2005).

[7] L. Zhu , J. Zhou, J. Y. Song, Z. L. Yan and Q. Q. Gu, A practical algorithm for learning scene information from monocular video, Optics Express 16, 1448-1459 (2008).

[8] C. H. Lee, Y. C. Su and L. G. Chen, An intelligent depth-based obstacle detection system for visually-impaired aid applications, in International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), (Academic, Dublin, Ireland, 2012).

[9] F. Oniga, and S. Nedevschi, Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection, IEEE Transaction On Vehicular Technology 59, 1172-1182 (2010).

[10] Y. Gao, X. Ai, Y.Wang, J. Rarity and N. Dahnoun, U-V-Disparity based Obstacle Detection with 3D Camera and steerable filter, in 2011 IEEE Intelligent Vehicles Symposium (IV), (Academic, Baden-Baden, Germany, 2011), pp. 957-962.

[11] X. Wei, S. L. Phung, and A. Bouzerdoum, Scene Segmentation and Pedestrian Classification from 3-D Range and Intensity Images, in 2012 IEEE International Conference on Multimedia and Expo, (Academic, Melbourne, VIC, 2012), pp. 103-108.

[12] S. Y. Kim, J. H. Cho, A. Koschan, and M. A. Abidi, Spatial and Temporal Enhancement of Depth Images Captured by a Time-of-flight Depth Sensor, in 2010 International Conference on Pattern Recognition, (Academic, Istanbul, Turkey, 2010).

[13] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, High-Quality Scanning Using Time-Of-Flight Depth Superresolution, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (Academic, Anchorage, AK, 2008).

[14] G. Rosenbush, T. Hong, and R. D. Eastman, Super-resolution enhancement of flash LADAR range data, in Unmanned/Unattended Sensors and Sensor Networks IV, Edward M. Carapezza, eds., Proc. SPIE 6736, (2007).

[15] J. Zhu, L. Wang, R. Yang, and J. Davis. Fusion of Time-of flight Depth and Stereo for High Accuracy Depth Maps, in IEEE Conference on Computer Vision and Pattern Recognition, (Academic, Anchorage, AK, 2008).

[16] L. Jovanov, A. Pižurica, and W. Philips, Fuzzy logic-based approach to wavelet denoising of 3D images produced by time-of-flight cameras, Optics Express 18, 22651-22676 (2010).

[17] A. Bleiweiss and M. Werman, Fusing Time-of-Flight Depth and Color for Real-Time Segmentation and Tracking, in Dynamic 3D Imaging, K. Andreas, and K. Reinhard eds. (Academic, Jena, Germany, 2009), pp. 58-69.

[18] R. Crabb, C. Tracey, A. Puranik, and J. Davis, Real-time Foreground Segmentation via Range and Color Imaging, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (Academic, Anchorage, AK, 2008).

[19] S. H. McClure, M. J. Cree, A. A. Dorrington, and A. D. Payne, Resolving depth measurement ambiguity with commercially available range imaging cameras, in Image Processing: Machine Vision Applications III, D. Fofi, K. S. Niel eds., Proc. SPIE 7538, (2010).

[20] A. D. Payne, A. P. P. Jongenelen, A. A. Dorrington1, and M. J. Cree1, Multiple frequency range imaging to remove measurement ambiguity, in Proceedings of at 9th Conference on Optical 3-D Measurement Techniques, (Academic, Vienna, Austria, 2009), pp. 139-148.

[21] L. A. Schwarz, D. Mateus, V. Castañda, and N. Navab, Manifold Learning for ToF-based Human Body Tracking and Activity Recognition, in Proceedings of the British Machine Vision Conference, (Academic, Aberystwyth, Britain, 2010).

[22] F. Wientapper, K. Ahrens, H. Wuest, U. Bockholt, Linear-Projection-Based Classification of Human Postures in Time-of-Flight Data, in Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, (Academic, San Antonio, TX, 2009), pp. 559-564.

[23] W. K. Liao and G. Medioni, 3D Face Tracking and Expression Inference from a 2D Sequence Using Manifold Learning, in IEEE Conference on Computer Vision and Pattern Recognition, (Academic, Anchorage, AK, 2008), pp. 1-8.

[24] Q. Zhang, J. L. Liu, X. J. Guo, and W. K. Gu, 3D Measurements from imaging laser sensor: problems and strategies, Journal of Zhejiang University (Natural Science) 32, 732-738 (1998).

[25] V. N. Balasubramanian, J. Ye, and S. Panchanathan, Biased manifold embedding: a framework for person-independent head pose estimation, in IEEE Conference on Computer Vision and Pattern Recognition, (Academic, Minneapolis, MN, 2007).

[26] C. BenAbdelkader, K. Daniilidis, P. Maragos, and N. Paragios, Robust Head Pose Estimation Using Supervised Manifold Learning, in Computer Vision-ECCV 2010, C. BenAbdelkader, K. Daniilidis, P. Maragos, and N. Paragios eds. (Academic, Crete, Greece, 2010), pp. 518-531.