AUTOMATIC CO-REGISTRATION OF TERRESTRIAL LASER SCANNER AND DIGITAL CAMERA FOR THE GENERATION OF HYBRIDS MODELS

Diego González Aguilera, Pablo Rodríguez Gonzálvez and Javier Gómez Lahoz

Land and Cartography Engineering Department, Univ. of Salamanca, Spain

daguilera@usal.es, pablorgsf@hotmail.com, fotod@usal.es

Commission V, WG V/3

KEY WORDS: Cultural Heritage, Laserscanning, Photogrammetry, Sensor Registration, Software Development.

ABSTRACT:

3D modeling and visualization of real world scenes is an important topic of research with applications in many areas such as virtual museums, game and entertainment, architecture description and restoration, virtual reality, archaeology, many industrial applications and last but not least important tourist applications. 3D modeling and visualization are the creation of a computer representation of real world environments that merges together data coming from one or more sensors. The representation of the geometric and texture information of a real scene is a very challenging task due to the acquisition of large-scale data, complexity of the geometry and difficulties to cope with reflectance properties of the objects and variations in the lighting in the scene. Two approaches, depending on the type of sensor (terrestrial laser scanner or digital cameras), are typically used to face the 3D reconstruction problem. Laser scanners provide 3D metric information in real time through an array of coordinates: range images. Digital cameras are used to acquire high-resolution images of the scenes. These images are 2D arrays of reflected light from objects but do not contain any explicit metric information. Further processing is necessary, to calibrate cameras and compute 3D models. This paper aims to demonstrate how active and passive sensors can be registered and combined through a hybrid approach to compute 3D models of complex scenes with photo-realistic quality. Particularly, the proposed approach tries to deal with two different images: a high-resolution image acquired with a digital camera and a range image obtained from a laser scanner model using collinearity condition. Our goal is to devise and implement a robust, automatic, and accurate hybrid-technique for registration of both sensors for efficient modeling (geometry) and rendering (radiometry) of complex environments. To this end, we have developed a novel application for laser scanning which allow us to test the approach developed over experimental results.

1. INTRODUCTION: A REVIEW

3D reconstruction of complex scenes is a very challenging task due to the variety of possible scenes to scan and difficulties to work with real data. Passive and active techniques used in 3D reconstruction have their limitations and, separately, none of these techniques can solve all the problems inherent to the modeling of real environments. To reinforce this need, next table (Table 1) illustrates a comparison based on the most important features with relation to laser scanner and digital camera.

Laser Scanner	Digital Camera
↓ Inaccurate lines and joints	↑ Accurate lines and joints
↓ Poor colour information	↑Good colour information
↑ Prompt and accurate metric information	↓ Hard-working and slow metric information
↑ Excellent technique for the description of complex and irregular surfaces	↓ Time-consuming technique for the description of complex and irregular surfaces
↓ High-cost technique	↑ Low-cost technique
↓ The 3D model is an entity disorganized and without topology	↑The 3D model is an entity organized and with topology
↑ Light is not required to work	↓ Light is required to work

Table 1: *Comparison of features: laser vs. camera.*

Up to now, several approaches have been developed trying to register both sensors. This problem of image-to-model registration is closely related to the problem of camera calibration, which finds a mapping between the 3D world (object space) and a 2D image. This mapping is characterized by a rigid transformation and a camera model, also referred to as the camera's extrinsic and intrinsic parameters. This rigid body transformation takes 3D points from object space to 2D points in the camera's reference frame, and the camera model describes how these are projected onto the image plane. The camera calibration problem is solved by matching features in the 3D model with features in the image. These features are usually points, lines or special designed objects that are placed in the scene. The matching process can be automatic or user driven, and the number of feature pairs required depend on whether we are solving for the intrinsic, extrinsic or both parameters sets. In the context of image registration for 3D modeling using dense laser scanner data, several approaches have been developed: (Rocchini et. al., 1999) develop a new approach for mapping and blending textures on 3D geometries. The system starts from a 3D mesh which represents a real object and texture detail acquired via a common photographic process. Both datasets are integrated based on initial rough registration. However, this approach requires manual interaction and is applied to small objects; (Lensch et. al., 2001) develop an image registration approach based on silhouette matching, where the contour of a rendered version of the object is matched against the silhouette of the object in the image. No user intervention is required, but their method is limited to cases where a single image completely captures the object; in other range of methods applied to large distances, dealing with outdoor scenes and based on locating invariant image features, (McAllister et. al., 1999) suggest correlating edges common to the color image and the range map's intensity component. However, care must be taken to place the camera's nodal point at the same physical location as the laser's center of rotation, and to rotate both devices about this point. This homographic relationship simplifies the registration to the camera's three rotations relative to the laser; (Elstrom, 1998) presents a novel stereo-based method for registering color and range images acquired from externally uncalibrated sensors. The multi-sensor alignment problem is solved by processing invariant features such as corner, edges or contours which are extracted from the raw data. The benefit of a feature-based approach is that it abstracts the data and thus simplifies the search for the registration parameters. Often, however, feature extraction leads to the loss of some information; (Stamos and Allen, 2001) present a semi-automatic method for image to model registration of urban scenes, where 3D lines are extracted from the point clouds of buildings and matched against edges extracted from the images. The method involves the utilization of parallelism and orthogonality constraints that naturally exist in urban environments. Therefore, their algorithm only operates in scenes which contain linear features with strong geometric constraints. Also a number of thresholds have to be manually be set by the user in order to customize the segmentation; (Ikeuchi et. al., 2003) in their Great Buddha work, use reflectance edges obtained form the 3D points and match them against edges in the image to obtain the camera position. They align edges extracted from reflectance images with those in color images so that the 3D position error of those edges is minimized by iterative calculation. Nevertheless, this approach has been focused on small and simple objects; (Allen et. al., 2003) present a novel method for 2D to 3D texture mapping using shadows as cues. They pose registration of 2D images with the 3D model as an optimization problem that uses knowledge of the Sun’s position to estimate shadows in a scene, and use the shadows produced as a cue to refine the registration parameters. However, they still have some limitations related to view planning and real-time model creation and visualization. More recently, (Aguilera and Lahoz, 2006) exploit the power of a single image-based modeling method to obtain an automatic co-registration of laser scanner and uncalibrated digital camera. Particularly, the problem of image registration is solved automatically through 2D and 3D points correspondences which are matched based on a search of spatial invariants: two distances and one angle. However, several input considerations such as especial targets, vanishing points and geometric constraints have to be taken into account; (Al-Manasir and Fraser, 2006) develop a strategy using a coded target placed on the object, which are registered by a calibrated digital camera, rigidly attached to the laser scanner. An automatic process is applied to solve the spatial position and orientation of the camera within the laser scanner coordinate system. The identified coded targets are used to apply a 3D similarity transformation. However, this approach needs a camera attached camera to laser scanner and placed some code target on the object; (Alshawabkeh et. al., 2006) propose a robust algorithm line detection within the high-resolution image, based on the mathematical properties of the mean which is invariant to arbitrary rotations and translation of surface. This segmentation is applied into 3D model, and can look for different types of edges. Finally, it uses the edge resulting to apply a matching between 2D and 3D datasheets.

The method that we have developed (Figure 1) exhibits significant improvements in flexibility, accuracy, and completeness over the approaches remarked above. Particularly, some relevant tasks have been automated; new strategies and algorithms integrating robust estimators have been adapted in each step guarantying more reliability and accuracy; considering that both sensors have been registered, a hybrid modeling process has been developed which allow us to complete and improve geometric and radiometric properties of the laser model. Finally, a novel laser scanning tool has been developed in order to test experimental results.

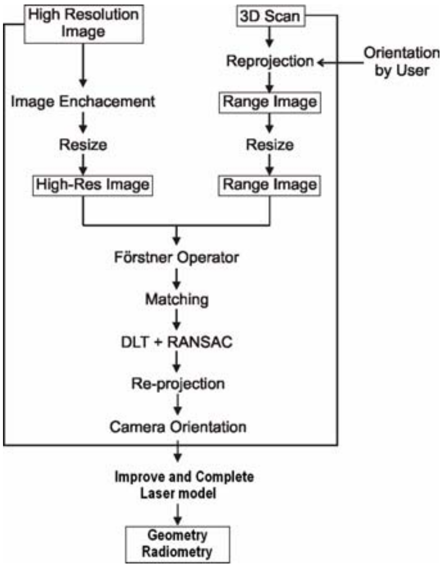

Figure 1. Full pipeline process.

The paper presents the following structure and organization: after this review, Section 2 develops the matching of both images: range and high-resolution. Section 3 explains in detail the co-registration approach developed for both sensors. Section 4 describes a hybrid modeling approach to improve and complete the laser model. Section 5 shows some experimental results using our own novel application of laser scanning. A final section is devoted to outline some conclusions and future works.

2. MATCHING RANGE AND HIGH-RESOLUTION IMAGES

Automated identification of image correspondences is a solved problem in aerial photogrammetry, since image geometry is more standard and relative camera rotations are small. However, in close range applications, each acquisition has its own image geometry, depending on the image scene, the baseline cannot be kept always constant and the rotations around the camera axis are significant. Moreover, more problems arise if we try to match two different types of images such as range and high resolution images. Therefore, the algorithm must be as robust as possible and extracted features should be invariant under different transformations to be re-detectable and useful in automatic procedures.

The matching strategy developed presents a robust and hierarchical approach with the aim of extracting and matching features (interest points) between high-resolution and range images. Firstly, an image pre-processing was applied based on some radiometric equalization and contrast enhancement. Particularly, a blue channel filter is applied in order to eliminate sky influence on high resolution image. With relation to range image, one problem is the ‘air’ or holes due to an insufficient density of points, so a bilinear interpolation is applied to reduce its influences. At last, both images are resized interactively based on some user information. The goal of the resizing is to apply a planar affine transformation to the range image to fit as well as possible to the high-resolution image. The resized images will be used in the matching process. Secondly, an interest point detector method based on Förstner operator (Förstner and Gülch, 1987) has been applied. Many interest point detectors exist in the literature but only a few satisfy accuracy requirements. Through Förstner operator, the detection and localization stages are separated, into the selection of windows, in which features are known to reside and feature location within the selected windows. The windows of interest are computed with a gradient operator (1) and the normal matrix.

$$\begin{aligned} g'_x &= g'_{x+1,y} - g'_{x-1,y} \\ g'_y &= g'_{x,y+1} - g'_{x,y-1} \end{aligned} \qquad (1)$$

The point of interest is determined as weighted centre of gravity of all points inside the window. Further statistics performed locally allow estimating automatically the thresholds for the classification, like trace of inverse matrix (Q) or form of the confidence ellipse. The algorithm requires a complicate implementation and is generally slower compared to other detectors.

Once interest points have been extracted, a hierarchical matching strategy based on ABM (Area Based Matching) and LSM (Least Square Matching) has been applied. At first, the cross-correlation coefficient is used to get a first approximation (2). Around the predicted position a search box is defined and scanned for searching the position which has the higher cross-correlation value. This position is considered a first approximation of the correct point to be matched.

$$\rho = \frac{\sigma_{HR}}{\sigma_H \sigma_R} \qquad (2)$$

where p is the cross-correlation coefficient, σ_{HR} is the covariance between the windows of high-resolution and range image; σ_H is the high-resolution image deviation and σ_R is range image deviation.

Then, the approximation found with cross-correlation is refined using LSM algorithm (Grün, 1985), which provides precise and sub-pixel location of the matching elements. The cross-correlation process uses a small window around each point in the high-resolution image and tries to correlate it against all points that are inside a search area in the range image. The search area is given considering the resize of both images. The final number of possible matches depends on the threshold parameters of the LSM and on the disparity between the two images pairs; usually it is around 40% of the extracted points.

3. CO-REGISTRATION OF SENSORS

Due to the unguided matching process, the found matched pairs might contain outliers. Therefore, a filtering of false correspondences must be performed. Taking into account the authors' experience testing some robust approaches (Aguilera et. al., 2004), several assumptions can be confirmed:

- Least squares adjustments are not robust estimation techniques as wrong observations (like false interest points) can lead to completely wrong results and might even prevent convergence of the adjustment.

- The classical approach to detect blunders in the observations based on the reliability theory or data-snooping technique, developed by (Baarda, 1968) could not solve some critical problems. Moreover, the blunder detection technique has a solid theoretical formulation but it is based on some hypothesis which can lead to unsuccessful results if not satisfied.

- When a large number of observations are available, robust estimators such as RANSAC or M-estimators can perform more efficiently to check for possible outliers. In robust estimations, gross errors are defined as observations which do not fit to the stochastic model used for the parameters estimation.

Taking into account the assumptions remarked upon above, a twofold approach for co-registration of both sensors has been developed:

I. An *estimation step* which allows us to obtain a first approximation of co-registration parameters based on Direct Linear Transformation (DLT) (Abdel-Aziz and Karara, 1971) combined with RANSAC (RANdom SAmpling Consensus) (Fischler and Bolles, 1981). As a result, the most important wrong matched interest points are rejected.

II. A *computation step*, which applies a re-weighted least square adjustment supported by modified Danish M-estimator (Domingo, 2000). A re-projection strategy based on collinearity condition and supported by Danish M-estimator allows us to refine the DLT solution and thus to compute accurate and reliable co-registration parameters.

In a first step, DLT is applied to solve camera orientation using pixel and terrain coordinates. Terrain coordinates are obtained from laser scanner file which relations every pixel of range image with its 3D point projection. This algorithm is used due to be a very well developed algorithm in computer vision, and it can obtain a first result without iterations. This process is upgraded with RANSAC in order to get a reliable camera position.

RANSAC computes several registrations based on a minimal subset of correspondences selected randomly. For each of these "random registrations", the technique searches for all supporting correspondences (correspondences with a DLT error below a given threshold). All correspondences are then used to compute a new camera registration. The process is repeated and the estimation that has the larger set of points and the minimum error is selected as the final registration. The algorithm needs three parameters: the maximum error (in pixels) to consider a correspondence pair as supporting a given registration (10 pixels); the subset of points used in each trial for the first evaluation (11 points); the number of trials (20-30). To estimate this last parameter, we need to know the number of outliers in the data. In our case, we do not know the percentage of outliers in the initial data; furthermore, it will depend on the type of image (the error will increase in poorly textured images where matching algorithms performance is worse). In order to provide more automatism and efficiency, adaptative thresholds have been introduced (Hartley&Zisserman, 2000).

In a second step, a final computation of the co-registration parameters has been obtained based on a re-projection strategy of range image into high-resolution image. An iterative process using collinearity-based approach has been applied to refine the DLT solution allowing us to improve the co-registration of both sensors. Particularly, a set of 2D range image points have been re-projected over the image based on colinearity condition principles and the approximated camera parameters provided by DLT step. Small discrepancies remain between the projected range points and the original extracted high-resolution image points. The 2D coordinates of the extracted points and the re-projected corresponding range image points constitute the input to compute a new registration. This iterative process continues until the Euclidean distance between the re-projected points and the original interest points gets to a minimum (threshold distance). The general idea is that at each iteration the distance between the two datasets is reduced, allowing a better computation of registration parameters. To ensure the convergence of the algorithm and the improvement of the initial camera model estimation, the registration error of each correspondence is computed and recorded. In each new iteration, only matching pairs for which the registration error decreases are updated, and the other are kept unchanged.

Particularly, the method consists of minimizing the Euclidean distance between the re-projected points and the original interest points. Nevertheless, the presence of accidental and "light" gross errors in observations will make that each interest point does not have the same degree of participation in the adjustment. In this way, a re-weighted least square adjustment supported by modified Danish estimator (3) is applied. The numeric solution for this adjustment follows an iterative re-weighted approach, in which the iteration starts with some initial values for the weights of observations and a conventional least square adjustment. In the following iterations, new weights are calculated for each observation based on the residuals obtained in the previous iteration, and a least square adjustment with these new weights is repeated.

$$\boldsymbol{W}(\boldsymbol{v}) = e^{(-|\boldsymbol{v}|^{2})} \tag{3}$$

where w represents the weight function and v the residual vector.

The iterative process continues until the convergence is achieved (usually 3 to 10 iterations). After the computation of co-registration, a full model for the digital camera with relation to laser scanner is available and ready to be exploited. One of the most important advantages of modified Danish estimator regarding to RANSAC is its continuous approach without total loss of observations, providing more accuracy and reliability in the result.

4. A HYBRID MODELING APPROACH

The idea developed in this section, is based on the use of registered high-resolution image not only to texture mapping but also to complete and improve 3D laser model geometry.

4.1 Geometry.

For some modeling applications, like building reconstruction, where the object is mainly described with straight lines, laser scanner technology does not provide a final solution.

The goal of surface modeling can be stated as follows: given a set of sample points *Pi* assumed to lie on or near an unknown surface *S*, create a surface model *S'* approximating *S*. A surface modeling procedure based only in laser scanner dataset cannot exactly guarantee the recovery of *S*, since principal straight lines of the building are not provided by laser scanner. Sometimes additional information of the surface (e.g. breaklines) can be available and thus the output result *S'* is more likely to be topologically correct and converges to the original surface *S*. A perfect scan system should be dense in detail area and sparse in featureless parts and performs automatically. But usually the measured points are unorganized and often noisy; moreover the surface can be arbitrary, with unknown topological type and with sharp features.

The approach presented in this section does not extract directly 3D information from a high-resolution image; it uses features detected in the high-resolution image to complete and improve segmentation in laser model. The process considers the laser model completely triangulated, and not the point's cloud, to make easier the detection and matching of 3D edges.

The co-registration of both sensors allows us to compute correspondences between 2D straight lines belonging to high-resolution image and 3D edges belonging to laser model. The final step of the correction process consists in the alignment of the laser model edges.

A first approximation to laser model has been obtained based on Delaunay triangulation, especially through an incremental strategy (Bourke, 1989). Furthermore, in order to make easier the process, laser model is filtered in order to isolate the main 3D edges. Topological information, as well as normal of triangles are used to isolate these features. As a result, two different types of 3D edges are isolated: *final edges*, those that constitute the surface perimeter and *breaklines edges*, those whose normal variation is greater than a predefined threshold (30-40º).

Afterwards, a robust straight line extraction is performed over the high-resolution image. Particularly, a combination of Canny (1986) and Burns (1986) operators is used to obtain accurate lines. Furthermore, a clustering of these segments based on the analysis of slope and orthogonal distance allow us to obtain the principal lines.

The hybrid modeling method that we propose should infer the correct geometry, topology and features based on the co-registration of both sensors. In this sense, the features extracted on the high-resolution image can provide geometric constraints as well as breaklines to model the object.
At this point, an iterative process starts in which the 3D edges isolated over the laser model start to be corrected based on the straight lines extracted over the high-resolution image. For each extracted segment in the high-resolution image, the algorithm selects the 3D edge points that are projected to the line defined by the segment (the orthogonal distance and the slope variation are used to validate the 3D edges candidates).

Finally, an algorithm corrects the 3D coordinates of points close to straight lines in the laser model. This is done in two steps. First, the parameterised equation of the 3D line is computed, and then 3D coordinates are modified so that the final points will lie in the computed line. The algorithm is applied iteratively with adaptative thresholds to correct as many edge points as possible. The algorithm stops when no correction is computed or when a user-defined number of iterations are reached.

4.2 Texture.

The visualization of a 3D model is often the only product of interest for the external world and remains the only possible contact with the model. Therefore, a realistic and accurate visualization is often required.

Through a texture mapping technique, our registered high-resolution image is mapped onto the 3D laser model in order to achieve photo-realistic virtual models. Knowing the parameters of interior and exterior orientation of the image, for each triangular face of the 3D laser model the corresponding image coordinates are calculated. This approach is performed based on the Anchor Points method developed by (Krauss, 1993). This method has three main steps:

(i) Firstly, laser model image-coordinates are computed through camera model and collinearity condition.

(ii) Next, a correspondence between each face of the laser model with each face of the high-resolution image is established.

(iii) Finally, a projection of the photographic texture between the face of the high-resolution image and its homologous in laser model is performed. In this sense, each triangular face receives a specific transformation model well known as affine transformation.

On the other hand, after the registration procedure, a full model for camera is available. Using this information, each 3D coordinate in the range image can be re-projected into the intensity high-resolution image according to the camera model. Since both images are directly registered, it is possible to establish an association between pixels in range and high-resolution images, and compute a new range image based on the high-resolution colour values. The final image is useful to evaluate the quality of the registration in an easy fast way. It can also be used directly to map texture on the 3D models, giving a much more realistic impression than for a model only textured with the range image. The 3D coordinates when projected in the high-resolution image, will not correspond normally to an integer pixel value. To avoid distortion in the colors, a bilinear interpolation is used to compute the resulting RGB value for the re-projected image. Furthermore, due to the different images resolution, the interpolation is used to compute the "extra" 3D positions.

Finally, regarding visualization, the VRML (Virtual Reality Modeling Language) format was the standard chosen to provide an interactive visualization of the results guaranteeing flexibility and scalability in the visualization at the same time, so different 3D laser models can be incorporated and managed easily. In this way, an automatic transformation of the reconstructed laser 3D model into a topological structure (points, lines and surfaces) sorted hierarchically in a nodes network was performed, allowing three different levels of visualization: wireframe, shaded and textured. Materials defined by their colours and radiometric properties (opaqueness, transparency, diffusion, reflection and emission) and high-resolution textures, are mapped through a uniform and continuous renderization supported internally by VRML. At last, in order to increase the level of realism and completeness of the scene, several basic primitives combined with spherical panoramas can be added.

5. EXPERIMENTAL RESULTS

In order to determine the accuracy, limitations and advantages of the hybrid approach proposed, a series of experiments are tested using our own tool developed.

5.1 The medieval wall of Avila

The medieval wall of Avila represents a fundamental reference point to the Spanish Cultural Heritage. Alfonso VI ordered the construction of this fortification after his conquest of Avila in 1090. Apparently, he used Moorish prisoners to build the wall.

5.1.1 Problem and goal. Two different sensors and images are used to put in practice the approach developed. Particularly, a time of flight laser scanner, Trimble GX200, is used to obtain range image, while a conventional digital camera, Nikon D70 is used to obtain high-resolution images.

The workspace is situated in the north of the medieval wall, in a popular place known as "Arco del Carmen". The principal problems with this experiment are related with its irregular patterns, battlements, which causes a lot of problems in matching phase, as well as its low density scan (about 300.000 points).

5.1.2 Methodology and results. The input data are constituted by a high-resolution image (3008 x2000 pixels) and a resized range image obtained from laser scanner point cloud and collinearity condition (Figure 2).

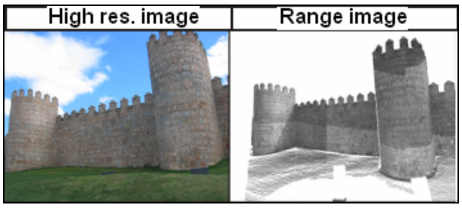

Figure 2. Input data: high-resolution and range image.

The Förstner detector (Figure 3) and a twofold matching strategy are applied to relate both images. A matching kernel of 35x35 pixels is used at first. The final deviation of the matching is 0.71 pixels.

Figure 3. Förstner detector applied to medieval wall.

Afterwards, a robust registration of both sensors is performed. In a first step, DLT and RANSAC are combined to obtain a first approximation of camera parameters (Table 2). Then, in a second step, a re-projection strategy supported by Danish robust estimator is applied iteratively (Table 3).

Sensor registration estimation: DLT + RANSAC

External Parameters (Unit: radians, metres)			
Axis: -1.16377	**X**: 1.195	σ_A : 0.163	σ_X: 0.70
Tilt: -1.0355	**Y**: -1.382	σ_T: 0.0178	σ_Y: 1.10
Swing: -0.3572	**Z**: -0.107	σ_S: 0.0037	σ_Z: 0.52

Table 2: Twofold registration process: estimation

Sensor registration computation: Collinearity + Danish estimator (7th iteration)

External Parameters (Unit: radians, metres)			
Axis: -1.1824	**X**: 1.607	σ_A : 0.0153	σ_X: 0.19
Tilt: -1.0521	**Y**: -2.035	σ_T: 0.0238	σ_Y: 0.14
Swing: -0.1687	**Z**: -0.204	σ_S: 0.0022	σ_Z: 0.13

Table 3: Twofold registration process: computation

Finally, once a complete camera model has been computed, an automatic high-resolution texture mapping is applied (Figure 4).

Figure 4. Mapping high-resolution textures.

5.2 The romanesque church of San Pedro

This romanesque church was founded on the XII century.
The main facade is considered to have important examples of architectural sculpture, even though somewhat damaged.

5.2.1 Problem and goal. Two different sensors and images are used to put in practice the approach developed. Particularly, a time of flight laser scanner, Trimble GS200, is used to obtain range image, while a conventional digital camera, Canon IXUS400 is used to obtain high-resolution images.

The workspace is the main façade of the church. In this case, the scan density is high (over 1.5 millions of points), which allows obtain of range-image with enough texture to apply the matching process.

5.2.2 Methodology and results. The input data are constituted by a high-resolution image (2272 x 1704 pixels) and a resized range image obtained from laser scanner point cloud and collinearity condition (Figure 5).

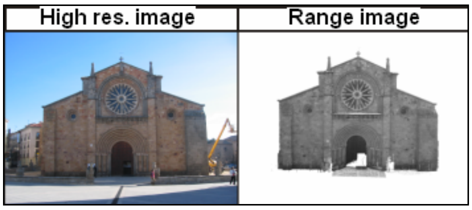

Figure 5. Input data: high-resolution and range image.

The Förstner detector and a twofold matching strategy are applied to relate both images (Figure 6). A matching kernel of 35x35 pixels is used at first. The final deviation of the matching is 0.51 pixels.

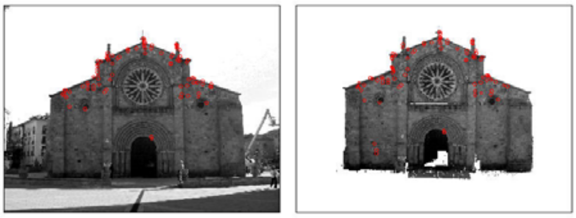

Figure 6. Twofold matching strategy

Afterwards, a robust registration of both sensors is performed. In a first step, DLT and RANSAC are combined to obtain a first approximation of camera parameters (Table 4). Then, in a second step, a re-projection strategy supported by Danish robust estimator is applied iteratively (Table 5).

Sensor registration computation: DLT + RANSAC

External Parameters (Unit: radians, metres)			
Axis: -1.429903	**X**: 8.048	σ_A : 0.0150	σ_X: 0.4433
Tilt: 0.09558	**Y**: -2.615	σ_T: 0.0226	σ_Y: 0.1928
Swing: 0.009088	**Z**: 0.277	σ_S: 0.0055	σ_Z: 0.6027

Table 4. Twofold registration process: estimation

Sensor registration computation: Collinearity + Danish estimator (4th iteration)

External Parameters (Unit: radians, metres)			
Axis: -1.451542	**X**: 8.456	σ_A : 0.0073	σ_X: 0.2143
Tilt: 0.098474	**Y**: -2.377	σ_T: 0.0111	σ_Y: 0.0981
Swing: 0.013599	**Z**: 0.585	σ_S: 0.0025	σ_Z: 0.2956

Table 5. Twofold registration process: computation

Finally, a hybrid modeling process has been developed in order to complete and improve laser model. Regarding geometry, several structural lines related to mesh model (breaklines and final edges) have been corrected based on the co-registration (Figure 7).

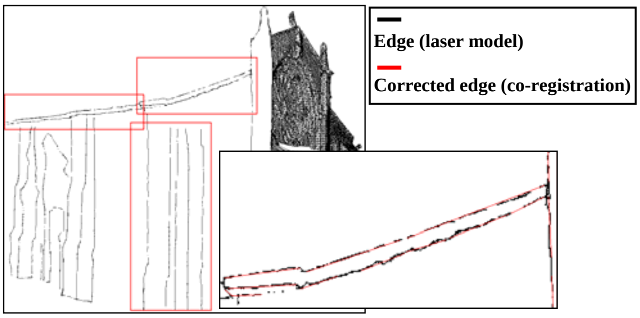

Figure 7. Hybrid modeling process in San Pedro Church.

The time effort required in both experimental results was irrelevant, only user interaction was required for providing an initial approximation of camera pose.

6. CONCLUSIONS AND FUTURE PERSPECTIVES

The presented paper has investigated and developed the automatic co-registration of two sensors: terrestrial laser scanner and high-resolution digital camera, as well as the hybrid modeling approach derived from this fusion. A consistent and reliable full process pipeline has been developed and presented. It was demonstrated with different practical examples tested through our own software.

With relation to the most relevant aspects of the proposed approach, we could remark on:

- Automation in the matching of both images has been achieved.
- No need for previous calibration.
- A robust registration of both sensors is obtained using RANSAC and Danish estimator.
- An alternative to improve and complete laser models is presented.

As for the most critical aspects, this approach has the following limitations:

- User interaction is required to provide a first approximation of the area of interest.
- High resolution scans is advisable to obtain fine texture and thus good quality in matching process.

We feel that we have attacked one of the most difficult problems in the laserscanning success. Nevertheless, several improvements could be considered in the next future. Focusing on geometry, the research could be extended to provide the improvement of complex geometries such as arcs and quadrics. Aiming on radiometry, also algorithms that allow handling the problem of occlusions, illumination properties and transition between junctions, could be developed.

Acknowledgements

We would like to thank to several local institutions, especially ‘Ayuntamiento de Ávila’ and ‘Institución Gran Duque de Alba’ for providing all technical and human assistance as well as for the partial economical support of the research work.

References from Journals:

Allen P. K., Troccoli A., Smith B., 2003. New methods for digital modeling of historic sites. *IEEE Journal*.

Burns, J.B., A.R. Hanson, E.M. Riseman, 1986. Extracting straight lines. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 8 (4).

Canny J.F. 1986. A computational approach to edge detection. *IEEE Journal*.

Grün, A., 1985. Adaptive least squares correlation: a powerful image matching technique. *S. Afr. J. of Photogrammetry, Remote Sensing and Cartography*, 14, 3.

References from Books:

Domingo Preciado Ana, 2000. *Tesis Doctoral*. Universidad de Cantabria.

Elstrom M., 1998. A Stereo-Based Technique for the Registration of Color and LIDAR images, *Master’s Thesis*, University of Tennessee, Knoxville.

Hartley R. and Zisserman A., 2000. *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press.

Kraus, K. 1993. *Advanced Methods and Applications*. Vol.2. *Fundamentals and Standard Processes*. Institute for Photogrammetry Vienna University of Technology.

References from Other Literature:

Abdel-Aziz YI y Karara HM, 1971. Direct linear transformation from comparator coordinates into space coordinates in close range photogrammetry. Symposium on close range photogrammetry.

Aguilera D. G., Lahoz J. G., 2006. Terrestrial laser scanner and high-resolution camera registration through single image-based modeling. *Eurographics Book*.

Aguilera D., Claro-I. M. A., Gomez Lahoz J., Finat J., Gonzalo-T. M., 2004. Development of a simulator for reliable and accurate 3D reconstruction from a single view. ISPRS Congress. Istambul.

Al-Manasir K., Fraser C. S., 2006. Automatic registration of terrestrial laser scanner data via imagery. *ISPRS Archives*. Dresden (Germany).

Alshawabkeh Yahya, Norbert Haala, Dieter Fritsch, 2006. 2D-3D feature extraction and registration of real world scenes. *ISPRS Archives*. Dresden (Germany).

Baarda, W., 1968. A Testing Procedure for use in Geodetic in Networks. *Publications on Geodesy, News Series*, 2(5), Netherlands Geodetic Commission.

Bourke P., 1989. Efficient Triangulation Algorithm Suitable for Terrain Modelling. Pan Pacific Computer Conference, Beijing, China.

Fischler, M. A., and R. C. Bolles, 1981. Random sample consensus. CM Communications.

Förstner, W. and Gülch, E., 1987. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. ISPRS Workshop, Interlaken.

Ikeuchi, Nakazawa, 2003. Creating virtual buddha statues through observation. IEEE Workshop.

Lensch H.P., W. Heidrich, 2001. A silhouette-based algorithm for texture registration and stitching, Graphical Models, vol. 63, n4.

Levoy M., K. Pulli, B. Curless, 2000. The digital Michelangelo project: 3D scanning of large statues. Siggraph, CG Proceedings.

McAllister D. K., L. Nyland, V. Popescu, A. Lastra, and C. McCue, 1999. Real-time rendering of real world environments. In Rendering Techniques, EuroGraphics Rendering Workshop.

Rocchini C., P. Cignomi, C. Montani. Multiple textures stitching and blending on 3D objects, 1999. Eurographics, New York.

Stamos I. And P. K. Allen. Automatic registration of 2-D with 3-D imagery in urban environments, 2001. Proceedings ICCV.