CLIPS Proceedings

Sebastian Tilch* and Rainer Mautz*
* Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.
Email: tilchs@ethz.ch

***Abstract*—This paper presents the latest developments related to an optical indoor positioning system named CLIPS, which consists of a projector and a mobile camera. With the projector, a pattern of laser spots is projected on any surface in an indoor environment. For the camera pose estimation in respect to the projector device, the camera has to capture the scene with the laser spots. The geometric relationship between the camera pose and the projector device can be established by the coplanarity constraint. This paper focuses on new findings regarding the real-time detection of laser spots in images and their unique identification. In addition, a novel method of system scale determination is presented.**

***Keywords* — Optical Indoor Positioning; Camera-based positioning, Camera Pose; Relative Orientation; CLIPS**

I. INTRODUCTION

The use of indoor positioning systems becomes more and more eminent in daily life. There is an increasing demand on positioning capabilities in applications such as healthcare, disaster management, metrology, logistics or housekeeping. A growing market in indoor positioning technologies appoints the research community as well as companies to satisfy this increasing demand. The successful launch of the IPIN (Indoor Positioning and Navigation) Conference in 2010 with its unprecedented number of presentations dedicated to indoor positioning techniques confirms this trend.

In addition to popular indoor positioning techniques such as those exploiting UWB- (Ultra-WideBand), WLAN- (Wireless Local Area Network) and ultrasound-signals, the alternative technique of optical indoor positioning can be found in areas of pedestrian navigation [2], [11], [14], [20], metrology [1], [4] and robot controlling [3], [4].

For a coarse classification of optical indoor positioning systems it is sufficient to distinguish between those systems that determine the pose of a mobile camera and those systems that determine the position of objects using a static camera. A more detailed classification of optical positioning systems is given by [13]. Thereby, the categories a) Reference from 3D Building Models, b) Reference from Images, c) Reference from Deployed Coded Targets, d) Reference from Projected Targets, e) Systems without Reference and f) Reference from other Sensors are distinguished. Other attempts for classification can be found in [7] or [19]. Despite of the differences among these systems, they have one common key problem to solve. The camera pose or the position of an object in three-dimensional space has to be determined whereas the only observation is a 2D measurement of image coordinates on a CCD or CMOS sensor.

This paper describes the optical indoor positioning system CLIPS (Camera and Laser-based Indoor Positioning System) which consists of a projector and a mobile camera. The reference information for the camera pose estimation is obtained by the projection of a laser spot pattern on any surface. The main focus of this paper is on novelties that have been implemented subsequent to our last report [18]. It is organized in three sections. The first section briefly repeats the basic concept of CLIPS. In the second section, the new research activities related to CLIPS are detailed. The third section is dedicated to a brief outlook and a conclusion.

II. BASIC CONCEPT

The basic concept of CLIPS is illustrated in Figure 1. In contrast to the photogrammetric relative camera orientation, CLIPS is based on the principle of an inverse camera, where one camera is substituted for a device that emits laser-beams from a virtual central point.

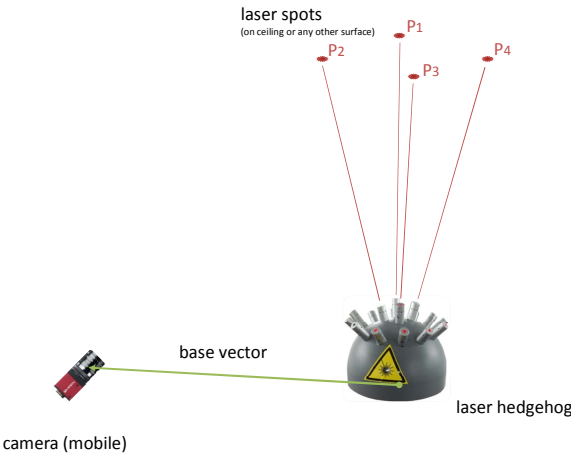

Figure 1. The camera pose can be estimated in respect to the laser hedgehog by capturing the projected pattern of red laser spots.

The laser beams project bright spots on any surface in an indoor environment. Through a one-time high precision calibration the 3D directions of these laser beams are known. Thereby, the laser device represents an inverse camera. The main functions of the projector can be summarized as, (a) the projection of flexible reference points and (b) the simulation of a second camera. When laser spots have been captured by a mobile camera, the relative orientation of the camera to the laser-device is

This work is funded by the Swiss National Fund (SNF).

determined by exploiting the concepts of stereo photogrammetry. As a major advantage no high-precision mechanics or sophisticated set-ups are required, making the system a low-cost, mobile and easy-to-use device for high precision positioning.

III. CURRENT RESEARCH ACTIVITIES

A. Point Detection

In a first step, the laser points need to be detected in the images. For this purpose a simple intensity threshold was applied, accepting that there is no adaption to varying illumination conditions from additional light sources or varying viewing angles

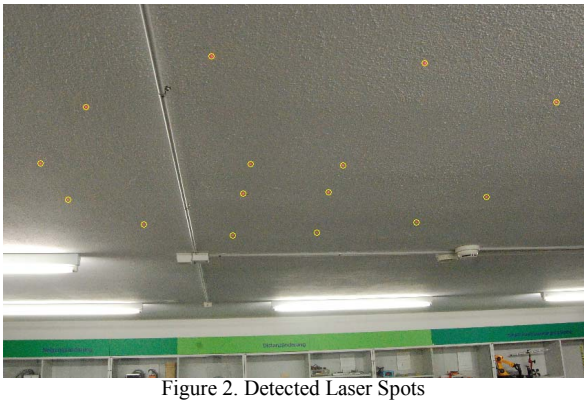

Figure 2. Detected Laser Spots

In order to solve the problem of disturbing light sources, we assumed that the saturation of natural and artificial light sources in the red and green channel is almost the same. Therefore, these unwanted light sources could be eliminated by taking the difference of both colour channels and applying an intensity threshold on the difference image. Furthermore, we choose a lower sensor exposure time to reduce the influence of ambient light. As a consequence, the laser spots can be more robustly detected at low computational cost.

B. Point Identification

In a second step, the detected laser spots have to be assigned to the corresponding laser beams of the hedgehog. This is a crucial task since the relative orientation algorithm requires at least five pairs of corresponding points to solve the coplanarity constraint.

The projected spot pattern consists of sixteen laser spots arranged on two concentric rings with four spots on the inner ring and twelve spots on the outer ring, as illustrated in Figure 4. The resulting symmetry is dissolved by an additional green laser spot that will also be used for scale determination as we will see later. For the purpose of point identification it is assumed that the inner four spots and at least four laser spots of the outer ring are visible in the image. Additionally, the detected spots of the outer ring must form a convex hull around the inner four spots. Under these assumptions, the following steps can be applied to uniquely identify all visible laser spots.

First, the algorithm identifies points on the outer ring. This means, that the algorithm determines all spots that belong to a convex hull. All spots on this hull are assigned to the outer hull and as a consequence, all points not being part of this hull can be regarded as inner points.

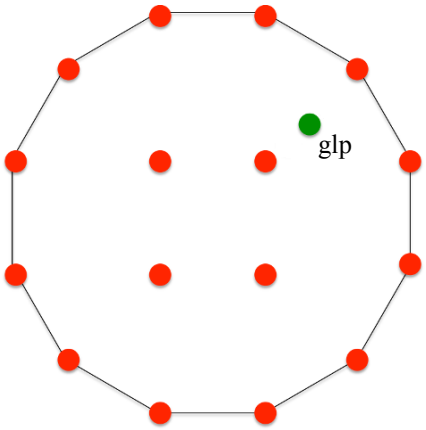

Figure 4. The projected laser spots pattern consists of two concentric rings and an additional green laser spot (marked with a 'glp') for scale introduction.

Second, the inner points are identified. For that purpose, the algorithm exploits the green laser spot as shown in Figure 4. The four inner points with the smallest distances to the green laser spot belong to the inner ring. One of these four spots is closest to the green spot compared to the other three spots. This laser spot is denoted with ID 3. The remaining three points are labeled clockwise with IDs 2, 1 and 4. Now, all spots on the inner ring are identified.

Third, the laser spots on the outer ring are identified. For this purpose, every spot of the outer ring is assigned

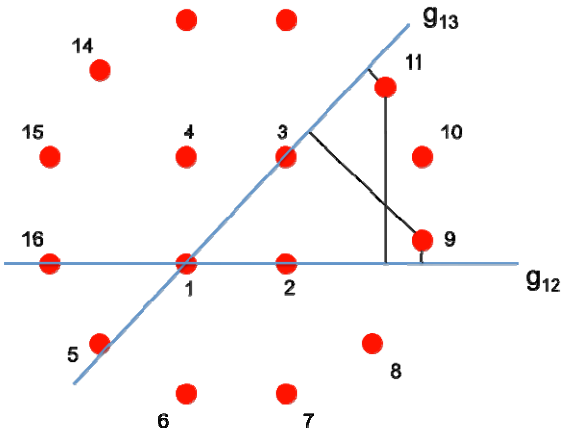

Figure 3. The fine identification queries the membership of a spot to a certain line defined by two of the four inner spots. This approach is based on a simple distance analysis.

to one of the six possible lines between the four spots of the inner circle. In Figure 3, two of these six lines are shown. According to Tilch [17], the distances between each spot $\mathbf{x}_i$ on the outer circle and each line g (that is defined by the vector $\mathbf{r}$ and an initial point $\mathbf{x}_{IP}$) can be calculated by

$$d_i = \|(\mathbf{E} - \mathbf{n}\mathbf{n}^{\mathrm{T}})(\mathbf{x}_i - \mathbf{x}_{IP})\|, \quad (1)$$

where

$$\mathbf{n} = \frac{\mathbf{r}}{\|\mathbf{r}\|}. \quad (2)$$

In our case, the initial point $\mathbf{x}_{\mathrm{IP}}$ is a spot of the inner circle and the vector $\mathbf{r}$ parallel to the direction of the line. Matrix $\mathbf{E}$ in (1) denotes a 2×2 identity matrix and vector $\mathbf{n}$ is the unity vector of the line direction.

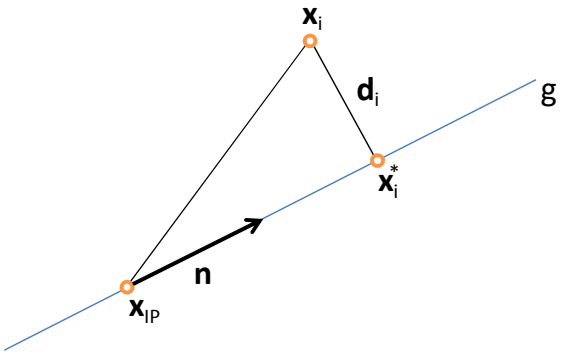

Figure 6. Illustration how the distance of a laser spot x_i to a line is found.

To better retrace (1) we assume that line

$$\mathrm{g}: \mathbf{x} = \mathbf{x}_{\mathbf{IP}} + \lambda * \mathbf{n} \quad (3)$$

is defined by point $\mathbf{x}_{\mathrm{IP}}$ on the line and direction vector $\mathbf{n}$. The distance vector of point $\mathbf{x}_{\mathrm{i}}$ to its projection $\mathbf{x}_{\mathrm{i}}^{*}$ on the line g can be expressed by

$$\mathbf{d}_{\mathrm{i}} = \mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{i}}^{*}. \quad (4)$$

In (4), the projection $\mathbf{x}_{\mathrm{i}}^{*}$ can be replaced by the term $\mathbf{x}$ in (3). The result

$$\mathbf{d}_{\mathrm{i}} = \mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{IP}} - \lambda * \mathbf{n} \quad (5)$$

contains the variable

$$\lambda = \mathbf{n}^{\mathbf{T}}(\mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{IP}}) \quad (6)$$

that corresponds to the distance between point $\mathbf{x}_{\mathrm{IP}}$ and the projected point $\mathbf{x}_{\mathrm{i}}^{*}$. The insertion of λ according to (6) into (5) leads to

$$\mathbf{d}_{\mathbf{i}} = \mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{IP}} - \mathbf{n}^{\mathbf{T}}(\mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{IP}}) * \mathbf{n} \quad (7)$$

that can be rewritten to the afore mentioned form

$$\mathbf{d}_{\mathbf{i}} = (\mathbf{E} - \mathbf{n}\mathbf{n}^{\mathbf{T}})(\mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{IP}}). \quad (8)$$

The distance d_i of point $\mathbf{x}_{\mathbf{i}}$ to its projection $\mathbf{x}_{\mathbf{i}}^{*}$ on line g can be determined by adding the norm as in (1). This equation can be regarded as a closed form to determine the distance of a point to a line, only by inserting the direction vector $\mathbf{n}$, the coordinates of point $\mathbf{x}_{\mathbf{i}}$ and the point on the line $\mathbf{x}_{\mathrm{IP}}$. Originally, this equation was derived for the intersection of lines in three-dimensional space. More details can be found in Tilch [17].

The distances of all spots can be stored in a distance matrix $\mathbf{D}$, where each row contains the distances of all spots to one line and each column contains the distances of a spot to all lines. Now, the minimum entry of a column represents the minimal distance of a spot to its nearest line. This way, to each of these six lines two points of the outer ring can be related. In order to solve this ambiguity, the lines are split at the center into two rays (half lines). Based on the rays, the ambiguity which spot is located on which side can be solved easily. Finally, all visible spots are labeled with their ID.

The benefits of this approach are twofold. Firstly, the real-time identification of all laser spots can be achieved from a single image. Secondly, just a fraction of the laser spots must be visible for their identification. The only precondition is that at least four spots on the outer ring are visible in the image and that they form a convex hull that fully encloses the inner ring.

C. Camera Pose Estimation

The camera pose can be estimated by solving the coplanarity constraint

$$\mathbf{\Delta} = \mathbf{d}_{\mathrm{RLP}}{}^{\mathrm{T}}(\mathbf{b}_{\mathrm{Cam}} \times \mathbf{d}_{\mathrm{Cam}}) = 0 \quad (9)$$

which can be rewritten as the determinant

$$\Delta = \begin{vmatrix} 1 & \mathrm{d_{RLP.x}} & \mathrm{d_{C.x}} \\ \mathrm{b^{*}_{C.y}} & \mathrm{d_{RLP.y}} & \mathrm{d_{C.y}} \\ \mathrm{b^{*}_{C.z}} & \mathrm{d_{RLP.z}} & \mathrm{d_{C.z}} \end{vmatrix} = 0. \quad (10)$$

This constraint describes the geometrical relationship between two convergent images, where the projection

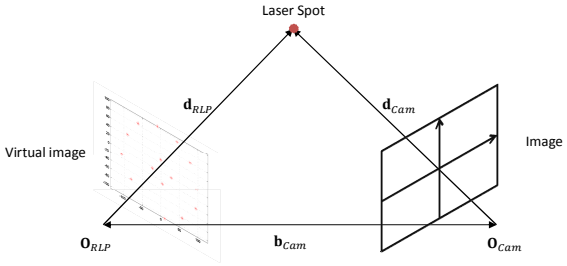

Figure 5. Epipolar Geometry (see Luhmann [12])

centres $\mathbf{O}_{RLP}$, $\mathbf{O}_{Cam}$ of the images and an object point define a so called epipolar plane, see Figure 5.

The vector $\mathbf{b}_{Cam}$ denotes the vector between the projection centre of the projector $\mathbf{O}_{RLP}$ and the projection centre $\mathbf{O}_{Cam}$ of the camera. The vector

$$\mathbf{d}_{Cam} = \mathbf{R}\mathbf{x}_{Cam} \tag{11}$$

represents the image projection vector $\mathbf{x}_{Cam}$ of a laser spot in the image rotated into the object coordinate system by the rotation matrix

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}. \tag{12}$$

An explicit expression of the camera rotation angles is given by

$$\omega = \arctan\left(\frac{-r_{23}}{r_{33}}\right), \tag{13}$$

$$\varphi = \arctan\left(\frac{r_{13}^{\,2}}{\sqrt{r_{11}^{\,2} + r_{12}^{\,2}}}\right) \text{ and}$$

$$\kappa = \arctan\left(\frac{-r_{12}}{r_{11}}\right).$$

The task of solving the coplanarity constraints leads to a nonlinear estimation problem and therefore requires good approximate values for the Relative Orientation (RO) parameters $\mathbf{b}_{Cam}$, ω, φ and κ in the linear model

$$v_\Delta = \frac{\partial\Delta}{\partial b_{C.y}} db_{C.y} + \frac{\partial\Delta}{\partial b_{C.z}} db_{C.z} + \frac{\partial\Delta}{\partial\omega} d\omega + \frac{\partial\Delta}{\partial\varphi} d\varphi + \frac{\partial\Delta}{\partial\kappa} d\kappa \tag{14}$$

of equation (9), see [12]. The initial parameters are improved by an iterative least-squares adjustment. For this approach an answer has to be found of how approximate values can be obtained.

One way is the exact algebraic solution of a set of polynomials that follow from the coplanarity constraints. Either the Gauss-Jordan elimination [15] or Gröbner bases [5], [10], [16] can be used to estimate the roots of the polynomials. Unfortunately, the solution is not unique. An RANSAC [8] framework can be used to determine the correct solution. An implementation of a 5-Point solver [16] for the CLIPS project was presented at the ISPRS Midterm Symposium in 2010 [18].

Another way of approximate value generation was proposed by Cronk et al. [9]. Hereby, the location of a second camera is simulated on a tessellated sphere around the first camera station. Starting from arbitrary initial values of the orientation angles (usually chosen to be zero) every set of initial values is improved by a least squares adjustment. The right solution can be found by considering the RMSE and the fact that the objects are always located in front of the projector and the camera. This can be easily done by the calculation of the laser spot

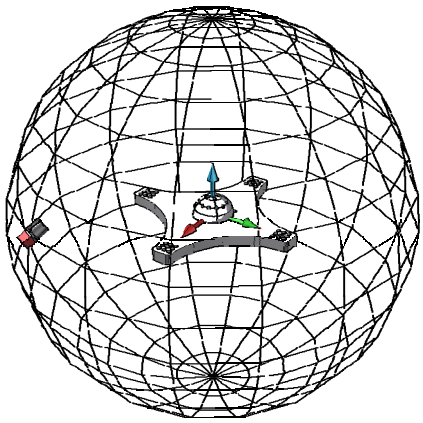

Figure 7. Approximate values are generated by the camera pose simulation on a tessellated sphere [9]

coordinates in object space via intersection of corresponding rays.

In order to spare the need for a RANSAC embedment we have chosen the approach by Cronk et al. Under the assumption that the camera pose of two consecutive frames only slightly changes we can use the refined RO parameters of the previous image as initial values for the RO determination of the subsequent image.

D. Introduction of the System Scale

A common approach is the scale estimation by extracting reference information in the image. For this purpose, we have added green laser pointers in an eccentric position to the laser hedgehog. Subsequent to the estimation of the relative orientation parameters, the system scale can be determined as shown in Figure 8. Since the direction $\mathbf{d}_{GLP}$ of the additional laser pointer and its base line $\mathbf{b}_{GLP}$ to the laser rig are known from a one-time calibration, the system scale becomes estimable.

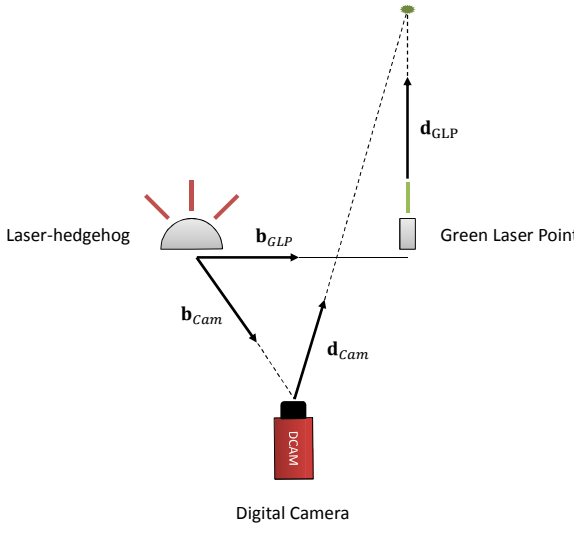

Figure 8. Introduction of the system scale

Please note that all vectors in Figure 8 are unit vectors. The factor b denotes the baseline's length between the laser-device and the additional green laser pointer.

An expression for the scale can be derived by intersecting plane Π spanned by the vectors $\mathbf{d}_{\mathrm{GLP}}$ and $\mathbf{d}_{\mathrm{Cam}}$ with the straight line l defined by the base vector $\mathbf{b}_{\mathrm{Cam}}$. The plane

$$\Pi\colon (\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{x} - b \cdot (\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{b}_{\mathrm{GLP}} = 0 \quad (15)$$

can be expressed by the Hessian normal form where the cross product $(\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})$ represents the normal vector and the base vector $\mathbf{b}_{\mathrm{GLP}}$ represents a vector to an initial point of that plane. The straight line

$$l\colon s \cdot \mathbf{b}_{\mathrm{Cam}} - \mathbf{x} = \mathbf{0} \quad (16)$$

is defined by the direction of vector $\mathbf{b}_{\mathrm{Cam}}$ and the hedgehog projection centre as one point.

Solving (16) to vector **x** and inserting the result in (15) leads to

$$s \cdot (\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{b}_{\mathrm{Cam}} - b \cdot (\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{b}_{\mathrm{GLP}} = 0. \quad (17)$$

After solving (17) for s we obtain

$$s = b \cdot \frac{(\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{b}_{\mathrm{GLP}}}{(\mathbf{d}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{Cam}})\mathbf{b}_{\mathrm{Cam}}}. \quad (18)$$

The system scale can be introduced by scaling the vector $\mathbf{b}_{\mathrm{Cam}}$ with the factor s.

Geometrically Unstable Cases

First tests have shown that the concept of using a single laser pointer for scale determination is geometrically unstable for some cases as shown in Figure 10. If the camera is located in the plane spanned by base vector $\mathbf{b}_{\mathrm{GLP}}$ and base vector $\mathbf{d}_{\mathrm{GLP}}$ then the plane Π contains line l and therefore they do not intersect. The scale determination is ill-posed.

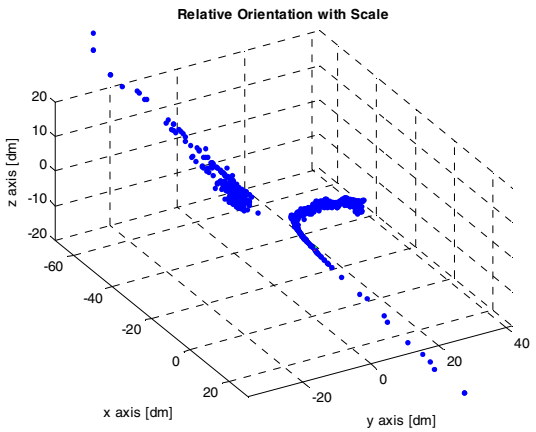

Figure 10. Example of a camera track revealing the asymptotic behaviour at a geometric unstable constellation

In order to avoid such a geometric constellation, a special hedgehog mount was constructed allowing us to aim 36 green lasers from the four main directions as shown in Figure 9.

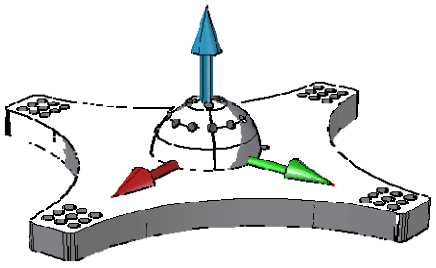

Figure 9. Introduction of the system scale

This means, that the scale can be estimated 36 times independently.

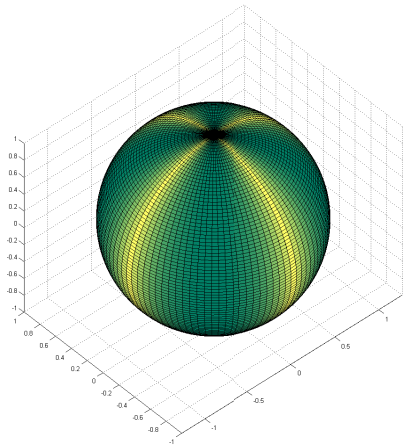

Figure 11. Contribution of each scale arm to the determination of the system scale

The sphere in Figure 11 illustrates the contribution of the scale arms to the system scale for each camera position around the laser-hedgehog. Whereas green zones indicate camera positions with good contributions of all arms to the system scale, the yellow zones reveal contributions of only two scale-cross arms. The black zones at the poles show camera positions around the laser-hedgehog where a scale determination is impossible. The contribution of each scale arm can be expressed by the SQM (Scale Quality Measure) which corresponds to a geometric measure and can be regarded as a weight to build the weighted average of the system scale.

Scale Quality Measure

In order to determine the stability of the scale determination, a scale quality measure

$$\mathrm{SQM} = \frac{\|(\mathbf{b}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{GLP}}) \times (\mathbf{b}_{\mathrm{Cam}} \times \mathbf{d}_{\mathrm{GLP}})\|}{\|\mathbf{b}_{\mathrm{GLP}} \times \mathbf{d}_{\mathrm{GLP}}\| \cdot \|\mathbf{b}_{\mathrm{Cam}} \times \mathbf{d}_{\mathrm{GLP}}\|} \quad (19)$$

is derived for each green laser pointer. SQM is a measure for the angle of incidence of the two planes defined by the vectors $\mathbf{b}_{Cam}$, $\mathbf{b}_{GLP}$ and $\mathbf{d}_{GLP}$. If these two planes are parallel then SQM approaches zero since the cross product of almost two collinear vectors decreases to zero. In our case the rejection criterion was chosen to be a threshold of SQM = 0.38. If the SQM is smaller than this threshold then the scale is rejected. In other words, a suitable scale determination can be obtained if the intersection angle of these two planes is greater than 22.5 degrees.

IV. Conclusions and Outlook

In contrast to the previous version, the updated CLIPS architecture enables real-time positioning and navigation on the level of a functional model. Now the laser spot identification and the introduction of the system scale are solved image by image. Furthermore, the capture of the whole point pattern is not a requirement anymore. These measures have been an essential leap forward towards building a practicable system.

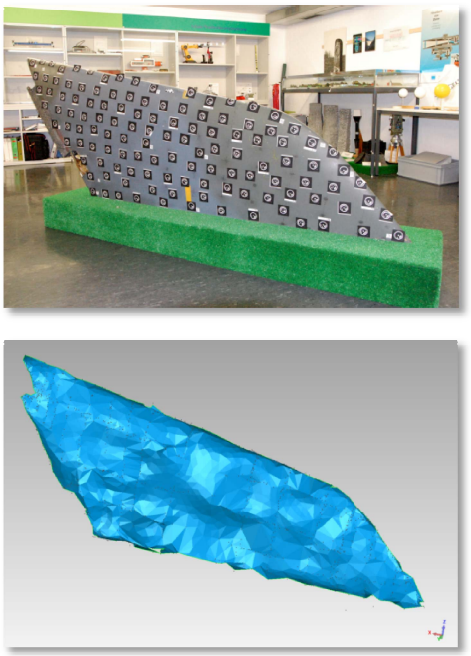

Figure 12: Measured test object (tail plane) and correspondent triangulated surface of this object

For a first test, a tail plane has been measured and reconstructed, as shown in Figure 12. In comparison to the test object the reconstruction models the shape and size of the object but it reveals a bumpy surface due to measurement uncertainty.

Future steps include the improvement and the investigation of the achievable accuracy and performance of CLIPS and the evaluation of possible applications.

V. References

[1] AICON 3D Systems (2010): http://www.aicon.de, last accessed December 2010.

[2] Aufderheide, D., & Krybus, W. (2010): Towards real-time camera egomotion estimation and three-dimensional scene acquisition from monocular image streams, (S. 1-10).

[3] Boochs, F., Schutze, R., Simon, C., Marzani, F., Wirth, H., & Meier, J. (2010): Increasing the accuracy of untaught robot positions by means of a multi-camera system, (S. 1-9).

[4] Breuckmann naviSCAN 3D. (2011): *http://www.breuckmann.com/industrie-technik/produkte/naviscan.htm*, last accessed March 2011.

[5] Buchberger, B. (2001): Gröbner Bases: A Short Introduction for Systems Theorists, http://www.risc.jku.at/publications/download/risc_323/2001-02-19-A.pdf, http://www.risc.uni-linz.ac.at/publications/download/risc_323/2001-02-19-A.pdf.last accessed March 2011.

[6] Creaform Products. (2011): *http://www.creaform3d.com/en/default.aspx*, last accessed March 2011.

[7] DeSouza, G. and Kak, A. (2002): "Vision for Mobile Robot Navigation: A Survey", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 2, pp. 237-267.

[8] Fischler, M. A., & Bolles, R. C. (1981): Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM, 24*, 381-395.

[9] Cronk, S., Fraser, C., Hanley, H.. (2006): Automated metric calibration of colour digital cameras. *The Photogrammetric Record, 21*, 355-372.

[10] Kalantari, M., Jung, F., Guedon, J.-P., & Paparoditis, N. (2009): The Five Points Pose Problem: A New and Accurate Solution Adapted to Any Geometric Configuration. In T. Wada, F. Huang, & S. Lin (Hrsg.), *Advances in Image and Video Technology* (Bd. 5414, S. 215-226). Springer Berlin / Heidelberg.

[11] Klopschitz, M., Schall, G., Schmalstieg, D., & Reitmayr, G. (2010): Visual tracking for Augmented Reality., (S. 1-4).

[12] Luhmann, T (2010): Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen, 3rd Edition, Wichmann Verlag, Berlin.

[13] Mautz, R. and Tilch, S. (2010): "Innenraumpositionierung mit optischen Methoden", AVN, no. 7, pp. 250-255.

[14] Mulloni, A., Wagner, D., Barakonyi, I., & Schmalstieg, D. (2009): Indoor Positioning and Navigation with Camera Phones. *Pervasive Computing, IEEE, 8*(2), 22-31.

[15] Nister, D. (2004): "An efficient solution to the five-point relative pose problem," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.26, no.6, pp.756-770

[16] Stewénius, H., Engels, C., Nistér, D. (2006): Recent developments on direct relative orientation, ISPRS *Journal of Photogrammetry and Remote Sensing, ISPRS, vol. 60,* 284-294

[17] Tilch, S. (2009): Calibration of CLIPS (German). Student Project Work, ETH Zurich, unpublished.

[18] Tilch, S., Mautz, R. (2010): Development of a new laser-based, optical indoor positioning system, ISPRS Com. V Mid-Term Symposium, no. 98, pp. 575-580.

[19] Trucco, E. and Plakas, K. (2006): "Video Tracking: A Concise Survey", IEEE Journal of Oceanic Engineering, vol. 31, issue 2, pp. 520-529, April 2006.

[20] Willert, V. (2010): Optical Indoor Positioning using a camera phone, Abstract Volume, IPIN 2010, pp. 289-290