REGISTRATION OF RANGE IMAGES USING A HISTOGRAM BASED METRIC

REGISTRO DE IMÁGENES DE RANGO MEDIANTE UNA METRICA BASADA EN HISTOGRAMA

MANUEL A. MALDONADO

M.Sc Departamento de Ciencias de la Computación y de la Decisión, Universidad Nacional de Colombia - Sede Medellín, Colombia, amaldona@gmail.com

GERMAN SÁNCHEZ TORRES

Ph.D Universidad del Magdalena - Santa Marta, Colombia, gsanchez@unimagdalena.edu.co

JOHN WILLIAN BRANCH

Ph.D Departamento de Ciencias de la Computación y de la Decisión, Universidad Nacional de Colombia - Sede Medellín, Colombia, jwbranch@unal.edu.co

ABSTRACT: In this paper, a new approach for coarse registration of triangular meshes for three-dimensional object reconstruction is proposed by using the shape-index measure for topological description and the Earth Mover's Distance (EMD) metric for region similarity estimation. Two of its main advantages are independence from initial pre-alignment and spatial transformation, obtaining a more accurate registration regardless of the initial viewing position.

KEYWORDS: Range Image registration, 3-D reconstruction, shape index, EMD.

RESUMEN: En este trabajo se propone un método de registro grueso de mallas triangulares para la reconstrucción de objetos tridimensionales, empleando el índice de forma para obtener información topológica de la superficie y la comparación de regiones con forma similar empleando la métrica Earth Mover's Distance (EMD). Las principales ventajas de este método es que permite independencia de la pre-alineación inicial y transformación espacial, permitiendo obtener un registro adecuado sin importar la posición inicial de cada vista.

PALABRAS CLAVE: Registro de imágenes de rango, reconstrucción 3-D, índice de forma, EMD.

1. INTRODUCTION

3D reconstruction is the process of generation of a computerized model from a digitalized object, keeping its physical features such as dimension, volume and shape [1]. This process plays a prominent role in computer-aided design systems (CAD) due to its robustness regarding complex objects with arbitrary topology [2]. The generation of these models has different application fields such as object recognition, facial reconstruction for identification and data exchange of archaeological findings.

The method proposed in this paper focuses on the registration stage. Problems raised in this stage include the non-alignment produced when taking two or more range images from different points of view and without any control of the relative positions between the object and the sensor. Once the acquisition stage is over by scanning the object's surface, and the visualization of the images in a unique system of coordinates is done, different overlapping views are obtained -without any coherence.

The aim of the registration is to transform all the acquired partial views to a common coordinate system. Because a simple range image can contain just the points of the scene that are visible to the sensor, it is necessary to link multiple range images in a single set of points to obtain the entire model of the scene.

Classification of registration methods is complex and it can be performed from different aspects. A common classification is done according to the precision level in which the method links both images. Thus, a coarse registration can exist as well as a fine one, whether the aim is to give an initial alignment or to link the images precisely.

This paper proposes a method of coarse registration of triangle mesh for three-dimensional object reconstruction, by using the shape-index measurement for the topological description of the surface. One of the most important advantages of this method is that it allows independence from initial pre-alignment and spatial transformation, making it possible to obtain an accurate registration regardless of the initial position of each view. The main difference regarding similar works lies on the fact that alignment is not based on analysis of pairs of points, like the classic model, but on regions. The matching of the regions is based on a similarity metric estimated using a histogram representing the geometric features of each region.

This paper is structured as follow: Section 2 presents relevant works of the state of the art. Section 3 describes the proposed method. Sections 4 and 5 show the obtained results and the conclusions, respectively.

2. RELATED WORKS

The main works in the state of the art are described below. The most used registration method to align pairs of images consists in the minimization of the difference between the set of points of each image, called *Iterative Closest Point* (ICP). Subsequently, an important number of works have followed this direction trying to improve different aspects related to the method, such as its efficiency, level of error, matching algorithm, among others. Additional works include those of Chua [3], who introduces the *Point Signature* method. This method is used to represent the structure of a specific neighborhood. For a point p and a sphere S centered at p with radius r, the method estimates the largest positive distance from a point p to the curve C. the curve C is a 3D curve formed by the intersection of S and the object surface. The correspondence between points is established between two points with similar point signatures. The matching process is fast but, sometimes, it requires interpolation, which affects quality; besides, calculating the intersection of the object's surface with the sphere is not easy. It is invariant to rotation and translation but it is sensitive to noise.

The *Spin Image* method, introduced by Johnson [4], is based on the creation of images called *Spin Images,* which locally represent the surface geometry with respect to a point, allowing characteristic zones on the surface to be obtained and to perform the registration with other surfaces. In general, this method is robust against noise but it depends of the resolution used (although, there are variations that make interpolations on every image), and it has some problems with symmetric images.

Chung and Lee [5] introduce the registration method *Principal Component Analysis* (PCA). Their aim is to reduce a wide set of variables to a new smaller set, without losing important parts of the original information. The direction of the main axis of the cluster of dots is used to align a sequence of range images. If the overlap region is large enough, the axes of the images must match so that the axes are aligned by applying just one transformation and therefore, the images are aligned too. This method implies the calculation of the co-variance matrix of each image, which implies high speed and precision with enough points.

DARCES (data-aligned rigidity-constrained exhaustive search) based on RANSAC (RANdom SAmple Consensus) was introduced by Chen *et al*. [6]. It is based on the search for the best match of three points between two range images, in order to calculate the rigid transformation. The points (main, secondary and auxiliary ones) are characterized by the distance between them. The method is robust against atypical data and noise but can only be used when the amount of points in the images is small, because calculation time increases considerably. Tarel *et al*. [7] proposed the *Algebraic Surface Model* (ASM) method, which estimates a transformation between surfaces represented as polynomials estimated by the *3L Fitting* method, which is a linear algorithm based on minimum squares. The matching method is faster, but it requires the estimation of the normal at each point; this can considerably increase the calculation time. However, it can bring good results with an overlap area lower than 15%.

Stamos and Leordeanu [8] propose a method calculating the match between images from straight lines. The algorithm was applied to structure registration with plane regions such as buildings, and it obtains good results, however the difficulty with this method lies in the fact that it is not automatic.

Registration using Genetic Algorithms (GA) was

introduced by Brunnström and Stoddart [9], and consists in the search of an individual whose chromosome is composed of 6 parameters determining the rigid transformation that aligns a couple of images: 3 elements of the translation vector, and 3 rotation angles. Variations of this method are presented using different GA aptitude functions. This method is robust against noise, local minimum and atypical data, although as a result of evolutionary techniques, it requires a long calculation time and, depending on the aptitude function, the cost of the matching calculation can be very high. Feldmar and Ayache [10] propose a method based on the difference of points characterized by their main curvatures as a geometrical feature for the matching process. This method requires just one match in order to calculate the transform. It calculates curvatures (k_1, k_2) at a point, where k_1 and k_2 are the maximum and the minimum curvature of the surface at the specific point. In addition, the normal vector and the direction of the main curvature are also considered. Even though it is not always possible to calculate curvatures and directions at a point, a parametrized surface is proposed (similar to Tarel's [7] proposal). The biggest problem of this method is that it is not robust; the algorithm stops when it finds a good match, presenting some false-positive.

Other works propose variants of some stages of the ICP classic method, such as [11], which proposes monitoring the sensor position, and the identification and indexing of surface features; or in [12], which propose using Minimum Variance Estimator (MVE) for the transform estimation; and [13], which propose using *Least Median of Squares* (LmedS), which aims to minimize the average of matching residuals.

One of the main weaknesses of registration methods lie with the algorithm executed when matching, that is, the stipulated metric to determine whether two points of different images constitute the same point, or the closest one, in the treated object. For this reason, the proposed approach is based on determining the matching of two points based on the geometry of the surrounding region.

3. REGISTRATION METHOD BASED ON GEOMETRIC FEATURES

The method proposed here consists of aligning the partial views from geometric features described by the shape-index measure. The process is carried out as follows:

1. The shape index is calculated for every vertex from the main curvatures values in each vertex.

2. Three vertexes are selected from the first partial view, and a histogram of shape-index of the neighborhood around each vertex is generated.

3. On the second view, a search for vertexes having similar histograms is done using a variant of the EMD (Earth's Mover Distance) method to calculate the distance between histograms. This search will generate a list of similar vertexes on the second view for each vertex selected on the first one.

4. By using the list of similar vertexes the aim is to find the match with less alignment error. Alignment is only made between the pairs formed by the vertexes selected from the first view, and their matches on the second view. That is why the matching search space is reduced.

5. The alignment of the complete views is done by using the pairs obtained in the previous step, and the alignment error, as well as the Surface Interpenetration Index, is calculated.

Next, each step is described.

3.1. Shape-index calculation

As mentioned before, in order to calculate the shape index it is necessary to calculate the main curvatures for each vertex. Curvature is an invariant measurement of a surface that indicates how much it is curved. They are said to be invariant because they remain constant against translations and rotations, although not for scale changes.

Curvature measures the tangent vector's change of direction in a curve. This vector corresponds to the reciprocal of the tangent sphere's ratio at the surface. It is obtained from the Second Fundamental Form, which represents the variation of the normal from point to point. For a continuous surface it applies:

$$k_n = (s \quad t)\begin{pmatrix} e & f \\ f & g \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = (s \quad t)II\begin{pmatrix} s \\ t \end{pmatrix} \qquad (1)$$

to any vector (s, t) in the tangent plane (expressed in terms of an orthonormal coordinates system centered on the point). Matrix II is known as *Weingarten matrix* or *Second Fundamental Form,* which can be diagonalized to obtain:

$$k_n = (s' \quad t')\begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}\begin{pmatrix} s' \\ t' \end{pmatrix} = k_1 s'^2 + k_2 t'^2 \qquad (2)$$

The maximum and minimum values of normal curvature at a point p of the surface S are called *principal curvatures* [14] of S at p, denoted by k_1 and k_2. Vectors (s', t'), in which the extreme values are found, are called *principal directions*.

$$II = (D_u n \quad D_v n)$$
$$= \begin{pmatrix} \dfrac{\partial n}{\partial u} \cdot u & \dfrac{\partial n}{\partial v} \cdot u \\ \dfrac{\partial n}{\partial u} \cdot v & \dfrac{\partial n}{\partial v} \cdot v \end{pmatrix} \qquad (3)$$

where (u, v) are the directions of the orthonormal coordinates system at the point. Multiplying II by a vector of the tangent plane, the derivative of the normal in that direction is obtained:

$$IIs = D_s n \qquad (4)$$

Although this expression is defined for continuous surfaces, an approximation can be carried out for the discrete case using finite differences: for a triangle having the direction of its edges defined, we have the differences of the normal in these directions [14]. This brings a set of linear restrictions for the calculation of II using least squares. Once the system is solved, we have the principal curvatures per vertex.

Once the curvatures are obtained, the associated shape index can be obtained. The shape index was proposed in [15], and it is an angular physics measurement that captures the intuitive notion locally at a surface's point:

$$s = \frac{2}{\pi} \arctan \frac{k_2 + k_1}{k_2 - k_1} \qquad (5)$$

The index range varies between -1 and 1. A point of a convex surface having the same principal curvatures has

a shape index of 1. A point of a concave surface having the same principal curvatures has a shape index of -1. A saddle point having the same magnitude of principal curvatures but different signs has a shape index of 0.

3.2. Search for vertexes having histograms of similar shape

Once the shape indexes per vertex of each view are obtained, a) at least three vertexes from the first view are manually selected; b) a histogram is generated with the shapes indexes of the neighborhood surrounding each vertex, and c) a search for the vertexes from the second view having similar histograms is carried out.

A histogram represents the frequency of a variable's values. The quantized variable is categorical, that is, its values are alphanumeric categories or labels, and the quantity of possible values is finite. Every value represented in the histogram corresponds to a class *(bin)*. In general, the quantity of classes in a histogram is the same as the quantity of values of the evaluated variable. But in this case, the shape index is a continuous variable, so it is not possible to represent it directly in a histogram; therefore, we group its possible values into sub-intervals. The interval [-1,1] was divided into 10 sub-intervals, each one associated with a histogram class.

Given two histograms, a quantitative measurement of their difference is required, reflecting the similarity perception as best as possible. Rubner and Guibas [16] propose the *EarthMover'sDistance* (EMD) metric to establish the difference between two distributions, which consists in the calculation of the minimal redistribution quantity needed to convert one distribution into the other. The EMD is calculated by using a linear optimization algorithm based on the solution to the transport problem [17] [18], which is a particular case of a linear programming problem in which the costs of supplying a series of demand points (destination), from a group of supply points (origin), must be minimized.

3.3. Similarity Calculation between Histograms

Given two histograms P, Q, EMD is defined as:

$$EMD(P,Q) = \frac{\left(\min_{f_{ij}} \sum_{i,j} f_{ij} d_{ij} \right)}{\sum_{i,j} f_{ij} d_{ij}}, \qquad (6)$$

Subject to:

$$f_{ij} \geq 0$$

$$\sum_{i,j} f_{ij} d_{ij} \leq P_i,$$

$$\sum_{i,j} f_{ij} d_{ij} \leq Q_j, \qquad (7)$$

$$\sum_{i,j} f_{ij} d_{ij} = \min\left(\sum_i P_i, \sum_j Q_j\right)$$

where f_{ij} denotes the network flows associated to transport problem, d_{ij} represents the transported quantity from class i of P to class j of Q, and d_{ij} represents the ground distance between class i and class j. [19].

3.4. Error Calculation and Surfaces Interpenetration Measure

Surfaces Interpenetration Measure (*SIM*) is based on the natural effect resulting from surface registration, due to small local distortions caused by the limitations in the sensor's precision, or by noise [20].

A more accurate evaluation can be established by quantifying the interpenetration, and it is possible to offer a more robust control of the registration, even though in this work it is used just as a measurement for evaluating the registration's quality. A high SIM value represents a good registration, a bad alignment produce a low value.

After the alignment of two views, A and B, identification of interpenetration points in A with respect to B is carried out. We will call them C_{AB}. For each point $p \in A$ a neighborhood N_p of radius n centered on p is defined. Defining q as a point in the neighborhood N_p, c as the corresponding point to p in B, obtained through a search of neighboring points by using a kd-tree [21], and y $\overrightarrow{n_c}$ the normal vector in c. Therefore, the set of interpenetration points in A is given by:

$$C_{A,B} = \{p \in A \mid [(\overrightarrow{q_i - c} \cdot \overrightarrow{n_c})(\overrightarrow{q_j - c} \cdot \overrightarrow{n_c})] < 0\} \qquad (8)$$

where $q_i, q_j \in N_p$ e $i \neq j$. $C_{A,B}$ correspond to points in A of neighborhoods which include at least a pair of points separated by a local tangent plane, obtained from corresponding points in B, as seen in Figure 1.
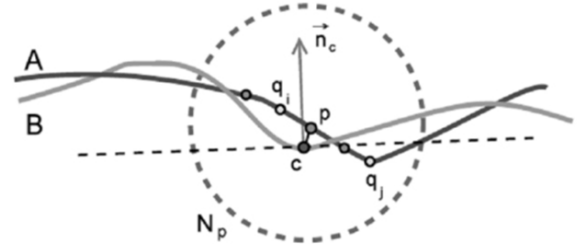

Figure 1. Interpenetration point *p* in A with respect to B [20]

3. RESULTS

To evaluate the method proposed here, the described steps were applied to pairs of views of the Stanford bunny and armadillo [22], and a pre-Columbian vase.

The first step consists in loading the triangular mesh of the views to be aligned, S_1 and S_2, and calculating the value of the shape index per vertex in each view. Because the shape index varies between -1 and 1, a color map per vertex can be created. Next, the initial selection of vertexes is made $\{v_{1i}\}$, where i is associated to each vertex selected from S_1, and an r radius neighborhood is established (a value of r=10 was chosen in this case) around $\{v_{1i}\}$. Figure 2 shows different selected neighborhoods.
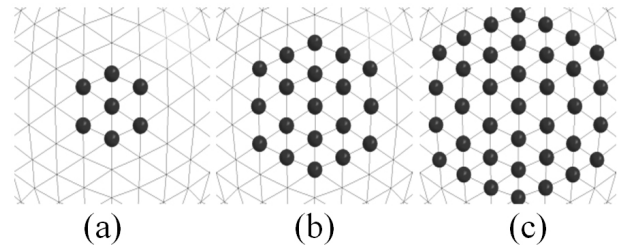

(a)   (b)   (c)

Figure 2. Neighborhood around the selected vertex, a) radius 1, b) radius 3 and c) radius 3.

From each neighborhood around $\{v_{1i}\}$ a histogram of the shape index is generated, and a search for neighborhoods having similar histograms is carried out in S_2. Figure 3 shows the result obtained from

two types of range image. The foot corresponds to synthetic data and the frog is from real range data. The results reported in Table 1 shows the benchmark of the implemented tests. The proposed method shows an error value comparable with the other methods. However, the time used by the proposed method was lower. All measurements were obtained using a registration range data toolbox [24] implemented using Matlab software.

Figures 4(a), 4(b) and 4(c) show the result of the registration process applied to the Stanford armadillo.
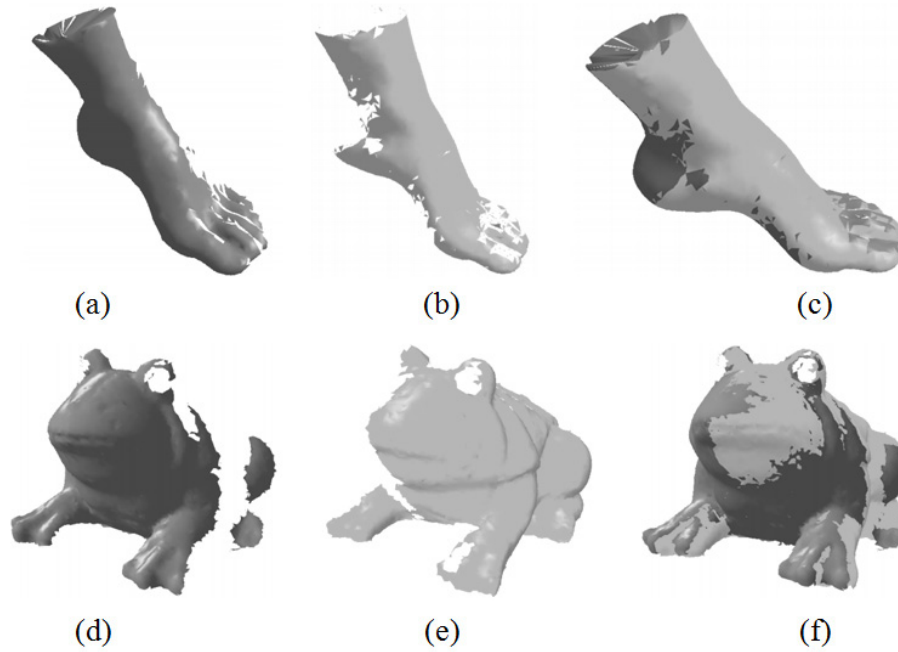

Figure 3. Different visual results, a-d, b-e) initial partial images, c-f) final registration result.
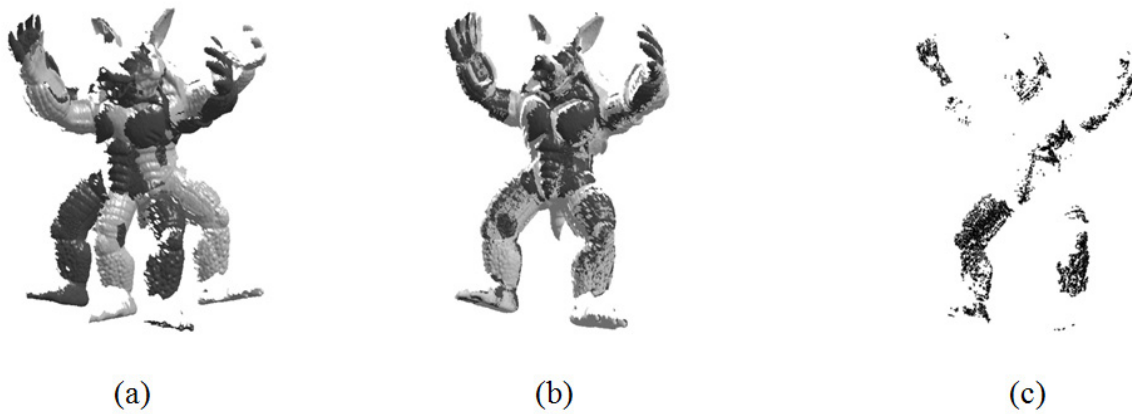

Figure 4. Registration of the armadillo (SIM= 27.24%, MSE=42.40): (a) partial views S_1 and S_2 with their corresponding selected and similar neighborhoods, (b) partial aligned views, and (c) interpenetration points.

Table 1. Results of different methods

Methods	Foot		Frog	
	RMS	Total time (seg)	RMS	Total time (seg)
Spin Image [4]	16.30	354.38	13.60	3437.10
Trucco [13]	0.62	9.78	5.37	27.12
Zinber [23]	1.96	5.67	0.75	25.37
Proposed method	1.83	4.98	2.38	2.73

5. CONCLUSIONS

In this work, a range image registration method is proposed, based on the geometric features described by the shape index per vertex.

This method is a starting point for the design of an automatic registration method, and presents some advantages such as its independence under rigid transformations. Furthermore, this method is deterministic, that is, with the same parameters the same result is obtained. One of its weaknesses is the fact that the result depends on the initial selection of points on S_1; a bad selection will lead to an incorrect registration, besides, a more robust method is needed to detect false matches for neighborhoods having similar histograms but in different regions.

As future work, it is proposed to establish a selection criterion of initial points, so that unique histograms can be obtained, and establish a correspondence validation method by using *Point Signature* [3], with which the matching neighborhood structure can be compared. Additionally, the registration can be implemented with scale independence, so that if any of the partial views presents a different scale, this can be corrected when carrying out the alignment.

REFRENCES

[1] Branch J. W., Reconstrucción de objetos de forma libre a partir de imágenes de rango empleando una red de parches NURBS, Ph.D. dissertation, Universidad Nacional de Colombia, 2007.

[2] Sanchez, G., Branch, J. y Boulanger, Pierre., Reconstrucción de objetos de topología arbitraria mediante selección de centros para la interpolación con FBR. Dyna: vol. 73, no. 150, pp. 189-201, 2006.

[3] Chua, C. J. R., Point signatures: a new representation for 3d object recognition, Int. J. Comput. Vision, vol. 25, (1), pp. 63–85, 1997.

[4] Johnson, A., Spin-Images: A Representation for 3-D Surface Matching, Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Aug. 1997.

[5] Chung, D. and Lee, Y. D. S., Registration of multiple-range views using the reverse-calibration technique, Pattern Recogn., vol. 31, (4), pp. 457–464, 1998.

[6] Song Chen, C., Ping Hung, Y. and Cheng. J., Ransac-based darces: A new approach to fast automatic registration of partially overlapping range images, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, pp. 1229–1234, 1999.

[7] Tarel, J., Civi, H. and Cooper, D., Pose estimation of free-form 3d objects without point matching using algebraic surface models, in Proceedings of IEEE Work shop on Model-Based 3D, pp. 13–21. 1998.

[8] Stamos, I. and Leordeanu, M., Automated feature-based range registration of urban scenes of large scale, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 555–561, 2003.

[9] Brunnstrom, K. and Stoddart, A., Genetic algorithms for free-form surface matching, in International Conference of Pattern Recognition, Vienna, pp. 689–693. 1996.

[10] Feldmar, J. and Ayache, N., Rigid and affine registration of smooth surfaces using differential properties, in ECCV (II), pp. 397–406. 1994.

[11] Faugeras, O., The Representation, Recognition, and Locating of 3-D Objects, Int. J. Robotic Res., vol. 5, (3), 1986.

[12] Dorai, C., Weng, J. and Jain, A., Optimal Registration of Object Views Using Range Data, Trans. PAMI, vol. 19, (10), 1997.

[13] Trucco, E., Fusiello, A. and Roberto, V., Robust motion and correspondences of noisy 3-D point sets with missing data, Pattern Recognition Letters, vol. 20, (9), pp. 889–898, Sep. 1999.

[14] Rusinkiewicz. S., Estimating curvatures and their derivatives on triangle meshes, in Symposiumon 3D Data Processing, Visualization, and Transmission, Sep. 2004.

[15] Koenderink, J. J. and Van Doorn, A. J., Surface shape and curvature scales, Image Vision Comput., vol. 10, (8), pp. 557–565, 1992.

[16] Rubner, Y., Tomasi, C. and Guibas, L. J., A metric for distributions with applications to image databases, in ICCV '98: Proceedings of the Sixth International Conference on Computer Vision. Washington, DC, USA: IEEE Computer Society, P. 59. 1998.

[17] Rachev, Themonge–Kantorovich. S. T., Mass transference problem and its stochastic applications, Theory of Probability and its Applications, vol. 29, (4), pp. 647–676, 1985.

[18] Hitchcock, F. L., The distribution of a product from several sources to numerous localities, Journal of Math. Phys., vol. 20, pp. 224–230, 1941.

[19] Pele O. and Werman, M., A linear time histogram metric for improved sift matching, in ECCV, 2008.

[20] Silva, L., Bellon, O. and Boyer. K., Precision range image registration using a robust surface interpenetration measure and enhanced genetical gorithms, vol. 27 (5), pp. 762–776, May 2005.

[21] Bentley, J. L., Multidimensional binary search trees used for associative searching, Commun. ACM, vol. 18 (9), pp. 509–517, September 1975.

[22] Turk G., The Stanford bunny, 2000, [Online; accessed May 14, 2007]. [Online]. Available:http://www.gvu.gatech.edu/people/faculty/greg.turk/bunny/bunny.html

[23] Zinber, T., Schmidt, J., Niemann, H., A refined ICP algorithm for robust 3D correspondence estimation. En: in IEEE International Conference on Image Processing, 2003.

[24] Salvi, J., Matabosch C.: Range Image Registration Toolbox. – [Online; accessed Dic 10, 2010].