Vision-based Interpolation of 3D Laser Scans

Henrik Andreasson[1], Rudolph Triebel[2] and Achim Lilienthal[1]

[1]Department of Technology
Örebro University, Örebro, Sweden
[2]Department of Computer Science
University of Freiburg, Freiburg, Germany
henrik.andreasson@tech.oru.se, triebel@informatik.uni-freiburg.de, achim.lilienthal@tech.oru.se

Abstract

3D range sensors, particularly 3D laser range scanners, enjoy a rising popularity and are used nowadays for many different applications. The resolution 3D range sensors provide in the image plane is typically much lower than the resolution of a modern colour camera. In this paper we focus on methods to derive a high-resolution depth image from a low-resolution 3D range sensor and a colour image. The main idea is to use colour similarity as an indication of depth similarity, based on the observation that depth discontinuities in the scene often correspond to colour or brightness changes in the camera image. We present five interpolation methods and compare them with an independently proposed method based on Markov Random Fields. The algorithms proposed in this paper are non-iterative and include a parameter-free vision-based interpolation method. In contrast to previous work, we present ground truth evaluation with real world data and analyse both indoor and outdoor data. Further, we suggest and evaluate four methods to determine a confidence measure for the accuracy of interpolated range values.

Keywords: 3D range sensor, laser range scanner, vision-based depth interpolation, 3D vision

1 Introduction

3D range sensors are getting more and more common and are found in many different areas. A large research area deals with acquiring accurate and very dense 3D models, potential application domains include documenting cultural heritage [1], excavation sites and mapping of underground mines [2]. A lot of work has been done in which textural information obtained from a camera is added to the 3D data. For example, Sequeira et al. [3] present a system that creates textured 3D models of indoor environments using a 3D laser range sensor and a camera. Früh and Zakhor [4] generate photo-realistic 3D reconstructions from urban scenes by combining aerial images with textured 3D data acquired with a laser range scanner and a camera mounted on a vehicle.

In most of the approaches that use a range scanner and a camera, the vision sensor is not actively used during the creation of the model. Instead vision data is only used in the last step to add texture to the extracted model. An exception is the work by Haala and Alshawabkeh [5], in which the camera is used to add line features detected in the images into the created model.

To add a feature obtained with a camera to the point cloud obtained with a laser range scanner, it is required to find the mapping of the 3D laser points onto pixel coordinates in the image. If the focus instead lies on using the camera as an active source of information which is considered in this paper, the fusing part in addition addresses the question of how to estimate a 3D position for each (sub) pixel in the image. The resolution that the range sensor can provide is much lower than obtained with a modern colour camera. This can be seen by comparing figure 1, created by assigning the intensity value of the projected laser point to its closest neighbors, with the corresponding colour image in figure 1.

To our knowledge the only approach that uses colour information from a camera image to obtain a high-resolution 3D point model from a low-resolution 3D range scan is the algorithm by Diebel et al. [6], where both colour information and the raw depth information are used. Their method is also compared with the methods suggested in this paper and is further described in section 3.

2 Suggested Vision-based Interpolation Approaches

The main idea is to interpolate low-resolution range data provided by a 3D laser range scanner under the assumption that depth discontinuities in the scene often correspond to colour or brightness changes in the camera image of the scene.

For the problem under consideration, a set of N laser range measurements $\mathbf{r}_1..\mathbf{r}_N$ is given where each measurement $\mathbf{r}_i = (\theta_i, \pi_i, r_i)$ contains a tilt angle θ_i, a pan angle π_i and a range reading $\mathbf{r}_i$ corresponding to 3D Euclidean coordinates (x_i, y_i, z_i).

Figure 1: Left: Image intensities plotted with the resolution of the 3D scanner. The laser range readings were projected onto the right image and the closest pixel regions were set to the intensity of the projected pixel for better visualisation. Right: Calibration board used for finding the external parameters of the camera, with a chess board texture and reflective tape (gray border) to locate the board in 3D using the remission / intensity values from the laser scanner.

The image data consists of a set of image pixels $\mathbf{P}_j = (X_j, Y_j, C_j)$, where X_j, Y_j are the pixel coordinates and $C_j = (C_j^1, C_j^2, C_j^3)$ is a three-channel colour value. By projecting a laser range measurement $\mathbf{r}_i$ onto the image plane, a projected laser range reading $\mathbf{R}_i = (X_i, Y_i, r_i, (C_i^1, C_i^2, C_i^3))$ is obtained, which associates a range reading $\mathbf{r}_i$ with the coordinates and the colour of an image pixel. An image showing the projected intensities can be seen in figure 1, where the closest pixel regions are set to the intensity of the projected pixel for better visualisation.

The interpolation problem can now be stated for a given pixel $\mathbf{P}_j$ and a set of projected laser range readings $\mathbf{R}$, as to estimate the interpolated range reading r_j^* as accurately as possible. Hence we denote an interpolated point $\mathbf{R}_j^* = (X_j, Y_j, r_j^*, C_j^1, C_j^2, C_j^3)$.

Five different interpolation techniques are described in this section and compared with the MRF approach described in section 3.

2.1 Nearest Range Reading (NR)

Given a pixel $\mathbf{P}_j$, the interpolated range reading r_j^* is assigned to the laser range reading r_i corresponding to the projected laser range reading $\mathbf{R}_i$ which has the highest likelihood p given as

$$p(\mathbf{P}_j, \mathbf{R}_i) \propto e^{-\frac{(X_j - X_i)^2 + (Y_j - Y_i)^2}{\sigma^2}} \tag{1}$$

where σ is the point distribution variance. Hence, the range reading of the closest point (regarding pixel distance) will be selected.

2.2 Nearest Range Reading Considering Colour (NRC)

This method is an extension of the NR method using colour information in addition. Given a pixel $\mathbf{P}_j$, the interpolated range reading r_j^* is assigned to the range
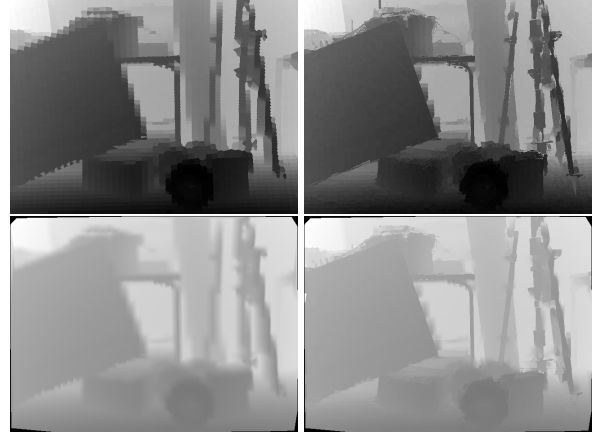

Figure 2: Top left: Depth image generated with the NR method. Top right: Depth image generated with the NRC method, small details are now visible. Note that a depth image generated from a similar viewpoint as the laser range scanner makes it very difficult to see flaws of the interpolation algorithm. Bottom left : MLI. Bottom right : LIC.
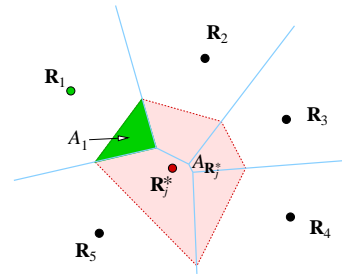

Figure 3: Natural neighbors $\mathbf{R}_1..\mathbf{R}_5$ of $\mathbf{R}_i^*$. The interpolated weight of each natural neighbor $\mathbf{R}_i$ is proportional to the size of the area which contains the points Voronoi cell and the cell generated by $\mathbf{R}_j^*$. I.e. the nearest neighbor $\mathbf{R}_1$ will have influence based upon the area of A_1.

value r_i of the projected laser range reading $\mathbf{R}_i$ which has the highest likelihood p given as

$$p(\mathbf{P}_j, \mathbf{R}_i) \propto e^{-\frac{(X_j - X_i)^2 + (Y_j - Y_i)^2}{\sigma_p^2} - \frac{||C_j - C_i||^2}{\sigma_c^2}} \tag{2}$$

where σ_p and σ_c is the variance for the pixel point and the colour respectively.

2.3 Multi-Linear Interpolation (MLI)

Given a set of projected laser range readings $\mathbf{R}_1..\mathbf{R}_N$, a Voronoi diagram V is created by using their corresponding pixel coordinates $[X, Y]_{1..N}$. The natural neighbors NN to an interpolated point $\mathbf{R}_j^*$ are the points in V, which Voronoi cell would be affected if $\mathbf{R}_j^*$ is added to the Voronoi diagram, see figure 3. By inserting $\mathbf{R}_j^*$ we can obtain the areas $A_{1..n}$ of the intersection between the Voronoi cell due to $\mathbf{R}_j^*$ and the Voronoi cell of $\mathbf{R}_i$ before inserting $\mathbf{R}_j^*$ and the area $A_{\mathbf{R}_j^*}$

as a normalisation factor. The weight of the natural neighbor $\mathbf{R}_i$ is calculated as

$$w_i(\mathbf{R}_j^*) = \frac{A_i}{A_{\mathbf{R}_j^*}}. \tag{3}$$

The interpolated range reading r_j^* is then calculated as

$$r_j^* = \sum_{i \in NN(\mathbf{R}_j^*)} w_i r_i. \tag{4}$$

This interpolation approach is linear [7]. One disadvantage is that nearest neighborhood can only be calculated within the convex hull of the scan-points projected to the image. However, this is not considered as a problem since the convex hull encloses almost the whole image, see figure 2.

2.4 Multi-Linear Interpolation Considering Colour (LIC)

To fuse colour information with the MLI approach introduced in the previous subsection, the areas $A_{\mathbf{R}_i}$ and $A_{\mathbf{R}_j^*}$ are combined with colour weights $w_{1..n}^c$ for each natural neighbor based on spatial distance in colour space.

Similar as in section 2.2, a colour variance σ_c is used:

$$w_i^c(\mathbf{R}_j^*) = e^{-\frac{\|C_i - C_j\|^2}{\sigma_c^2}}. \tag{5}$$

The colour based interpolated range reading estimation is then done with

$$r_j^* = \sum_{i \in NN(\mathbf{R}_j)} \frac{w_i w_i^c}{W^c} \tag{6}$$

where $W^c = \sum_{i=1}^n w_i^c$ is used as a normalisation factor.

2.5 Parameter-Free Multi-Linear Interpolation Considering Colour (PLIC)

One major drawback of the methods presented so far and the approach presented in the related work section is that they depend on parameters such as σ_c, for example. To avoid the need to specify colour variances, the intersection area $A_{\mathbf{R}_i}$ defined in Section 2.3 is used to compute a colour variance estimate for each nearest neighbor point $\mathbf{R}_i$ as

$$\sigma_c = \frac{1}{n_i - 1} \sum_{j \in A_i} \|\mu_i - C_j\|^2 \tag{7}$$

where $\mu_i = \frac{1}{n_i} \sum_{j \in A_i} C_j$ and n_i is the number of pixel points within the region A_i.

This results in an adaptive adjustment of the weight of each point. In case of a large variance of the local surface texture, colour similarity will have less impact on the weight w_i.

3 Related Work

To our knowledge, the only work using vision for interpolation of 3D laser data is [6] where a Markov Random Field (MRF) framework is used. The method works by iteratively minimising two constraints: ψ stating that the raw laser data and the estimated depth should be similar and ϕ stating that the depth estimates close to each other with a similar colour should also have similar depths.

$$\psi = \sum_{i \in N} k(r_i^* - r_i)^2 \tag{8}$$

where k is a constant and the sum runs over the set of N positions which contain a laser range reading r_i and r_i^* is the interpolated range reading for position i. The second constraint is given as

$$\phi = \sum_i \sum_{j \in NN(i)} e^{(-c\|C_i - C_j\|^2)}(r_i^* - r_j^*)^2 \tag{9}$$

where c is a constant, C is the pixel colour and $NN(i)$ are the neighborhood pixel around position i.

The function to be minimised is the sum $\psi + \phi$.

4 Evaluation

Experimental evaluation is a crucial point and has been done using both simulated and real data. All data sets D were divided into two equally sized parts D_1 and D_2. One dataset, D_1, is used for interpolation and D_2 is used as the ground truth where each laser range measurement is projected to image coordinates. Hence for each ground truth point $\mathbf{R}_i$ we have the pixel positions $[X,Y]_i$ and the range r_i. The pixel position $[X,Y]_i$ is used as input to the interpolation algorithm and the range r_i is used as the ground truth. The performance of the interpolation algorithms is analysed based on the difference between the interpolated range and the range from the ground truth.

5 Experimental Setup

5.1 Hardware

The scanner used is a 2D SICK LMS-200 mounted together with a 1 MegaPixel (1280x960) colour CCD camera on a pan-tilt unit from Amtec where the displacement between the optical axis is approx 0.2 m. The scanner is located on our outdoor robot, see figure 4, a P3-AT from ActivMedia. The SICK scanner has a larger spot size compared to many other laser scanners and often gives wrong range estimates close to edges where the laser spot covers multiple objects at different distances, see figure 4. Of course, this flaw of the sensor will be reflected in the ground truth as well. The angular resolution of the laser scanner is 0.5
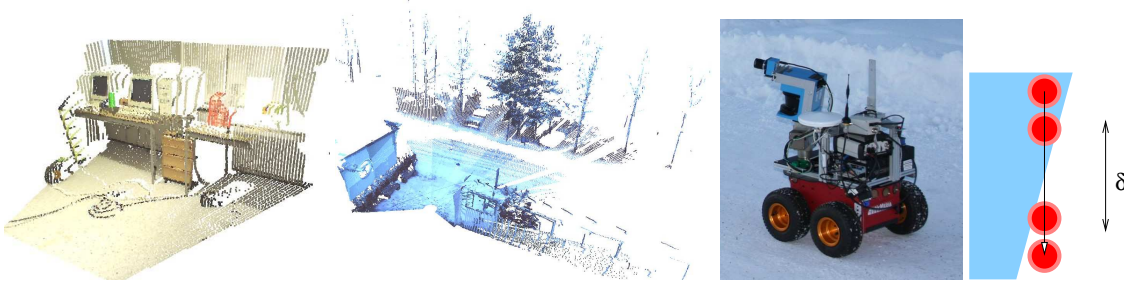
Figure 4: Left: The third indoor evaluation scan, *Indoor₃*. Middle left: Scans taken in winter time with some snow. (The figure consist of 7 fused scans. Only 3 of them are used in the evaluation presented in this paper). Middle right: Our outdoor robot with the SICK LMS scanner and a colour CCD camera mounted on a pan tile unit from Amtec, that were used in the experiments. Right: When the laser range finder spot covers an area which contain different depths (blue and white areas), the range reading returned might be unreliable and vary anywhere between the closest to the furthest range (shown as the region δ).

degrees. Half of the readings were used as ground truth, so the resolution for the points used for interpolation is 1 degree.

6 Results - Interpolation

The most common colour spaces were also compared to evaluate if better illuminance/shading invariance could be useful. The colour spaces compared were standard RGB, Normalised RGB, HSV and YUV. Since a consistent improvement could not be observed for neither of the colour spaces tested, only results based on standard RGB normalised to [0,1] are presented in this paper.

In all experiments the colour variance $\sigma_c = 0.05$ and the pixel distance variance $\sigma_d = 10mm$ where used, which were found empirically. The parameters used within the MRF approach described in section 3, where obtained by extensive empirical testing and were set to $k = 2$ and $C = 10$. The optimisation method used for this method was the conjugate gradient method described in [8]. In all experiments the full resolution (1280x960) of the camera image was used.

All the interpolation algorithms described in this paper were tested on real data consisting of three indoor scans and two outdoor scans. The outdoor scans were taken in winter time with snow, which presents the additional challenge that most of the points in the scene have very similar colours.

The results are summarised in table 1 and table 2, which show the mean error with respect to the ground truth $\bar{e}$, and the percentage of outliers o_t for different thresholds t. The percentage of outliers is the percentage of points for which the interpolated range value deviates from the ground truth value by more than a threshold t (specified in meters in table 1 and table 2).

Table 1: Results from $Indoor_1 - Indoor_3$ data sets.

	NR	NRC	MLI	LIC	PLIC	MRF
$\bar{e}$	0.092	0.087	0.076	**0.072**	0.073	0.074
$o_{0.1}$	0.139	**0.119**	0.155	0.123	0.126	0.141
$o_{0.2}$	0.096	0.084	0.091	**0.079**	0.080	0.083
$o_{0.5}$	0.050	0.048	0.029	0.036	0.037	**0.030**
$o_{1.0}$	**0.008**	0.022	0.011	0.012	0.012	0.011
$o_{3.0}$	0.003	0.004	0.003	0.003	**0.002**	0.003

For the indoor data sets, which comprise many planar structures, the lowest mean error was found with the multi-linear interpolation methods, particularly LIC and PLIC, and MRF interpolation. LIC and PLIC produced less (but larger) outliers.

With the outdoor data the results obtained were more diverse. For the data set *Outdoor₁*, which contains some planar structures, a similar result as in the case of the indoor data was observed. For data sets with a very small portion of planar structures such as *Outdoor₂*, the mean error was generally much higher and the MRF method performed slightly better compared to the multi-linear interpolation methods. This is likely due to the absence of planar surfaces and the strong similarity of the colours in the image recorded at winter time. It is noteworthy that in this case, the nearest neighbor interpolation method *without* considering colour (NR) performed as good as MRF. The interpolation accuracy of the parameter-free PLIC method was always better or comparable to the parameterised method LIC.

7 Confidence Measure

The interpolated range reading r_j^* may be a good estimate of the actual range or it might deviate substantially from the true value. Therefore a confidence measure for the correctness of the interpolated range reading estimate is desirable, allowing to detect and handle errornous measures appropriately.
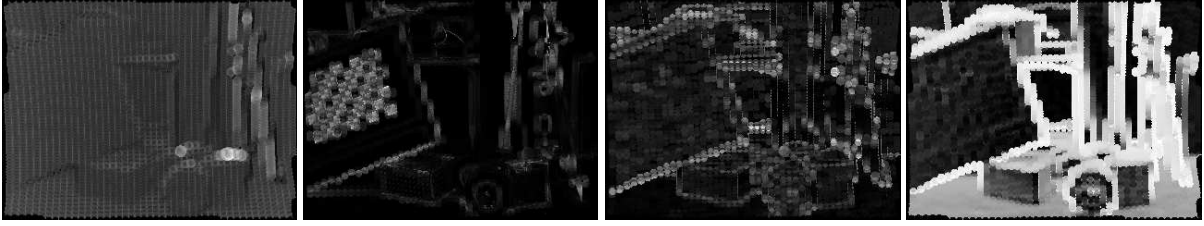
Figure 5: Visualisation of the confidence measures suggested. From left to right: NLR showing the distance to the closest point, NLRC using colour distance, PS showing the plane factor of the neighborhood of the interpolated point and AON showing the angle difference between the normal of the extracted local plane and the camera axis. The parameter free method (PLIC) were used.

Table 2: Results from *Outdoor*$_1$ and *Outdoor*$_2$.

	NR	NRC	MLI	LIC	PLIC	MRF
$\bar{e}$	0.067	0.068	0.056	0.059	**0.054**	**0.054**
$o_{0.1}$	0.147	0.160	0.156	0.146	**0.138**	0.150
$o_{0.2}$	0.076	0.080	0.078	0.073	**0.068**	0.076
$o_{0.5}$	0.032	0.032	0.016	0.020	**0.015**	0.016
$o_{1.0}$	0.005	0.002	0.001	**0.001**	0.002	**0.001**
$o_{3.0}$	**0.000**	**0.000**	0.001	0.001	**0.000**	**0.000**
$\bar{e}$	0.219	0.294	0.235	0.322	0.275	**0.218**
$o_{0.1}$	0.196	0.240	0.242	0.269	0.264	**0.187**
$o_{0.2}$	**0.096**	0.152	0.140	0.168	0.160	0.098
$o_{0.5}$	**0.047**	0.088	0.077	0.094	0.083	0.051
$o_{1.0}$	0.036	0.057	0.043	0.059	0.049	**0.030**
$o_{3.0}$	**0.016**	0.023	0.019	0.028	0.022	0.017

In this paper we suggest and evaluate four different confidence measures.

7.1 Proximity to the Nearest Laser Range Reading (NLR)

This confidence measure is based on the distance between the pixel position of the interpolated point $\mathbf{R}_j^*$ to the nearest projected laser range reading $\mathbf{R}_i$. The idea is that if the interpolated pixel point is close to a point where a range measurement is available the interpolation is considered more trustworthy.

$$NLR(\mathbf{R}_j^*, \mathbf{R}_i) = e^{-\sqrt{(X_j^* - X_i)^2 + (Y_j^* - Y_i)^2}} \qquad (10)$$

7.2 Proximity to the Nearest Laser Range Reading Considering Colour (NLRC)

This confidence measure is based on the distance between the colour of the pixel of the nearest projected laser range reading $\mathbf{R}_i$ and $\mathbf{R}_i^*$. This confidence measure works better since it takes into account that confidence in the interpolated value should decrease if the two points have different colour.

$$NLRC(\mathbf{R}_j^*, \mathbf{R}_i) = e^{-||C_j - C_i||} \qquad (11)$$

7.3 Degree of Planar Structure (PS)

Our confidence in the range interpolation also depends on how well a planar surface can be fitted to the local neighbors $NN(\mathbf{R}_j^*)$ of the interpolation point $\mathbf{R}_j^*$ since planar surfaces support a linear interpolation technique very well. The neighbors are either determined from the grid defined by the projected laser range readings or the nearest neighbors found in the Voronoi tessellation. The parameters of the planar surface are obtained from the 3D covariance matrix of $NN(j)$ where the two main eigenvectors are extracted which span a planar surface S_j with the normal vector $\mathbf{n}_j$. The confidence measure is then calculated from the average distance of the local neighbors to the fitted plane as

$$PS(\mathbf{R}_j^*) = e^{-\frac{1}{NN}\sum_{i \in NN(\mathbf{R}_j^*)} ||\mathbf{r}_i \cdot \mathbf{n}_j - d_j||} \qquad (12)$$

where d_j is the distance of the plane S_j to the origin and $\mathbf{r}_i = (x_i, y_i, z_i)$ is the 3D position of point i.

7.4 Angle Between the Optical Axis and the Fitted Plane Normal (AON)

This confidence measure considers the orientation of the planar surface S_j described in the previous section relative to the optical axis of the camera $\mathbf{z}_{cam}$. If the angle between the normal vector $\mathbf{n}_j$ and the optical axis is small, the confidence should be high since we expect only one reflexion from the laser scanner and the displacement between the laser and the camera will have a negligible impact.

$$AON(\mathbf{R}_j^*) = \mathbf{z}_{cam} \cdot \mathbf{n}_j \qquad (13)$$

8 Result - Confidence Measure

With the exception of the NLR method, a distinct negative correlation was found for all the confidence measures suggested in this section. Due to the experimental setup where the evaluation points were taken from the laser scanner in an evenly spaced grid the parallax errors caused by the displacement correspond to a low distance between $\mathbf{R}_j^*$ and $\mathbf{R}_i$ which gave the proposed
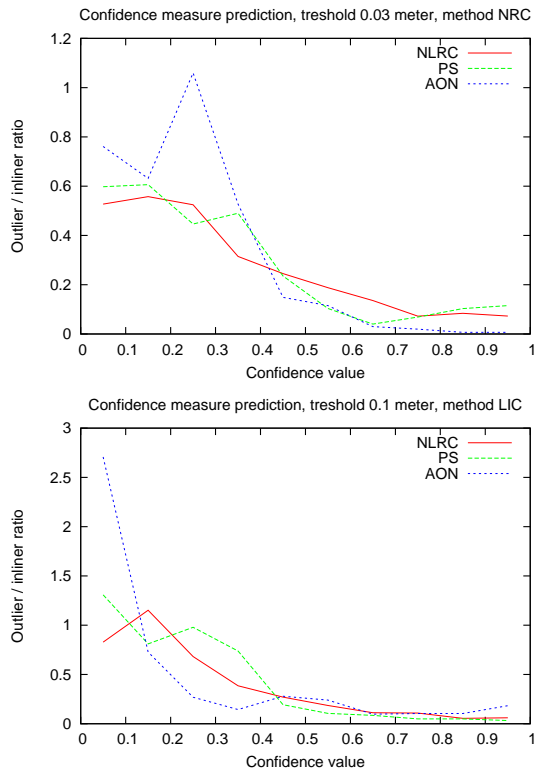
Figure 6: Suitability of the confidence measures introduced in this paper. The graphs show the number of outlier / number of inlier ratio, depending on the confidence in the interpolated points. All points with a depth error > 0.03 meter are considered outliers in the upper image and in the lower graph the threshold was 0.1 meter. Top: $Indoor_1$ data set with method NRC. Bottom : $Outdoor_2$ with method LIC.

NLR method to give high confidence correlated with parallax errors.

Figure 6 shows the inlier/outlier ratio depending on the confidence calculated with the NLRC, PS, and AON method. Interpolated range values were classified as outliers if the deviation from the ground truth value was larger than approximately a third of the mean error obtained with the particular interpolation method. The same general trend of a clear negative correlation, however, was observed with all interpolation methods and for all data sets.

9 Conclusions

This paper is concerned with methods to derive a high-resolution depth image from a low-resolution 3D range sensor and a colour image. We suggest five interpolation methods and compare them with an alternative method proposed by Diebel and Thrun [6]. In contrast to previous work, we present ground truth evaluation with real world data and analyse both indoor and outdoor data. The results of this evaluation do not allow to single out one particular interpolation method that provides a distinctly superior interpolation accuracy, indicating that the best interpolation method depends on the content of the scene. Altogether, the MRF method proposed in [6] and the PLIC method proposed in this paper provided the best interpolation performance. While providing basically the same level of interpolation accuracy as the MRF approach, the PLIC method has the advantage that it is a parameter-free and non-iterative method, i.e. that a certain processing time can be guaranteed. We further suggest and evaluate four methods to determine a confidence measure for the accuracy of interpolated range values. The confidence value calculated showed a distinct negative correlation with the occurrence of outliers. This was observed independent of the scene content and the interpolation method applied.

10 References

[1] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade and D. Fulk, "The Digital Michelangelo Project: 3D scanning of large statues", *Proceedings SIGGRAPH*, pp 131–144 (2000).

[2] S. Thrun, D. Hähnel, D. Ferguson, M. Montemerlo, R. Triebel, W. Burgard, C. Baker, Z. Omohundro, S. Thayer and W. Whittaker, "A system for volumetric robotic mapping of abandoned mines", *Proceedings ICRA*, Taipei, Taiwan, pp 4270–4275 (2003).

[3] V. Sequeira, J. Goncalves and M.I. Ribeiro, "3d reconstruction of indoor environments", *Proceedings ICIP*, Lausanne, Switzerland, pp 405–408 (1996).

[4] C. Früh and A. Zakhor, "3D model generation for cities using aerial photographs and ground level laser scans", *Proceedings CVPR*, Hawaii, USA, pp 31–38 (2001).

[5] N. Haala and Y. Alshawabkeh, "Application of photogrammetric techniques for heritage documentation", *Proceedings 2nd Int. Conf. on Science & Technology in Archaeology & Conservation*, Amman, Jordan (2003).

[6] J. Diebel and S. Thrun, "An application of markov random fields to range sensing", *Proceedings NIPS*, Vancover, Canada, pp 291–298 (2005).

[7] R. Sibson, "A brief description of natural neighbour interpolation", in V. Barnet (ed.), *Interpolating Multivariate Data*, John Wiley & Sons, Chichester, pp 21–36 (1981).

[8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge (UK) and New York (1992).