Automatic One Step Extrinsic Calibration of a Multi Layer Laser Scanner relative to a Stereo Camera

Gheorghe Lişca, Pangyu Jeong, Sergiu Nedevschi
Technical University of Cluj Napoca, Computer Science Department
Cluj-Napoca, Romania
{gheorghe.lisca, pangyu.jeong, sergiu.nedevschi}@cs.utcluj.ro

Abstract— **This paper describes a new extrinsic calibration method to estimate the extrinsic parameters of a multi layer laser scanner relative to a stereo camera. Unlike the existing methods, this method has two main advantages: it is automatic and single stepped. The user imprecise selections of calibration features are avoided by using the features of a special designed object as calibration features. The extraction and matching of calibration object features are accurate and automatic. Non coplanar distribution of calibration object features introduces strong constraints on extrinsic parameters. By using these constraints the computation of extrinsic parameters is simultaneously done in a single step. In addition the assumptions on extrinsic parameters and error minimization steps required by existing methods are avoided. The proposed method is tested and the obtained results prove its accuracy.**

Keywords: stereo camera, multilayer laser scanner, extrinsic calibration

I. INTRODUCTION

Indoor and outdoor mobile platforms navigation requires consistent environmental information in order to achieve environment mapping, obstacle detection and collision avoidance. Consistence of the required information depends on the richness of environment features sets (corners, edges, planes, etc.) perceived by sensorial system. The perceptions set of a sensorial system sensor has a specific type (environment features intensity for mono camera, environment feature coordinates for stereo camera or multi layer laser scanner, etc.) given by sensor perceiving model. Also the environment perceptions set is relative to the sensor Coordinate System (CS).

The perceptions sets can be processed in two different ways: 1) by processing separately each sensor's perceptions set or 2) by merging all sensors' perception sets into a common perception set and consequently process only the common perception set. The environment set of features obtained by 2) is more complex than the environment set of features obtained by 1). Therefore the consistence of information extracted by processing the common perceptions set is greater than the consistence of information extracted by processing separately each sensor perceptions set. The common perceptions set is obtained by transforming the coordinate sets of each sensor perceptions sets from the sensor CS into the Common Coordinate System (CCS) which is considered the sensorial system CS. For each of the sensor's perception set a unique transformation is required.

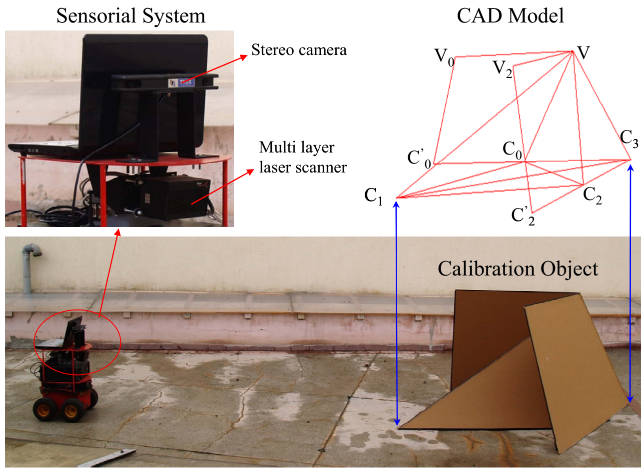


Figure 1. Extrinsic calibration.

The transformation domain contains the perception's coordinates relative to sensor CS and the transformation co-domain contains the perception's coordinates relative to the CCS. On the other hand the required transformation is described by sensor CS pose relative to CCS. The pose parameters are called *extrinsic parameters* and the pose parameters estimation is called *extrinsic calibration*. The calibration uses a mathematical closed form over two sets of parameters. The fist set of parameter contains the extrinsic parameters. The second set of parameters contains the pairs of coordinates obtained by mapping the transformation domain coordinates to the transformation co-domain coordinates. The mapping process is according to a special environment features set, called *calibration feature set*. The set of domain coordinates contains the coordinates of the calibration features relative to the sensor CS. The set of co-domain coordinates contains the coordinates of the calibration features relative to the CCS.

The extrinsic calibration has two parts: 1) calibration features extraction and coordinate pairs set building; 2) extrinsic parameters estimation. In the first part the set of calibration features is chosen and the two sets of calibration coordinates are computed relative to the sensor CS and the CCS. This step is very important because the computation accuracy of the two sets of calibration features coordinates strongly influences the estimation accuracy of extrinsic parameters obtained by solving the mathematic closed form. In the second part the set of coordinate pairs is used to solve the closed form and to find the unknown set of extrinsic

978-1-4244-8230-6/10/$26.00 ©2010 IEEE

parameters. Therefore this is strongly influenced by the distribution of calibration features. A coplanar distribution of calibration features leads to a set of coordinate pairs which can not completely solve the closed form without assuming some extrinsic parameters as known. Unlike the coplanar distribution, a spatial distribution of calibration features is able to solve the mathematic closed form, without any assumption. The mobile platform CS or one of the sensorial system's sensors CS can be chosen as the CCS. According to the above choice there are two types of extrinsic calibration: extrinsic calibration relative to mobile platform's CS and extrinsic calibration relative to one of the sensorial system's sensor CS.

This paper considers extrinsic calibration relative to one of the sensorial system's sensor CS. The stereo camera CS is considered as the CCS and the multi layer laser scanner is extrinsically calibrated relative to it. There are many approaches for the extrinsic calibration of a single layer or a multi layer laser scanner relative to a mono or stereo camera. The differences between these methods are relative to the two extrinsic calibration steps. According to these two steps it is possible to classify the existing extrinsic calibration methods as: manual or automatic relative to 1). Also, relative to 2), the methods can be classified as "multi step" or "single step". In manual mode, the user selects all calibration features so that any environment feature simultaneously perceived by the stereo camera and the multi layer laser scanner [13] and [14] can be considered as a calibration feature [2]. During the selection procedure, the user simultaneously builds the set of coordinates pairs. The accuracy of the coordinates pairs is low because the calibration features selection process is influenced by feature selection errors. ([3] and [10]). In automatic mode, the calibration features belong to a special designed object named Calibration Object (CO). Unlike manual mode, the calibration features can not be any of the environment features. The calibration features must belong to a CO in order to be automatically extractable from the sets of the sensors' perceptions ([4] and [5]). Also by using the CO features, the building of calibration features coordinates pairs is automatic. In the case of multi step methods, two subsets of extrinsic parameters exist [8]. The first subset is represented by the inestimable extrinsic parameters and the second subset is represented by the estimable extrinsic parameters. Inestimable extrinsic parameters are required to be known by the computation of estimable extrinsic parameters. Therefore, the inestimable extrinsic parameters are assumed to have fixed values. Assumptions on the inestimable extrinsic parameters introduce errors in the computation of estimable extrinsic parameters [7], [8], [9]. An error minimization algorithm is required to minimize the introduced errors. Unlike the case of "multi step" methods, in "single step" methods no assumption on inestimable extrinsic parameters is required [12]. All extrinsic parameters are simultaneously computed and no error minimization is required [6] and [11]. The extrinsic calibration method proposed in this paper uses the CO features as calibration features. The CO features are automatically extracted from the stereo cameras' set of perceptions and from the multi layer laser scanner's set of perceptions. Also the set of coordinate pairs is automatically built and the mathematic closed form, which uses this set, is solved in a single step.

This paper consists of five chapters. The second chapter describes the sensorial system and the properties of CO model. The third chapter presents the proposed calibration method. The experimental results are presented in the forth chapter and the conclusions are stated in the fifth chapter.

II. Sensorial System Alignment

The mobile platforms' sensorial systems include different types of sensors. The size of the environment features set and the accuracy of environment features perceptions depends on the fields of view of the sensorial system's sensors and also their orientation. A sensorial system which has separate fields of view for each sensor, perceives a set of environment features larger than the one perceived by a sensorial system which has overlapped fields of view. In contrast, the accuracy of perceptions sets returned by a sensorial system which has overlapped fields of view is higher than the accuracy of perceptions sets returned by the same sensorial system which has separate fields of view.

The sensorial system considered in this paper (a stereo camera and a multi layer laser scanner) has its sensors' fields of view overlapped.

A. Sensorial System

Stereo cameras and the multi layer laser scanner are omnipresent in all mobile platforms' sensorial systems because using them together considerably improves the accuracy of the environment features perceptions. The mobile platform considered in this paper is equipped with a sensorial system which includes a stereo camera and a multi layer laser scanner (Fig.1). The stereo camera returns a set of raw image pairs $<Img^C_{Left}, Img^C_{Right}>$ (from the left and right camera), and a set of 3D stereo reconstructed points $P^C=\{\ p^C_i \mid p^C_i \ \textit{stereo reconstructed}\}$. The stereo reconstruction process often fails in the case of surfaces with homogenous texture and leads to a wrong computation of these surfaces depths. The multi layer laser scanner is able to perceive just four slices of environment features because the set of laser beams is distributed over four distinct planes called *scanning layers*. The set of 3D scanned points $P^L=\{\ p^L_i \mid p^L_i \ \textit{scanned}\}$ accurately approximates the four intersection lines between the scanning layers and environment features. Processing together both sensors' perception sets, it is possible to take advantages like the correction of stereo depth with the scanned depth in the case of surfaces with homogenous texture, or the correction of far distance environment features. In order to benefit more advantages, it is required that the stereo camera's perception coordinates and multi layer laser scanner's perception coordinates to be relative to a unique CCS. The stereo camera's CS is considered as the CCS and the coordinates of multi layer laser scanner's perception set are transformed relative to the CCS. The multi layer laser scanner CS pose relative to the stereo camera CS describes the required transformation, denoted with T^{LC}. This pose contains a rotation matrix and a translation vector. The estimation of transformation T^{LC} is done thanks to the CO which has special geometric features highlighted in below paper section.

B. Calibration Object

The main goal of CO is to facilitate obtaining the corners coordinate pairs set $C^{LC} = \{< c^L, c^C >, c^L \in C^L \wedge c^C \in C^C\}$ required in order to solve the mathematic closed form. The C^{LC} set is obtained by mapping the set of CO corner coordinates $C^L = \{c^L, c \in C\}$ which describes the CO pose CO^L relative to multi layer laser scanner CS to the set of CO corners coordinates $C^C = \{c^C, c \in C\}$ which describes the CO pose CO^C relative to stereo camera. The CO corners coordinates mapping procedure is done according to CO corners set $C = \{C_i \mid C_i \text{ is CO corner}\}$. On the other hand corners coordinates mapping procedure is described by the transformation T^{LC}. The C^L set represents the transformation domain and the C^C set represents the transformation co-domain. Both CO poses CO^L and CO^C are estimated by processing the CO features sets perceived by each sensor. The proposed CO has different features sets: four lateral faces and a base $F=\{f \mid f \text{ is plane}\}$, four lateral edges and four base edges $E=\{e \mid e \text{ is segment}\}$ and five corners $C=\{c \mid c \text{ is point}\}$ perceivable by both sensors. Thanks to the CO features sets and CO pyramidal geometry, the two CO poses CO^L and CO^C are uniquely determined and sensitive to any small changes of CO position. Stereo camera directly perceives and estimates the set of CO corners coordinates C^C, but multi layer laser scanner indirectly computes the set of CO corners coordinates C^L, because multi layer laser scanner perceives only CO faces and CO edges. Both coordinate sets (C^L and C^C) are affected by stereo reconstruction errors (in C^C set case) and an error from corners coordinates indirectly computations (in C^L set case). Different CO corners coordinates C^C and C^L are unequal affected by stereo reconstruction errors and computing errors. Therefore the two CO poses CO^L and CO^C which correspond to the two sets of coordinates C^L and C^C are distorted and as consequence the geometric relations between the set of CO^C corners coordinates are different than the geometric relations between the set of CO^L corners coordinates. The required transformation T^{LC} preserves same geometric relations between domain coordinates and co-domain coordinates. Therefore in order to solve transformation T^{LC}, the required set of coordinates pairs must be built by mapping two sets of coordinates with the same geometric relations. The set of C^{LC} pairs is built from two sets of coordinates C^L and C^C which have different geometric relations. Therefore it is impossible to use it in order to solve the required transformation T^{LC}. The different geometric relations are induced by distorted CO^C and CO^L. By eliminating CO^C and CO^L pose distortions, the C^{LC} pair set is usable in order to solve mathematic closed form. In order to eliminate the CO^C and CO^L distortions and to equalize the stereo reconstruction errors and corners computing errors a CO CAD model fitting procedure based on RANSAC idea is used. The CO CAD model (CO pyramidal geometry and CO edges lengths) is known from CO construction (like checkerboard squares dimensions which are known for camera calibration). The fitting procedure and its applications for stereo camera and multi layer laser scanner perceptions are detailed in the third part of this paper.

III. Proposed Calibration Method

Generally a pair of coordinates $<P^L, P^C>$ corresponds to an environment feature P. The features coordinates $P^L=[x^L\ y^L\ z^L]^T$ are relative to the multi layer laser scanner CS and the feature coordinates $P^C=[x^C\ y^C\ z^C]^T$ are relative to stereo camera CS. The P^C coordinates are estimated from P^L coordinates using the mathematic closed form (1) and the transformation T^{LC}. Also the closed form (1) is used to estimate the transformation T^{LC} if at least four pairs of coordinates $<P^L, P^C>$ are known. This transformation T^{LC} includes the rotation R^{LC} matrix and the translation vector t^{LC}.

$$P^C = R^{LC} \cdot P^L + t^{LC} \quad (1)$$

The CO corner C_i represents the feature P and the pair of CO corner coordinates $<C^L_i, C^C_i>$ represents the pair of feature coordinates $<P^L, P^C>$. By using CO corner coordinates $P^L=[x^L\ y^L\ z^L\ 1]^T$ and $P^C=[x^C\ y^C\ z^C\ 1]^T$ in homogenous representation, closed form (1) is equivalently written (2).

$$C_i^C = T^{LC} \cdot C_i^L \quad (2)$$

Multi layer laser scanner CS orientation relative to the stereo camera CS is described by three rotation angles: γ representing the rotation about x axis, β representing the rotation about y axis and α representing the rotation about z axis and three relative position components noted with t_x, t_y and t_z. By using this notations the transformation T^{LC} has the matrix form (3) ($c_\alpha=\cos\alpha$ and $s_\alpha=\sin\alpha$) which its further simplified to the form (4).

$$T^{LC} = \begin{bmatrix} c_\alpha c_\beta & c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma & c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma & t_x \\ s_\alpha c_\beta & s_\alpha s_\beta s_\gamma - c_\alpha c_\gamma & s_\alpha s_\beta c_\gamma - c_\alpha s_\gamma & t_y \\ -s_\beta & c_\beta s_\gamma & c_\beta c_\gamma & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$T^{LC} \overset{Noted}{=} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

For a CO corner c_i, knowing its coordinates C^L_i relative to multi layer laser scanner CS and the transformation T^{LC}, it is possible to estimate the coordinates C^C_i relative to the stereo camera CS (2). The set of coordinates pairs $<C^L_i, C^C_i>$ is obtained by matching the two sets of CO corners coordinates C^C and C^L relative to the CO corners set C. The rotation matrix R^{LC} and the translation vector t^{LC} elements are computed using the mathematic closed form (2) and the constraints set induced by the set of corner coordinates pairs C^{LC} over transformation T^{LC} elements. In order to compute

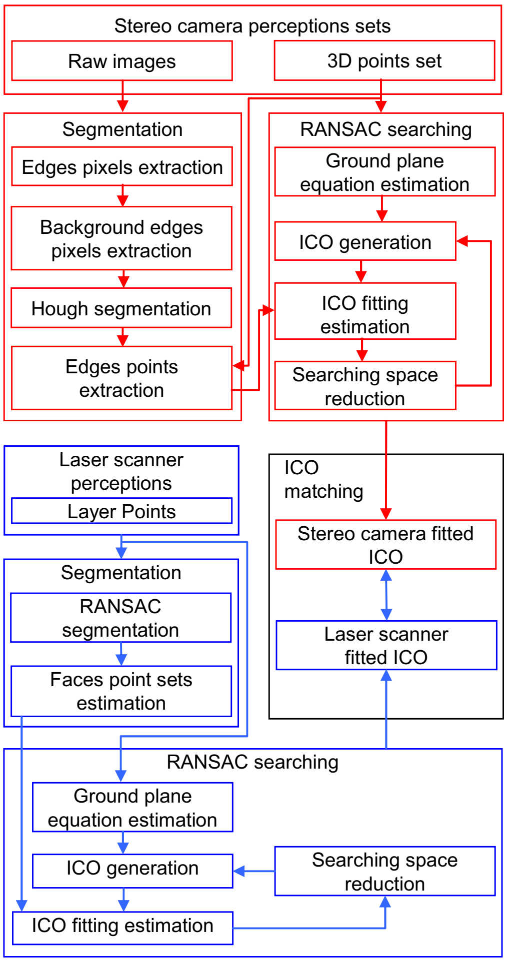


Figure 2. Extrinsic calibration scheme.

all twelve T^{LC} elements it is necessary to use at least four pairs of coordinates. The corners must be no coplanar because if they are coplanar some induced constraints become redundant. The CO corners coordinates $C^{C}{}_{i}$ are directly computed by stereo reconstruction procedure. Multi layer laser scanner can perceive only the CO faces so that the CO corners coordinates $C^{L}{}_{i}$ are computed by using the CO geometry and the perceived CO faces.

The corner coordinates $C^{C}{}_{i}$ stereo reconstructed are strongly affected by noise so that it is impossible to extract them with high accuracy. Multi layer laser scanner cannot directly estimate the coordinates $C^{L}{}_{i}$, it can percept only CO faces so that the coordinates $C^{L}{}_{i}$ are indirectly estimated. As consequences the two CO poses CO^{C} and CO^{L} are distorted making the set of coordinates pairs C^{LC} unusable in order to solve the closed form (2).

The main steps of proposed extrinsic calibration are depicted in Figure 2. The key idea is to find two CO CAD models instances (*ICO*) one which is the best fitted into stereo reconstructed points set C^{C} and another one which is the best fitted into scanned points set C^{L}. The best fitted two instances $ICO^{C}{}_{Final}$ and $ICO^{L}{}_{Final}$ aren't distorted and as immediate consequences their sets of coordinates (C^{C} and C^{L}) mapping is straight forward and the resulted coordinate pairs set C^{LC} is usable in order to solve the closed form (2). Finding the best fitted two CO CAD models is done by searching into stereo reconstructed points set C^{C} area and into scanned points set C^{L} area the best fitted $ICO^{C}{}_{Final}$ and $ICO^{L}{}_{Final}$. The searching is done by randomly generating multiple $ICO^{C}{}_{Random}$ and $ICO^{L}{}_{Random}$ into the searching spaces. An *ICO* random generation requires to known a minimum CO CAD elements set called *CO generation elements set*. This set consist of CO edges lengths, CO base supporting plane equation, a CO base corner's coordinates and the direction of CO base diagonal containing the known CO base corner. The searching start from approximations of CO generation elements sets. The three dimensional searching spaces size $|C^{C}|$ and $|C^{L}|$ strongly influence the searching convergences. The CO base plane surface coincides with the ground plane surface which is estimable from each sensor perceptions sets. Based on this observation the two three-dimensional searching spaces are reduced at two searching planes $\pi^{C}{}_{Support}$ and $\pi^{L}{}_{Support}$ which represent the ground surface relatively to stereo camera CS and multi layer laser scanner CS. The searching spaces reduction continues by considering only two limited searching areas $A^{C}{}_{Support}$ and $A^{L}{}_{Support}$ from the searching planes $\pi^{C}{}_{Support}$ and $\pi^{L}{}_{Support}$. These searching areas $A^{C}{}_{Support}$ and $A^{L}{}_{Support}$ which are obtained by considering the two interest volumes V^{C} and V^{L} and intersecting them with the corresponding supporting planes $\pi^{C}{}_{Support}$ and $\pi^{L}{}_{Support}$. In order to find the best fitted two $ICO^{C}{}_{Final}$ and $ICO^{L}{}_{Final}$ it is required to estimate for each two $ICO^{C}{}_{Random}$ and $ICO^{L}{}_{Random}$ how well is fitted. The fitting errors estimation $e^{C}{}_{Random}$ and $e^{L}{}_{Random}$ are done computing the average distance from $ICO^{C}{}_{Random}$ and $ICO^{L}{}_{Random}$ features (CO edges set E and CO faces F) to their correspondent stereo reconstructed 3D point subsets $E^{C}{}_{i}$ and scanned 3D points $F^{L}{}_{i}$. The stereo reconstructed 3D points subsets $E^{C}{}_{i}$ and scanned 3D points $F^{L}{}_{i}$ are obtained by segmenting the CO stereo reconstructed points set $P^{C}{}_{CO} \subset P^{C}$ and the CO scanned points set $P^{L}{}_{CO} \subset P^{L}$. Stereo camera points set $P^{C}{}_{CO}$ and multi layer laser scanner points set $P^{L}{}_{CO}$ segmentation procedures assume different intermediate steps which are particularly described below. Searching the best fitted two CO CAD convergences and searching space reductions are also explained below. Searching ends when the fitting errors $e^{C}{}_{Random}$ and $e^{L}{}_{Random}$ are smaller than sensors errors. The best fitted $ICO^{C}{}_{Final}$ and $ICO^{L}{}_{Final}$ are automatically matched and the set of coordinate pairs C^{LC} is built. Using the set of coordinates C^{LC}, the mathematic closed form is solved and the extrinsic calibration is done.

A. *Stereo camera model fitting*

This procedure has the goal to find the best fitted $ICO^{C}{}_{Final}$ into the stereo camera points set $P^{C}{}_{CO}$. Its components are depicted in Fig. 2 and their intermediate steps are described below.

1) Estimation of ground plane equation: In order to enlarge the size of stereo reconstructed ground points set, the CO is kept out of calibration scene and only the ground surface is perceived by sensorial system. A volume of interest $V^{C}{}_{Grond\ Interest}$ is defined and the stereo reconstructed point set $P^{C}{}_{Scene}$ is filtered relative to it. The resulted ground points set $P^C{}_{Ground}$ contains only ground points. Least square method is used to find the support plane $\pi^C{}_{Support}$ which is the best fitted into the ground points set $P^C{}_{Ground}$.

2) Segmentation of object points cloud: The CO is fixed in frond of sensorial system. Stereo camera returns the stereo reconstructed points set P^C. This points set is filtered relative to an interest volume $V^C{}_{CO}$ so that the resulted object points set $P^C{}_{CO}$ contains just CO stereo reconstructed points. The segmentation purpose is to identify stereo reconstructed points subsets $E^C{}_i$ corresponding to CO edges $e^C{}_i$. Segmentation begins with CO edges image obtained by applying canny algorithm to left camera image $Img^C{}_{Left}$. As result the pixels set $P^{Image}{}_{SceneAndObject}$ corresponding to all CO edges is obtained. The background pixels from edges pixels set $P^{Image}{}_{SceneAndObject}$ are eliminated and the resulted image pixels set $P^{Image}{}_{CO}$ it is divided in edge pixels subsets $P^{Image}{}_{CO}(e_i)$ using the Hough Line Detector. Each set $P^{Image}{}_{CO}(e_i)$ represents the projection of an *ICO* edge $e^C{}_i$ on camera left image $Img^C{}_{Left}$. The 3D points sets $E^C{}_i$ corresponds to $P^{Image}{}_{CO}(e_i)$ is found by identifying the 3D stereo reconstructed point $p^c{}_i$ corresponding to each $P^{Image}{}_{CO}(e_i)$ pixel. The obtained 3D points subsets $E^C{}_i$ are used to approximate the CO generation elements set required in order to start the $ICO^C{}_{Final}$ searching. Trough each 3D points subsets $E^C{}_i$ a 3D edge $e^C{}_i$ is fitted using the least square method. The intersections between lateral edges and the support plane $\pi^C{}_{Support}$ represent the four CO base corners approximations $B^C{}_i$. Knowing them the CO base diagonals are found straight forward. Therefore the searching procedure described below can begin.

3) RANSAC searching: In order to find the best fitted $ICO^C{}_{Final}$ many random $ICO^C{}_{Random}$ are generated starting from the CO generation elements set obtained before.

A circle $C^C{}_0(B^C{}_0, r^C)$ with the base corner $B^C{}_0$ as center and radius r^C is considered. Inside $C^C{}_0$ circle a random point $B^C{}_{0\ random}$ is generated. In a similar way is considered the second circle $C^C{}_1(B^C{}_1, r^C)$ with the opposite base corner $B^C{}_1$ as center and same radius r^C and inside it a second random point $B^C{}_{1\ random}$ is generated. The segment direction $\overline{B^C_{0\,random}B^C_{2\,random}}$ gives the necessary diagonal direction. Knowing these elements the corresponding *ICO* is constructed.

In order to estimate the $ICO^C{}_{Random}$ fitting error $e^C{}_{Random}$ it is necessary to find the correspondences pairs $<e^C{}_i, E^C{}_i>$ between the $ICO^C{}_{Random}$ edges $e^C{}_i$ and their stereo reconstructed points subsets $E^C{}_i$. The pairs $<e^C{}_i, E^C{}_i>$ are obtained by ordering the $e^C{}_i$ edges set relative to their middle and the $E^C{}_i$ subsets relative to their centroid. The edge $e^C{}_i$ fitting error is computed as the average distance $m^c{}_i$ (5) of

$$m_i^c = \frac{1}{p^c}\sum_{j=1}^{j=p^C} d(e_i^c, p_j^c) \tag{5}$$

each point $p^c{}_i \in E^C{}_i$ relative to the edge $e^C{}_i$, where $d(e^C{}_i, p^c{}_i)$ is the point $p^c{}_i$ distance to the $ICO^C{}_{Random}$ edge $e^C{}_i$ and $p^c=|P^C{}_i|$ is the set $E^C{}_i$ cardinal. The $ICO^C{}_{Random}$ fitting error is all edge $e^C{}_i$ fitting error average (6).

$$e^c_{Random} = \frac{1}{e^c}\sum_{i=1}^{i=e^C} m_i^c \tag{6}$$

If $ICO^C{}_{Random}$ fitting error $e^C{}_{Random}$ is smaller than current $ICO^C{}_{Final}$ fitting error $e^C{}_{Final}$ then the $ICO^C{}_{Final}$ is updated with the current $ICO^C{}_{Final}$. The base corner $B^C{}_0$ approximation is replaced by $B^C{}_{0\ random}$ and the base corner $B^C{}_1$ approximation is replaced by $B^C{}_{1\ random}$. Also the two circles ($C^C{}_0$ and $C^C{}_1$) radius r^C are reduced with the distance from $ICO^C{}_{Random}$ centroid to $ICO^C{}_{Final}$ centroid. This reduction contributes to the searching procedure convergence (Fig.3). If the estimated error is still bigger than sensor error, the searching continues with new $ICO^C{}_{Random}$ random generation otherwise the searching stops and the kept $ICO^C{}_{Final}$ is considered the searched one.

B. Multi layer laser scanner model fitting

Finding the best fitted $ICO^C{}_{Final}$ into the stereo camera points set $P^C{}_{CO}$ is the goal of this procedure. Its components are depicted in Fig. 2 and their intermediate steps are described below.

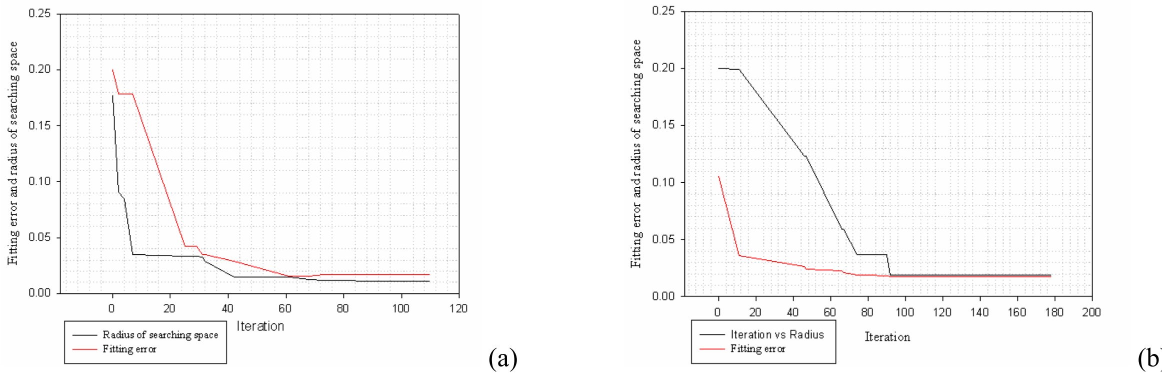
Multi layer laser scanner scanned points set P^L is obtained by merging the scanned points $P^{L(i)}$ sets $P^L = P^{L(0)} \cup P^{L(1)} \cup P^{L(2)} \cup P^{L(3)}$ corresponding to the four laser scanner layers. Even that P^L points set is obtained from four scanning layers its size is reduced determining a difficult information extraction process.

1) Estimation of ground plane equation: In order to increase the ground surface scanning points set size $P^L{}_{Scene} = P^{L(0)}{}_{Scene} \cup P^{L(1)}{}_{Scene} \cup P^{L(2)}{}_{Scene} \cup P^{L(3)}{}_{Scene}$, the CO is kept out from the calibration scene. The scanning points which don't correspond to ground surface are filtered out by the defined interest volume $V^L{}_{Grond\ \mathrm{Interest}}$. Trough the resulted points set $P^C{}_{Ground}$ the plane $\pi^L{}_{Support}$ is fitted using least square method.

2) Segmentation of object points cloud: The CO is fixed in front of the sensorial system and multi layer laser scanner returns four scanning points sets $P^{L(i)}{}_{SceneAndCO}$, $\overline{i=0\ldots3}$ corresponding to the four CO faces slices perceived. The points subsets $P^{L(i)}{}_{CO}$, $\overline{i=0\ldots3}$ results by filtering relative to an interest volume $V^L{}_{CO\ Interest}$ the four scanning points sets.

The segmentation purpose is to identify the scanned points subsets $F^L{}_j$ corresponding to the four CO faces $f^L{}_j$. Therefore, four points subsets $F^L{}_{ij} \subset P^{L(i)}{}_{CO}$ corresponding to the four CO faces $f^L{}_j$ scanned by layer *i* are obtained by applying four times RANSAC search on layer *i* points set $P^{L(i)}{}_{CO}$. The points sets $F^L{}_j$ corresponding to CO faces $f^L{}_j$ are obtained by cumulating the resulted points subsets $F^L{}_j = \cup F^L{}_{ij}$ $\overline{j=0\ldots3}$.

The obtained points sets $F^L{}_i$ are used to approximate the CO generation elements set required in order to begin $ICO^L{}_{Final}$ searching. Therefore trough each points sets $F^L{}_i$ using least square method CO face plane $f^L{}_j$ coefficients are computed. The CO lateral edges $e^L{}_k$ equations are computed by intersecting two by two the CO faces. Finally the CO base



(a) (b)

Figure 3. Searching space reduction and fitting convergences: (a) stereo camera, (b) multi layer laser scanner

corners approximations B^L_i are computed intersecting the CO lateral edges with the ground surface plane $\pi^L_{Support}$. Estimation of CO base diagonals directions is straightforward so that the searching procedure can begin.

3) RANSAC searching: Many random ICO^L_{Random} are generated inside the area of CO generation elements set in order to find the best fitted ICO^L_{Final}.

Inside each circle of the two considered $C^L_0(B^L_0, r^L)$ and $C^L_1(B^L_1, r^L)$ a point $B^L_{0\ Random}$ and $B^L_{0\ Random}$ is randomly generated. Knowing the CO generation elements set, first ICO^L_{Random} is randomly generated.

The e^L_{Random} fitting error is estimated relative to correspondences pairs $<f^L_j, F^L_j>$. This correspondence set is obtained by mapping the CO faces f^L_j to (ordered relative to their centers) to the points F^L_i sets (ordered relative to their centroids). The face f^L_j fitting error is computed as average distance m^L_i (7) of each point $p^L_i \in F^L_i$ relative to its corresponding face f^L_j, where $d(p^L_j, f^L_j)$ is the point p^L_j distance to ICO^L_{Random} face f^L_j and $p^L=|P^L_i|$ is set F^L_i cardinal.

$$m^L_k = \frac{1}{p^L}\sum_{j=1}^{j=p^L} d(f^L_k, p^L_j) \tag{7}$$

The ICO^C_{Random} fitting error e^L_{Random} represents all faces f^L_j fitting errors average (8).

$$e^L_{Random} = \frac{1}{p^L}\sum_{k=1}^{k=p^L} m^L_k \tag{8}$$

Each ICO^L_{Random} is compared with ICO^L_{Final} and is replaced by ICO^L_{Random} only if the fitting error e^L_{Random} is smaller than e^L_{Final}. The space reduction and its contribution to searching procedure convergence are depicted in (Fig.3 (b)). Finally by searching, the e^L_{Final} become lower than multi layer laser scanner error and the searching is end.

C. Matching fitted instances and solving the closed form

The two undistorted ICO^C_{Final} and ICO^L_{Final} best fitted into CO stereo reconstructed points set P^C and CO scanned points set P^L are matched and their sets of corners coordinates C^C and C^L are used in order to build the pairs of corners coordinates $<C^L_i, C^C_i>$ required in order to solve the mathematic closed form (2).

Equation system form (9) is obtained by explicitly writing the mathematic closed form (2) in the terms of transformation T^{LC} elements and the elements of CO corners coordinates pairs $<C^L_i, C^C_i>$. For equation system (9) the pairs of coordinates represent the known system coefficients and the transformation T^{LC} elements (4) represent the unknown parameters. A coordinate pair $<C^L_i, C^C_i>$ introduces three equations in equations system (9) but in order to find all twelve unknown elements is necessary to use at least four coordinate pairs corresponding to four non coplanar CO corners.

$$\begin{cases} r_{00}\cdot x^L_i + r_{01}\cdot x^L_i r_{02}\cdot x^L_i + t_0 = x^C_i \\ r_{10}\cdot x^L_i + r_{11}\cdot x^L_i r_{12}\cdot x^L_i + t_1 = y^C_i \\ r_{20}\cdot x^L_i + r_{21}\cdot x^L_i r_{22}\cdot x^L_i + t_2 = z^C_i \end{cases} \tag{9}$$

The equation system (9) is of type (10). The coordinate set C^L elements represents the A matrix elements (11), the coordinate set C^C elements represent the B matrix elements (12) and the transformation T^{LC} represent the X matrix elements (13).

$$A\cdot X = B \tag{10}$$

$$A=\begin{bmatrix} x^L_0 & y^L_0 & z^L_0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x^L_0 & y^L_0 & z^L_0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x^L_0 & y^L_0 & z^L_0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x^L_{m-1} & y^L_{m-1} & z^L_{m-1} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x^L_{m-1} & y^L_{m-1} & z^L_{m-1} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x^L_{m-1} & y^L_{m-1} & z^L_{m-1} & 1 \end{bmatrix} \tag{11}$$

$$B = \begin{bmatrix} x^C_0 & y^C_0 & z^C_0 & 1 & \ldots & x^C_{m-1} & x^C_{m-1} & x^C_{m-1} & 1 \end{bmatrix}^T \tag{12}$$

$$X=[\, r_{00}\, r_{01}\, r_{02}\, t_0\;\; r_{10}\, r_{11}\, r_{12}\, t_1\;\; r_{20}\, r_{21}\, r_{22}\, t_2 \,]^T \tag{13}$$

The equations system (10) is solved using CVD matrix decomposition. Multi layer laser scanner CS position vector components relative to stereo camera CS are directly computed (3), (4). Multi layer laser scanner CS orientation angles relative to stereo camera CS are computed using the mathematic relations (14), (15), (16).

$$\alpha = \tan^{-1}(r_{10}/r_{00}) \tag{14}$$

$$\beta = \tan^{-1}(-r_{20}/\sqrt{r^2_{21}+r^2_{22}}) \tag{15}$$

$$\gamma = \tan^{-1}(r_{21}/r_{22}) \tag{16}$$

Therefore all three rotations angles and three position vector components of multi layer laser scanner CS relative to stereo camera CS are simultaneously computed.

TABLE I. VARIANCES OF ESTIMATIONS OF RELATIVE POSITION

CO Position	Nr Samples	$t1_{L2C}(m)$	$t1_{C2L}(m)$	$\Delta_{t1}(m)$	$t2_{L2C}(m)$	$t2_{C2L}(m)$	Δ_{t2} (m)	$t3_{L2C}(m)$	$t3_{C2L}(m)$	Δ_{t3} (m)
1	20	-0.1039	0.1094	0.0055	0.3213	-0.3161	0.0052	-0.0966	0.0943	0.0023
2	20	0.0606	-0.0625	0.0018	0.3646	-0.3612	0.0034	-0.1280	0.1219	0.0061
3	20	0.1551	-0.1513	0.0038	0.2915	-0.2956	0.0041	-0.1493	0.1421	0.0072

TABLE II. VARIANCES OF ESTIMATIONS OF RELATIVE ORIENTATION

CO Position	Nr Samples	$Yaw_{L2C}(°)$	$Yaw_{C2}(°)$	$\Delta_{Yaw}(°)$	$Pitch_{L2C}(°)$	$Pitch_{C2L}(°)$	$\Delta_{Pitch}(°)$	$Roll_{L2C}(°)$	$Roll_{C2L}(°)$	$\Delta_{Roll}(°)$
1	20	-0.7097	0.5718	0.1379	1.9640	-2.0787	0.1147	-0.9693	0.9193	0.0500
2	20	-0.1265	0.0941	0.0324	0.6347	-0.6681	0.0334	-0.7131	0.7000	0.0131
3	20	-0.9339	0.9184	0.0155	-0.6672	0.6658	0.0014	-1.6548	1.6374	0.0174

TABLE III. CALIBRATION OBJECT DIMENSIONS

VC_0 (m)	VC_1 (m)	VC_2 (m)	VC_3 (m)	C_0C_1 (m)	C_1C_2 (m)	C_2C_3 (m)	C_3C_0 (m)
1.30	2.38	1.30	1.31	2.0	2.0	1.25	1.25

VV_0 (m)	$V_0C'_0$ (m)	C'_0C_0 (m)	VV_2 (m)	$V_2C'_2$ (m)	C'_2C_2 (m)
1.30	1.30	0.80	1.30	1.30	0.80

IV. EXPERIMENTAL RESULTS

The proposed method extrinsically calibrates the IBEO LUX multi layer laser scanner relative to the TYZX stereo camera by using the proposed CO. All fourteen lengths of CO edges are shown in the Table III.

The set of stereo reconstructed CO points is affected by stereo reconstruction errors which reduce the robustness of stereo camera model fitting procedure. In addition, the size of a CO scanned points set reduce the robustness of multi layer laser scanner model fitting procedure. The influences of stereo reconstruction errors and the size of scanned points set in the CO model fitting procedures are diminished by considering 20 accumulated sets of CO stereo reconstructed points and 20 accumulated sets of CO scanned points at the same CO fixed position.

The convergence graph of the stereo camera model fitting is presented in Fig.3 (a) which shows how the searching space reduction influence the minimization of fitting error. As the convergence is obtained in less than 60 searching iterations, the fitting error is reduced from 17 centimeters to less than 3 centimeters (under the accuracy of stereo reconstruction). Similarly, Fig. 3 (b) shows how the searching space reduction in the case of multi layer laser scanner model fitting influences the minimization of fitting error. The fitting convergence is obtained in less than 100 iterations and the fitting error is reduced from 11 centimeters to less than 2 centimeters (lower than the accuracy of laser scanner which is 4 centimeters).

The multi layer laser scanner CS pose relative to stereo camera CS is three times estimated. The CO is fixed in three different positions relative to sensorial system. From each fixed position, the all three orientation angles (*yaw=α, pitch=β and roll=γ*) and the all three components of position vector are computed. The computed values and their variances according to the three CO fixed positions are shown in Table I and Table II. The variances of the all three orientation angles of multi layer laser scanner CS relative to the stereo camera CS are less than 2 degrees. Also, the variances of the all three components of multi layer laser scanner CS position vector relative to the stereo camera CS are less than 7 centimeters.

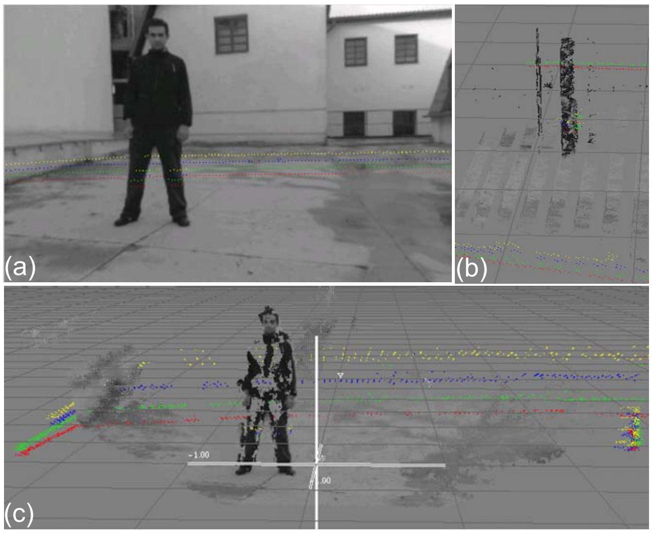


Figure 4. New calibrated sensorial system perceiving a pedestrian: (a) scanned points projected on stereo camera left image, (b) lateral view and (c) front view of scanned points overlapped on stereo reconstructed points.

The new coordinates of scanned points relative to stereo camera CS are computed by applying the computed transformation to the scanned points, which are initially relative to multi layer laser scanner CS. The new calibrated sensorial system is tested in two scenarios (in cases of pedestrian and car). In first scenario the sensorial system perceives a pedestrian (Fig. 4). The transformed scanned points are projected with different colors according to the scanning layers on stereo camera left image (Fig. 4 (a)). The projections of scanned points are correctly overlapped on stereo camera left image. This overlapping is correct thanks to the transformation computed by proposed extrinsic calibration. Fig, 4(b) depicts the front view and Fig. 4(c)

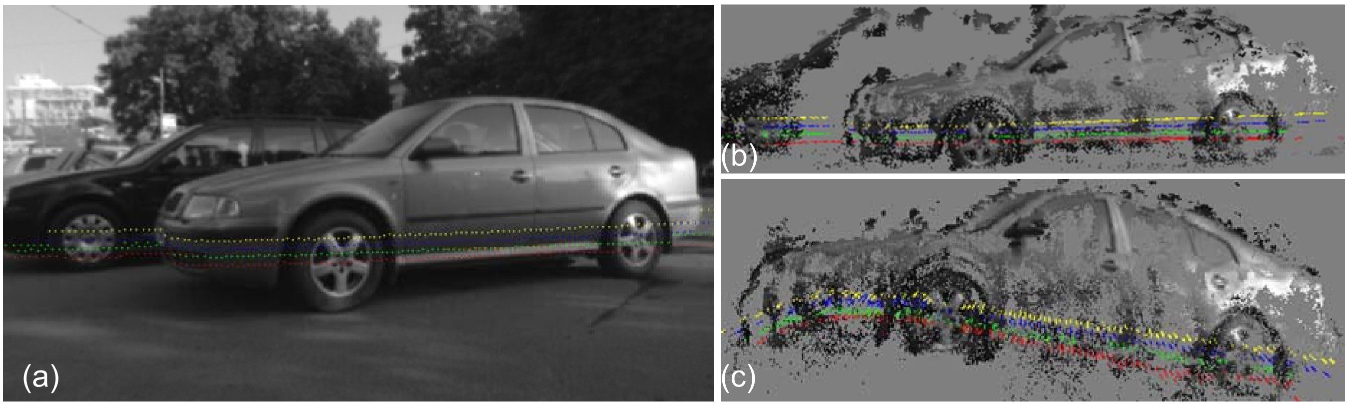


Figure 5. New calibrated sensorial system perceiving a car: (a) scanned points projected on stereo camera left image, (b, c) different front views of scanned points overlapped on stereo reconstructed points.

depicts the lateral view of stereo reconstructed points set and the scanned points set. In same manner correct overlapping is obtained for the second scenario which is depicted in Fig. 5. In many cases, the stereo depth estimation is wrong because homogenous texture Therefore incomplete stereo depths are obtained (Fig. 4b, 4c, 5b, 5c). Stereo reconstructed points are depicted with gray colors and the four scanning layers' points are depicted with different colors.

V. CONCLUSIONS

This paper presented a method for extrinsic calibration of a multi layer laser scanner relative to a stereo camera. The coordinates of scanned point and the coordinates of stereo reconstructed point corresponding to the same environment features are impossible to be automatically matched. Unlike the existing methods which require user interaction to solve the impossibility of coordinates matching, the proposed method avoids this impossibility by using a special designed calibration object. The corners coordinates of the calibration object was used in order to avoid impossibility of coordinates matching. The fitting procedures of calibration object CAD model eliminated the calibration object distortion introduced by stereo reconstruction errors and the reduced size of scanned points set. The obtained set of corners coordinate pairs generates strong constraints over the extrinsic parameters. These constraints make possible simultaneous estimation of all extrinsic parameters. The estimation of extrinsic parameters is accurate and no error minimization algorithm is mandatory.

REFERENCES

[1] Antone, M., Friedman, Y., 2007. Fully Automated Laser Scanner Range Calibration. *British Machine Vision Conference (BMVC).*

[2] Zhao, H., Chen, Y., Shibasaki, R., 2007. An efficient extrinsic calibration of multiple laser scanners and cameras' sensor system on a mobile platform, *Inteligent Vehicle Symposium*, pp 422-427J.

[3] Scaramuzza D. Harati A. Siegward R., 2007. Extrinsic Self Calibration of a Camera and a 3D Laser Range Finder from Natural Scenes, *International Conference on Intelligent Robots and Systems*, pp 4164-4169.

[4] Naikal, N., Zakhor, A., Kua, J., 2009. Image Augmented Laser Scan Matching for Indoor Localization. *Thechnical Report No. UCB/EECS-2009-35.*

[5] Zang, Q., Pless. R., 2004. Extrinsic Calibration of a Camera and Laser Range Finder (improves camera calibration), *Proceedings of IEEE/RSJ ICIRS*, pp 2301-2306.

[6] Andreasson, H. Triebel, R., Burgard, W., 2005. Improving Plane Extraction from 3D Data by Fusing Laser Data and Vision, *Intelligent Robots and Systems (IROS),* pp 2656-2661.

[7] Wang, T., McLauchlan, P. F., Palmer, P., Hilton, A., Wang, L., 2001. Calibration of an Integrated Measurement System of Camera and Laser and its Application, *The 5th World Multiconference on Systemics, Cybernetics and Informatic.*

[8] Kämpchen, D., 2007. Feature-Level Fusion of Laser Scanner and Video Data for Advanced Driver assistance *Phd dissertation.*

[9] Mühlbauer, Q., Kühnlenz K., Buss, M., 2008. Fusing Laser and Vision Data with a Generic ICP Algorithm, *Control, Automation, Robotics and Vision*. pp 1844-1849.

[10] H. Aliakbarpour, P. Núñez, J. Prado, K. Khoshhal and J. Dias 2009. An efficient algorithm for extrinsic calibration between a 3D laser range finder and a stereo camera for surveillance. *International Conference on Advanced Robotic.*

[11] Huang, L., Barth M., 2009. A novel multi-planar LIDAR and Computer Vision Calibration procedure using 2D patterns for Automated navigation, *Intelligent Vehicle Symposium* pp 117-122.

[12] Santolaria, J., Pastor, J. J., Brosed, F.J., Aguilar J. J. 2009. A one-step intrinsic and extrinsic calibration method for laser line scanner operation in coordinate measuring machines. 2009. *IOP Electronic Journals, Measurement Science and Technology*.

[13] Früh, C., Zakohor A., 2004. An Automated Method for Large-Scale, Ground-Based City Model Acquisition. *International Journal of Computer Vision 60(1), 5-24.*

[14] Rhee, S., Kim, T., Yun, D., Sung, J., 2007. Automated 3D road model generation by Fusion of circular-scanning Laser Data and CCD Images.