Side-Scan Sonar Based SLAM for the Deep Sea

Philipp Woock

Vision and Fusion Laboratory
Institute for Anthropomatics
Karlsruhe Institute of Technology (KIT), Germany
woock@ies.uni-karlsruhe.de

Technical Report IES-20xx-xx

Abstract: In order to robustly perform SLAM (Simultaneous Localization and Mapping), places need to be recognized when they are visited again. In the deep-sea environment SLAM-assisted navigation based on side-scan sonar data benefits from using three-dimensional features of the environment as they are much less view-dependent than classic 2D features. Obtaining these features requires processing of the sonar data as the side-scan sonar sensor readings contain three dimensional information only indirectly. To extract that information the ensonification process needs to be inverted. This inversion is an ill-posed inverse problem and therefore regularization is needed before a unique solution can be found. Once the true seabed shape is reconstructed, wide area SLAM techniques can be applied.

1 Introduction and Existing Work

The process of robustly navigating an autonomous underwater vehicle (AUV) through unknown terrain using side-scan sonar data and an inertial measurement unit (IMU) is a multi-layered process. The first stage involves transforming the side-scan sonar data to a 3D representation of the environment. Afterwards, significant features of the environment have to be identified. Using the extracted features and the data of the inertial sensors a SLAM method can be employed.

The paper is structured as follows. First, related work to underwater SLAM is presented. In Section 2 the challenge in side-scan sonar reconstruction is explained and it is shown how data is preprocessed for the regularization steps (Section 3.1) where a state-of-the-art method to tackle the problem is explained. Section 4 shows a hybrid SLAM concept that is suitable for deep-sea applications and gives an outlook to further developments. In Section 5 a short peek of the used hardware is given.

1.1 Loop-closing

An essential task of every SLAM system is the closing of trajectory loops, e.g., detecting that exactly the same spot has been visited before. Johnson-Roberson et al. [JPWM10] perform loop closing in a shallow water environment with a stereo camera. With the great amount of information obtained by high resolution color camera images, the resulting loop closures are very reliable and accurate.

For a similar approach in the deep-sea it would be necessary to artificially illuminate the environment. However, turbid water and energy constraints make classic lighting impossible. In the future, gated viewing solutions using a very short laser pulse and precise camera shutter timing may enable image-assisted loop-closing also in the deep sea.

It is difficult to apply this technique to side-scan sonar data only as the side-scan sonar provides gray level image lines which contain much less information compared to high-resolution 2D color images.

1.2 Dead-reckoning

In order to build a robust SLAM solution one has to deal with cases where the AUV is unable to detect features in the environment either because of sensor error or simply because the environment is lacking significant features. Without environmental features the dead-reckoning navigation of the vehicle is the only source of ego-motion estimation and therefore has to be calibrated carefully.

Jakuba et al. show in [JPW10] that with sophisticated calibration methods very small navigation errors can be achieved. To correctly calibrate the compass they performed star-shaped test dives. The trajectory obtained through camera-based visual SLAM served as ground truth. That way they could calibrate the compass readings according to that ground truth. With their calibration they can bridge several minutes without observing external features. They also investigated the influences of the tides on the depth measurements.

1.3 SLAM in Man-made Environments

SLAM for underwater man-made environments (e.g., a marina) using sonar sensors was investigated in [RRN08]. They use a mechanically scanning imaging sonar sensor (MSIS) that is rotating while taking measurements. That way the vehicle repeatedly gains a full 360° view of the environment. Depending on the sensor settings it takes only several seconds for a full 360° scan. Due to the sensor principle the measurements are carried out in a polar coordinate frame. Applying the Hough-Transform to the data, they are able to detect lines that correspond to the man-made walls. Eventually, their environment map in which the vehicle localizes itself and navigates consists of these lines.

The approach was refined to a scan matching approach in [MRRH10] that does not rely on line-shaped structures any more.

2 Data preprocessing

The main difficulty in performing SLAM on side-scan sonar data is the ambiguity in the sensor readings itself: the side-scan sonar records an echo amplitude over time, i.e., with no spatial information *where* that echo came from but only *when* it arrived (two-way travel time). Furthermore, a certain echo tells only whether there is a reflector or not and how strong the reflectance is. Unfortunately, the reflectance strength is only partially linked to the reflector's geometry. Amongst other things the echo strength is also dependent on the sediment material, water absorption and the grazing angle (denoted β in Figure 2.3). More information about this problem and related research in that field is given in [WF10].

Before the three-dimensional reconstruction of the environment based on the side-scan sonar data can take place the sonar data has to be pre-processed: The first bottom return has to be found and the sonar data is mapped to ground coordinates (slant range correction).

2.1 First bottom return

Before the side-scan sonar pulses reach the ground they propagate through the water beneath the AUV which results in a silent period after the sending peak at the beginning of each line as water does not produce an echo. After the silent period the first echo is received. The time until the first echo arrives is the two-way travel time of the sonar signal and thereby gives information about the distance to the nearest reflecting object which is assumed to be the seabed perpendicularly below the vehicle. That assumption does not always hold as can be seen in Figure 2.1. This fact is mostly neglected as the seafloor shape can often be approximated by a plane and the error is usually not too large. Hence, this first echo is called the First Bottom Return (FBR) and its detection can be seen as an additional altitude sensor which can be combined with altitude measurements of another sensor.

Before the detection of the FBR the sonar line is filtered by a median filter to reduce the influence of speckle noise while preserving discontinuities for a more robust

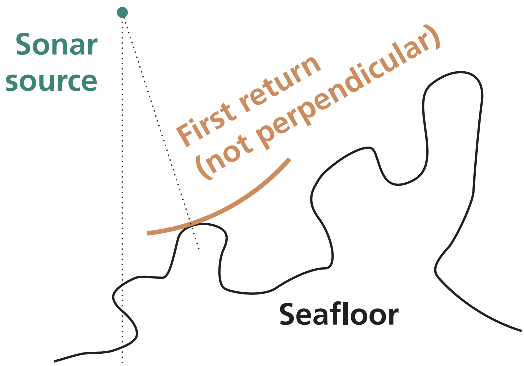


Figure 2.1: The FBR is not necessarily an echo from perpendicularly below the vehicle and therefore not always equal to the altitude of the vehicle. However, in a deep-sea environment it is assumed to be the case most of the time.

detection. Speckle noise stems from the side-scan sonar being a coherent recording method. Rank value filters, for example a median filter, are very effective against this type of noise [RVRV95].

In Fig 2.2 the FBR is detected as the first occurrence of two adjacent sonar samples that are more than 1.7 standard deviations apart from the mean value of the sonar line. The threshold value is determined empirically for every sensor configuration in advance. However, an automatic derivation and adaptive behavior of the threshold is possible and will be added in the future.

The FBR detection applied is strictly per-line. Under the assumption that the first return always stems from perpendicularly beneath and that the surface geometry is not changing too quickly one could also apply Kalman filtering in time to smooth the detection and limit the influence of outliers as occasionally a stronger echo that arrives later could be mistaken for the FBR.

2.2 Slant Range Correction

In order to obtain ground coordinates (e.g., to perform the aforementioned regularization) the sonar image has to be slant range corrected when the sonar altitude is known either from detection of the FBR or from additional sensor readings. The geometric configuration is depicted in Figure 2.3. The original data can be thought of lying on the r_s ray (brown). The data of the water column (before the FBR) is mapped to the height h and the part starting at the FBR until the end is mapped to the ground r_g (green).

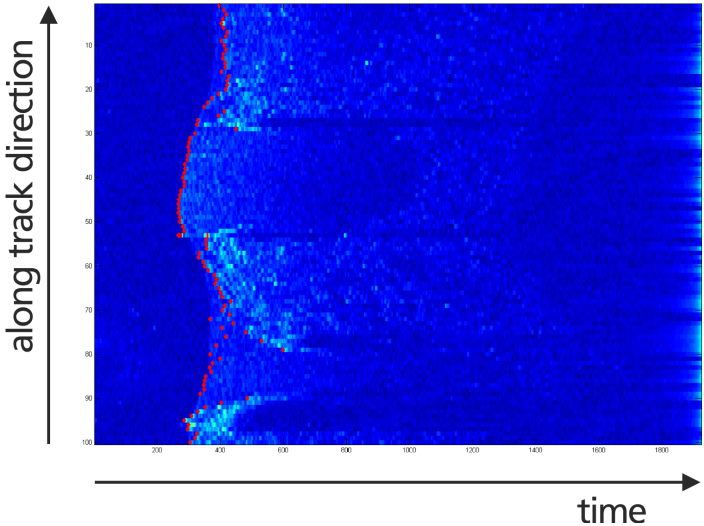


Figure 2.2: Per-line detection of the FBR (red) on side-scan sonar data. The plot shows the echo intensity over time.

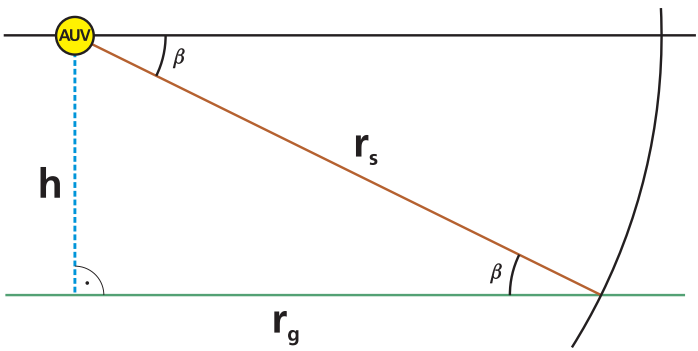
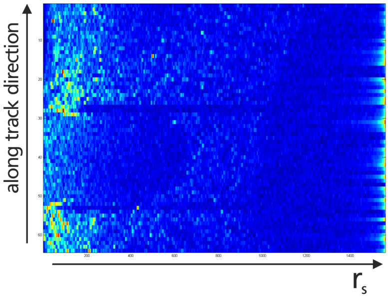
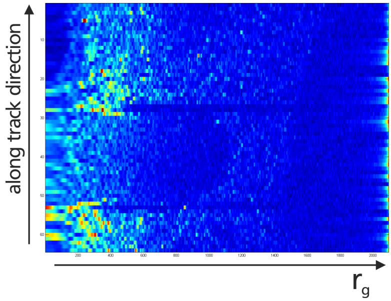


Figure 2.3: Geometry of slant range correction



(a) Before slant range correction.



(b) After slant range correction.

Figure 2.4: Sonar image slant range correction. The part containing the water column up to the FBR has been removed.

This correction is a non-linear distortion and its effect on a sonar image is depicted in Figure 2.4 where the signal part from the FBR to the end are shown. The strongest effect is seen for the near-nadir parts of the sonar image (left hand side in Figure 2.4b)). For longer ranges the difference between slant range and ground range diminishes as the grazing angle β is decreasing.

All sonar lines in the image are corrected with the same depth value, derived from the median of the FBR detections in the whole observation window, assuming constant altitude and a flat seabed. This is necessary to apply the regularization process of [CPL07] which is detailed in Section 3.1.

$$r_g = \sqrt{r_s^2 - h^2}$$
$$\cos\beta = \frac{r_g}{r_s}$$

In Figure 2.3 and in the above equation r_s denotes the slant range, r_g denotes the ground range and h the sensor altitude.

3 Estimating Seabed Shape from Side-Scan Sonar Data

To obtain the seabed shape that created a particular echo one has to *invert* the ensonification process. As it is known how sonic waves propagate in water and how they are reflected on objects, inverting that process means estimating what kind of shape may be responsible for a certain echo. This is done through a so-called forward model that performs a simulation of the ensonifying process and the echo generation.

However, there is a multitude of environmental parameters that alter a sonic echo, e.g., sediment type, water temperature, salinity, sonar beam form, etc. It is not feasible to incorporate all of them into such a forward model as there are only few measurements and infinitely many possibilities to adjust parameters. Besides, computation time restrictions call for a simple but sufficiently accurate model.

An overview over different sonar simulation models is given in [CPL07].

3.1 Iterative Optimization

Incorporation of a priori knowledge or assumptions about the seabed to make the problem invertible are called regularization. Coiras et al. proposed a powerful method for the inversion process, where they investigated different regularization strategies ([CPL07] and [CG09]).

Their inversion method is an iterative approach, the basic principle is shown in Figure 3.1. Initially, in [CPL07] a flat seabed is assumed and the sonar data is slant range corrected (see Section 2.2) before the estimation starts. Within an Expectation-Maximization (EM) framework they optimize the seabed shape to be close to the truth. With a forward model they simulate an echo response for a given seabed shape and compare it to the measurements taken. The difference between simulation and measurement indicates whether the model represents reality already correctly and where it needs further adaptation. Then, parameters are adjusted accordingly. Then, the simulation with the forward model is done with the updated parameter set and the difference is examined again. This is repeated until the difference between the simulated echo and the measured echo falls below an error threshold or until the model is unable to diminish the difference further.

3.2 Regularization techniques

In this iterative cycle the regularization is performed after each iteration step: For example in [CPL07] it is assumed that the sensor's beam form is constant over time. That means that every surface patch pointing to the sensor at a certain angle is subject to the same beam form model parameter even across sonar scan lines.

In [Woo10] it was illustrated that no information about the surface shape is obtained in areas that are not reached by the sonic waves (so-called sonar shadows). It is assumed that in such areas the seabed does not behave differently than in the directly adjacent areas. It is therefore plausible to assume the model parameters describing the reflectance properties of the sediment are similar both in the shadowed areas and the surrounding non-shadowed areas.

Additionally, as another regularization restriction it could be assumed that sediment composition in general changes rather slowly and not with each sonar sample which corresponds to smoothing the reflectance map.

It is also possible to introduce smoothness constraints to be imposed on the reconstructed surface. The rationale is that between two scan lines and two sonar samples the changes of the surface height are assumed to be minor (i.e., no needle-shaped seafloor). Thus, the connection between the surface points should also be continuous and may have matching derivatives (first and higher derivatives, depending on the level of intended smoothing). However, this method is prone to smooth out interesting surface details. Coiras et al. try to avoid this effect using different surface constraints, e.g., based on a deformable mesh with forces between the mesh points [CG09].

Further, Coiras et al. achieved much better reconstruction results with a hierarchical approach using an image pyramid. That way, the coarser surface parts are reconstructed first while the details are recovered later in the stages with finer resolution. This helps to avoid local minima in the optimization process. In addition, the subsampling reduces the image noise which helps the optimization process to converge quicker in the low resolution stages.

Using synthetic aperture sonar (SAS) data the method can be modified [CG09] to work in polar coordinates. With SAS it is not necessary to estimate the beam form of the sonar sensor in contrast to the side-scan approach [CPL07] as it can be assumed as being uniform. In the procedure for SAS data the surface is modeled as a deformable elastic mesh where forces between the grid points make sure that smoothness requirements are met.

The method in [CPL07] is related to [LH91] who worked their way from nadir towards the peripheral areas by summing up differential surface patches. A detailed description of this Propagation Shape-from-Shading and a comparison to a Fourier-based approach is done by [DBL04].

3.3 Extension

In order to perform SLAM with the reconstructed seabed surface, the mentioned inversion methods need to be extended so that curved trajectories and varying AUV altitudes may be also considered as AUVs drive turns and are unable to perfectly

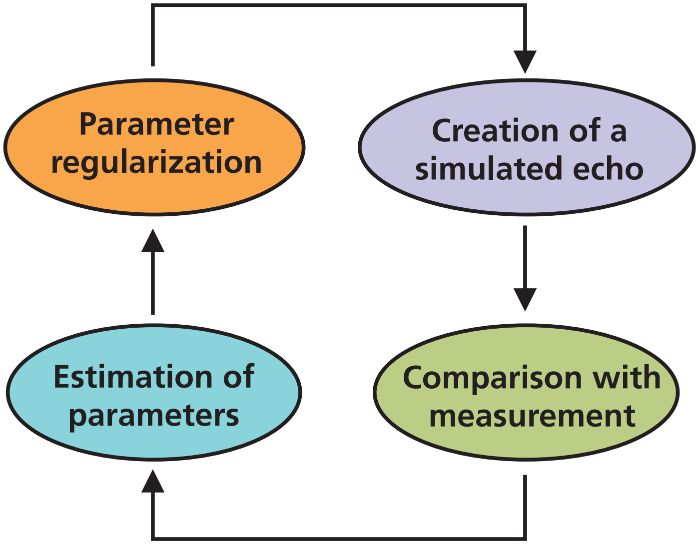


Figure 3.1: Iterative estimation process of the seabed surface shape.

maintain a constant flying altitude. With an ego-motion estimation based on inertial sensors the reconstruction may be modified accordingly to obtain an undistorted map.

4 SLAM concept

4.1 3D features

SLAM approaches that use higher-level landmarks opposed to plain point cloud data feature a more robust recognition and distinction. Spline features ([PDM$^+$07] [PRLM$^+$09]) are an example for such landmarks. They are more robust because formulated as a spline one landmark comprises multiple points and as a result small variations of the input points in general do not alter the overall shape of the spline much. Three-dimensional landmarks may be compiled of several one-dimensional splines in space that have one point in common or may even consist of full spline surface patches.

A comprehensive overview of other features that may be used for three-dimensional SLAM applications is given in [Run10].

4.2 Hybrid SLAM Concept

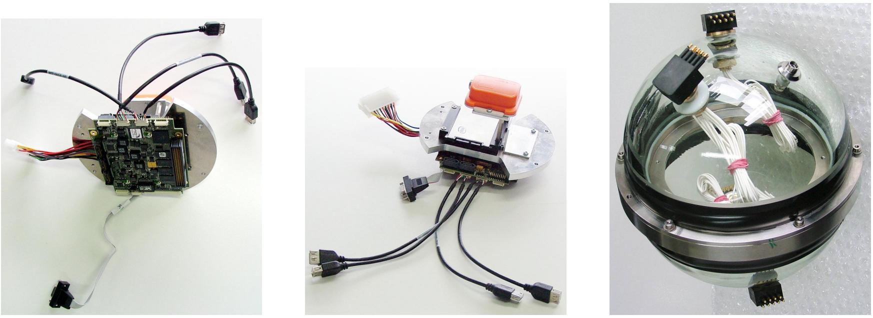
In order to achieve long underwater mission durations, the used SLAM algorithm needs to be carefully chosen. Furthermore, it is important to balance computing requirements and the stability of the SLAM solution.

In [BB08] Brooks and Bailey propose a combined EKF/FastSLAM framework termed HybridSLAM. It tries to retain the best of both SLAM approaches while mitigating their respective weaknesses. FastSLAM works by approximating the true probability distribution of the vehicle paths and maps by particles and can therefore inherently track multiple hypotheses about the path taken and the respective map of the environment. However, the approach suffers from so-called *particle depletion* which means that the longer the filter runs the more certain it becomes about the beginning of the path trajectory and the oldest parts of the map. It more and more ‘forgets’ about other possibilities in the past. On the other hand, the EKF can only maintain a single hypothesis and wrong associations between measurements and landmarks are ‘remembered’ forever and may eventually lead to divergence of the filter. The idea is to use a FastSLAM front-end and an EKF back-end. The rationale behind the concept is that for the most recent measurements the FastSLAM algorithm can deal better with wrong associations by maintaining multiple path and map hypotheses. The parts of the map about which the FastSLAM filter has become ‘more confident’ over time are placed inside an EKF as submaps. Until then the risk of wrong submap associations has become very low and the EKF will most probably remain stable.

Other submapping approaches that tackle the same problem are Fairfield’s SegSLAM [Fai09] and Bosse’s ICP-based map matching [BZ08].

5 Hardware

The computing devices are placed inside a glass sphere that can withstand water pressure up to 10000 m. The main difficulty is to integrate a powerful processing unit while at the same time managing heat transfer and placing all components in the small volume of the sphere. For the on-board sonar data processing, feature extraction and the SLAM algorithm we will use an embedded PCI/104 board using an Intel Core2Duo$^{\text{TM}}$ SP9300 processor. The inertial measurements are provided by an Xsens MTi. Data logging is done with an Intel 1,8″ Postville SSD as SSD technology provides high speed and shock resistance. The hardware is pictured in Figure 5.1. Not shown are the two side-scan sonar sensors that are connected via USB.



(a) View from below: PCI/104-Module with heat spreader baseplate.

(b) Top view: inertial measurement unit and SSD data storage.

(c) Pressure hull made of glass resistant to 10000 m water depth.

Figure 5.1: Hardware of the sensor data processing module.

The cooling solution consists of an aluminum baseplate that serves as heat spreader to conduct the heat to the titan flange of the sphere which in turn is cooled by the surrounding seawater. All metal-metal connections have thermal grease applied to feature better thermal conductance.

6 Conclusion

The systematic process of creating an underwater SLAM solution that is suitable for deep-sea applications has been shown using inertial measurements and side-scan sonar. The challenge to invert the side-scan sonar image creation into a 3D representation has been discussed. A suitable SLAM architecture for long deep-sea missions and the chosen hardware layout have been outlined.

Bibliography

[BB08] Alex Brooks and Tim Bailey. HybridSLAM: Combining FastSLAM and EKF-SLAM for reliable mapping. In *WAFR 2008: The Eighth International Workshop on the Algorithmic Foundations of Robotics*, 12 2008.

[BZ08] Michael Bosse and Robert Zlot. Map matching and data association for large-scale two-dimensional laser scan-based slam. *I. J. Robotic Res.*, 27(6):667–691, 2008.

[CG09] Enrique Coiras and Johannes Groen. 3D target shape from SAS images based on a deformable mesh. In *Proceedings of the 3rd International Conference on Underwater Acoustic Measurements (UAM), Nafplion, Greece*, 2009.

[CPL07] Enrique Coiras, Yvan Petillot, and David M. Lane. Multiresolution 3-D Reconstruction From Side-Scan Sonar Images. *IEEE Transactions on Image Processing*, 16(2):382–390, February 2007. Heriot-Watt University Edinburgh.

[DBL04] Esther Durá, Judith M. Bell, and David M. Lane. Reconstruction of textured seafloors from side-scan sonar images. *IEE Proceedings – Radar, Sonar and Navigation*, 151(2):114–126, April 2004.

[Fai09] Nathaniel Fairfield. *Localization, Mapping, and Planning in 3D Environments*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, January 2009.

[JPW10] Michael V. Jakuba, Oscar Pizarro, and Stefan B. Williams. High resolution, consistent navigation and 3D optical reconstructions from AUVs using magnetic compasses and pressure-based depth sensors. In *Proceedings of MTS/IEEE OCEANS 2010, Sydney, Australia*, 2010.

[JPWM10] Matthew Johnson-Roberson, Oscar Pizarro, Stefan B. Williams, and Ian Mahon. Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. *Journal of Field Robotics*, 27(1):21–51, 2010.

[LH91] Dirk Langer and Martial Hebert. Building Qualitative Elevation Maps From Side Scan Sonar Data For Autonomous Underwater Navigation. In *Proc. IEEE International Conference on Robotics and Automation*, pages 2478–2483, April 9–11, 1991.

[MRRH10] Angelos Mallios, Pere Ridao, David Ribas, and Emili Hernàndez. Probabilistic sonar scan matching SLAM for underwater environment. In *Proceedings of the Oceans IEEE*, Sydney, Australia, May 2010.

[PDM+07] Luis Pedraza, Gamini Dissanayake, Jaime Valls Miró, Diego Rodríguez-Losada, and Fernando Matía. Bs-slam: Shaping the world. In Wolfram Burgard, Oliver Brock, and Cyrill Stachniss, editors, *Robotics: Science and Systems*. The MIT Press, 2007.

[PRLM+09] Luis Pedraza, Diego Rodriguez-Losada, Fernando Matía, Gamini Dissanayake, and Jaime Valls Miró. Extending the limits of feature-based slam with b-splines. *Trans. Rob.*, 25(2):353–366, 2009.

[RRN08] David Ribas, Pere Ridao, and José Neira. *Underwater SLAM for Structured Environments Using an Imaging Sonar*. PhD thesis, Departament of Computer Engineering, Universitat de Girona, September 2008.

[Run10] Adrian Runte. Robuste Merkmale auf 3D-Daten. Term Paper, Vision and Fusion Laboratory (IES), Karlsruhe Institute of Technology, July 2010.

[RVRV95] P. V. Narasimha Rao, M. S. R. R. Vidyadhar, T. Ch. Malleswara Rao, and L. Venkataratnam. An adaptive filter for speckle suppression in synthetic aperture radar images. *International Journal of Remote Sensing*, 16(5):877–889, 1995.

[WF10] Philipp Woock and Christian Frey. Deep-sea auv navigation using side-scan sonar images and slam. In *Proceedings of the IEEE Oceans Conference 2010*, may 2010.

[Woo10] Philipp Woock. Deep sea navigation using SLAM. *Proceedings of the 2009 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*, 4, 2010.