3D RECONSTRUCTION SCENE ERROR ANALYSIS

Rogério Yugo Takimoto
Marcos de Sales Guerra Tsuzuki
Thiago de Castro Martins
Edson Kenji Ueda
Computational Geometry Laboratory, Escola Politécnica da USP
takimotoyugo@gmail.com, mtsuzuki@usp.br, thiago@usp.br, ueda.edson@gmail.com

Toshiyuki Gotoh
Seiichiro Kagei
Yokohama National University, Japan
gotoh@sci.ynu.ac.jp,kagei@ynu.ac.jp

Abstract. *The analysis and estimation is an essential step towards to a reliable reconstruction process and for a proper use of 3D data. Therefore it is interesting to analyze which error sources influence the final result, and what is the sensitivity of each of these error sources. The objective of this work is to perform the error analysis of a reconstructed surface. The 3D reconstructed scene has the precision improved by combining point clouds obtained from different viewpoints using structured light. The main task is the point cloud registration algorithm that matches two point clouds. A well-known algorithm for point cloud registration is the ICP that determines the rotation and translation that when applied to one of the point clouds, place both point clouds in accordance. An improved ICP is used in the error analysis. The quality of the 3D reconstruction is described as a function of input data and the correctness of the reconstructed scene is compared with the real surface measurement.*

Keywords: *Structured Light;3D Reconstruction;ICP;Error analysis.*

1. INTRODUCTION

3D reconstruction is an important subject in computer vision. 3D models generation of static real world objects has reached increased importance in many fields of application, such as manufacturing, medicine, reverse engineering, prototyping and also in the design of virtual worlds for movies and games. To provide a reliable reconstruction process the error analysis and estimation is required for a proper use of 3D data. The analysis of which error source influences the final result, and what is the sensitivity of each of these error sources are of particular interest.

A well known algorithm for point cloud registration is the Iterative Closest Point algorithm (ICP) (Besl and McKay, 1992), it has become the most widely-used approach for aligning 3D surfaces, especially for surfaces created from range images. The ICP determines the rotation and translation that when applied to one of the point clouds, place both point clouds in accordance. In the ICP, two main steps are executed iteratively: point correspondence and transformation determination. Once the mapping is established, a variety of tasks can be performed using the aligned object representation including model based localization and 3D object recognition.

If the point correspondence determination is not executed properly the ICP might converge to a local minimum and the registration of two surfaces is not guaranteed to be successful. The failure in the registration process can occur for some reasons, such as when the overlap between the two surfaces is insufficient, or when the range measurement error is too large, or when the geometric constraint on the 3D rigid-body transformation between the two surfaces is insufficient (for example, when a plane is being registered to another plane) or when the initial relative pose between the two surfaces is too large. A conventional root-mean-squared error, coupled with the criterion measure from constraint analysis provides the ability to estimate true registration accuracy online and can be used to guide the registration data by determining when additional data is required to satisfy accuracy requirements and is built upon the registration constraint analysis (Simon, 1996). In Simon (1996) the relative amount of constraint on the 3D rigid-body transformation exerted by the shape of a surface is evaluated. In his approach, a set of surface points, a value that represents the amount of constraint on the 3D rigid-body transformation when the surface is being aligned with itself is computed. The computed value does not indicate the absolute accuracy that can be achieved with each point set during registration. It is only useful for comparing the registration between the data sets, so as to determine which point set has the best constraint on the transformation (Low and Lastra, 2007). The absolute accuracy of the 3D model is evaluated considering the sensor accuracy.

Since the error analysis is a necessary step to know whether an estimated transformation satisfies accuracy requirements at the time of registration, this work intends to evaluate the 3D reconstruction error using the technique presented by Takimoto *et al.* (2012). This paper is organized as follows, the techniques used to capture the point cloud in Section 2. The modified ICP algorithm are detailed in Section 3. Next, the error analysis is explained in Section 4. Finally, some

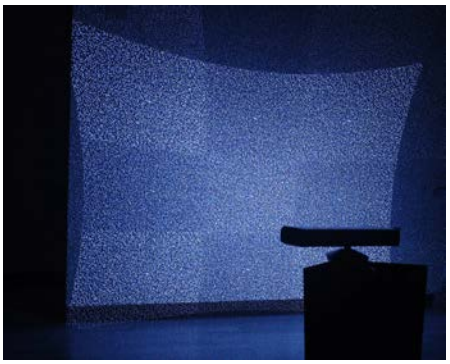

Figure 1. Kinect structured-light pattern, a known pattern is projected onto the scene and depth is inferred from the deformation of that pattern.

results are shown in Section 5 and Section 6 presents the conclusions and future works.

2. Point Cloud Acquisition Using Structured-Light

The structured-light technique adds additional information for the 3D reconstruction. A coded structured-light system is based on the projection of a single pattern or a sequence of patterns onto the measuring scene which is imaged by a single camera or a set of cameras. In systems that use spatial encoding, the patterns are specially designed so that codewords are assigned to a set of pixels. Every coded pixel has its own codeword, so there is a direct mapping from the codewords to the corresponding coordinates of the pixel in the pattern. The codewords are simply numbers, which are mapped in the pattern by using grey levels, color or geometrical representations.

As the number of points (codewords) that must be coded increases, the mapping of such codewords to a pattern is more difficult. The problems are that they typically need complex patterns or colors to encode positional information. To determine the spatial codes uniquely, the size of a code becomes large. Such patterns are easily affected by textures, shape discontinuities, image compression caused by tilted surfaces. Therefore, density of patterns should be inevitable low, and thus, 3D reconstruction tends to be sparse and unstable in reality (Sagawa *et al.*, 2009).

To deal with this issue, fast methods (Rusinkiewicz *et al.*, 2002; Koninckx and Gool, 2006; Zhang *et al.*, 2002; Je *et al.*, 2004) that can measure the shape at a high frame rate were proposed. Among them, the methods proposed by Zhang *et al.* (2002) and Je *et al.* (2004) can extract the shape from a single image then are suitable for real-time acquisition. Also, several researches reduced the required number of patterns using both temporal and spatial changes (Young *et al.*, 2007; Hall-Holt and Rusinkiewicz, 2001). However, the main difficulty is the simultaneous texture acquisition because visible light patterns are projected to measure the shape. Two solutions were proposed for the problem: time-sharing methods (Raskar *et al.*, 1998; Waschbüsch *et al.*, 2005; Weise *et al.*, 2007) and band-separating methods (Frueh and Zakhor, 2005; Hiura *et al.*, 1996; Microsoft, 2012).

The time-sharing method uses structured-light not only for capturing shape but also for illuminating objects to obtain texture. If the structured-light patterns are removed from the images, they can be used as texture images. This is accomplished by computing the average of the images of time-multiplexing patterns, or low-pass filtering of fringe patterns. The advantage of this approach is that texture and shape can be obtained using the same camera. However this method cannot be used with other illumination in the environment because structured-light is used for both shape and texture. Consequently, the illumination using this approach is restricted. Additionally, if an object in the image moves fast, misalignments between the texture and the shape may appear since the duration of the shape and the texture capture are different.

The band-separating methods uses different light wavelengths for shape and texture. The shape is obtained by using infra-red (IR) light, while the texture image is captured with visible light. Since the structured-light patterns for IR light are not detected by the camera for visible light, the acquisition of shape and texture can be accomplished simultaneously. Frueh and Zakhor (2005) use vertical stripes and a horizontal line of IR. The vertical stripes are used for shape reconstruction and the horizontal line for identifying the vertical stripes. Hiura *et al.* (1996) proposed a laser range finder that obtains a range image at 30 Hz by fast scanning of IR light stripes. The texture image is captured simultaneously with the system. Kinect (Microsoft, 2012) can capture a shape (depth data) by using IR projector and a IR camera and at the same time a texture (color data) by another CCD camera. Fig. 3 shows the static spatially encoded IR pattern projected onto the scene generated by the Kinect. The pattern is deformed as it falls onto the objects and the IR camera then captures an image of the scene and decodes the result. Although the low resolution of the generated 3D scene, the depth can be calculated with only one capture.

The pattern is deformed as it falls onto the objects and the IR camera then captures an image of the scene and decodes the result. Although the low resolution of the generated 3D scene, the depth can be calculated with only one capture.

Figure 2. Kinect camera. Depth resolution: 640×480 px, RGB resolution: 1600×1200 px.

The modified ICP algorithm proposed by Takimoto *et al.* (2012) uses data from both the color and depth cameras in the Kinect (see Fig. 4). The IR projector illuminates the scene with a structured dot pattern, in which the IR camera records to produce a depth map in VGA resolution with eleven bits of precision per sample. The depth reconstruction is robust in the absence of strong IR interference from other sources like other Kinects or direct sunlight. The camera in the middle records visible-color images at 30 frames per second. It has fixed focal-length lens with low radial lens distortion and automatic brightness adjustment.

3. Modified ICP Registration

In the ICP algorithm proposed by Besl and McKay (1992), each point in one data set is paired with the closest point in the other data set to form correspondence pairs. Then, using a point-to-point error metric, the sum of the squared distance between points in each correspondence pair is minimized by rigidly transforming one of the data sets. The process is iterated until the error becomes smaller than a threshold or it stops changing. Thus, two views of a surface are said to be in registration when any pair of points (p_i, q_i) from the two views representing the same surface point can be brought into coincidence by one rigid transformation (Chen and Medioni, 1991). That is, there exists a rigid transformation $\mathbf{T}$, such that

$$\forall\ p_i \in P, \exists\ q_j \in Q\ |\ \|\mathbf{T} \cdot p_i - q_j\| = 0 \tag{1}$$

where P and Q are two views of the same surface and $\mathbf{T} \cdot p_i$ is the result of applying $\mathbf{T}$ to p_i.

Although the ICP algorithm is widely used to register 3D point clouds, some limitations must be addressed. The algorithm has a local minima very close to the global minimum and getting trapped at one of these local minima can lead to sub-optimal registration accuracy. In (Besl and McKay, 1992) an approach to sample the 6-dimensional space of the rigid motion parameters, and run the ICP registration from each sample pose is suggested. This method does not really address the problem of local minima close to the global one because the subset of initial poses that will converge to the global minimum can be very small, and therefore the pose sampling may miss them completely. In the ICP algorithm described by Besl and McKay (1992) a point-to-point error metric is used in which the sum of the squared distance between points in each correspondence pair is minimized, therefore each point in one data set is paired with the closest point in the other data set to form correspondence pairs. The use of this ICP algorithm is not suitable because there are some outliers in both point clouds due to sensor movement. The conventional ICP algorithm (Besl and McKay, 1992) iteratively executes the point correspondence by determining the points from one set that are closest to the points in the other set. The correspondence cannot be one-to-one, because outliers can exist because of occlusion, appearance and disappearance. Actually, it is necessary to execute subset-subset point correspondence.

The modified ICP algorithm that improves the point correspondence algorithm is proposed by Takimoto *et al.* (2012). The algorithm uses robust statistics to generate a dynamic distance threshold on the distance allowed between closest points. This dynamic distance threshold is used to relax the requirement that one data set be a proper subset of the other, so that partially overlapping data sets can be registered. Instead of matches all points, this technique uses a distance parameter D and a statistical analysis to remove some of them.

The algorithm uses a maximum tolerable distance $D^1_{\max}$ to choose the matches that will be used, thus each point $\{p_i\}$ in the first data set whose distance to its closest point is less than $D^1_{\max}$ is retained, together with its closest point $\{q_i\}$ and their distance $\{d_i = p_i - q_i\}$. The mean μ and the sample deviation is given by

$$\mu = \frac{1}{N}\sum_{i=1}^{N} d_i \tag{2}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - \mu)^2}. \tag{3}$$

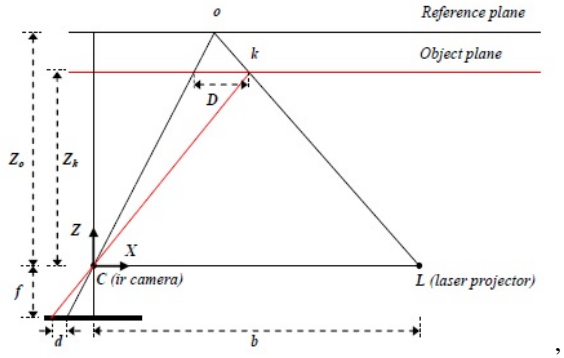

Figure 3. Relative depth and measured disparity relationship.

The maximum tolerable distance $D^1_{\max}$ is adaptively set based on statistics of the distance evaluated values. The mean μ is compared with the distance parameter D to update $D^1_{\max}$ using the distances statistics. The iteration ends when the change in the translation and rotation is less than 1%.

The modified algorithm also adds a weight to the points that are closer to the sensor. Since the weight is used in the registration precision ($d_k - d_{k+1}$), the algorithm will force a better registration of the points that are near to the sensor. As the sensor can perform translation and rotation movement, considering only the rotation movement the points that are near will have a smaller displacement when compared with the points that are far.

3.1 Some Improvements

To improve the robustness of the modified algorithm an implementation based on the color information is also proposed. The proposed algorithm analyze and uses the color information in the closest point evaluation. Let P and Q be two data sets with colors $cp = (c_p, c_p, c_p)$ and $cn = (c_n, c_n, c_n)$ respectively. The Euclidean distance with the color information is described by

$$d^2 = d^2_{pq} + \alpha_1(c_p - c_n)^2 + \alpha_2(c_p - c_n)^2 + \alpha_1(c_p - c_n)^2 \tag{4}$$

where α_1, α_2 and α_3 are parameters used to weight the color component importance. To avoid the specular reflection influence, the distance using the color information is not used if the color difference is greater than a defined threshold.

4. Error Analysis

The following errors are going to be considered: point cloud acquisition error and registration error. The translational and rotational alignments are separately considered in the registration error analysis.

4.1 Point Cloud Acquisition Error

Since Kinect is a recent development - it was released in November 2010 - little information about the geometric quality of its data is available thus an analysis of the systematic and random errors of the data is necessary. The correction of systematic errors is a prerequisite for the alignment of the depth and color data, and relies on the identification of the mathematical model of depth measurement and the calibration parameters involved. The characterization of random errors is important and useful in further processing of the depth data. The Kinect error analysis uses the mathematical model discussed in (Khoshelham and Elberink, 2012). Khoshelham and Elberink (2012) provided an insight into the geometric quality of the Kinect depth data through calibration and an analysis of the accuracy and density of the points. The Kinect error analysis is evaluated from the relation between the distance of an object point k to the sensor relative to a reference plane and the measured disparity d (see Fig. 5).

From the similarity of triangles, the distance (depth) Z_k of the point k in object space is

$$Z_k = \frac{Z_o}{1 + \frac{Z_o}{fb}d} \tag{5}$$

where b is the base length, f is the focal length of the infrared camera and d is the observed disparity in image space. The Z coordinate of a point together with f defines the imaging scale for that point and the world coordinate of each point can be calculated using the image coordinate and the scale, as follows

$$X_k = \frac{Z_k}{f}(x_k - x_o + \delta x) \tag{6}$$

$$Y_k = \frac{Z_k}{f}(y_k - y_o + \delta y) \tag{7}$$

where x_k and y_k are the image coordinates of the point, x_o and y_o are the coordinates of the principal point, and δx and δy are corrections for lens distortion. Since there is a shift in the x direction between the disparity and infrared images (Khoshelham and Elberink, 2012), (6) and (7) must be normalized. Therefore, d should be replaced with $md + n$ with d as the normalized disparity and m, n the parameters of a linear normalization. The normalized distance Z_k can be written as

$$Z'_k = \left(\frac{m}{fb}\right) d' + \left(Z'_o + \frac{n}{fb}\right), \tag{8}$$

in which expresses a linear relation between the inverse depth of a point and its corresponding normalized disparity. By observing the normalized disparity for a number of object points (or planes) at known distances to the sensor the coefficients of this linear relation can be estimated in a least-squares fashion. The calibration parameters mentioned above completely define the relation between the image measurements (x, y, d) and object coordinates (X, Y, Z) of each point. Once estimated, they enable the generation of a point cloud from each disparity image. Assuming that the calibration parameters are determined accurately and that d is a random variable with a normal distribution it is possible to propagate the variance of the disparity measurement to obtain the variance of the depth measurement as follows

$$\sigma_Z^2 = \left(\frac{\delta Z}{\delta d}\right)^2 \sigma_d^2 \tag{9}$$

After simplification this yields the following expression for the depth standard deviation

$$\sigma_Z = \left(\frac{m}{fb}\right) Z^2 \sigma_d \tag{10}$$

with σd and σZ respectively the standard deviation of the measured normalized disparity and the standard deviation of the calculated depth. Expression (10) basically expresses that the random error of depth measurement is proportional to the square distance from the sensor to the object. Since depth is involved in the calculation of the world coordinates, it is expected the error in X and Y to be also a second order function of depth. By propagating the errors, and assuming that the random error of image coordinates x, y can be ignored, the random error of X and Y is obtained

$$\sigma_x = \left(\frac{mx}{f^2 b}\right) Z^2 \sigma_d \tag{11}$$

$$\sigma_y = \left(\frac{my}{f^2 b}\right) Z^2 \sigma_d \tag{12}$$

What shows that the error of the point cloud acquisition is directly proportional to the distance of the object to the Kinect.

4.2 Registration Error

In the registration error analysis, translation and rotation are considered separately as they are parameterized by different entity types. The translation is parameterized by distance and rotation by angle. For the rotation analysis, the arbitrary scale of the object is not a problem since the analysis will be done separately. The registration error analysis can be done using the first-order approximation of the true point-to-surface distance

$$D(x) = \frac{F(x)}{\|\nabla F(x)\|} \tag{13}$$

where $F(x) = 0$ is the implicit equation of the surface, $\|\nabla F(x)\|$ is the magnitude of the gradient to the surface, x is a point in which may not lie on the surface and $D(x)$ is the approximate distance. The application of a differential transformation T on point x_s that lies on the surface, $D(x_s) = 0$ can be written as

$$V(x_s) = \frac{\partial}{\partial t} D(T(x_s)) = \begin{bmatrix} n_{x_s} \\ x_s \times n_{x_s} \end{bmatrix} \tag{14}$$

where T can be represented by a homogeneous transformation in which is a function of the six parameters (t_x, t_y, t_z, ω_x, ω_y, ω_z), rotations and translations along the x, y and z axes and n_{x_s} is the unit normal to the surface evaluated at the point x_s.

Figure 4. Scene to be analyzed.

4.2.1 Translational Alignment Error

Let $[x_i, y_i, z_i]^T$ be the 3D coordinates of the ith true surface point, where $i = 1, 2, \ldots, M$. Suppose there are two sets of measurements of the surface points, producing the point set A with coordinates $[x_i + u_{Ai}, y_i + v_{Ai}, z_i + w_{Ai}]^T$ and the point set B with coordinates $[x_i + u_{Bi}, y_i + v_{Bi}, z_i + w_{Bi}]^T$, where (u_{Ai}, u_{Bi}), (v_{Ai}, v_{Bi}) and (w_{Ai}, w_{Bi}) are measurements errors in the x, y and z directions, respectively. After the ICP, the point set B is translated so that it is aligned with point set A. The alignment uses the least-squares (least-sum-of-squares) error metric, where the translation vector $\tau = [t_x, t_y, t_z]^T$ to translate the point set B is used to minimize

$$SSE = \sum_{i=1}^{M} \left\| \begin{bmatrix} x_i + u_{Ai} \\ y_i + v_{Ai} \\ z_i + w_{Ai} \end{bmatrix} - \begin{bmatrix} x_i + u_{Bi} \\ y_i + v_{Bi} \\ z_i + w_{Bi} \end{bmatrix} - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \right\|^2 . \tag{15}$$

Given that the measurement errors u_{Ai}, u_{Bi}, v_{Ai}, v_{Bi}, w_{Ai} and w_{Bi} are independent and normally distributed with mean 0 and standard deviations not greater than e_{RMS}, it is possible to conclude that the error alignment is directly related to the translation parameter.

4.2.2 Rotational Alignment Error

Using an analogous approach for the rotation component, the alignment uses the least-squares (least-sum-of-squares) error metric, where the rotation vector $\theta = [\omega_x, \omega_y, \omega_z]^T$ to translate the point set B is used to minimize

$$SSE = \sum_{i=1}^{M} \left\| \begin{bmatrix} x_i + u_{Ai} \\ y_i + v_{Ai} \\ z_i + w_{Ai} \end{bmatrix} - R(\omega_x, \omega_y, \omega_z) \cdot \begin{bmatrix} x_i + u_{Bi} \\ y_i + v_{Bi} \\ z_i + w_{Bi} \end{bmatrix} \right\|^2 \tag{16}$$

where $R(\omega_x, \omega_y, \omega_z)$ is the composite rotation matrix for rotations of ω_x, ω_y and ω_z radians about the $x - axis$, $y - axis$ and $z - axis$, respectively. Since the rotations will be small, the matrix $R(\omega_x, \omega_y, \omega_z)$ can be approximated by using the approximations $\sin\theta \approx \theta$ and $\cos\theta \approx 1$ when $\theta \approx 0$.

Given that the measurement errors u_{Ai}, u_{Bi}, v_{Ai}, v_{Bi}, w_{Ai} and w_{Bi} are independent and normally distributed with mean 0 and standard deviations not greater than e_{RMS}, it is possible to conclude that the error alignment is directly related to the rotation parameter.

5. Results

The error analysis and the parameters influence were verified in the acquisition of real scenes. For the registration error analysis, two cases were considered the translational and the rotational movement. The Fig. 6 shows the image used to analyze the translational movement. The point cloud generated in the ICP process in two different situations. The Figs. 7 and 8 show translational acquisitions using small and large translational movement of the sensor respectively. The results showed that it was possible to get good results when the movement of the sensor is lower than 5 cm.

The reconstructed image obtained after processing the Figs. 7 and 8 can be viewed in Figs. 9 and 10. From the images it was possible to verify the point cloud acquisition error. The images show that as the distance from the sensor increases the acquisition error increases. Moreover, the point cloud density has great influence over the reconstruction quality, as

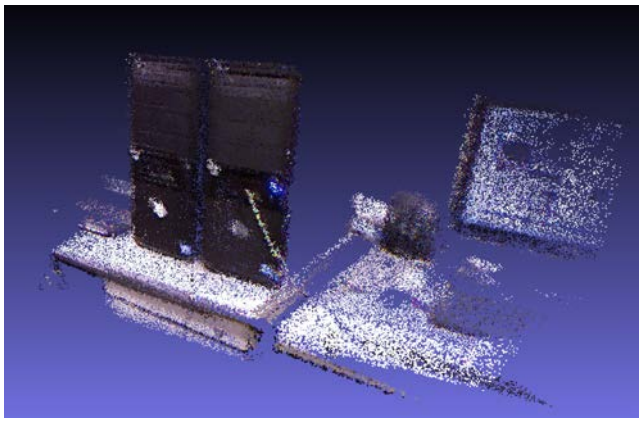

Figure 5. Point Cloud generated by the ICP for a small translational (3 cm) movement of the sensor.

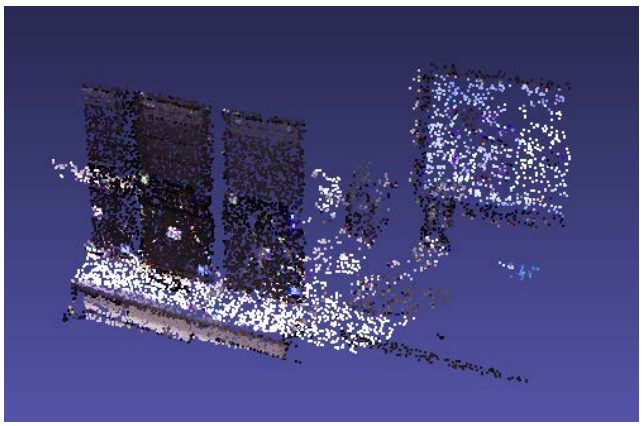

Figure 6. Point Cloud generated by the ICP for a large translational (9 cm) movement of the sensor.

the point cloud density increases the reconstruction quality also increases. The Fig. 8 ilustrates this behaviour, as the point cloud density is low in some regions, the surface reconstruction failed in those regions.

For the rotational movement analysis, two different situations were also considered, small and large angles displacement. In this analysis the face reconstruction were considered and the point cloud result using different angles of acquisitions can be viewed in Figs. 11 and 12. The results showed that it was possible to get good results when the rotation of the sensor is lower than 70 degrees. The angle used in Fig. 12 was 80 degrees.

Using first-order approximation of the point-to-point distance between the point clouds, it was possible to evaluate the error considering two point clouds. In the error analysis the sum all distances errors were considered, therefore the inliers and the outliers distance errors wer included. Although the error between the points is small and on the order of millimeters, the sum of all distance errors is higher and on the order of centimeters. The error behavior of the modified ICP (Takimoto *et al.*, 2012) to align two point clouds can be viewed in Figs. 13. and 14. The Fig. 13 shows the distance error for the ICP applied to translated point clouds, it is possible to notice that as the translation movement increases the error also increases. However, the distance between acquisitions of maximum of $3cm$ is a good balance (number of acquisition $\times$ error) as the error increases too much after that. The Fig. 14 shows the distance error for the ICP applied to rotated point clouds, it is possible to notice that as the rotation movement increases the error also increases. The rotation has a different behavior when compared to translation, as from 0 to 50 the error is kept almost stable. Based on this, it is possible to conclude that rotations have smaller errors when compared to translations.

6. Conclusion

The results showed that the point cloud acquisition must be performed within a certain range to ensure the quality of the reconstruction because as the distance from the sensor increases the error and the point cloud density decreases. From the results, it is possible to verify that the ICP is more robust to rotation movement when compared to the translation movement. The reconstruction quality is satisfactory when the translational movement is low and the rotational movement is lower than 50 degrees.

Figure 7. Reconstructed scene for a small translational (3 cm) movement of the sensor.

Figure 8. Reconstructed scene for a large translational (9 cm) movement of the sensor.

Figure 9. Point Cloud generated by the ICP for a small rotational (40 degrees) movement of the sensor.

Figure 10. Point Cloud generated by the ICP for a small rotational (90 degrees) movement of the sensor.

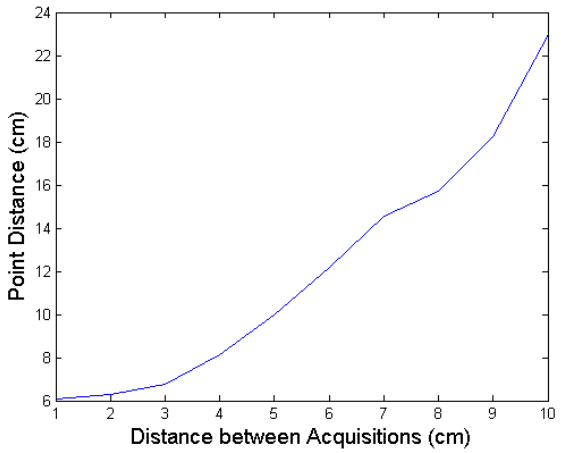

Figure 11. Distance Error between point clouds considering only translation.

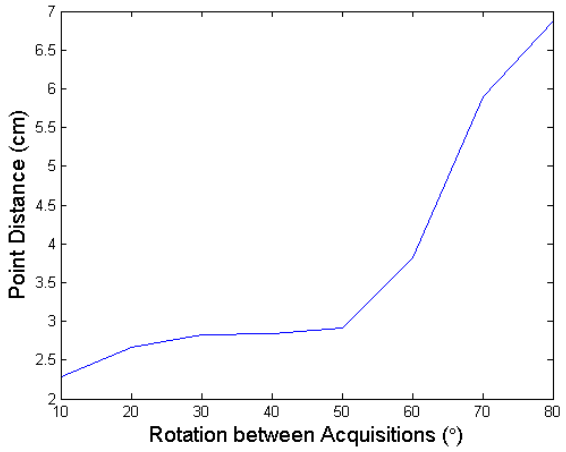

Figure 12. Distance Error between point clouds considering only rotation.

7. ACKNOWLEDGEMENTS

Marcos Sales Guerra Tsuzuki and Thiago de Castro Martins were partially supported by CNPq (grants 309570/2010-7 and 06415/2012-7). Rogério Yugo Takimoto is supported by FAPESP (grant 2011/22402-8).

8. REFERENCES

Besl, P. and McKay, N., 1992. "A method for registration of 3D shapes". *IEEE T Pattern Anal Machine Intelligence*, Vol. 14, pp. 239–256.

Chen, Y. and Medioni, G., 1991. "Object modeling by registration of multiple range images". In *Proc IEEE Int Conf Robotics and Automation*. Sacramento, USA, pp. 2724–2729.

Frueh, C. and Zakhor, A., 2005. "Capturing 2 $\frac{1}{2}$ D depth and texture of time-varying scenes using structured infrared light". In *Proc 5^{th} Int Conf 3D Digital Imaging and Modeling*. Washington, USA, pp. 318–325.

Hall-Holt, O. and Rusinkiewicz, S., 2001. "Stripe boundary codes for real-time structured-light range scanning of moving objects". In *Proc 8^{th} Int Conf Computer Vision*. Vancouver, Canada.

Hiura, S., Yamaguchi, A., Sato, K. and Inokuchi, S., 1996. "Real-time object tracking by rotating range sensor". In *Proc Int Conf Pattern Recognition*. Washington, USA, pp. 825–829.

Je, C., Lee, S.W. and Park, R.H., 2004. "High-contrast color-stripe pattern for rapid structured-light range imaging". In *Proc 8^{th} Euro Conf Computer Vision*. pp. 95–107.

Khoshelham, K. and Elberink, S.O., 2012. "Accuracy and resolution of kinect depth data for indoor mapping applications". *Sensors*, Vol. 12, No. 2, pp. 1437–1454. ISSN 1424-8220. doi:10.3390/s120201437. URL `http://www.mdpi.com/1424-8220/12/2/1437`.

Koninckx, T.P. and Gool, L.V., 2006. "Real-time range acquisition by adaptive structured light". *IEEE T Pattern Anal Machine Intelligence*, Vol. 28, No. 3, pp. 432–445.

Low, K.L. and Lastra, A., 2007. "Predetermination of icp registration errors and its application to view planning". *3D Digital Imaging and Modeling, International Conference on*, Vol. 0, pp. 73–80. ISSN 1550-6185. doi: http://doi.ieeecomputersociety.org/10.1109/3DIM.2007.41.

Microsoft, 2012. "Xbox 360 kinect. @ONLINE". URL `http://www.xbox.com/en-US/kinect`.

Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L. and Fuchs, H., 1998. "The office of the future: a unified approach to image-based modeling and spatially immersive displays". In *Proc 25^{th} A Conf Comput Graph Interactive Techniques*. Orlando, USA, pp. 179–188.

Rusinkiewicz, S., Hall-Holt, O. and Levoy, M., 2002. "Real-time 3D model acquisition". In *Proc 29^{th} A Conf Comp Graph Inter Techn*. San Antonio, USA, pp. 438–446.

Sagawa, R., Ota, Y., Yagi, Y., Furukawa, R., Asada, N. and Kawasaki, H., 2009. "Dense 3D reconstruction method using a single pattern for fast moving object". In *Proc IEEE 12^{th} Int Conf Computer Vision*. Kyoto, Japan, pp. 1779–1786.

Simon, D., 1996. *Fast and Accurate Shape-Based Registration*. Ph.D. thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Takimoto, R.Y., Vogelaar, R., Ueda, E.K., Tsuzuki, M., Gotoh, T. and Kagei, S., 2012. "3d reconstruction of large point clouds with a new point correspondence algorithm." In *Proceedings of the The 16th IASTED International Conference on Software Engineering and Applications*. IASTED, ACTA Press, pp. 247–254.

Waschbüsch, M., Würmlin, S., Cotting, D., Sadlo, F. and Gross, M., 2005. "Scalable 3D video of dynamic scenes". *The Visual Computer*, Vol. 21, No. 8, pp. 629–638.

Weise, T., Leibe, B. and Van Gool, L.J., 2007. "Fast 3D scanning with automatic motion compensation". In *Proc IEEE Conf Comput Vision Pattern Recognition*. Minneapolis, USA, pp. 1–8.

Young, M., Beeson, E., Davis, J., Rusinkiewicz, S. and Ramamoorthi, R., 2007. "Viewpoint-coded structured light". In *Proc IEEE Conf Comp Vision Pattern Recognition*. Minneapolis, USA, pp. 1–8.

Zhang, L., Curless, B. and Seitz, S.M., 2002. "Rapid shape acquisition using color structured light and multi-pass dynamic programming". In *Proc 1^{st} IEEE Int Symp 3D Data Proc, Vis and Trans*. Padova, Italy, pp. 24–36.

9. RESPONSIBILITY NOTICE

The author(s) is (are) the only responsible for the printed material included in this paper.