Co-registration of Surfaces by 3D Least Squares Matching

Devrim Akca

Abstract

A method for the automatic co-registration of 3D surfaces is presented. The method utilizes the mathematical model of Least Squares 2D image matching and extends it for solving the 3D surface matching problem. The transformation parameters of the search surfaces are estimated with respect to a template surface. The solution is achieved when the sum of the squares of the 3D spatial (Euclidean) distances between the surfaces are minimized. The parameter estimation is achieved using the Generalized Gauss-Markov model. Execution level implementation details are given. Apart from the co-registration of the point clouds generated from spaceborne, airborne and terrestrial sensors and techniques, the proposed method is also useful for change detection, 3D comparison, and quality assessment tasks. Experiments using terrain data examples show the capabilities of the method.

Introduction

With the availability of the various sensors and automated methods, the production of large numbers of point clouds is no longer particularly notable. In many cases, the object of interest is covered by a number of point clouds, which are referenced in different spatial or temporal datums. Therefore, the issue of co-registration of point clouds (or surfaces) is an essential topic in 3D modeling.

In terrestrial laser scanning practice, special targets provided by the vendors, e.g., Zoller+Fröhlich, Leica, and Riegl, are mostly used for the co-registration of point clouds. However, such a strategy has several deficiencies with respect to fieldwork time, personnel and equipment costs, and accuracy. In a recent study, Sternberg *et al.* (2004) reported that registration and geodetic measurements comprise 10 to 20 percent of the total project time. In another study, a collapsed 1,000-car parking garage was documented in order to assess the damage and structural soundness of the building. The laser scanning took three days, while the conventional survey of the control points required two days (Greaves, 2005). In a project conducted by our research group at Pinchango Alto (Lambers *et al.*, 2007), two persons set the targets in the field and measured them using the real-time kinematic GPS technique in one and one-half days.

As well as fieldwork time, accuracy is another important concern. The target-based registration methods may not exploit the full accuracy potential of the data. The geodetic measurements naturally introduce some error, which might exceed the internal error of the scanner instrument. In addition, the targets must be kept stable during the whole of the data acquisition campaign. This might be inconvenient when the scanning work stretches over more than one day. On the other hand, one important advantage of the target based methods should not be ignored. Targets are essentially required in projects where the absolute orientation to an object coordinate system is needed.

The surface based registration techniques stand as efficient and versatile alternatives to the target-based techniques. They simply bring the strenuous additional fieldwork of the registration task to the computer in the office, at the same time optimizing the project cost and duration, and achieving a better accuracy. In the last decade, surface-based registration techniques have been studied extensively. The large number of research activities on the topic demonstrates the relevance of the problem. For an exhaustive literature review, we refer to Gruen and Akca (2005).

The co-registration is crucially needed wherever spatially related data sets can be described as surfaces and has to be transformed to each other. Examples can be found in medicine, computer graphics, animation, cartography, virtual reality, industrial inspection and quality control, change detection, spatial data fusion, cultural heritage, photogrammetry, etc.

We treat the co-registration problem as a Least Squares matching of overlapping surfaces. Least Squares matching is a mathematical concept, which was originally developed for automatic point transfer on stereo or multiple images (Ackermann, 1984; Pertl, 1984; Gruen, 1985). More recently, it has been extended to many problem-specific cases, e.g., 3D voxel matching (Maas and Gruen, 1995) and the line feature extraction techniques (Gruen, 1996).

This work, called 3D Least Squares surface matching (LS3D), is another straightforward extension of the 2D Least Squares image matching and has the same underlying ideas and concepts. The next section introduces the basic estimation model. The execution aspects and the implementation details are extensively elaborated. Particular attention is given to the surface-to-surface correspondence search, outlier detection, and the computational acceleration. In previous work, examples covering the co-registration of the terrestrial laser scanning data sets were given (Gruen and Akca, 2005; Akca, 2007a). In the work presented here, the experimentation concentrates on the topographic data sets, mostly generated by use of photogrammetric and airborne lidar techniques. Apart from the co-registration of point clouds, the proposed method can be utilized for many types of geo-data

Department of Information Technologies, Isik University, 34980 Sile, Istanbul, Turkey, and formerly with the Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland (akca@isikun.edu.tr).

analyses. The experiments given in the third section show examples of how the 3D surface matcher performs more advantageously than the conventional methods for change detection and quality assessment tasks.

Least Squares 3D Surface Matching

The Basic Estimation Model

Two 3D surfaces are subject to a co-registration procedure. The search surface $g(x, y, z)$ is going to be transformed to the spatial domain of the template surface $f(x, y, z)$. Both $f(x, y, z)$ and $g(x, y, z)$ are piecewise discrete representations of the continuous function of the object surface. In the current implementation, a triangular mesh or a grid mesh type of representation is used. In the case of the triangular mesh representation, the piecewise surface is composed of planar triangle elements; in the same manner, the grid mesh representation is composed of bi-linear grid surface elements.

Surface topology is established simply by loading the data files, e.g., range scanner point clouds or photogrammetric digital elevation models (DEM), in row or column order. The point spacing is by definition irregular, wherefore the regular grid DEM is an under-capability case.

Every template surface element is matched to a conjugate search surface element to establish the surface-to-surface correspondence. This is achieved by a correspondence operator. Occlusions and outliers are the perturbation cases, which are excluded from the system automatically. While all template surface elements are sought by the operator, some of the search surface elements might not coincide at all.

If a matching is established between the two surface elements $f(x, y, z)$ and $g(x, y, z)$, the following equation holds:

$$f(x, y, z) - e(x, y, z) = g(x, y, z) \qquad (1)$$

where $e(x, y, z)$ is a true error vector covering the random errors of the template and search surfaces, which are assumed to be uncorrelated. Equation 1 is the observation equation, which is set up for each template element that has a valid surface match. The transformation parameters of the search surface $g(x, y, z)$ are variables to be estimated.

Here, we have a peculiar case where the search surface $g(x, y, z)$ is not analytically continuous; rather it is composed of discrete finite elements in the form of planar triangles or bilinear grids. As a consequence, the mathematical derivation operation cannot be performed analytically.

Equation 1 is non-linear. It is linearized by the Taylor Series expansion:

$$-e(x, y, z) = \frac{\partial g^0(x, y, z)}{\partial x}\,\mathrm{d}x + \frac{\partial g^0(x, y, z)}{\partial y}\,\mathrm{d}y$$
$$+ \frac{\partial g^0(x, y, z)}{\partial z}\,\mathrm{d}z - (f(x, y, z) - g^0(x, y, z)) \qquad (2)$$

with notations:

$$g_x = \frac{\partial g^0(x, y, z)}{\partial x}, g_y = \frac{\partial g^0(x, y, z)}{\partial y}, g_z = \frac{\partial g^0(x, y, z)}{\partial z} \qquad (3)$$

where the terms g_x, g_y, and g_z are the numerical first derivatives of the function $g(x, y, z)$, which are defined as the components of the local surface normal vector $\boldsymbol{n}$. Their calculation depends on the analytical representation of the search surface elements, i.e., planar triangles or bilinear grids. The derivative terms are given as the x-y-z

components of the local normal vectors, which are computed at the exact matching location on the respective search surface elements:

$$[g_x \ g_y \ g_z]^T = \boldsymbol{n} = [n_x \ n_y \ n_z]^T. \qquad (4)$$

The terms $\mathrm{d}x$, $\mathrm{d}y$, and $\mathrm{d}z$ are the differentiation terms of the selected 3D transformation model. The geometric relationship is established with a seven-parameter 3D similarity transformation whose differentiation gives:

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix} = \begin{bmatrix} \mathrm{d}t_x \\ \mathrm{d}t_y \\ \mathrm{d}t_z \end{bmatrix} + \mathrm{d}m \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \mathrm{d}\omega \\ \mathrm{d}\varphi \\ \mathrm{d}\kappa \end{bmatrix} \qquad (5)$$

where a_{ij} are the coefficient terms. Their expansions are given in Akca (2007b). The vector $[t_x \ t_y \ t_z]^T$ is the translation vector, the scalar m is the uniform scale factor, and the angles ω, φ, and κ are the elements of the orthogonal rotation matrix $\mathbf{R}$. Depending on the characteristics of the template and search surfaces, any other higher order transformation model, e.g., a 3D affine or polynomial model, can be chosen.

By substituting Equations 3 and 5, Equation 2 gives the following form:

$$-e(x, y, z) = g_x\,\mathrm{d}t_x + g_y\,\mathrm{d}t_y + g_z\,\mathrm{d}t_z$$
$$+ (g_x a_{10} + g_y a_{20} + g_z a_{30})\mathrm{d}m$$
$$+ (g_x a_{11} + g_y a_{21} + g_z a_{31})\mathrm{d}\omega \qquad (6)$$
$$+ (g_x a_{12} + g_y a_{22} + g_z a_{32})\mathrm{d}\varphi$$
$$+ (g_x a_{13} + g_y a_{23} + g_z a_{33})\mathrm{d}\kappa$$
$$- (f(x, y, z) - g^0(x, y, z))$$

where $g^0(x, y, z)$ is the coarsely aligned search surface. The coarse alignment is performed using the initial approximations of the transformation parameters (t_x^0, t_y^0, t_z^0, m^0, ω^0, φ^0, and κ^0). The term $f(x, y, z) - g^0(x, y, z)$ denotes the Euclidean distance between the template and the corresponding search surface elements.

Equation 6 gives in matrix notation:

$$-\boldsymbol{e} = \mathbf{A}\boldsymbol{x} - \boldsymbol{l}, \ \mathbf{P} \qquad (7)$$

Where $\mathbf{A}$ is the design matrix, $\boldsymbol{x} = [\mathrm{d}t_x \ \mathrm{d}t_y \ \mathrm{d}t_z \ \mathrm{d}m \ \mathrm{d}\omega \ \mathrm{d}\varphi \ \mathrm{d}\kappa]^T$ is the parameter vector, $\mathbf{P} = \mathbf{P}_{ll}$ is the *priori* weight coefficient matrix, and $\boldsymbol{l} = f(x, y, z) - g^0(x, y, z)$ is the discrepancy vector.

With the statistical expectation operator E and the assumptions

$$\mathrm{E}(\boldsymbol{e}) = 0, \qquad \mathrm{E}(\boldsymbol{e}\boldsymbol{e}^T) = \sigma_0^2 \mathbf{P}^{-1}, \qquad (8)$$

the system (Equation 7) is a Gauss-Markov estimation model. The unknown parameters are introduced into the system as fictitious observations:

$$-\boldsymbol{e}_b = \mathbf{I}\boldsymbol{x} - \boldsymbol{l}_b, \qquad \mathbf{P}_b \qquad (9)$$

where $\mathbf{I}$ is the identity matrix, $\boldsymbol{l}_b$ is the (fictitious) observation vector, and $\mathbf{P}_b$ is the associated weight coefficient matrix. By selecting a very large weight value (($\mathbf{P}_b)_{ii} \to \infty$), an individual parameter can be assigned as constant. In commonly used topographic data sets, scale differences, even in some cases the rotational differences do not occur. This extension is especially useful to avoid such over-parameterization problems and for the flexible selection of the appropriate degree of freedom (DOF).

The joint system of Equations 7 and 9 is a Generalized Gauss-Markov model. The Least Squares solution gives the

estimated parameter vector and the variance factor as in the following equations:

$$\hat{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} + \mathbf{P}_b)^{-1}(\mathbf{A}^{\mathrm{T}}\mathbf{P}l + \mathbf{P}_b l_b) \qquad (10)$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} + \mathbf{v}_b^{\mathrm{T}}\mathbf{P}_b\mathbf{v}_b}{r} \qquad (11)$$

where ^ stands for the Least Squares estimator, r is the system redundancy, and the residual vectors $\mathbf{v}$ and $\mathbf{v}_b$ are the Least Squares estimation of the true error vectors $-\mathbf{e}$ and $-\mathbf{e}_b$, respectively.

The solution is iterative. At the end of each iteration, the search surface is transformed to a new state using the updated set of transformation parameters:

$$[t_x \quad t_y \quad t_z]^{\mathrm{T}} = [t_x^0 \quad t_y^0 \quad t_z^0]^{\mathrm{T}} + [\mathrm{d}\hat{t}_x \quad \mathrm{d}\hat{t}_y \quad \mathrm{d}\hat{t}_z]^{\mathrm{T}} \qquad (12)$$

$$m = m^0 + \hat{\mathrm{d}}m \qquad (13)$$

$$[\omega \quad \varphi \quad \kappa]^{\mathrm{T}} = [\omega^0 \quad \varphi^0 \quad \kappa^0]^{\mathrm{T}} + [d\hat{\omega} \quad d\hat{\varphi} \quad d\hat{\kappa}]^{\mathrm{T}}. \qquad (14)$$

At each iteration, the design matrix $\mathbf{A}$ and the discrepancies vector l are re-evaluated. The iteration stops if each element of the vector $\hat{x}$ falls below a certain limit:

$$|\mathrm{d}p_i| < c_i, \quad \mathrm{d}p_i \in \{\mathrm{d}t_x, \mathrm{d}t_y, \mathrm{d}t_z, \mathrm{d}m, \mathrm{d}\omega, \mathrm{d}\varphi, \mathrm{d}\kappa\} \qquad (15)$$

where the criteria c_i is selected as 1 ppm ($=10^{-6}$) for the scale factor, between 1/10 and 1/100 of the least significant digit for the translation elements, and 10^{-3} grad for the rotation elements.

The functional model is non-linear; thus, the initial approximations of the parameters are required. The initial approximations should be given or should be computed prior to the matching. In this paper, the experiments use geo-datasets, which are crudely aligned. Thus, the initial approximations of the rotation and translation parameters are assumed to be zero. This might not be the case for the terrestrial laser scanning point clouds. In such cases, the initial approximations can be provided by interactively selecting three common points on both surfaces before the matching.

Correspondence Search
For every template surface element, the correspondence operator seeks a minimum Euclidean distance location on the search surface. The template surface elements are represented by the data points. Accordingly, the procedure becomes a point-to-plane distance or point-to-bilinear surface distance computation. When a minimum Euclidean distance is found, in a subsequent step the matching point is tested to determine whether it is located inside the search surface element (point-in-polygon test). If not, this element is disregarded and the operator moves to the next search surface element with the minimum distance. Hypothetically, the correspondence criterion searches a minimum magnitude vector that is perpendicular to the search surface element and passes through the template point.

In the most straightforward case, the computational complexity is of order $O(N_t N_s)$, where N_t is the number of template elements and N_s is the number of search elements. This computational expense is reduced by constricting the search space within 3D boxes. The details are given in the *Computational Acceleration* sub-section.

Outlier Detection and Reliability Aspects
Detection of the 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 k-neighborhood points are calculated. In our implementation, k is selected as 8. If most of those k-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 that are greater than a given distance threshold.

In the course of iterations, a simple weighting scheme adapted from the robust estimation methods is used:

$$(\mathbf{P})_{ii} = \begin{cases} 1 & \text{if} & |(v)_i| < K\hat{\sigma}_0 \\ 0 & \text{else} \end{cases} \qquad (16)$$

The constant value K can be altered according to the task. If it is an ordinary surface co-registration task, it should be set to a high value (e.g., $K > 8$ or > 10) to reduce type I errors confidently. Because of the high redundancy of a typical data set, a certain number of occlusions and/or smaller outliers, i.e., type II errors, do not have significant effects on the estimated parameters. If it is a change detection or deformation study, the constant value K should be selected based on the *a priori* knowledge in order that the changed or deformed parts are excluded from the estimation.

Finally, the correspondences coinciding with mesh boundaries are excluded from the estimation. The mesh boundaries represent the model borders, and in addition the data holes inside the model. The data holes are possibly due to occlusions. Rejecting the surface correspondences on the mesh boundaries effectively eliminates the occlusions.

Precision
The quality of the estimated parameters can be checked through the *a posteriori* co-variance matrix.

The theoretical precisions of the transformation parameters are optimistic, mainly due to the stochastic properties of the search surfaces that have not been considered as such in the estimation model, as is typically done in Least Squares matching (Gruen, 1985). The omissions are expected to be minor and do not disturb the solution vector significantly. However, the *a posteriori* covariance matrix will be affected by the neglected uncertainty of the function values $g(x, y, z)$. This causes deterioration in the realistic precision estimates. More details on this issue can be found in Gruen (1985), Maas (2002), Gruen and Akca (2005), and Kraus *et al.* (2006).

Computational Acceleration
Searching for correspondence is guided by an efficient boxing algorithm (Chetverikov, 1991), which partitions the search space into voxels. For a given surface element, the correspondence is searched only in the box containing this element and in the adjacent boxes (Figure 1a). The original publication concerned 2D point sets. It is straightforwardly extended here to the 3D case.

Let points $\mathbf{a}_i = \{x_i, y_i, z_i\} \in S$, $i = 0, 1, \ldots, N_s - 1$, represent the object $S \in \mathfrak{R}^3$, and be kept in list $\mathbf{L}_1$ in spatially non-ordered form. The boxing data structure consists of a rearranged point list $\mathbf{L}_2$ and an index matrix $\mathbf{I} = I_{u, v, w}$ whose elements are associated with individual boxes: $u,v,w = 0,1,\ldots, M - 1$. The items of $\mathbf{L}_2$ are the coordinates of N_s points placed in the order of boxes. The index matrix $\mathbf{I}$ contains integers indicating the beginnings of the boxes in $\mathbf{L}_2$ (Figure 1b).
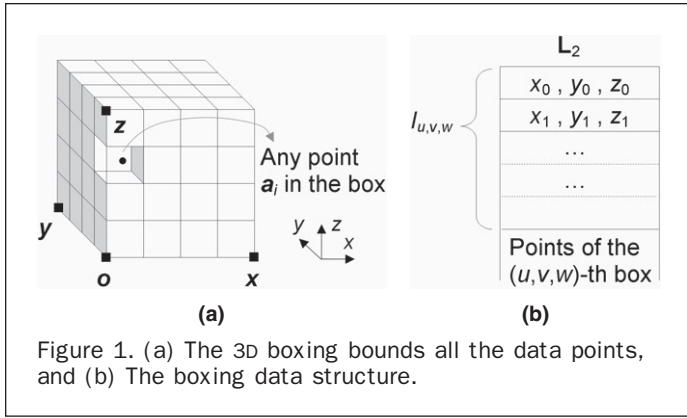
Figure 1. (a) The 3D boxing bounds all the data points, and (b) The boxing data structure.

Initialization: Defining the Box Size

Step 1. Recall min, max$\{x_i, y_i, z_i\}$ of data volume.

Step 2. Define the number of boxes along the $x-y-z$ axes. For the sake of simplicity, they are given the same (M) here.

Pass 1: Computing **I**

Step 1. Allocate an $M \times M \times M$ size accumulator array $\mathbf{B} = B_{u,v,w}$, which is to contain the number of points in each box.

Step 2. Scan $\mathbf{L}_1$ and fill $\mathbf{B}$. For any point $\boldsymbol{a}_i$, the box indices are as follows:

$$u_i = \left\lfloor \frac{x_i - x_{\min}}{D_X} \right\rfloor, v_i = \left\lfloor \frac{y_i - y_{\min}}{D_Y} \right\rfloor, w_i = \left\lfloor \frac{z_i - z_{\min}}{D_Z} \right\rfloor \quad (17)$$

where $\lfloor \rfloor$ stands for the truncation operator, and D_X, D_Y, and D_Z are dimensions of any box along the $x-y-z$ axes, respectively.

Step 3. Fill **I** using the following recursive formula: $\boldsymbol{I}_{0,0,0} = 0$. For all $(u,v,w) \neq (0,0,0)$

$$I_{u,v,w} = \begin{cases} I_{u,v,w-1} + B_{u,v,w-1} & \text{if} & w > 0 \\ I_{u,v-1,M-1} + B_{u,v-1,M-1} & \text{else if} & v > 0 \\ I_{u-1,M-1,M-1} + B_{u-1,M-1,M-1} & \text{else} \end{cases} \quad (18)$$

Pass 2: Filling $\mathbf{L}_2$

Step 1. For all u, v, and w, set $B_{u,v,w} = 0$.

Step 2. Scan $\mathbf{L}_1$ again. Use Equation 17, **I** and **B** to fill $\mathbf{L}_2$. In $\mathbf{L}_2$, the first point of the $(u,v,w)^{\text{th}}$ box is indexed by **I** while the address of the subsequent points is controlled using **B** whose value is incremented each time a new point enters the box. Finally, release the memory area of **B**.

The memory requirement is of order $O(N_s)$ for $\mathbf{L}_2$ and $O(M^3)$ for **I**. For the sake of clarity of the explanation, $\mathbf{L}_2$ is given as a point list containing the x-y-z coordinate values. If one wants to keep the $\mathbf{L}_1$ in the memory, then $\mathbf{L}_2$ should only contain the access indices to $\mathbf{L}_1$ or pointers, which directly point to the memory locations of the point coordinates.

Access Procedure

Step 1. Using Equation 17, compute the indices u_i, v_i, and w_i of the box that contains point $\boldsymbol{a}_i$.

Step 2. Use the boxing structure to retrieve the points bounded by the $(u,v,w)^{\text{th}}$ box. In $\mathbf{L}_2$, **I** indexes the first point, while the number of points in the box is given by the following formula:

$$\begin{cases} I_{u,v,w+1} - I_{u,v,w} & \text{if} & w < M-1 \\ I_{u,v+1,0} - I_{u,v,M-1} & \text{else if} & v < M-1 \\ I_{u+1,0,0} - I_{u,M-1,M-1} & \text{else if} & u < M-1 \\ N_s - I_{M-1,M-1,M-1} & \text{else} \end{cases} \quad (19)$$

The access procedure requires $O(q)$ operations, where q is the average number of points in the box. One of the main advantages of the boxing structure is a faster and easier access mechanism than the tree search-based methods provide. Construction time of the boxing method $O(N_s)$ is less than what the tree search methods need, i.e., order of $O(N_s \log N_s)$ for a k-D tree (Greenspan and Yurick, 2003; Arya and Mount, 2005). On the other hand, the tree search methods obviously need less storage space, which is only order of $O(N_s)$.

The boxing structure, and in general all search structures, are designed for searching the nearest neighborhood in the static point clouds. In the LS3D surface matching case, the search surface, for which the boxing structure is established, is transformed to a new state by the current set of transformation parameters. Nevertheless, there is no need either to re-establish the boxing structure or to update the **I** and $\mathbf{L}_2$ in each iteration. Only the positions of those four points (Figure 1a) are updated in the course of iterations: $\boldsymbol{o} = \{x_{min}, y_{min}, z_{min}\}$, $\boldsymbol{x} = \{x_{max}, y_{min}, z_{min}\}$, $\boldsymbol{y} = \{x_{min}, y_{max}, z_{min}\}$, $\boldsymbol{z} = \{x_{min}, y_{min}, z_{max}\}$. They uniquely define the boxing structure under the similarity transformation. The access procedure is the same, except the following formula is used for the calculation of indices:

$$u_i = \left\lfloor \frac{\boldsymbol{oa}_i \cdot \boldsymbol{ox}}{\|\boldsymbol{ox}\| D_X} \right\rfloor, v_i = \left\lfloor \frac{\boldsymbol{oa}_i \cdot \boldsymbol{oy}}{\|\boldsymbol{oy}\| D_Y} \right\rfloor, w_i = \left\lfloor \frac{\boldsymbol{oa}_i \cdot \boldsymbol{oz}}{\|\boldsymbol{oz}\| D_Z} \right\rfloor \quad (20)$$

Where '·' stands for a vector dot product. If the transformation is a similarity rather than a rigid body, the D_X, D_Y, and D_Z values must also be updated in the iterations:

$$D_X = \frac{\|\boldsymbol{ox}\|}{M}, D_Y = \frac{\|\boldsymbol{oy}\|}{M}, D_Z = \frac{\|\boldsymbol{oz}\|}{M}. \quad (21)$$

In the current implementation, the correspondence is searched in the boxing structure during the first few iterations, and at the same time, 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.

Algorithmic Extensions
Multiple Surface Matching

When more than two point clouds with multiple overlaps exist, a two step solution is adopted. First, the pairwise LS3D matchings are run on every overlapping pair and a subset of point correspondences is saved to separate files. In the global registration step, all these files are passed to a block adjustment by the independent models procedure (Ackermann *et al.*, 1973), which is a well known orientation procedure in photogrammetry. More details can be found in Akca (2007b).

Simultaneous Matching of Surface Geometry and Intensity

When the object surface lacks sufficient geometric information, i.e., homogeneity or isotropicity of curvatures, the basic algorithm will either fail or will find a side minimum. In this extension, available attribute information, e.g., intensity, color, temperature, etc., is used to form quasi-surfaces in addition to the actual ones. The matching is performed by simultaneous use of surface geometry and attribute information under a combined estimation model (Akca, 2007a).

Further Conceptual Extensions

The further conceptual extensions are given as: the Least Squares matching of 3D curves, matching of 3D curves or 3D sparse points (e.g., ground control points) with a 3D surface, and a general framework, which can perform the multiple surface matching, the combined surface geometry and intensity matching, and georeferencing tasks simultaneously (Akca, 2007b).

Experimental Results

The algorithm was implemented as a stand-alone MS Windows™ application with a graphical user interface. The software package was developed with the C/C++ programming language. The presented examples use solely the basic model, not the algorithmic extensions.

Tucume

The first experiment is the matching of two photogrammetrically derived digital terrain models (DTM) of an area in Tucume (Peru). The horizontal resolution of the DTMs was 5 m. The DTMs were manually measured as two independent models from 1:10 000 scale B/W aerial images in one strip with an overlap of 60 percent in the flight direction. More details are given in Sauerbier *et al.* (2004).

Although it is only a 2.5D model, it is a good example of the weak data configuration case since the overlapping area is relatively narrow (along the *Y*-axis) with little information regarding the surface geometry (Figure 2a). The LS3D algorithm was run in 6-DOF mode with three translation and three rotation parameters. This version showed a large correlation coefficient 0.99 between the t_y and κ angle, which is an over-parameterization case. Thus, κ angle was excluded from the system, and the second version of the computation was run in 5-DOF mode. The results are successful (Table 1). The computation takes 1.9 and 2.5 seconds for the plane surface (P) and bi-linear surface (B) representation versions, respectively.

The ratio between the standard deviations of ω and φ angles is by factor 14. This difference in angular uncertainty is due to difference in overlapping areas along the *X* and *Y* axes. The residuals between the fixed and transformed surfaces show a random distribution pattern, except for some occasional measurement and modeling errors (Figure 2b).

-7.0 m	-4.9	-2.8	-0.7	+0.7	+2.8	+4.9	+7.0 m

(d)
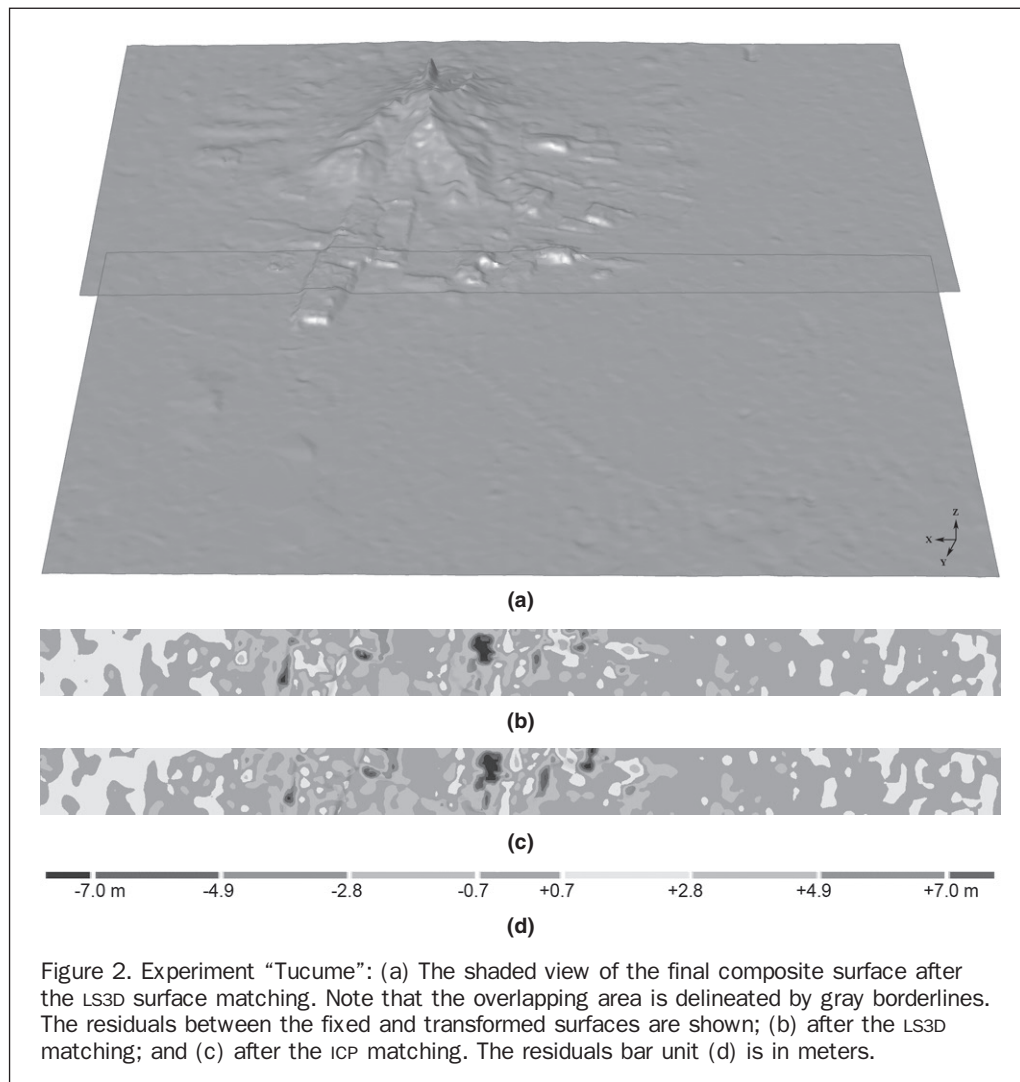
Figure 2. Experiment "Tucume": (a) The shaded view of the final composite surface after the LS3D surface matching. Note that the overlapping area is delineated by gray borderlines. The residuals between the fixed and transformed surfaces are shown; (b) after the LS3D matching; and (c) after the ICP matching. The residuals bar unit (d) is in meters.

was investigated in a cooperative venture between the Cartographic Institute of Catalonia (ICC, in Spain) and the Institute of Geodesy and Photogrammetry (IGP) of ETH Zurich. Further details can be found in Zhang *et al.* (2006).
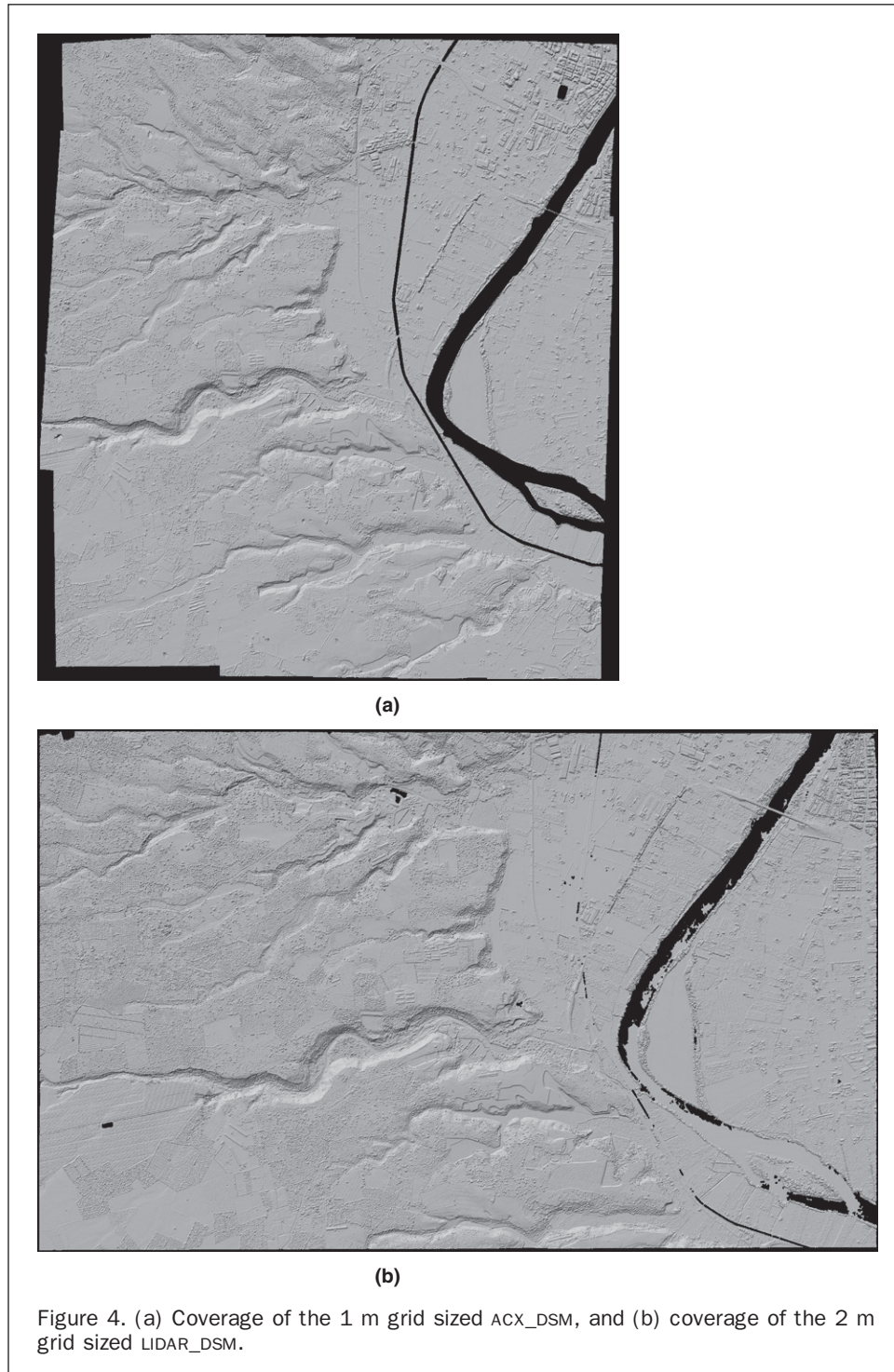
Image Data

The image data consisted of 28 DMC images with a ground sampling distance (GSD) of 22 cm arranged in four parallel flight strips in the E/W direction, each of seven images. The forward and side overlap of the DMC images were 60 percent and 75 percent, respectively. They covered an area of about 5 × 5 km, located close to the Ebro delta south-west of Barcelona with variable land-cover. The DMC images were acquired together with the lidar data from 2,500 m altitude ASL (1:21 000 image scale).

The aerial triangulation was carried out with the ACX/GeoTex software package developed at the ICC. The mean theoretical precisions of the object point coordinates were $X = \pm 0.04$, $Y = \pm 0.05$, $Z = \pm 0.10$ m.

A DSM of the test site was generated with a 1 m grid size, called the ACX_DSM (Figure 4a). The automated DSM generation was performed using the SAT-PP software

(a)

(b)

Figure 4. (a) Coverage of the 1 m grid sized ACX_DSM, and (b) coverage of the 2 m grid sized LIDAR_DSM.

developed at the IGP. The SAT-PP uses a coarse-to-fine hierarchical matching method with an effective combination of several image matching algorithms and automatic quality control. It generates very dense and highly accurate point clouds by matching of multiple image primitives (feature points, grid points, edges) and by making use of the sensor model, network structure, image content, and geometrical constraints such as the epipolar geometry constraint (Zhang and Gruen, 2004).

Lidar Data as Reference
The lidar system was an Optech ALTM 3030. There were some gaps in the data due to water surfaces and low reflectance objects (there were rice fields, possibly flooded, in the region). The lidar data consisted of four strips and two cross strips at 45 and 90 degrees, which were used for in-flight calibration. The accuracy was about 50 cm in planimetry and better than 15 cm in height. The average point density was 1.2 points/m^2. The first pulse of the unfiltered laser data was interpolated to a 2 m regular grid (called LIDAR_DSM) to be used as the reference data.

The ACX_DSM was compared with respect to the LIDAR_DSM (Figure 4b) by use of the LS3D surface matching method.

DSM Comparison and Analysis
For quality evaluation of DSMs, often a reference DSM is interpolated in the DSM to be checked. Evaluation is done based on the height differences. This approach is sub-optimal, since (a) at surface discontinuities surface modeling errors may lead to large height differences although the measurements are correct (Poli *et al.*, 2004), and (b) if the reference frames of the two DSMs differ (e.g., shifts and tilts), then again large differences occur, especially at discontinuities although the heights may be correct. The LS3D surface matching method was used to avoid both these shortcomings.

Table 3 shows the numerical results of the LS3D matching where five transformation parameters (three translations, ω and φ angles) were used in 5-DOF mode. The computation takes 392.4 seconds. In an initial step without applying any transformation the *a priori* sigma was computed as 0.95 m.

The sigma *a posteriori*, the standard deviation of the Euclidean distances (shown in Plate 1) between the search

TABLE 3. THE NUMERICAL RESULTS OF THE DSM COMPARISON

Temp. Srch.[1]	No. of match. points	No. of iter.	$\hat{\sigma}_0$ X/Y/Z[2] (m)	t_x $\hat{\sigma}_{tx}$ (m)	t_y $\hat{\sigma}_{ty}$ (m)	t_z $\hat{\sigma}_{tz}$ (m)	ω $\hat{\sigma}_{\omega}$ (grad)	φ $\hat{\sigma}_{\varphi}$ (grad)	κ $\hat{\sigma}_{\kappa}$ (grad)
LIDAR_ ACX_	4249512	3	0.79 0.32/0.31/0.65	0.04 0.01	−0.61 0.01	−0.06 0.01	0.0010 0.0001	0.0050 0.0001	N.A. N.A.

[1]) The surface representation is in planar (P) triangle mesh form.
[2]) Decomposition of the *a posteriori* sigma $\hat{\sigma}_0$ into X, Y and Z components.
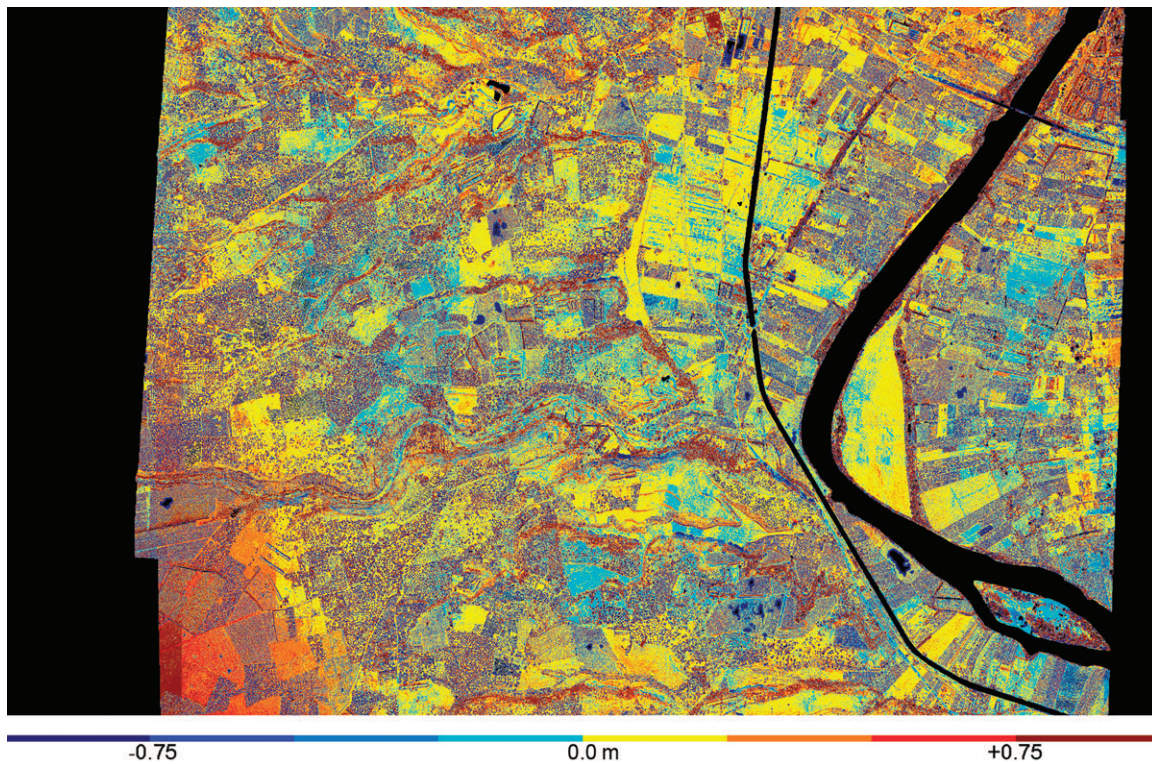Temp.: Template surface; Srch.: Search surface.

Plate 1. Colored residuals (3D spatial differences between the template and search DSMS) after co-registration of the ACX_DSM with respect to the reference LIDAR_DSM.

and the template DSMs after performing the transformation, was computed as 0.79 m. Its decomposition into X, Y, and Z components is 0.32/0.31/0.65 m. The X and Y values are similar and about 0.3 m, while in height it is about 0.65 m. This is actually the difference between the two DSMs after removing the reference frame differences.

The transformation parameters represent the reference frame differences between the two DSMs. In planimetry, this bias is due to the different orientations of the images and the lidar, and is significant only in the Y (N/S) direction. In height, the bias is possible due to the same reason and additionally to partial penetration of vegetation by the lidar data (note that the ACX_DSM is higher than the lidar height values, as the negative sign of t_z shows).

The differences are higher at surface discontinuities, possibly also due to the modeling errors, e.g., in the upper right urban area (Plate 1). The image matching DSMs show a jump in the bottom left portion. The possible explanation is the change of matching from the use of four images to the use of three images at this position.

The achieved standard deviation of height differences was about 0.65 m. With a better orientation and DSM generation and without the differences due to vegetation penetration, we expect this difference to drop significantly below 0.5 m. It should be noted that the lidar data is not good enough to be the reference.

Assessing Changes of Forest Areas and Shrub Encroachment
This is an assessment study of change (1997 to 2002) of forest and other wooded areas in a mire biotope in the Pre-alpine zone of Central Switzerland using airborne remote sensing data. The study is a cooperative project between the IGP and the Department of Landscape Inventories of the Swiss Federal Research Institute WSL. More details are given in Baltsavias *et al.* (2007).

Study Area and the Data Sets
The study area is located on a small plateau to the east of the Lake of Zug, a sensitive environmental territory in the Pre-alpine zone of Central Switzerland, covering an area of approximately 2.61 km^2. Two sets of aerial images (August 1997 and July 2002) and one airborne laser scanning point cloud (year 2001) data were used.

Both sets of aerial images were taken with a Leica RC30 analogue camera with color infrared film (IR-R-G). The 1997 data set contains four images (1:10 000 image scale) in one strip with a 75 percent forward overlap. The 2002 data set is 12 images (1:5 500 image scale) in two strips with a 75 percent forward and a 30 percent lateral overlap.

All images were digitized with a Vexcel UltraScan® scanner with a 15 micron pixel size, which results in a GSD of 15 cm and 8.25 cm for the 1997 and 2002 images, respectively. The 1997 film images had severe scratches on the emulsion side, causing artifacts in the digitized images and DSM errors in the automated DSM generation (Figure 5a).

The image orientation was established with 15 ground control points measured by a differential GPS survey and using the bundle adjustment of SocetSet v.5.2. The sigma naught of orientation was 0.20 pixels and 0.23 pixels for the 1997 and 2002 blocks, respectively. Two DSMs were generated automatically from the above images, called 1997_DSM and 2002_DSM, using the SAT-PP software (Figure 5). The DSMs have a grid spacing of 0.5 m. The national lidar data of the Swiss Federal Office of Topography (Swisstopo) was acquired in 2001 when leaves were off the trees. The average density was 1 to 2 points/m^2, a the height accuracy (1 sigma) 0.5 m for open areas, and 1.5 m for terrain with vegetation. The first pulse point cloud was interpolated to a regular grid with 2.5 m grid spacing, called 2001_DSM.
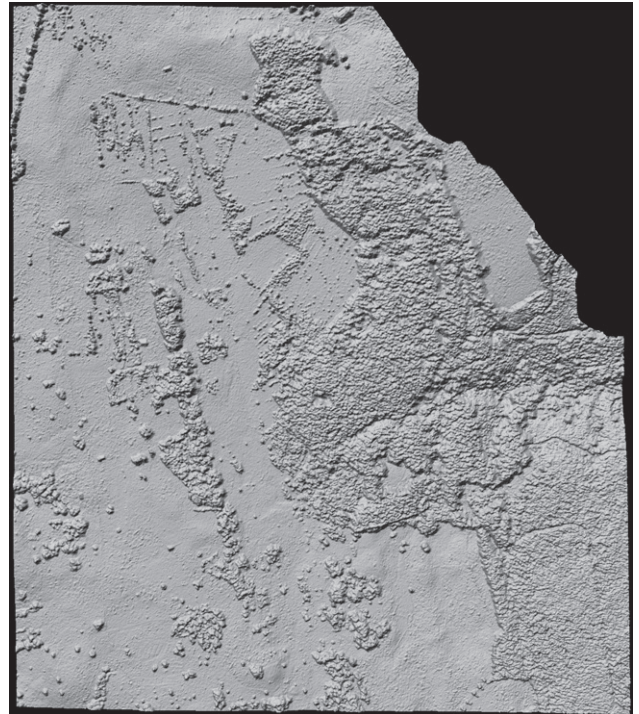
(a)
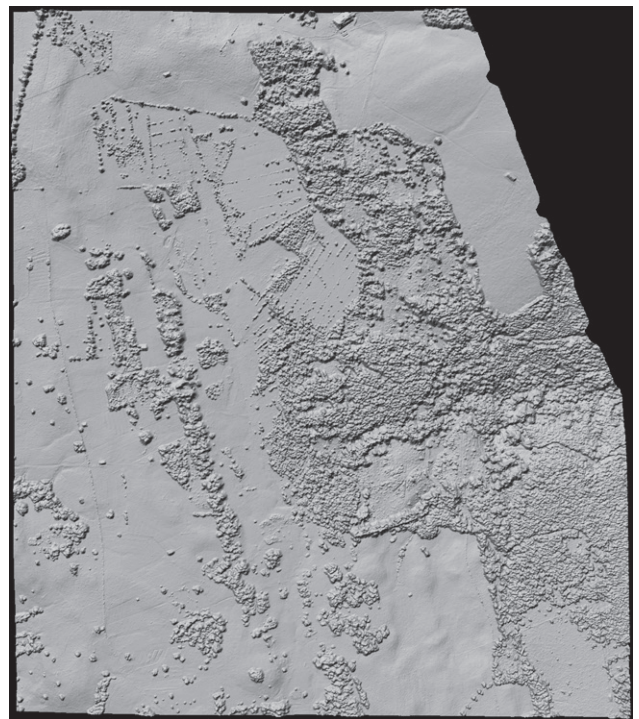
(b)

Figure 5. The gray-shaded visualization of (a) 1997_DSM, and (b) the 2002_DSM. Significant deforestation in 2002 is apparent, especially in the bottom right parion of the model.

Co-registration and Change Detection
The 1997_DSM, 2002_DSM, and 2001_DSM were co-registered using the LS3D surface matching method. The matcher was run in 5-DOF mode to accommodate the translation and tilt.

Temp. Srch.[1]	No. of match. points	No. of iter.	$\hat{\sigma}_0$ X/Y/Z[2] (m)	t_x $\hat{\sigma}_{tx}$ (m)	t_y $\hat{\sigma}_{ty}$ (m)	t_z $\hat{\sigma}_{tz}$ (m)	ω $\hat{\sigma}_\omega$ (grad)	φ $\hat{\sigma}_\varphi$ (grad)	κ $\hat{\sigma}_\kappa$ (grad)
2002_ 1997_	8694077	8	4.56 2.09/1.97/3.54	−0.25 0.01	−0.16 0.01	0.13 0.01	−0.0195 0.0001	−0.0270 0.0001	N.A. N.A.
2002_ 2001_	8491073	9	1.42 0.86/0.86/0.73	0.92 0.01	0.85 0.01	0.60 0.01	−0.0186 0.0001	−0.0276 0.0001	N.A. N.A.

[1] The surface representation is in planar (P) triangle mesh form.

For the estimation of these parameters, the control surfaces, i.e., DSM parts that did not change in the two datasets (bare ground), were used, and also the large differences due to image matching errors were removed with the robust re-weighting factor. After the co-registration, the Euclidian distances between the two datasets were computed, as well as the X, Y, Z components (Table 4), the latter being more important for these investigations.

Using the outputs of the co-registration, different products could be generated and conclusions drawn. The difference between the 2002_DSM and 1997_DSM gives the changes between the two epochs, especially regarding vegetation. After co-registration, the Z-component of the Euclidian distances (sigma *a posteriori*) was 3.54 m, showing a clear reduction of vegetation from 1997 to 2002. In Plate 2a, the red areas show the deforestation, and the blue areas show the growth of vegetation.

The lidar data generated 2001_DSM was also compared to the 2002_DSM in spite of the small time difference. This could give a comparison between the two DSMs and also an indication of the extent that lidar penetrates the tree canopy. After co-registration, the Z-component of the Euclidian distances (sigma a posteriori) was 0.73 m. In Plate 2b, the blue areas show the growth of the vegetation including the partial penetration of the lidar. The small red spots, mostly inside the forest, show the loss of individual trees during the one year time difference.

At the top and bottom of Plate 2b, the effect of the stripes in the 2002_DSM, possibly due to film scanner (geometric) mis-calibration, is visible. The orange areas at the top left are due to differences in image orientation between the two flight strips and within each strip causing discontinuities in the 2002_DSM. These areas are also visible in Plate 2a but have less sharp boundaries due to the noise of the 1997_DSM.

Conclusions

The basic estimation model is a generalization of the Least Squares matching concept. It is an algorithmic extension of the early 2D image matching version.

The current implementation uses a 3D similarity transformation model for the geometric relationship. The unknown transformation parameters are treated as observables with proper weights, so that sub-versions of the 7-parameter model can be run, i.e., rigid body (6-DOF), tilt and translation (5-DOF), translation (3-DOF), horizontal shift (2-DOF), and depth (1-DOF).

The LS3D method fully considers the 3D information by evaluating the Euclidean distances, while the 2.5D surface matching algorithms (Ebner and Strunz, 1988; Rosenholm and Torlegard, 1988; Karras and Petsa, 1993; Pilgrim, 1996; Mitchell and Chadwick, 1999; Postolov et al., 1999; Xu and Li, 2000; Maas, 2002; Kraus *et al.*, 2006) evaluate the height differences, which is sub-optimal even for the terrain applications (see the *DSM Comparison and Analysis* sub-section for the critical remarks). They cannot consider the surface modeling errors. The estimation model is the Generalized Gauss-Markov model. It is a strict and rigorous formulation, which describes the physical nature of the problem mathematically. It provides a flexible basis, which makes further algorithmic extensions possible.

The quality of any individual parameter can be checked using the *a posteriori* variance-covariance matrix, but must account for correlations between the estimated parameters. This feature can be highly important when the data set does not contain sufficient surface information along one or more coordinate directions in order to support the computation of all transformation parameters. The parameters with low precision values help to diagnose and to explain the configuration and content of the data (see the Tucume example).

The capability to match surfaces of different quality and resolution is another positive aspect of the proposed method. The SRTM TerrainScape work, in cooperation with Swissphoto AG, proves this capability. Swissphoto AG has matched many DEMs around the world with the SRTM C-SAR DEMs. The local DEMs are of any accuracy, point spacing, and production technique. According to their report, a complete failure case has not happened, except for some software debugging cases.

The co-registration of the point clouds of different quality and resolution is only one capability. The proposed 3D Least Squares surface matcher can also perform 3D comparison. This is especially useful for quality assessment and change detection tasks as discussed in the Experimental Results Section.

References

Ackermann, F., H. Ebner, and H. Klein, 1973. Block triangulation with independent models, *Photogrammetric Engineering & Remote Sensing*, 39(9):967–981.

Ackermann, F., 1984. Digital image correlation: Performance and potential application in photogrammetry, *The Photogrammetric Record*, 11(64):429–439.

Akca, D., 2007a. Matching of 3D surfaces and their intensities, *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(2):112–121.

Akca, D., 2007b. *Least Squares 3D surface Matching*, Ph.D. thesis, Institute of Geodesy and Photogrammetry, ETH Zurich,
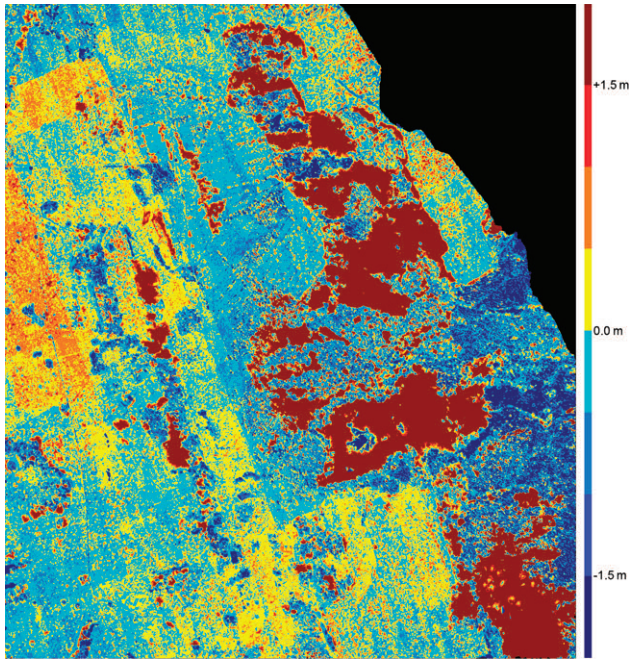
(a)
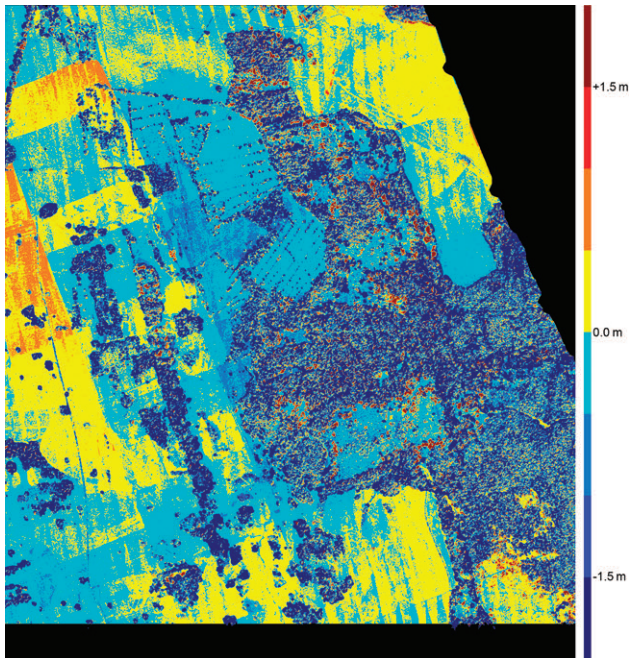
(b)

Plate 2. (a) The Z component of the Euclidian distances between the 1997_DSM (search surface) and 2002_DSM (template surface) clearly shows areas of deforestation and shrub encroachment, and (b) The Euclidian distances between the 2001_DSM (search) and the 2002_DSM (template) show that the lidar measures tree canopies lower than the image matching.

Switzerland, Mitteilungen, Nr.92, 78 p, URL: *http://www.photogrammetry.ethz.ch/general/persons/devrim_publ.html* (last date accessed: 15 December 2009).

Arya, S., and Mount, D.M., 2005. *Computational Geometry: Proximity and Location, The Handbook of Data Structures and Applications* (D. Mehta and S. Sahni, editors), Chapman & Hall/CRC, Boca Raton, Florida, pp. 63.1–63.22.

Baltsavias, E., H. Eisenbeiss, D. Akca, L.T. Waser, M. Küchler, C. Ginzler, and P. Thee, 2007. Modeling fractional shrub/tree cover and multi-temporal changes using high-resolution digital surface models and CIR-aerial images, *Proceedings of the Dreiländertagung SGPBF, DGPF und OVG*, 19–21 June, Muttenz, Switzerland, DGPF Tagungsband 16/2007, pp. 287–297.

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

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

Chetverikov, D., 1991. Fast neighborhood search in planar point sets, *Pattern Recognition Letters*, 12(7):409–412.

Ebner, H., and G. Strunz, 1988. Combined point determination using digital terrain models as control information, *International Archives of Photogrammetry and Remote Sensing*, 27(B11/3): 578–587.

Greaves, T., 2005. Laser scanning shaves weeks from failure analysis of parking garage collapse, Spar View™, 3(14), URL: *http://www.sparllc.com/archiveviewer.php?vol=03&num=14&file=vol03no14–01* (last accessed date: 15 December 2009).

Greenspan, M., and Yurick, M., 2003. Approximate K-D tree search for efficient ICP, *Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling (3DIM 2003)*, Banff, Canada, 06–10 October, pp. 442–448.

Gruen, A., 1985. Adaptive least squares correlation: A powerful image matching technique, *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3): 175–187.

Gruen, A., 1996. Least squares matching: A fundamental measurement algorithm, *Close Range Photogrammetry and Machine Vision* (K. Atkinson, editor), Whittles, pp. 217–255.

Gruen, A., and D. Akca, 2005. Least squares 3D surface and curve matching, *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3):15–174.

Karras, G.E., and E. Petsa, 1993. DEM matching and detection of deformation in close-range photogrammetry without control, *Photogrammetric Engineering & Remote Sensing*, 59(9):1419–1424.

Kraus, K., C. Ressl, and A. Roncat, 2006. Least squares matching for airborne laser scanner data, *Proceedings of the 5th International Symposium Turkish-German Joint Geodetic Days*, 29–31 March, Berlin, Germany, unpaginated CD-ROM.

Lambers, K., H. Eisenbeiss, M. Sauerbier, D. Kupferschmidt, Th. Gaisecker, S. Sotoodeh, Th. Hanusch, 2007. Combining photogrammetry and laser scanning for the recording and modelling of the late intermediate period site of Pinchango Alto, Palpa, Peru, *Journal of Archaeological Science*, 34(10):1702–1712.

Maas, H.G., and A. Gruen, 1995. Digital photogrammetric techniques for high resolution three dimensional flow velocity measurements, *Optical Engineering*, 34(7):1970–1976.

Maas, H.-G., 2002. Methods for measuring height and planimetry discrepancies in airborne laserscanner data, *Photogrammetric Engineering & Remote Sensing*, 68(9):933–940.

Mitchell, H.L., and R.G. Chadwick, 1999. Digital photogrammetric concepts applied to surface deformation studies, *Geomatica*, 53(4):405–414.

Norris, G., 2005. Jeppesen puts synthetic vision system on display, *Flight International*, 25–31 October, URL: *http://www.photogrammetry.ethz.ch/research/srtm/images/a_JEPPESEN_200510_flight_international_svs.pdf* (last date accessed: 15 December 2009).

Pertl, A., 1984. Digital image correlation with the analytical plotter Planicomp C-100, *International Archives of Photogrammetry and Remote Sensing*, 25(3B):874–882.

Pilgrim, L., 1996. Robust estimation applied to surface matching, *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(5):243–257.

Poli, D., Zhang, L., and A. Gruen, 2004. SPOT-5/HRS stereo images orientation and automated DSM generation, *International*

Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B1):421–432.

Postolov, Y., A. Krupnik, and K. McIntosh, 1999. Registration of airborne laser data to surfaces generated by photogrammetric means, *International Archives of Photogrammetry and Remote Sensing*, 32(3/W14):95–99.

Rosenholm, D., and K. Torlegard, 1988. Three-dimensional absolute orientation of stereo models using digital elevation models, *Photogrammetric Engineering & Remote Sensing*, 54(10):1385–1389.

Sauerbier, M., M. Kunz, M. Fluehler, and F. Remondino, 2004. Photogrammetric reconstruction of adobe architecture at Tucume, Peru, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5/W1), unpaginated CD-ROM.

Sternberg, H., Th. Kersten, I. Jahn, and R. Kinzel, 2004. Terrestrial 3D laser scanning – Data acquisition and object modeling for industrial as-built documentation and architectural applications, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B7):942–947.

Xu, Z., and Z. Li, 2000. Least median of squares matching for automated detection of surface deformations, *International Archives of Photogrammetry and Remote Sensing*, 33(B3):1000–1007.

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

Zhang, L., and A. Gruen, 2004. Automatic DSM generation from linear array imagery data, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B3):128–133.

Zhang, L., S. Kocaman, D. Akca, W. Kornus, and E. Baltsavias, 2006. Test and performance evaluation of DMC images and new methods for their processing, *Proceedings of the SPRS Commission I Symposium*, 03–06 July, Paris, France, unpaginated CD-ROM.