Computer Vision and Graphics for Heritage Preservation and Digital Archaeology

Luciano Silva [1]
Olga R.P. Bellon [1]
Kim L. Boyer [2]

Abstract: The goal of this work is to provide attendees with a survey of topics related to Heritage Preservation and Digital Archaeology, which are challenging and motivating subjects to both computer vision and graphics community. These issues have been gaining increasing attention and priority within the scientific scenario and among funding agencies and development organizations over the last years. Motivations to this work are the recent efforts in the digital preservation of cultural heritage objects and sites before degradation or damage caused by environmental factors or human development. One of the main focuses of these researches is the development of new techniques for realistic 3D model building from images, preserving as much information as possible. We intend to introduce and discuss several emerging topics in computer vision and graphics related to the proposed theme while highlighting the major contributions and advances in these fields.

Keywords: digital archaeology, heritage preservation, range images, multiview registration, 3D modeling of physical objects.

1 Introduction

Building realistic 3D models from sensor data is still a challenging problem. In recent years the demand for reconstruction and modeling of objects is increasing and it is widely used in many research areas, including medical imaging, robotic vision and archaeology [23]. Most applications focus on developing techniques to construct 3D object models of physical objects, preserving as much information as possible.

In 2003 important conferences have organized sections to discuss that issue, such as the IEEE/CVPR Workshop on Applications of Computer Vision in Archaeology (ACVA'03) in association with the IEEE Conference on Computer Vision and Pattern Recognition and the Special Session on Heritage Applications of 4th International Conference on 3-D Digital Imaging and Modeling (3DIM'03).

[1]Departamento de Informática, Universidade Federal do Paraná - Brazil, {luciano,olga@inf.ufpr.br}, IMAGO - Computer Vision, Graphics and Image Processing Research Group (www.inf.ufpr.br/imago).
[2]Department of Electrical and Computer Engineering, OSU - USA, {kim@ee.eng.ohio-state.edu}, SAMPL - Signal Analysis and Machine Perception Laboratory (sampl.eng.ohio-state.edu).

Recently, projects in digital archaeology have presented new challenges and are gaining popularity in computer vision and graphics community such as the Digital Michelangelo project [27] and the Great Buddha project [23]. A primary objective of these efforts is the digital preservation of cultural heritage objects before degradation or damage caused by environmental factors, erosion, fire, flood, or human development. Thus, archaeologists can also share their sites for in-depth investigation by other research groups or make them available in virtual museums promoting social inclusion, cultural diversity, and issues around.

Also, one can generate small scale replicas of memorable statues and hube sculptures from different places around the world to make them available in other museums. Furthermore, research on collaborative projects have supported the repair and restoration of historic buildings [33] and the analysis of complex structures by their 3D models [2]. We can list several relevant contributions in this area [2, 17, 16, 22, 23, 26, 27, 29, 33, 45].

Typically, a 3D model is built by the alignment (*i.e.* registration process) and integration of multiple views of an object or scene [5, 14, 22, 33, 41]. Since a physical object cannot be completely scanned with a single image, multiple scans from different views are required to supply the information needed to construct the 3D model. Usually, three dimensional geometry of the objects are obtained by a range scanner [3], which measure the distance from the scanner to each point on the surface of the object. The most important issues in the 3D modeling process are to minimize the number of views to reduce error accumulation in the 3D model and because data acquisition is expensive [23]. Therefore, it is fundamental to adopt a proper and robust registration technique to align the views in a common coordinate frame, *i.e.* a multiview registration process, to avoid model distortion in subsequent surface reconstruction stage. Indeed, the registration of views is the most important stage in the 3D modeling process.

When the object's views are effectively aligned one can obtain the entire 3D model by merging the views using a variety of well-known approaches for 3D model representation [37]. Usually, a triangular mesh is used and further simplifications are performed to reduce the number of points in the model while preserving the shape of the object [27]. In addition, texture mapping of the reconstructed 3D surfaces is also a challenging problem in creating realistic 3D models [23, 53].

The remainder of this paper is organized as follows: range images are described in Section 2. Section 3 presents the background of range image registration problem and related works. Also we introduces a novel evaluation measure, called the Surface Interpenetration Measure (SIM), to assess the quality of registration results and to guide the registration process. In Section 4 we present the 3D modeling process from range images with a critical analysis of several approaches. Finally, in Section 5 we conclude by summarizing the paper and discussing directions for future work.

2 Range images

In recent years, range scanners have been improved, allowing an increased number of applications in important areas, such as digital archaeology [2], building reconstruction and restoration [33], and medicine [49].

There is a number of range scanner models with distinct methods of acquisition and varying accuracy [3]. Some can acquire different images by combining time-of-flight (range image) and amplitude (reflectance image) of laser beams. These images can be combined to improve image processing stages, such as segmentation [47] or edge-based representations [46].

Most up-to-date range scanners use a laser beam to precisely measure the distance from the sensor to points in the surface of the object or scene [3], typically in a regular grid pattern as shown in Figure 1. This grid can be defined as a *range image* in which each point corresponds to a range sample. Each sampled point has 3D information, (x_i, y_i, z_i), where i is a position on the grid. By defining the resolution of the range image one can obtain the 3D coordinates for each sampled point in the surface of the object. The range image is also known as a $2\frac{1}{2}D$ representation because the 3D information relates only to the visible surface of the object as seen from a given view point.

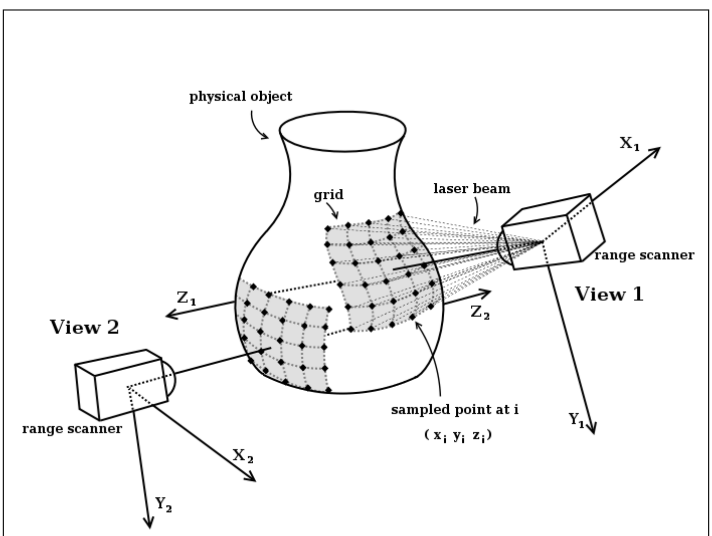
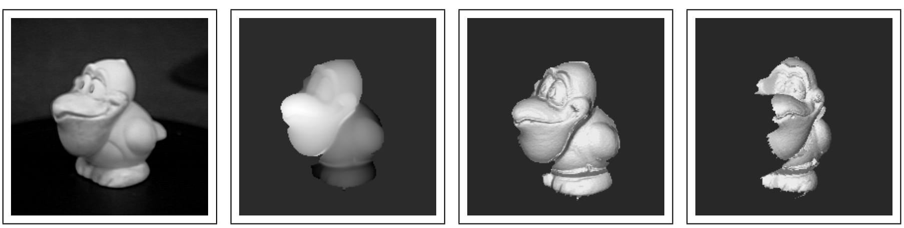


Figure 1. Diagram of the range image acquisition process for two distinct views.

More formally, a range image can be defined as a set of K discrete samples of a scalar function $j : \mathcal{I}^2 \rightarrow \mathcal{R}$, with $r_i = j(u_i)$, where $u_i \in \mathcal{I}^2$ is the index of the 2D grid (as shown in Figure 1), $r_i \in \mathcal{R}$ and $i = \{1, ..., K\}$. A range image gives the distances between the image plane and points on the objects surfaces in the scene. By consulting a lookup table that indicates the relationship between the image coordinate system and the range scanner coordinate system, a range image can be further converted to range data. These are defined as a set of K discrete samples points of a vector function $h : \mathcal{I}^2 \rightarrow \mathcal{R}^3$, with $d_i = h(u_i)$, where $d_i \in \mathcal{R}^3$ and $i = \{1, ..., K\}$. Then, each sampled point has 3D coordinates (x_i, y_i, z_i).

Since that only part of the object can be seen from any given view point, multiple views are needed to obtain the entire 3D surface of a physical object (see Figure 1). Also, it is necessary that these views have some overlap to allow their registration. As can be seen in Figure 1 there is no overlap between views 1 and 2. Figure 2 shows an example of a range image in which it is possible to note unseen regions on the object surface when it is observed from differents points of view.



(a) (b) (c) (d)

Figure 2. An example of range image acquired with the Minolta Vivid 700 scanner: (a) the picture of the object; (b) the range image, in which the lighter pixels are closer to the sensor than the dark ones; (c) and (d) two rendered views of the range image of (a) observed from different points of view.

There are several different range image databases, some are available on the Internet and others can be obtained from the research groups, as listed bellow:

- OSU range image database: Maintained by the Signal Analysis and Machine Perception Laboratory (SAMPL) at The Ohio State University - USA, coordinated by Prof. Kim L. Boyer. There are several images from different small objects acquired with the Minolta Vivid 700 range scanner. Each object was imaged in 18 views acquired at 20 degree intervals using a computer-controlled turntable (`sampl.eng.ohio-state.edu`).

- IMAGO range image database: Maintained by the IMAGO group at the Universidade Federal do Paraná - Brazil, coordinated by Prof. Olga R.P. Bellon. The images of small objects with high resolution and precision were acquired using the Roland MDX-15 scanner (`www.inf.ufpr.br/imago`).

- The Digital Michelangelo Project: Developed at Stanford University - USA, coordinated by Prof. Marc Levoy. This database has a number of range images from many statues created by Michelangelo in Italy. The images were acquired with high resolution using a special scanner fabricated for them by Cyberware (`graphics.stanford.edu/projects/mich`).

- The Great Temple in Petra: obtained by Prof. Frederic Leymarie of the SHAPE lab at Brown University - USA. The images were acquired with a ShapeGrabber laser scanner in June 2002 at the site of the Great Temple in Petra, Jordan (`www.lems.brown.edu/vision/extra/SHAPE`).

- The Cathedral of Saint Pierre: obtained by Prof. Peter Allen of the Robotics Group at Columbia University - USA. The images were acquired with a Cyrax 2400 scanner in 2001 for a project to model the Cathedral of Saint Pierre, Beauvais, France (`www1.cs.columbia.edu/~allen/BEAUVAIS`).

- The Thomas Hunter Building: obtained by Prof. Ioannis Stamos of Hunter College - USA. The external views of the Thomas Hunter building in New York were acquired with a Cyrax 2400 scanner.

- Stuttgart range image database: This database contains a collection of synthetic range images with different view points taken from high-resolution polygonal models available on the web (`range.informatik.uni-stuttgart.de`).

3 Range image registration

The registration process for two views consists of finding the best geometric transformation that, when applied to one view, aligns it with the other in a common coordinate system. This transformation represents the relative displacement between two surfaces having a common overlapping area. In fact, how to efficiently estimate this transformation is one of the main issues in registration [35].

There are many methods to perform the registration of views to create 3D models, including calibrated pose measurement and manual registration and verification [23, 52]. By using mechanical equipment (*e.g.* robot arm or controlled turntable) to obtain the absolute poses for the scanner views, the calibrated pose methods generally are limited to small objects [5]. In manual registration and verification, time is the main problem because the user must search for corresponding feature points in the view pair by hand.

When registering partially overlapped views, one of the most important issues is to develop methods to deal with low-overlapped views that can guarantee a precise alignment [35]. Recently, some papers have directly addressed the problem of calculating the overlapping area between views and measuring the registration quality [13, 22, 48, 45].

For the registration of two range images, A and B, the goal is to estimate the 3D rigid motion, represented by the transformation T, that best aligns the points in A with the points in B. We can define T by a 3×3 rotation matrix R and a 3D translation vector t, with $T = (R, t)$. A rigid motion of a 3D point p is given by a linear transformation as $T(p) = Rp + t$. The transformation applied to a set of n points $P = \{p_i\}$ with $i = \{1, ..., n\}$ is represented by $T(P) = \{Rp_i + t\}$.

We can parameterize T with a vector $\nu = [\theta_x, \theta_y, \theta_z, t_x, t_y, t_z]$, where θ_x, θ_y and θ_z are rotation angles about the x, y and z axes (the X-Y-Z Euler angles) and t_x, t_y and t_z are the components of the translation vector.

3.1 Registration approaches

The main differences among range image registration methods are the techniques to find corresponding points between views and to estimate the transformation to align the views. Also, some methods make use of image features other than points, such as edge maps and surface curvatures, to guide the registration process [39].

Clearly, if there is no overlap between the views, it is impossible to align them by only analyzing their 3D points (see Figure 1). Therefore, it is necessary to provide a significant overlapping area between views in the acquisition process to allow them to be registered.

Several registration approaches have been proposed during the last decade. They can be classified basically into two distinct groups: *coarse registration* and *fine registration* techniques. In coarse registration the goal is to find a set of approximate registration transformations without prior knowledge of the relative spatial positions of the views. Most of these methods are based on finding correspondences between distinctive features that may be present in the overlapping area. The basic procedure involves the identification of features, assignment of feature correspondences, and alignment based on these correspondences. There are many different features that can be explored: edge maps [39], lines and planes [18], bitangent curves [54], surface curvatures [9, 55], surface orientation [24] and invariant features, such as moments and curvature [41].

Sappa *et al.* [39] presented a method using an edge-based segmentation technique to guide the registration process. Chua and Jarvis [9] used principal curvatures and Darboux frames to compute invariant features. Also, Feldmar and Ayache [19] proposed a method to estimate rigid displacements by using principal curvatures of surfaces.

Stein and Medioni [50] proposed the splash structure, which is a local map describing the distribution of surface normals along a geodesic circle. The *spin image* presented by Johnson and Hebert [24], which is a data level shape descriptor, have been used in registration [22]. Lucchese *et al.* [29] exploit the geometric regularity obtained by the Fourier transform as a frequency domain-based method for range image registration. The RANSAC-based DARCES of Chen*et al.* [7] is a robust method based on exhaustive search that can check all possible data alignments between two given sets of points to register two partially overlapped views.

Because of the difficulty in calculating precise transformations, the coarse registration methods generally supply only rough alignments. In contrast, fine registration approaches are based on the assumption that a good initial transformation (*i.e.* close to the solution) was previously obtained. Then, precise alignments may be obtained with reliable criteria to measure the quality of the refined transformations. Additionally, a number of registration approaches in the literature proposed to combine both techniques, a coarse registration followed by a fine registration, to achieve automatic and precise registration results [7, 22, 39].

The best-known methods for fine range image registration are variations on the Iterative Closest Point (ICP) algorithm [4]. ICP is an iterative procedure minimizing the mean squared error (MSE), computed by the sum of the squared distances between points in one view and the closest points, respectively, in the other view. In each ICP iteration, the best geometric transformation that aligns the two images is calculated.

The ICP algorithm is composed of two basic procedures: one to find correspondences between points and another to estimate the transformations iteratively from the point correspondences until some termination criteria are satisfied. As reported by Besl and McKay [4], the correspondence search consumes about 95% of the runtime. Then, given two range images $A = \{a_i\}$ and $B = \{b_j\}$ with $i = \{1, ..., n\}$ and $j = \{1, ..., m\}$ the goal of the iterative process of the algorithm is to minimize the following function:

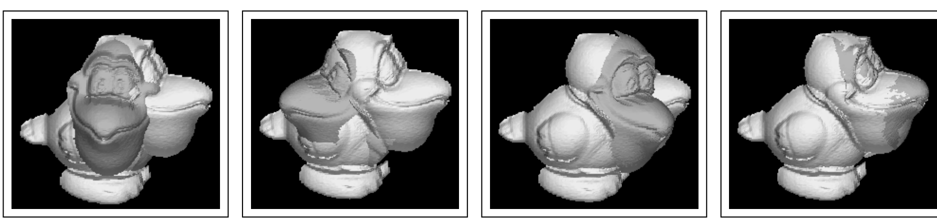
$$f(R,t) = \frac{1}{n}\sum_{i=1}^{n}||a_i - Rb_j - t||^2 \quad (1)$$

where b_j is the closest point in B to the point $a_i \in A$, R the rotation matrix and t the translation vector.

To find R and t one can use singular value decomposition or orthonormal matrices. The correspondence between points is usually performed by a nearest-neighbor search using *k*-d tree structure for optimization. The *k*-d tree is a spatial data structure originally proposed to allow efficient search on orthogonal range queries and nearest neighbor queries. Recently, Greenspan and Godin [20] proposed a significant improvement in the nearest neighbor queries by using correspondences of previous iterations of the ICP and searching only in their small neighborhood to update the correspondences. Another important strategy to speed up the registration process uses sampling techniques to reduce the number of points in the views [36].

Nevertheless, the proper convergence of ICP is guaranteed only if one of the datasets is a subset of the other; otherwise, erroneous alignments can result. Although ICP is efficient, with average case complexity of $O(n \log n)$ for n point images, it converges monotonically to a local minimum.

Another drawback of ICP is that it requires a good pre-alignment of the views to converge to a correct solution. Many variants of the ICP have been proposed [8, 27, 28, 36, 41] to overcome these limitations. For registration with partial overlap, heuristics have been proposed to ignore non-overlapped regions [36] and consequently to obtain more effective transformations. Section 3.2 addresses these methods. Figure 3 shows two registration results by an ICP-based method [27] using the same pair of range views, but in one a pre-alignment was applied before the registration procedure. As can be seen, without pre-alignment ICP converged to a completely erroneous result.



(a) (b) (c) (d)

Figure 3. Registration results by ICP for a pair of range images: (a) initial views positions without any pre-alignment; (b) the registration result starting from the views positioned as in (a); (c) pre-aligned views; and (d) the registration result from views positioned as in (c).

The main differences among these various proposals are in the evaluation functions to measure the quality of the alignments in each iteration and outlier rejection rules, such as discarding boundary points since they are more likely to yield false matches [27]. Chen and Medioni [8] developed an approach which minimizes the sum of squared distance between points in one view to a local tangent plane in the other view. Correspondences are formed by projecting the points onto the other view in the direction of their normal vectors (point-to-plane search) rather than searching for their closest point (point-to-point search). This approach is relatively faster than traditional ICP and usually the final results are better if a good initial pre-alignment is provided. However, this approach presents some numerical and search limitations since some correspondences may not be found [8].

Zhang [56] proposed a sophisticated set of modifications to ICP and his method has been used in many recent registration systems. The method automatically computes a threshold that is used to classify a point as an outlier if its distance from the corresponding point exceeds this threshold. By removing outliers from the registration calculations, one can estimate more precise transformations.

Masuda and Yokoya [32] proposed a robust method for registering a pair of dense range images, that integrates ICP, random sampling, and the least median squares estimator (LMS). They also proposed a modification of the *k*-d tree structure to improve the search and speed up the method.

A careful analysis of these various approaches reveals their true strengths and weaknesses [28]. In addition, comparative studies of variants of the ICP have been done [13, 36], but it is difficult to evaluate these comparisons because there is neither a common image database, nor well defined metrics. Recently, a special issue on the registration and fusion of range images [35] presents a great overview of different registration approaches.

The common lesson in these comparative studies is that, while ICP continues to undergo development and extension, progress on this front has become asymptotic; further improvements are likely to be incremental, at best. The characteristics of ICP that limit its effectiveness in domains where robustness is required and pre-alignment is infeasible remain. Occasionally, the strategies to overcome the pre-alignment limitation may guide the convergence process to an erroneous solution. What is needed is a fundamentally different search approach, including a robust assessment of alignment quality.

Another promising approach to register two range images is to find the geometric transformation through a pose-space search [34, 48], rather than the correspondence-based search of ICP-based methods. In this case, the goal is to find, in a huge search space of geometric transformations, a solution that can be used to align precisely two views, in a reasonable time. One way to perform this search is through efficient stochastic optimization techniques, such as Genetic Algorithms (GAs) [30] and Simulated Annealing (SA) [25]. This approach is generally considered to provide coarse registration. However, one can combine different operators, such as local search heuristics, to obtain precise alignments during the convergence process. In that way, the final registration results can be compared with traditional fine registration approaches.

GAs have been applied to image registration problems in several areas, including remote sensing and medical imaging. However, there exist many difficulties in developing reliable and automatic registration methods based on GAs to obtain precise alignment.

Brunnstrom and Stoddart [6] proposed a simple method for free-form surface matching combining GA and ICP approaches. First, an initial alignment is estimated by a traditional GA followed by ICP refinement to obtain a final registration. Their results show that, in some cases, this method becomes stuck in a local optimum; the authors concede the need for some modifications to their approach. In fact, GAs are efficient at finding promising areas of the search space, but are not so efficient at fine scale search within those areas. In practical terms, GAs can reach a solution close to the global optimum in a reasonable time, but a great deal of time may be required to improve the solution significantly beyond this point.

Despite their effectiveness, GAs are generally expensive to compute and have many problem-dependent parameters to adjust, which are empirically and carefully determined to avoid premature convergence. Since GAs can be performed using parallelism, many proposals for parallel genetic algorithms [34] have been developed to overcome the speed problem.

Recently, an alternative range image registration algorithm based on a GA was proposed [34] using a parallel evolutionary registration approach. The experimental results show that the method better avoids premature convergence. However, while this parallel GA approach reduces the computational time, it does not yield improved solutions, because the inter-point error between aligned views is not reduced.

We perform an extensive study using enhanced GA approaches for range image registration. We focus on the problem of obtaining precise alignments of range views through a robust registration method using views that are only partially overlapped. We explored the hybridization of GAs and hillclimbing heuristics as applied to range image registration and obtained better results in far less time than traditional GAs [48]. To support this, we define and use a new robust measure, the Surface Interpenetration Measure (SIM), to calculate the interpenetration of two registered range views, presented in Section 3.4.

3.2 Outlier rejection rules

The iterative search for corresponding points, as in ICP-based approaches, has become the dominant method for aligning range images because of its simplicity and efficiency. Given their main limitations, such as pre-alignment and local convergence, comparisons between different methods suggest a combination of heuristics to obtain a more robust ICP variant [36]. Most of these heuristics are related to outlier rejection rules to discard erroneous corresponding points.

In aligning partially overlapped range images, the points in regions without any correspondence must be labeled as outliers to avoid imprecise estimation of the transformation and consequently to prevent erroneous registration results [27].

One of the most effective outlier rejection rules is to exclude corresponding pairs that include points on the boundaries of the images [27] because correspondences including points on boundaries can introduce a systematic bias into the alignment. This rule is especially useful for avoiding erroneous pairings because in 3D modeling from range images the overlap between views is incomplete. The computational cost to apply this rule is usually low and has few drawbacks.

Another traditional rule applied in most registration approaches is to use a threshold to eliminate pairs with larger corresponding distance. In [40] the authors defined the threshold empirically based on the separation of the centers of mass of the images. Zhang [56] computes the threshold dynamically for each iteration based on the mean point pair distance and the standard deviation of the distances. Masuda and Yokoya [32] defines the threshold empirically to be 2.5σ, where σ signifies the standard deviation of residuals estimated by a robust estimator based on least median of squares. Also, in [27] the threshold is defined as a percentage of pairs with the largest corresponding distances. The main problem in defining the rule is how to set the threshold to deal with different range images and to guarantee precise alignments. In fact, this important issue has been explored elsewhere in computer vision and image processing research [21, 38].

Recently, Huber and Hebert [22] explored a combination of outlier rejection rules to define the overlapping region between views. They define a set of constraints to discard erroneous corresponding pairs, such as maximum distance between pairs and the maximum angle between the normal vectors of the pair. They also used the rule of eliminating points on the boundaries.

Generally, the rejection of outliers does not help with initial convergence, but it may improve the accuracy and stability with which the correct alignment is determined [36]. Usually, because of the difficulty of threshold selection the resulting registration may present local misalignment. Additionally, we have observed that even with a low threshold a precise alignment may not be guaranteed.

3.3 Registration quality measures

One of the main difficulties in evaluating registration results is acquiring the "ground truth" registration of two views. Some papers present quantitative evaluation metrics applying randomly chosen transformations to one view from a pair of aligned views (usually obtained from synthetic range data), and performing the registration [36]. After registration a mean distance is computed between the points of the transformed view before the transformations and after the registration, since the true corresponding points are known.

In registering real range images the true point correspondences are difficult to obtain even through calibrated ranging systems [5]. Therefore, erroneous alignments can generate a small inter-point error, giving a misleading (optimistic) measure because of the incorrect estimation of corresponding points. This error is usually computed by the mean squared error (MSE) between corresponding points of two images after the registration process.

Another problem is to define the minimum overlapping area between two views needed to guarantee a precise alignment. If there exists inadequate overlap between the views the registration process may generate erroneous solutions. In fact, to address this problem it is necessary to evaluate the registration method used before undertaking an exhaustive acquisition of multiple views [23].

A similar situation occurs in multiview registration in identifying overlapped views to be registered to generate a full 3D model. Huber and Hebert proposed a "brute force" method [22] performing exhaustive registrations between each pair of views and searching for a network of views with the lowest global error. They define a set of evaluation metrics to identify incorrect alignments based on the visibility of the registration and overlapping areas. By analyzing the z-buffer of the registration scene and the direction of the scanning process they can identify if there are occluding areas of a view and in this case the alignment is identified as an erroneous registration result.

Also, they define the concept of overlapped points in two range images based on three conditions: 1) the closest point distance between the point in one view and its corresponding point in the other view is less than a threshold t_1; 2) the points cannot be on the boundary of either image; and 3) the angle between the surface normal at the point in one view and the surface normal at its corresponding point in the other view is less than a threshold t_2. In their implementation, t_1 is estimated from the mean distance residuals at each iteration and t_2 is set at 45 degrees. The problem in using t_1 is that for imprecise alignment some parallel regions have a very low inter-point error and may be incorrectly labeled as overlapped regions. Also, threshold t_2 and condition 2 can only eliminate a small fraction of points. Consequently, the most important condition, and the main drawback, is to define a reasonable t_1. Usually, only clearly outlying points are eliminated with this criterion; it cannot define overlapping regions precisely.

Visual comparisons can provide qualitative evaluation of registration using real range images. Dalley and Flynn [12] suggest that a good registration must present a large "splotchy" surface, which is the visual result of two surfaces, each rendered in a different color, crossing over each other repeatedly in the overlapping area. This effect can be described as the *interpenetration* of the two surfaces. However, it is impossible to measure the degree of interpenetration by visual inspection alone because the resulting image depends on a variety of factors, such as rendering resolution, illumination, image sampling, surface representation, etc. At best we gain a qualitative assessment. Moreover, as a qualitative assessment, "splotchiness" provides no useful control mechanism to guide the registration process. Figure 4 shows two registrations obtained with ICP using range images, presented in [12], to illustrate the "splotchiness" effect.

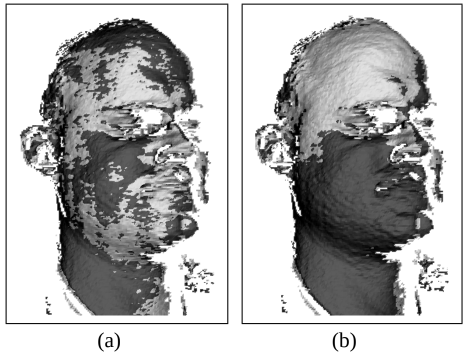


Figure 4. Examples of registration to illustrate the “splotchiness” effect (from [12]): (a) good “splotchiness” and (b) poor “splotchiness”.

3.4 Surface Interpenetration Measure (SIM)

We define the surface interpenetration by analyzing different visual results of two aligned surfaces, each rendered in a different color, crossing over each other repeatedly in the overlapping area. The interpenetration effect results from the nature of real range data, which presents slightly rough surfaces with small local distortions caused by limitations in the precision of the acquiring sensor or by noise. Because of this, even flat surfaces (*e.g.* polished plate made of wood or metal) present a “roughness” in range images. We also observed that two views acquired from the same object surface with the same scanner position and parameters provide two different range images.

By performing the registration of two range images of flat surfaces, we can confirm that there are interpenetrating regions between them. With this, we can assume that independently of the surfaces’ shapes the interpenetration will always occur.

By quantifying interpenetration, we can more precisely evaluate registration results and provide a highly robust control. To do this we developed the following measure based on the surface normal vector, computed by a local least squares planar fit, at each point. After the alignment of two images, A and B, we identify the set of interpenetrating points in A with respect to B. For each point $p \in A$ we define a neighborhood N_p to be a small $n \times n$ window centered on p. We choose $n = 5$ based on the observation that the interpenetration is a local effect.

With q denoting a point in the neighborhood N_p, c the corresponding point of p in image B (computed by a point-to-point nearest neighbor search using a *k*-d tree structure for optimization) and $\vec{n}_c$ the local surface normal at c, we define the set of interpenetrating points as:

$$C_{(A,B)} = \{p \in A \mid [(\overrightarrow{q_i - c}) \cdot \vec{n}_c][(\overrightarrow{q_j - c}) \cdot \vec{n}_c] < 0\} \tag{2}$$

where $q_i, q_j \in N_p$ and $i \neq j$. This set comprises those points in A whose neighborhoods include at least one pair of points separated by the local tangent plane, computed at their correspondents in B, as can be seen in the diagram of Figure 5.

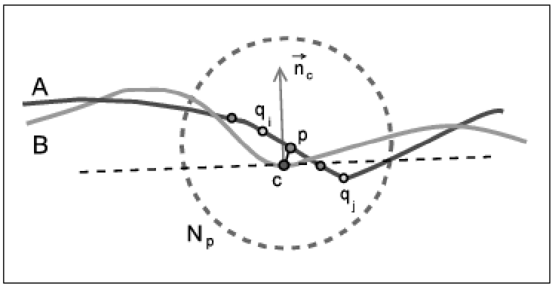


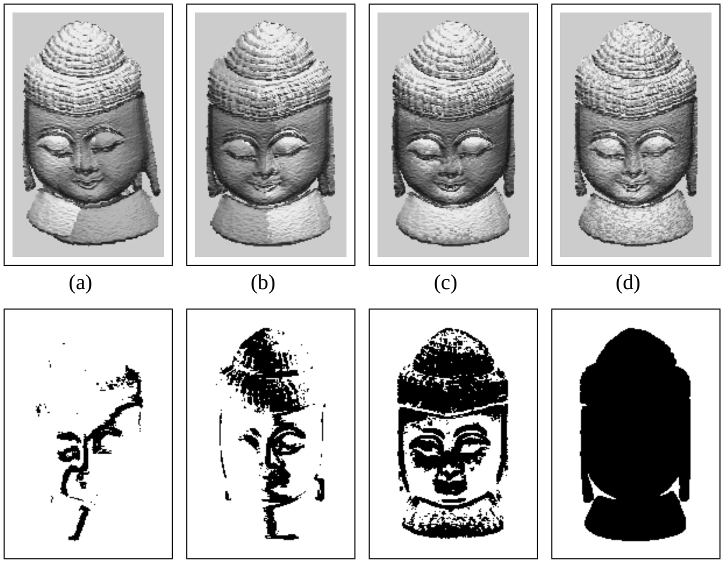
Figure 5. Interpenetrating point p in A with respect to B.

With this, we then define the Surface Interpenetration Measure as the fraction of interpenetrating points in A:

$$SIM_{(A,B)} = \frac{|C_{(A,B)}|}{|A|} \tag{3}$$

We also performed experiments using point-to-plane search [8] in Eq. 2, and the results are quite similar for precise alignments, but with some numerical limitations for erroneous alignments. A registration of two surfaces that presents a good interpenetration has a high SIM value. We also can analyze the distribution of the interpenetrating points by generating a binary image from the SIM calculation. Figure 6 offers an example of different alignments of two views with 100% of overlap. In this example both images are of the same view, but small rigid transformations were applied to one of them to simulate a misalignment in the registration process. After that, for each point the normal vectors are computed and the search for corresponding points was performed to give the necessary inputs for the SIM calculation (see Eq. 2).

Our experimental results show that erroneous alignments produce low SIM values and that *small differences in MSE can yield significant differences in SIM*. Furthermore, alignments with high SIM present a very low interpoint-error between surfaces. That is, the interpenetration measure is a far more sensitive indicator of alignment quality when comparing “reasonable” alignments. Also, the SIM is simple to compute and have linear time complexity $O(n)$ independently on the number of points in the image.



(a) (b) (c) (d)

(e) 11.97% (f) 30.67% (g) 60.77% (h) 100%

Figure 6. Binary images of interpenetrating points for different alignments: (a)-(d) are simulated alignments to compute the SIM and (e)-(h) are their binary images and SIM values obtained from the Eq. 3, respectively. The alignment shown in (d) was obtained with no transformation, *i.e.* both images A and B are the same range image.

We performed a number of experiments to evaluate the registration results in terms of SIM and MSE. The idea was to confirm that the SIM is a robust sensitive indicator of alignment quality for different alignments using partially overlapped views. One of the main advantages of the SIM is that one can identify precise registrations for good alignments. For instance, a correct alignment with low MSE may have no interpenetration if the aligned surfaces are parallel. However, by using the SIM we can reach a precise alignment with a high SIM while preserving a low MSE. Besides, many authors have reported that there may exist several local solutions for good alignments when the evaluation of registrations are computed by MSE [22, 41].

To evaluate and compare SIM *vs.* MSE measures, we performed a number of registrations using different combinations of aligned views with unequal overlapping areas as in [44]. Also, we explored different parameters and constraints to be applied to the SIM formulation and present a number of experiments to show its stability and performance on noisy surfaces.

4 3D modeling from range images

In summary, the 3D modeling process can be described as registration and integration of range views. These basic steps can be performed sequentially or simultaneously. Usually, sequential methods [8, 52] result in imprecise object models since the transformations errors accumulate and propagate from one iteration to another. However, if one can guarantee precise transformations, this method is more attractive and requires less computation resources (*i.e* memory) than others.

The simultaneous method generally is a more robust way to reach precise 3D models. In this category, a global registration between all pairs of views is performed, followed by their integration [14, 23, 31, 42]. In this process the accumulation error between the previously registered views is distributed among all alignments, avoiding model distortions while preserving the geometry.

Stoddart and Hilton [51] proposed one of the first global registration methods based on a physical equivalent model. A similar approach was proposed by Eggert *et al.* [15] based on a multi-resolution framework. Bergevin *et al.* [1] minimized the registration error of all views simultaneously using a well-balanced network of views.

Huber and Hebert [22] proposed a similar approach to find the minimum spanning tree in a graph, which represents possible model hypotheses for a set of views. The method assumes no knowledge of initial pose, nor even of which views overlap. However, the problem of inaccurate alignment between some views is evident, owing to the local convergence of ICP in the first stage. Nevertheless, their experimental results show that for some objects the resulting 3D model preserves topology without significant distortion.

In the multiview registration process, the precise alignment of two views is the fundamental stage. Therefore, it is important to adopt a robust registration technique to properly avoid incorrect alignments and further model distortion. It is also important to minimize the number of views to be aligned and, consequently, the registration method must be able to deal with low-overlapped views.

To reduce the complexity of the problem, most methods proposed for multiview registration perform an initial registration stage between each pair of overlapped views before the global registration process [23]. However, to obtain a reliable solution, a good alignment is required for each combination of overlapping views before attempting global registration.

Based on these observations, we developed a new method for multiview registration using GAs. Our approach combines a robust GA-based method with the SIM to obtain a precise alignment of low-overlap views [43]. The objective is to provide good alignment among more than two overlapping views before global registration to generate a precise 3D model. The results show that our multiview approach distributes the error between the alignments while preserving a good SIM between views. This is crucial for the construction of 3D models from multiple views, minimizing error accumulation in the final model.

After all views are precisely aligned, views can be fused to generate the 3D model. This process, also known as surface reconstruction, is usually accomplished by a volumetric approach [11], in which the data from multiple views are combined into a single and compact volumetric representation of the object. One of the main problems in surface reconstruction is the need to eliminate the redundancy of overlaps between view pairs. Generally, a weighted average is performed for nearby points that are considered to be from the same region. However, it is essential to have precise registrations, or alignment consistency, to avoid 3D model distortion [23].

Considering that we usually have many regions with a common overlapping area, if the views are not sufficiently close, the surface reconstruction result may generate overly smooth surfaces or even small distortions in the 3D model. Thus, the registration results, even for pair of views, must be robust to generate precise alignments. Our proposed approach [43] enjoys a great advantage in this process, because we use the SIM to drive the final refinement and the results contain large interpenetrating surfaces. Consequently, in the surface reconstruction stage one can accept small errors between views and still obtain precise 3D models.

Motivated by this observation, we performed some experiments using the SIM as a measure of alignment consistency for subsequent surface reconstruction. After the global multiview registration is complete we can locate in each view those points having no interpenetration with neighboring views. Thus, we can discard all these points to supply consistent overlapping regions in the global registration.

Our experiments have shown that noise regions present in many views of the objects in the OSU range image database can be eliminated through this simple rule using the SIM. Eventually, some gaps may appear in the "noise free" surface of the object, where the views are aligned but have parallel regions with no interpenetration.

Figure 7 shows the result of eliminating non-interpenetrating points of a global alignment of 9 views. As can be seen, some gaps appear in the object surface, indicating that in these regions it is necessary to add other overlapping views to provide a more accurate surface covering. Also one can use procedures for filling of gaps [10] to provide more realistic 3D models. Despite the small gaps, the alignment is very precise and reveals good overall interpenetration.

The experimental results presented in Figure 7 illustrates that many inconsistent regions (noise regions) were eliminated with our approach. We intend to investigate the use of the SIM in a surface reconstruction method for precise 3D model generation in the near future.

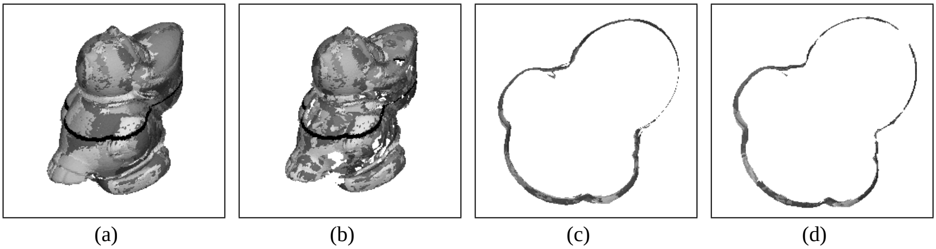


Figure 7. Comparison results of global alignments: (a) the global registration of 9 views; (b) the result by eliminating non-interpenetrating points; (c) and (d) are cross sections (black line) of the registrations in (a) and (b), respectively.

5 Final remarks

This paper has presented an overview of topics related to 3D model building for digital archaeology and heritage preservation. We emphasize that the most important stage in this problem is the registration process. We also introduce the problem of obtaining precise alignments between range views and how to identify and quantify them.

We have several ideas to extend the work and promising applications to pursue in the near future. Since we have a robust multiview registration method, we intend to develop surface reconstruction methods based on our work to generate precise 3D models from range images.

We also intend to make our methods available to the content-based image retrieval system of 3D images and models (CBIRS3D) under development in the IMAGO group (www.inf.ufpr.br/imago) at Universidade Federal do Paraná. The idea is to integrate the registration methods in the system to guide both reconstruction and visualization processes.

Acknowledgements

The authors would like to thank to CAPES and CNPq for financial support.

References

[1] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau. Towards a general multi-view registration technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):540–547, 1996.

[2] F. Bernardini, I. Martin, J. Mittleman, H. Rushmeier, and G. Taubin. Building a digital model of michelangelo's florentine pietà. *IEEE Computer Graphics & Applications*, 22(1):59–67, 2002.

[3] P.J. Besl. *Advances in Machine Vision*, chapter 1. Active Optical Range Imaging Sensors, pages 1–63. Springer-Verlag, New York, NY, 1989.

[4] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[5] G. Blais and M.D. Levine. Registering multiview range data to create 3D computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, 1995.

[6] K. Brunnstrom and A.J. Stoddart. Genetic algorithms for free-form surface matching. In *Proceedings of the 13th International Conference on Pattern Recognition*, volume 4, pages 689–693, 1996.

[7] C.S. Chen, Y.P. Hung, and J.B. Cheng. RANSAC-based DARCES: A new approach to fast automatic registration of partially overlapping range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1229–1234, 1999.

[8] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[9] C.S. Chua and R. Jarvis. 3D free-form surface registration and object recognition. *International Journal of Computer Vision*, 17(1):77–99, 1996.

[10] W. Corrêa, M.M. Oliveira, C.T. Silva, and J. Wang. Modeling and rendering of real environments. *Revista de Informática Teórica e Aplicada*, 9(2):127–156, 2002.

[11] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23th Annual Conference on Computer Graphics and Interactive Techniques*, pages 303–312, 1996.

[12] G. Dalley and P. Flynn. Range image registration: A software platform and empirical evaluation. In *Proceedings of the 3th International Conference on 3-D Digital Imaging and Modeling*, volume 1, pages 246–253, 2001.

[13] G. Dalley and P. Flynn. Pair-wise range image registration: A study in outlier classification. *Computer Vision and Image Understanding*, 87(1-3):104–115, 2002.

[14] C. Dorai, G. Wang, A.K. Jain, and C. Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):83–89, 1998.

[15] D.W. Eggert, A.W. Fitzgibbon, and R.B. Fisher. Simultaneous registration of multiple range views for use in reverse engineering of CAD models. *Computer Vision and Image Understanding*, 69(3):253–272, 1998.

[16] S.F. El-Hakim, J.A. Beraldin, M. Picard, and G. Godin. Detailed 3d reconstruction of large-scale heritage sites with integrated techniques. *IEEE Computer Graphics and Aplications*, 24(3):21–29, 2004.

[17] S.F. El-Hakim, J.A. Beraldin, M. Picard, and A. Vettore. Effective 3d modeling of heritage sites. In *Proceedings of the 4th International Conference on 3-D Digital Imaging and Modeling*, volume 1, pages 302–309, 2003.

[18] O.D. Faugeras and M. Hebert. The representation, recognition, and locating of 3D objects. *International Journal of Robotics Research*, 5(3):27–52, 1986.

[19] J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *International Journal of Computer Vision*, 18(2):99–119, 1996.

[20] M. Greenspan and G. Godin. A nearest neighbour method for efficient ICP. In *Proceedings of the 3th International Conference on 3-D Digital Imaging and Modeling*, volume 1, pages 161–168, 2001.

[21] J.F. Haddon. Generalized threshold selection for edge detection. *Pattern Recognition*, 21:195–203, 1988.

[22] D. Huber and M. Hebert. 3D modeling using a statistical sensor model and stochastic search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 858–865, 2003.

[23] K. Ikeuchi and Y. Sato, editors. *Modeling From Reality*. Kluwer Academic Publishers, 2001.

[24] A.E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, 1999.

[25] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[26] H.C.G. Leitao and J. Stolfi. A multiscale method for the reassembly of two-dimensional fragmented objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1239–1251, 2002.

[27] 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. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pages 131–144, 2000.

[28] Y. Liu and M.A. Rodrigues. Geometrical analysis of two sets of 3D correspondence data patterns for the registration of free-form shapes. *Journal of Intelligent and Robotic Systems*, 33(4):409–436, 2002.

[29] L. Lucchese, G. Doretto, and G.M. Cortelazzo. A frequency domain technique for range data registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(11):1468–1484, 2002.

[30] K.F. Man, K.S. Tang, and S. Kwong. Genetic algorithms: concepts and applications. *IEEE Transactions on Industrial Eletronics*, 43(5):519–534, 1996.

[31] T. Masuda. Registration and integration of multiple range images by matching signed distance fields for object shape modeling. *Computer Vision and Image Understanding*, 87(1-3):51–65, 2002.

[32] T. Masuda and N. Yokoya. A robust method for registration and segmentation of multiple range images. *Computer Vision and Image Understanding*, 61(3):295–307, 1995.

[33] M.K. Reed and P.K. Allen. 3D modeling from range imagery: An incremental method with a planning component. *Image and Vision Computing*, 17(2):99–111, 1999.

[34] C. Robertson and R.B. Fisher. Parallel evolutionary registration of range data. *Computer Vision and Image Understanding*, 87(1-3):39–50, 2002.

[35] M. Rodrigues, R. Fisher, and Y. Liu. Special issue on registration and fusion of range images. *Computer Vision and Image Understanding*, 87(1-3):1–7, 2002.

[36] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings of the 3th International Conference on 3-D Digital Imaging and Modeling*, volume 1, pages 145–152, 2001.

[37] S. Rusinkiewicz and M. Levoy. Streaming QSplat: a viewer for networked visualization of large, dense models. In *Symposium on Interactive 3D Graphics*, pages 63–68, 2001.

[38] P.K. Sahoo, S. Soltani, A.K.C. Wong, and Y.C. Chen. A survey of thresholding techniques. *CVGIP - Graphical Models and Image Processing*, 41(2):233–260, 1988.

[39] A.D. Sappa, A. Restrepo-Specht, and M. Devy. Range image registration by using an edge-based representation. In *International Symposium on Intelligent Robotic Systems*, pages 167–176, 2001.

[40] C. Schutz, T. Jost, and H. Hugli. Multi-feature matching algorithm for free-form 3D surface recognition. In *Proceedings 14th International Conference on Pattern Recognition*, pages 982–984, 1998.

[41] G.C. Sharp, S.W. Lee, and D.K. Wehe. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):90–102, 2002.

[42] H.Y. Shum, M. Hebert, K. Ikeuchi, and R. Reddy. An integral approach to free-form object modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(12):1366–1370, 1997.

[43] L. Silva, O.R.P. Bellon, and K.L. Boyer. Robust multiview range image registration. In *Proceedings of the 16th Brazilian Symposium on Computer Graphics and Image Processing*, volume 1, pages 80–88, 2003.

[44] L. Silva, O.R.P. Bellon, and K.L. Boyer. *Robust Range Image Registration*. World Scientific Publishing, Imperial College Press, 2004. (in press).

[45] L. Silva, O.R.P. Bellon, K.L. Boyer, and P.F.U. Gotardo. Low-overlap range image registration for archaeological applications. In *Proceedings of IEEE/CVPR Workshop on Applications of Computer Vision in Archaeology*, 2003.

[46] L. Silva, O.R.P. Bellon, and P.F.U Gotardo. Edge-based image segmentation using curvature sign maps from reflectance and range images. In *Proceedings of 8th IEEE International Conference on Image Processing*, volume 1, pages 730–734, 2001.

[47] L. Silva, O.R.P. Bellon, and P.F.U. Gotardo. A global-to-local approach for robust range image segmentation. In *Proceedings of 9th IEEE International Conference on Image Processing*, volume 1, pages 773–776, 2002.

[48] L. Silva, O.R.P. Bellon, P.F.U. Gotardo, and K.L. Boyer. Range image registration using enhanced genetic algorithms. In *Proceedings of 10th IEEE International Conference on Image Processing*, volume 2, pages 711–714, 2003.

[49] T.K. Sinha, D.M. Cash, R.J. Weil, R.L. Galloway, and M.I. Miga. Laser range scanning for cortical surface characterization during neurosurgery. In *Proceedings of SPIE Medical Imaging*, volume 5029, pages 98–107, 2003.

[50] F. Stein and G. Medioni. Structural indexing: Efficient 3-D object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):125–145, 1992.

[51] A.J. Stoddart and A. Hilton. Registration of multiple point sets. In *Proceedings IEEE International Conference on Pattern Recognition*, pages 40–44, 1996.

[52] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Proceedings of the 21st Annual Conference on Computer Graphics*, pages 311–318, 1994.

[53] J. Wang and M.M. Oliveira. Improved scene reconstruction from range images. *Computer Graphics Forum*, 3(21):512–530, 2002.

[54] J.W. Wyngaerd and L.V. Gool. Automatic crude patch registration: Toward automatic 3D model building. *Computer Vision and Image Understanding*, 87(1-3):8–26, 2002.

[55] S. Yamany and A. Farag. Free-form surface registration using surface signatures. In *Proceedings of International Conference on Computer Vision*, pages 1098–1104, 1999.

[56] Z.Y. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.