Abstract: We present an algorithm for the least squares matching of overlapping 3D surfaces. It estimates the transformation parameters between two or more fully 3D surfaces, using the Generalized Gauss-Markoff model, minimizing the sum of squares of the Euclidean distances between the surfaces. This formulation gives the opportunity of matching arbitrarily oriented 3D surfaces simultaneously, without using explicit tie points. Besides the mathematical model of the procedure, we discuss the computational aspects. We give practical examples to demonstrate the method.

1. Introduction

For 3D object modeling data acquisition must be performed from different standpoints. The derived local point clouds must be transformed into a common system. This procedure is usually referred to as registration. In the past, several efforts have been made concerning the registration of 3D point clouds. One of the most popular methods is the Iterative Closest Point (ICP) algorithm developed by Besl and McKay [3], Chen and Medioni [8], and Zhang [40]. The ICP is based on the search of pairs of nearest points in the two sets and estimates the rigid body transformation, which aligns them. Then, the rigid body transformation is applied to the points of one set and the procedure is iterated until convergence.

In Besl and McKay [3] and Zhang [40] the ICP requires every point in one surface to have a corresponding point on the other surface. Alternatively, the distance between the transformed points in one surface and corresponding tangent planes on the other surfaces was used as a registration evaluation function [2,8,31]. The point-to-tangent plane approach gives a better registration accuracy than the point-to-point approach.

The outliers due to erroneous measurements (e.g. points on the object silhouette) and occlusions may significantly impair the quality of the registration. The following strategies have been proposed for localization and elimination of outliers and occlusions: rejection of pairs based on predefined (constant) distance threshold [4,11,21,38,40] or variable distance thresholds adapted from Robust Estimation Methods [15,20,25,26], rejection of pairs based on the orientation threshold for surface normals [21,40], rejection of pairs containing points on mesh boundaries [21,31,38], rejection of pairs based on the reciprocal correspondence [28], rejection of the worst $n\%$ of pairs [31], employing the Least Median of Squares [25] and the Least Trimmed Squares [10] estimators.

The parameters of the rigid body transformation are generally estimated by use of closed-form solutions, mainly singular value decomposition and quaternion methods. For an extensive review and comparison we refer to Williams et al. [39] and Eggert et al. [14]. The closed-form solutions cannot fully consider the statistical point error models. Zhang [40] and
Dorai et al. [12] weighted the individual points based on a priori noise information. In references [21,27,39] methods that can model the anisotropic point errors were proposed.

The gradient descent type of algorithms can support full stochastic models for measurement errors, and assure substantially less number of iterations than the ICP variants [15,26,37]. The Levenberg-Marquardt method is usually adopted for the estimation.

The ICP, and in general all surface registration methods, require heavy computations. Strategies, mainly employed to reduce the computation time are: reduction of the number of iterations, reduction of the number of employed points, and speeding up the correspondence computation. Extensive surveys on commonly used methods are given in [1,23,29].

Several reviews and comparison studies about surface registration methods are available in the literature [7,20,22,34,39].

In Photogrammetry, the problem statement of surface patch matching and its solution method was first addressed by Gruen [16] as a straight extension of Least Squares Matching (LSM).

There have been some studies on the absolute orientation of stereo models using Digital Elevation Models (DEM) as control information [13,33]. This work is known as DEM matching. This method basically estimates the 3D similarity transformation parameters between two DEM patches, minimizing the sum of squares of differences along the z-axes. Schenk et al. [36] showed the clear advantage of minimization of distances along surface normals against to minimization of elevation differences. Beside the many other applications it has been used for the registration of airborne laser scanner strips as well [24,30]. The DEM matching corresponds mathematically to Least Squares Image Matching, but can only be applied to 2.5D surfaces, which is of limited value in case of generally formed objects.

In our previous work an algorithm for least squares matching of overlapping 3D surfaces was given [18,20]. It estimates the transformation parameters between two or more fully 3D surfaces, using the Generalized Gauss-Markoff model, minimizing the sum of squares of the Euclidean distances between the surfaces. This formulation gives the opportunity of matching arbitrarily oriented 3D surfaces simultaneously, without using explicit tie points. Our mathematical model is a generalization of the least squares image matching method, in particular the method given by Gruen [16]. We gave further extensions of the basic model: simultaneous matching of multi sub-surface patches, and matching of surface geometry and its attribute information, e.g. reflectance, color, temperature, etc. under a combined estimation model [19].

In this study we focus on the computational aspects with regard to outlier and occlusion detection and the optimization of the run-time of the correspondence computation. The details of the mathematical modeling of the proposed method and the execution aspects are explained in the following section. The two strategies for the fast correspondence computation are given in the third section. Practical examples to demonstrate the feasibility of the method are presented in the fourth section.

2. Least Squares 3D Surface Matching (LS3D)

2.1. The basic estimation model

Assume that two different partial surfaces of the same object are digitized/sampled point by point, at different times (temporally) or from different viewpoints (spatially), \( f(x, y, z) \) and \( g(x, y, z) \) are conjugate regions of the object in the left and right surfaces respectively. In other words \( f(x, y, z) \) and \( g(x, y, z) \) are discrete 3D representations of the template and search surfaces. The problem statement is estimating the final location, orientation and shape of the search surface \( g(x, y, z) \), which satisfies minimum condition of Least Squares Matching with respect to the template \( f(x, y, z) \). In an ideal situation one would have
Taking into account the noise and assuming that the template noise is independent of the search noise, Equation (1) becomes

\[ f(x, y, z) - e(x, y, z) = g(x, y, z) \]  

where \( e(x, y, z) \) is a true error vector. Equation (2) are observation equations, which functionally relate the observations \( f(x, y, z) \) to the parameters of \( g(x, y, z) \). The matching is achieved by least squares minimization of a goal function, which represents the sum of squares of the Euclidean distances between the surfaces. The final location is estimated with respect to an initial position of \( g(x, y, z) \), the approximation of the conjugate search surface \( g^0(x, y, z) \).

To express the geometric relationship between the conjugate surface patches, a 7-parameter 3D similarity transformation is used:

\[
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix}^T = \begin{bmatrix} t_x & t_y & t_z \\ 0 & 0 & 0 \end{bmatrix}^T + m R_0 \begin{bmatrix} x_0 \\
  y_0 \\
  z_0 \end{bmatrix}^T
\]

where \( R = R(\omega, \phi, \kappa) \) is the orthogonal rotation matrix, \( [t_x, t_y, t_z]^T \) is the translation vector, and \( m \) is the uniform scale factor. This parameter space can be extended or reduced, as the situation demands it.

In order to perform least squares estimation, Equation (2) must be linearized by Taylor expansion.

\[
f(x, y, z) - e(x, y, z) = g^0(x, y, z) + \frac{\partial g^0(x, y, z)}{\partial x} d x + \frac{\partial g^0(x, y, z)}{\partial y} d y + \frac{\partial g^0(x, y, z)}{\partial z} d z
\]

with

\[
dx = \frac{\partial x}{\partial p_i} d p_i, \quad dy = \frac{\partial y}{\partial p_i} d p_i, \quad dz = \frac{\partial z}{\partial p_i} d p_i
\]

where \( p_i \in \{t_x, t_y, t_z, m, \omega, \phi, \kappa\} \) is the \( i \)-th transformation parameter in Equation (3). Differentiation of Equation (3) gives:

\[
d x = d t_x + a_{10} d m + a_{11} d \omega + a_{12} d \phi + a_{13} d \kappa
\]

\[
d y = d t_y + a_{20} d m + a_{21} d \omega + a_{22} d \phi + a_{23} d \kappa
\]

\[
d z = d t_z + a_{30} d m + a_{31} d \omega + a_{32} d \phi + a_{33} d \kappa
\]

where \( a_{ij} \) are the coefficient terms, whose expansions are trivial. Using the following notation

\[
g_x = \frac{\partial g^0(x, y, z)}{\partial x}, \quad g_y = \frac{\partial g^0(x, y, z)}{\partial y}, \quad g_z = \frac{\partial g^0(x, y, z)}{\partial z}
\]

and substituting Equations (6), Equation (4) results in the following:

\[
-e(x, y, z) = g_x d t_x + g_y d t_y + g_z d t_z + (g_x a_{10} + g_y a_{20} + g_z a_{30}) d m + (g_x a_{11} + g_y a_{21} + g_z a_{31}) d \omega + (g_x a_{12} + g_y a_{22} + g_z a_{32}) d \phi + (g_x a_{13} + g_y a_{23} + g_z a_{33}) d \kappa - (f(x, y, z) - g^0(x, y, z))
\]
In the context of the Gauss-Markoff model, each observation is related to a linear combination of the parameters, which are variables of a deterministic unknown function. The terms \( \{g_x, g_y, g_z\} \) are numeric first derivatives of this function \( g(x,y,z) \). Equation (8) gives in matrix notation

\[
-e = Ax - l, \quad P
\]  

where \( A \) is the design matrix, \( x^T = [d_x, d_y, d_z, dm, d\omega, d\varphi] \) is the parameter vector, and \( l = f(x,y,z) - g(x,y,z) \) is the discrepancy vector that consists of the Euclidean distances between the template and correspondent search surface elements. The template surface elements are approximated by the data points, on the other hand the search surface elements are represented in two different kind of piecewise surface forms (planar and bi-linear) optionally. In general both surfaces can be represented in any kind of piecewise form.

With the statistical expectation operator \( E\{\} \) and the assumptions \( E\{e\} = 0, E\{ee^T\} = \sigma_0^2 P_{ll}^{-1} \) Equation (9) is a Gauss-Markoff estimation model, where \( P = P_{ll} \) is a priori weight matrix.

The unknown transformation parameters are treated as stochastic quantities using proper a priori weights. This extension gives advantages of control over the estimating parameters. We introduce the additional observation equations on the system parameters as

\[
-e_b = Ix - l_b, \quad P_b
\]  

where \( I \) is the identity matrix, \( l_b \) is the (fictitious) observation vector for the system parameters, and \( P_b \) is the associated weight coefficient matrix. The least squares solution of the joint system Equations (9) and (10) gives as the Generalized Gauss-Markoff model the unbiased minimum variance estimation for the parameters

\[
\hat{x} = (A^T PA + P_b)^{-1}(A^T PL + P_b l_b)
\]

(solution vector)  

\[
\hat{\sigma}_0^2 = (v^T P v + v_b^T P_b v_b)/r
\]

(variance factor)  

\[
v = A\hat{x} - l
\]

(residuals vector for surface observations)  

\[
v_b = I\hat{x} - l_b
\]

(residuals vector for parameter observations)  

where ^ stands for the Least Squares Estimator, and \( r \) is the redundancy. Since the functional model is non-linear, the solution is obtained iteratively. In the first iteration the initial approximations for the parameters must be provided. After the solution vector (Equation 11) is solved, the search surface \( g^0(x,y,z) \) is transformed to a new state using the updated set of transformation parameters, and the design matrix \( A \) and the discrepancies vector \( l \) are re-evaluated. The iteration stops if each element of the alteration vector \( \hat{x} \) in Equation (11) falls below a certain limit: \( |d_{pi}| < c_i \).

The numerical derivative terms \( \{g_x, g_y, g_z\} \) are defined as local surface normals \( n \). Their calculation depends on the analytical representation of the search surface elements. Two first degree \( C^0 \) continuous surface representations are implemented: triangle mesh form, which gives planar surface elements, and optionally grid mesh form, which gives bi-linear surface elements. The derivative terms are given as \( x-y-z \) components of the local normal vectors:

\[
[g_x \ g_y \ g_z]^T = n = [n_x \ n_y \ n_z]^T.
\]

For the details of the method we refer to [20].

2.2. Error detection and execution aspects

The standard deviations of the estimated transformation parameters and the correlations between themselves may give useful information concerning the stability of the system and quality of the data content [16]:
where $Q_{xx}$ is the cofactor matrix for the estimated parameters.

Detection of false correspondences with respect to the outliers and occlusions is a crucial part of every surface matching method. We use the following strategies in order to localize and eliminate the outliers and the occluded parts.

A median type of filtering is applied prior to the matching. For each point the distances between the central point and its 8-neighbourhood points are calculated. If some of those 8 distance values are much greater than the average point density, the central point is likely to be an erroneous point on a poorly reflecting surface (e.g. window or glass) or a range artifact due to surface discontinuity (e.g. points on the object silhouette). The central point is discarded according to the number of distances $n$, which are greater than a given distance threshold.

In the course of iterations a simple weighting scheme adapted from Robust Estimation Methods is used:

$$\hat{P}_{ii} = \begin{cases} 1 & \text{if } |(v)_{i}| < K\sigma_{0} \\ 0 & \text{else} \end{cases}$$

(16)

In our experiments $K$ is selected as $>10$, since it is aimed to suppress only the large outliers. It can be changed according to a given confidence level. Finally, we reject the correspondences containing points on the mesh boundaries. Because of the high redundancy of a typical data set, a certain amount of occlusions and/or smaller outliers do not have significant effect on the estimated parameters.

The convergence behaviour of the proposed method basically depends on the quality of the initial approximations and quality of the data content. For a good data configuration case it usually achieves the solution after 5 or 6 iterations.

Two first degree $C^0$ continuous surface representations are implemented. In the case of multi-resolution data sets, in which point densities are significantly different on the template and search surfaces, higher degree $C^1$ continuous composite surface representations, e.g. bi-cubic Hermit surface, should give better results, of course increasing the computational expense.

3. Acceleration Strategies

3.1. Fast correspondence computation with boxing structure

The computational effort increases with the number of points in the matching process. The main portion of the computational complexity is to search the corresponding elements of the template surface on the search surface, whereas the parameter estimation part is a small system, and can quickly be solved using Cholesky decomposition followed by back-substitution. Searching the correspondence is guided by an efficient boxing structure [9], which partitions the search space into cuboids. For a given surface element, the correspondence is searched only in the box containing this element and in the adjacent boxes. In the original publication [9] it was given for 2D point sets. We straightforwardly extend it to the 3D case. For the implementation details we refer to [1]. The access procedure requires $O(q)$ operations, where $q$ is the average number of points in the box. It is easy to implement and time-effective for accessing the data.

In our implementation, the correspondence is searched in the boxing structure during the first few iterations, and in the meantime its evolution is tracked across the iterations. Afterwards the searching process is carried out only in an adaptive local neighborhood according to the
previous position and change of correspondence. In any step of the iteration, if the change of correspondence for a surface element exceeds a limit value, or oscillates, the search procedure for this element is returned to the boxing structure again.

3.2. Simultaneous multi-subpatch matching

The basic estimation model can be implemented in a multi-patch mode, that is the simultaneous matching of two or more search surfaces \( g_i(x, y, z) \), \( i = 1, \ldots, k \) to one template \( f(x, y, z) \).

\[-e_i = A_i x_i - l_i \, , \, P_i \]  

Equation (17)

Since the parameter vectors \( x_1, \ldots, x_k \) do not have any joint components, the sub-systems of Equation (17) are orthogonal to each other. In the presence of auxiliary information those sets of equations could be connected via functional constraints, e.g. as in the Geometrically Constrained Multiphoto Matching [16,17] or via appropriate formulation of multiple (>2) overlap conditions.

An ordinary point cloud includes enormously redundant information. A straightforward way to register such two point clouds could be matching of the whole overlapping areas. This is computationally expensive. We propose multi-subpatch mode as a further extension to the basic model, which is capable of simultaneous matching of sub-surface patches, which are interactively selected in cooperative surface areas. They are joined to the system by the same 3D transformation parameters. This leads to the observation equations

\[-e_i = A_i x - l_i \, , \, P_i \]  

Equation (18)

with \( i = 1, \ldots, k \) subpatches. They can be combined as in Equation (9), since the common parameter vector \( x \) joints them to each other. The individual subpatches may not include sufficient information for the matching of whole surfaces, but together they provide a computationally effective solution, since they consist of only relevant information rather than using the full data set.

4. Experimental Results

Two practical examples are given to show the capabilities of the method. All experiments were carried out using own self-developed C/C++ software that runs on an Intel® P4 2.53Ghz PC. In all experiments the initial approximations of the unknowns were provided by interactively selecting 3 common points on both surfaces before matching. The scale factor \( m \) was fixed to unity by infinite weight value \((P_{bi})_i \rightarrow \infty\).

The first example is the registration of two point clouds of a newspaper page (Fig. 1). The scanning was performed by using the stereoSCAN\(^3\)D system developed by Breuckmann GmbH (Germany). It is a high accurate scanner system based on the fringe projection technique [6].

The average point spacing is 150~170 microns. The surface of Figure 1a was matched to the one in Figure 1b by use of LS3D surface matching. The iteration criteria values \( c_i \) were selected as 1 micron for the translation vector and \( 10^4 \) for the rotation angles. Although it is a difficult example due to very little changes in surface curvature, the matching is successful (Fig. 1c). Totally 377234 points were used for matching. The \textit{a posteriori} sigma value was 11.3 microns, with 13 iterations in 36.7 seconds. Interestingly, the letters are clearly visible on the surface model (Fig. 1c). However, they are due to range artifacts created by the pixel-discretization on the chip level, leading to intensity discontinuities. For a detailed discussion of the range artifacts we refer to Blais et al. [5].
A comparison against the non-accelerated version was made. The non-accelerated version exhaustively searches the correspondence in a large portion of the search surface during the first few iterations. In the following iterations it uses the same adaptive local neighborhood search as in the accelerated version. For a fair comparison same number of points were employed in the matching. The non-accelerated version found the same solution in 106.1 seconds. As seen in this experiment, the accelerated version speeds the computation up typically by factor 2 to 3 [1]. This is the sole effect of the space partitioning technique.

Figure 1: Example “newspaper”. (a) Search and (b) template surfaces, (c) shaded view of the final composite surface after the LS3D surface matching. Zoom-in part of (c) shows the range artifacts due to intensity discontinuities. Note that the scanner derived intensities are back-projected onto the surfaces (a) and (b) only for the visualization purposes, they are not used in matching.

The second experiment refers to the matching of point clouds of a wooden Buddha statue (ca 30x40x20 cm$^3$). It has a very shiny polished surface (Fig. 2a), which is not an optimal surface reflectivity case. In a recent study Remondino et al. [32] have modeled the same object using the photogrammetric technique and additionally the BIRIS laser scanner. In our experiment
the point clouds were acquired by the triTOS system, which is another structured light-based scanner product of Breuckmann GmbH. It is mainly used for art and cultural heritage applications [6].

The data set contains 15 scans, each of which has ca. 1.4 million points. The average point spacing is 0.3~0.5 millimeters. Nineteen consecutive matching processes were performed using the simultaneous multi-subpatch approach of the LS3D matching method. The iteration criteria values $c_i$ were selected as 0.1 micron for the translation vector and $10^{-5}$ for the rotation angles. The average numerical results of the matching are given in Table 1.

![Figure 2: Example “wooden Buddha”. (a) The wooden Buddha statue, (b) shaded view of the generated 3D model.](image)

The first scan was selected as the reference, which defines the datum of the common coordinate system. Since it is a closed object, there is need for a global registration, which distributes the residuals evenly among all the scans, and also considers the closure condition, i.e., matching of the last scan to the first one. For this purpose we used the block adjustment by independent model solution, which was formerly proposed for global registration of laser scanner point clouds, but for the case of retro-reflective targets as tie points [35]. In the LS3D matching processes, the final correspondences were saved to separate files. Then all these files were given as input to a block adjustment by an independent model procedure, which concluded with 42 microns $a$ posteriori sigma value. At this step the model contains ca. 9.5 million triangles. Using Geomagic Studio v.6 (Raindrop Geomagic) all surfaces were merged as one manifold, in parallel reducing the number of triangles to ca. 3 million and applying a low level noise reduction (Fig. 2b).

<table>
<thead>
<tr>
<th>Average no. of employed points</th>
<th>Average no. of iterations</th>
<th>Average CPU times (sec.)</th>
<th>Average $a$ posteriori sigma values (micron)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100~400K</td>
<td>4~11</td>
<td>9~52</td>
<td>66~105</td>
</tr>
</tbody>
</table>
5. Conclusions
An algorithm for the least squares matching of overlapping 3D surfaces is presented. Our proposed method, the Least Squares 3D Surface Matching (LS3D), estimates the transformation parameters between two or more fully 3D surfaces, using the Generalized Gauss Markoff model, minimizing the sum of squares of the Euclidean distances between the surfaces. The mathematical model is a generalization of the least squares image matching method and offers high flexibility for any kind of 3D surface correspondence problem. The least squares concept allows for the monitoring of the quality of the final results by means of precision and reliability criterions.

The practical example shows that our proposed method can provide successful matching results in reasonable processing times. The use of our space partitioning technique alone leads to a speed up of computing times by factor 2-3. Another aspect of our experiment is that registration task can be performed automatically without using retro-reflective or other special kinds of targets even for the surfaces with little geometric information.

Acknowledgements
The data sets are courtesy of Breuckmann GmbH. The authors would like to thank Hans Woerner and Fabio Remondino for the data acquisition. The first author of this paper is financially supported by an ETHZ Research Grant, which is gratefully acknowledged.

References:


