





Restoration of Range Images by the Gaussian Pyramid Method, Testing Different Interpolation Techniques to Select the Best Performance

Enrique Chavira Calderón  and Alejandra Cruz-Bernal 

Universidad Politécnica de Guanajuato, Cortazar, Gto, Mexico
{echavira,acruz}@upgto.edu.mx

Abstract. The inpainting method implemented in this work was used to estimate the missing information in a range image from a single image, achieved independence of the RGB image of the scene or multiple range images to perform the restoration. The proposal is based on improving the results of restoring range images using the Gaussian Pyramid method. This, finding the best interpolation technique to use in this method to estimate the missing information. Different interpolation techniques were computed and applied in order to know the best option to implement. This is carried out considering the amount of information that can be estimated, processing time and the total of information missing in the image to be restored. The method was tested with five different databases, one of which was created specifically for this work. These databases include different interior scenarios with several objects. A qualitative and quantitative comparative analysis of the obtained results was performed.

Keywords: Range image · Image restoration · Gaussian Pyramid method

1 Introduction

The range image is obtained from a stereo system or a depth sensor, with which a disparity map or a depth image is formed, respectively. These images have attractive characteristics, such as invariance to lighting, rotation or scaling as presented by Creusot [1], allowing to work with data without prior processing. However, the stereo system and depth sensors such as the Kinect present images with areas without information, because infrared light does not find a surface where it is reflected. These areas can be presented as small spaces scattered throughout or as large regions of null information (without intensity levels).

There are different works in the literature focused on the restoration of range images, which we can classify into four groups according to the information they use to perform this task. The first consists of techniques that make use of one or several RGB images of the same scene for restoration, as presented by Pertuz et al. [2] and Torres and Dudek [3]. Wang et al. [4] use multiple range and RGB images. Additionally, the inpainting method is used to estimate depth and color information together. The second group is

conformed of proposals that implement multiple range images obtained by the same sensor. Examples are found in Kolmogorov and Zabih [5] and Lin et al. [6] use a couple of images to carry out the restoration, both use the graphic cutting method for this task. Lin et al. [7] use a sequence of depth images obtained by the Xtion pro live sensor, achieving both spatial and temporal noise reduction in depth images. The third group makes use of several images, but these come from different sensors. The works presented by Zhu et al. [8] and Gudmundsson et al. [9], use images obtained by the ToF sensor, as well as, a stereo system. The authors take advantage of both systems, obtaining higher quality range images. However, prior alignment and calibration of these sensors is required. Finally, the fourth group consists of those works that use only the range image to restore. Sruthy et al. [10] and Chen et al. [11] manage to restore range images without using RGB images or multiple images. Both works detect regions without information, classify or reduce these regions in a preprocessing, and use linear and non-linear filters to restore the image.

This proposal implements an inpainting method that allows restoring the range image from the same information, without using other images (RGB or range). The processing is carried out by the Gaussian Pyramid method proposed by Ogden [12], which is mainly based on processes of reduction, enlargement and interpolation. In addition to implementing the method for range images, the work is based on improving the results of the Gaussian Pyramid by modifying the interpolation technique to be used. Usunariz in [13] uses this method but applying the 4-directions interpolation technique. The proposed work aims to improve the Gaussian Pyramid method in order to carry out the interpolation technique to be used, this to know which can be the most appropriate technique to achieve the best results. The obtained results are analyzed qualitatively and quantitatively, evaluating aspects such as the percentage of estimation achieved and the processing time. The structure of the work presented is described below. Section 2 describes the proposed methodology, as well as the interpolation techniques that were implemented. Section 3 exposes the obtained experimental results and Sect. 4 shows the conclusions based on the results.

2 Methodology

In this section, both the inpainting method for range images and the different interpolation techniques to be evaluated are described. In addition, in Sect. 2.3, it is detailed how the pixel estimation is performed applying an interpolation technique.

2.1 Restoration by Gaussian Pyramid Method



Fig. 1. Block diagram representative of the Gaussian Pyramid method to restored range image.

The proposed method for restoring range images is illustrated in Fig. 1. The first block is made up of the initialization processes, in which different matrices are created to store the corresponding information to be used in the following stages (Fig. 2, Block 1). A counter is also initialized to carry the number of reductions that make up the Gaussian reduction pyramid and a variable called “flag”, is declared with a value of one that will serve as the start and end the indicator for the next stage.

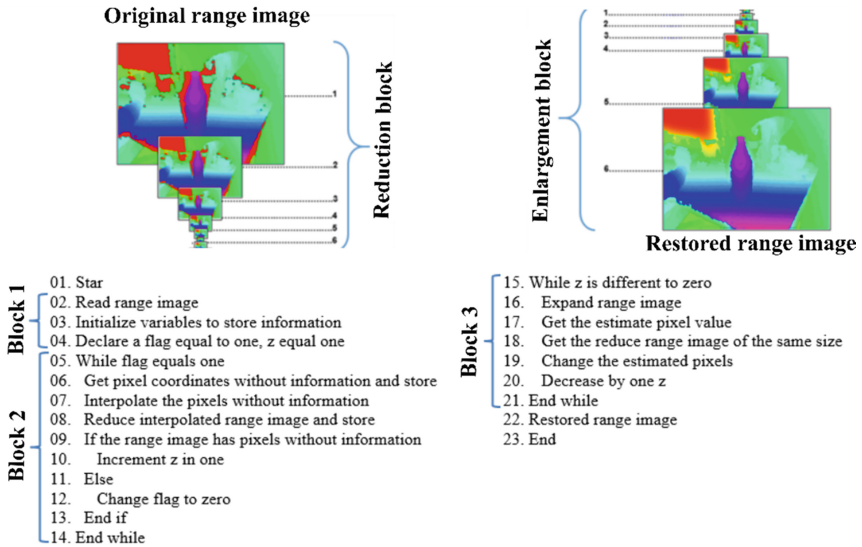


Fig. 2. Visual diagram and pseudocode of the process of restoring a range image. Block 1 corresponding to the initialization of the process; Block 2 refers to compute of the Reduction process and Block 3 aims to realize the Enlargement process.

Second block consists of a cycle for the processes of estimating information and removes areas without information (Fig. 2, Block 2). First process uses a function named “Fill_holes”. This function estimates the values of the empty pixels contain in the outline of the regions without information on the image being processed. For this task, it is proposed to implement four different interpolation techniques (see Sect. 2.2), in order to know which technique allows us to obtain the best restoration. Second process involves reducing regions without information present in the original image (Fig. 2, Reduction block). The cycle ends when there are no more zones without information, indicating the last reduction. Finally, third block consists of a cycle to carry out the enlargement of the restored but reduced image; as well as the replacement of the estimated information in each image. Applying the Gaussian Pyramid technique, the enlargement process is realized. By means of several substitution processes (Fig. 2, Block 3), an increasingly restored image is obtained (Fig. 2, Enlargement block). The cycle repeats until an image of the same size as the initial image is reached, indicating that the restored image has been obtained. The result is the restored image, which is the initial image, but with the estimated information inserted in the areas that initially did not contain information.

2.2 Interpolation Techniques

4-directions. This interpolation technique was implemented according to the technique proposed by Usunariz in [13]. Starting from a pixel without information, it is necessary to confirm that, the two consecutive neighboring pixels have information. This can happen to the right, left, up or down. Figure 3a shows a representative diagram of this technique, where the possible information to use in the interpolation process (purple area) is observed. The black box represents the pixel without information and the red point is the position at which its value is interpolated.

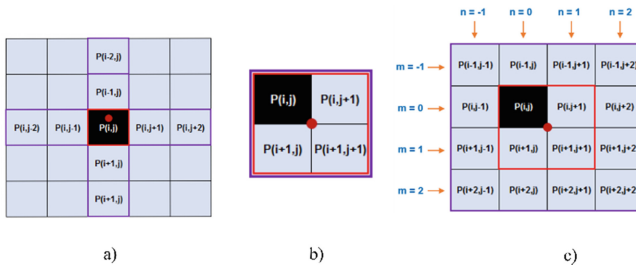


Fig. 3. Representative diagram of a) 4-directions, b) Bilinear and c) Bicubic technique. (Color figure online)

4-directions 2. Here, is presented a 4-directions technique variation. The difference with respect to the original technique consists in the values that are going interpolated and these are not consider for the following interpolations. Therefore, all interpolated values, are stored in a vector, until the interpolation process is finished, these new values are substituted in the corresponding pixels. Figure 4 shows a comparison of the two techniques.

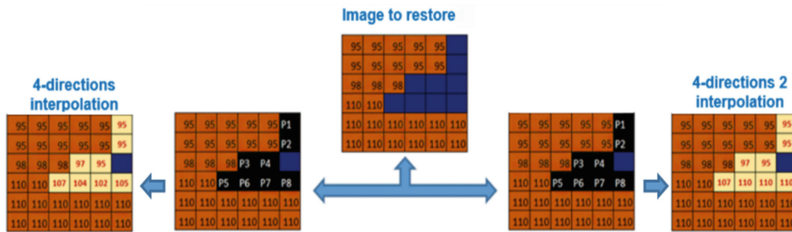


Fig. 4. Representative diagram of the estimation of pixels without information by 4-directions and 4-directions 2 interpolation techniques. The estimated value of the pixels P6, P7 and P8 changes according to the technique implemented.

Bilinear. In digital image processing, Bilinear interpolation makes use of the four pixels closest to the position to be interpolated [14]. A representative diagram of the Bilinear

interpolation for pixel estimation without information is shown in Fig. 3b. The purple area encloses the 2×2 neighborhood pixels. The black box represents the pixel without information, which can take any of the four positions within the red area. The point is the position of which the pixel value is interpolated. The estimated value is calculated with (1), which ponders the influence of the neighborhood pixels and their corresponding weights in both directions. The parameters a and b are the distances in the vertical and horizontal direction, respectively, to the point interpolated from the pixel $P(i, j)$. The function $H(x)$ is the core function of the Bilinear interpolation given by (2).

$$P(i, j) = H(-a) * H(b) * P(i, j) + H(-a) * H(-(1-b)) * P(i, j+1) \\ + H(1-a) * H(b) * P(i+1, j) + H(-(1-a)) * H(-(1-b)) * P(i+1, j+1) \quad (1)$$

$$H(x) = \begin{cases} 1 - |x|, & x \in [-1, 1] \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Bicubic. This technique considers the pixels within a 4×4 neighborhood to interpolate the desired value [15]. A representative diagram of this interpolation technique for pixel modification without information is shown in Fig. 3c. The purple area encloses 16 pixels within the 4×4 neighborhood; the black box represents the empty pixel, which can take any of the four positions within the red area. The red point is the position at which it interpolates in value of the empty pixel. The variables m and n indicate the rows and columns, respectively. This contains the neighborhood pixels in a 4×4 region for the point to be interpolated. Equation (3) allows us to calculate the value to be estimated; which is the pixel interpolated by a double sum of the product between each pixel, that is contained in the neighborhood and this is multiplied by the value of the core function $H_c(x)$, calculate with (4), in both directions.

$$P(i, j) = \sum_{m=-1}^{m=2} \sum_{n=-1}^{n=2} P(i+m, j) * H_c(m-a) * H_c(n-b) \quad (3)$$

$$H_c(x) \begin{cases} (\alpha + 2) * |x|^3 - (\alpha + 3) * |x|^2 + 1, & x \in [0, 1) \\ \alpha * |x|^3 - 5\alpha * |x|^2 + 8\alpha * |x| - 4\alpha, & x \in [1, 2) \\ 0, & x \geq 2 \end{cases} \quad (4)$$

The process for the estimation of the pixel without information through this interpolation technique considers that, several pixels within the neighborhood will take the value of zero. Therefore, the contribution within the double sum will be null by these pixels. For the above, the processing time is reduced due to redundant calculations must be avoided. That is, if should know what are the pixels that really have information within the neighborhood, then these are the ones that contribute their weight to obtain the estimate. For this, the position of the pixel $P(i, j)$ must be known, since it will act as a pivot in (3), which allows us to know according to the location of the pixel without information, the values that m and n must take.

2.3 Fill_holes Function

All scheduled versions of the “Fill_holes” function follow the same execution sequence. Figure 5 presents a diagram with the steps performed by the “Fill_holes” function in

general. The diagram begins with a matrix that represents an image that contains a region without information (the blue pixels make up the region without information while the orange pixels contain information). First step that takes place is an inverse binarization, identifying the regions without information. Second step is carried out is an emptying of all the regions present in the binary image. That is, all the pixels that do not correspond to the contour of each region will take a value of zero. In the third step, the interpolation process is performed to estimate the value of the pixels corresponding to the contour of the region without information. Here the different interpolation techniques described above are implemented, where each uses information from the neighborhood surrounding the pixel to be interpolated.

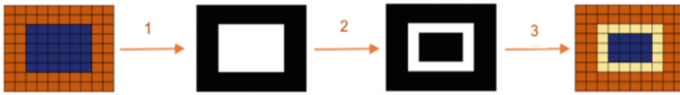


Fig. 5. Representative diagram of the process carried out by the “Fill_holes” function.

2.4 Analysis of the Amount of Missing Information in Range Images

Additionally, an analysis was carried out to know if the amount of missing information in the range images influences with respect to the estimated information rate. For this, all the images of the databases were used, being a total of 159. A statistical analysis of frequency distribution was performed, to know the average value of information estimation at different intervals of missing information. Eight classes were obtained, with a class width of 9. However, classes 5–8 were grouped into one (new class 5), because they had few elements in comparison with the other classes, making the results insignificant. The five classes, identified by C1, C2, C3, C4 and C5, contain 51, 43, 37, 14 and 14 images, respectively.

3 Results

To validate and know the performance of the proposed method using the different interpolation techniques, it was tested with five image range databases. Database B1 [16] with 43 images of 581×421 pixels, B2 generated for this research work, that contains 26 images of 627×464 pixels, B3 comprehends 26 images of the generic Middlebury database [17, 18] of 1282×1110 pixels, B4 [19, 20], with 24 images of 581×421 pixels, and B5 [21] with 40 images of 581×421 pixels. Figure 6 shows the qualitative results of applying our proposed method, using different interpolation techniques to restore range images. Two images of each databases are presented. The images in the first column presents the RGB image as reference only of the scene. Column 2 illustrates the range images to be restore. From the third to the sixth columns, the images restored by interpolation techniques are shown, corresponding to the techniques 4-directions, 4-directions 2, Bilinear and Bicubic, respectively. According to the qualitative results, since the implementation of the proposed method with interpolation in 4-directions, an

image restoration is obtained (Fig. 6, column 3). However, this technique can add non-existent artifacts (Fig. 6, row 5, column 3) due to the how interpolation is performed. The second technique, 4-directions 2, does not present this problem (Fig. 6, row 5, column 4); this is due to not considering the interpolated pixels for the following interpolations. In the same way, an improvement is qualitatively obtained in several estimated regions, this with respect to the 4-directions (Fig. 6, row 2, column 4). Bilinear and Bicubic techniques offer the best qualitative results; because increasing the number of estimated regions with a major homogeneity with respect to intensity values (Fig. 6, row 1, 3–4, column 5–6).

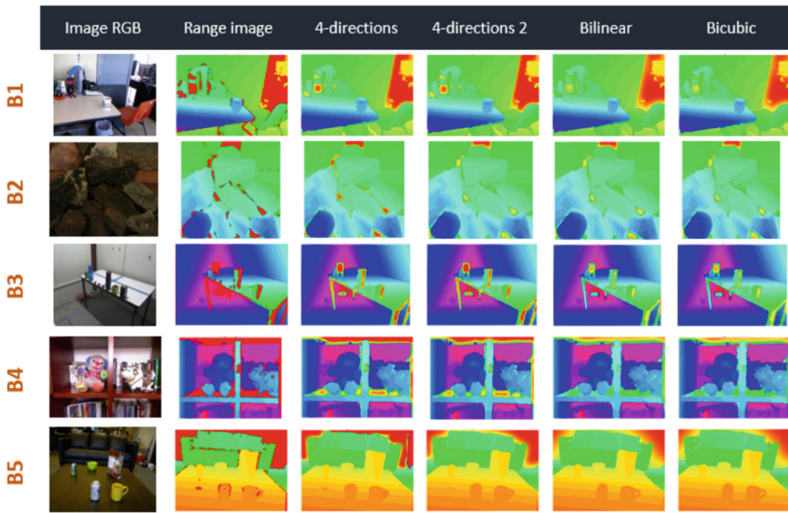


Fig. 6. Qualitative results of the restoration applying our proposed method with different interpolation techniques using several generic and own databases. Range images are presented in the HSV color space, where regions in red represent pixels without information.

To compare and determine which of these techniques is the most appropriate to use in our estimation method, three metrics were calculated to evaluate them: the mean square error (MSE), the estimated information rate and the processing time. Figure 7 presents the obtained results for the estimated information rate (Fig. 7b) and the processing time (Fig. 7a), applying the four interpolation techniques in each database. Considering the individual average of each technique in the different databases, we calculate a global average to compare them with each other and know the overall result of implementing these techniques in range images. First, we present the global results regarding the percentage of estimated information: For the 4-directions technique, 79.33% is presented, for 4-directions 2 with of 79.9%, for Bilinear and Bicubic, 83.76% and 83.57%, respectively. According to the above, the 4-directions, the original technique, is the one that presents the lowest estimated information rate. The technique 4-directions 2, Bilinear and Bicubic, presents 0.71%, 5.58% and 5.34%, respectively, of increase with respect to the 4-directions. Now we present the global results regarding the processing time:

For the 4-directions a 0.84 s, 4-directions 2 an average of 1.04 s, for the Bilinear and Bicubic interpolation, there are 30.60 s and 12.23 s, respectively. Therefore, the original technique is the one that presents the best processing time, followed by the 4-directions 2, then the Bicubic and finally, the Bilinear. For the MSE, all techniques have a value greater than 97%, so it is ensured that the original information is maintained in the depth image, regardless of the interpolation technique implemented.

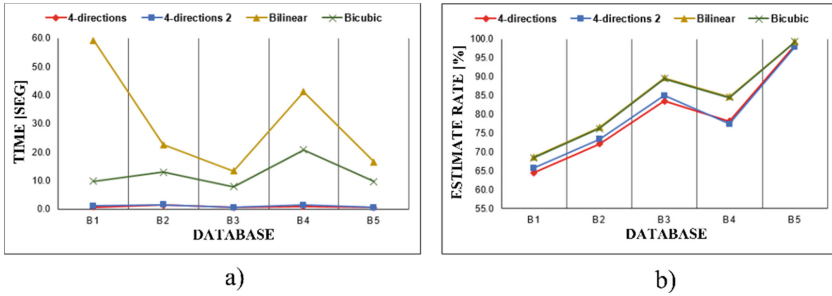


Fig. 7. Quantitative results in the 5 databases, applying the different interpolation techniques. a) Results of the processing time. b) Results of the estimated information rate.

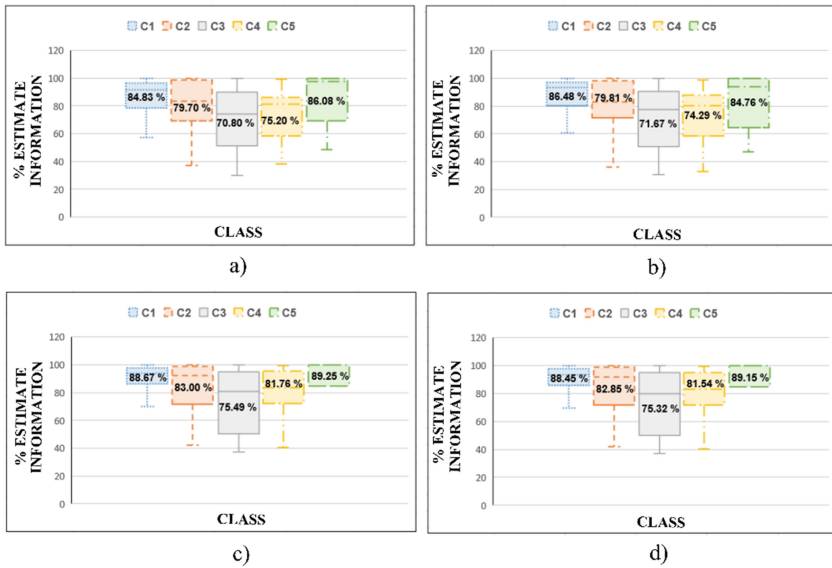


Fig. 8. Estimated information rate by class, applying the different interpolation techniques. a) 4-directions, b) 4-directions 2, c) Bilinear and d) Bicubic.

The results of the estimated information rate with respect to the percentage of missing information presented by the range images are presented below. First, for the 159

images, with a maximum missing information of 68.2% and a minimum of 0.5%. The analysis was carried out in all classes, applying several interpellation techniques, the results are shown in Fig. 8. The 4-directions, Bilinear and Bicubic techniques have a maximum percentage of estimated pixels for C5, images with the greatest lack of information. C3 is the one that, regardless of the interpolation technique implemented, presents the minimum amount of estimated information. The 4-directions 2 technique allows obtaining a higher estimation rate for images belonging to class C1. Therefore, the present proposal offers repetitiveness and consistency with respect to the different interpolation techniques here tested.

4 Conclusions

The implementation of the Gaussian Pyramid method allows the restoration of range images from the same image, without making use of multiple range images or RGB image. The interpolation technique used in this method does affect the result of the restoration of the range images. The original interpolation technique, 4-directions, can present non-existent artifacts in the restored image. However, this problem does not occur when applying the other interpolation techniques. According to the amount of information missing in the range images, the interpolation techniques in 4-directions, Bilinear and Bicubic, have a maximum of estimated pixels for the C5. The 4-directions 2 technique has a value maximum estimate for C1. Finally, it is concluded that, if the processing time is not an important point to consider, the Bilinear interpolation technique is the recommended technique to implement. If the processing time is important, it is recommended to implement the 4-directions 2 technique, since, although it has a longer time than the original technique, it allows to increase the amount of estimated pixels and does not show the problem of unwanted artifacts.

Future work implies the restored range images will be used to perform object detection tests, whose results in processing time and detection accuracy will be compared with those obtained from the detection of objects in range images without process.

References

1. Creusot, C.: A machine-learning approach to keypoint detection and landmarking on 3D meshes. *Int. J. Comput. Vision* **102**, 146–179 (2013)
2. Pertuz, S.: Region-based depth recovery for highly sparse depth maps. In: *Proceedings - International Conference on Image Processing*, pp. 2074–2078 (2018)
3. Torres, L.: Reconstruction of 3D models from intensity images and partial depth. In: *AAAI*, pp. 476–481 (2004)
4. Wang, L.: Stereoscopic inpainting: joint color and depth completion from stereo images. In: *26th IEEE Conference on Computer Vision and Pattern Recognition* (2008)
5. Kolmogorov, V.: Computing visual correspondence with occlusions using graph cuts. In: *Proceedings of the Eighth IEEE International Conference on Computer Vision*, vol. 2, pp. 508–515 (2001)
6. Lin, M.: Surfaces with occlusions from layered stereo. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2003)

7. Lin, B.: Temporal and spatial denoising of depth maps. *Sensors (Basel, Switzerland)* **15**, 18506–18525 (2015)
8. Zhu, J.: Fusion of time-of-flight depth and stereo for high accuracy depth maps. In: 26th IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
9. Gudmundsson, S.: Fusion of stereo vision and Time-Of-Flight imaging for improved 3D estimation. *Int. J. Intell. Syst. Technol. Appl.* **5**, 425 (2008)
10. Sruthy, T.: Novel method for depth map pre-processing in depth image based rendering. *Int. J. Eng. Sci. Res. Technol.* **6**, 185–190 (2017)
11. Chen, W.: Efficient depth image based rendering with edge dependent depth filter and Interpolation. In: IEEE International Conference on Multimedia and Expo, pp. 1314–1317 (2005)
12. Ogden, J.: Pyramid-based computer graphics. *RCA Eng.* **30**(5), 4–15 (1985)
13. Usunariz, I.: Tesis Restauración de imágenes mediante Pirámide Gaussiana y técnicas de reducción y ampliación. Universidad Publica de Navarra (2014)
14. Han, D.: Comparison of commonly used image interpolation methods. In: Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering, vol. 1, pp. 1556–1559 (2013)
15. Qifang, X.: Super-resolution reconstruction of satellite video images based on interpolation method. *Procedia Comput. Sci.* **107**, 454–459 (2017)
16. Bernal, A.: Object detection from range imagen using the sparse keypoint detector technique. *IEEE Latin Am. Trans.* **16**(5), 1532–1538 (2018)
17. Scharstein, D.: High-accuracy stereo depth maps using structured light. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 195–202 (2003)
18. Scharstein, D.: Learning conditional random fields for stereo. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2007)
19. Yaman, M.: Multimodal stereo vision using mutual information with adaptive windowing. In: 13th IAPR Conference on Machine Vision and Applications (2013)
20. Yaman, M.: An iterative adaptive multi-modal stereo-vision method using mutual information. *J. Vis. Commun. Image Represent.* **26**, 115–131 (2015)
21. Lai, K.: Unsupervised feature learning for 3D scene labeling. In: IEEE International Conference on Robotics and Automation, pp. 3050–3057 (2014)