We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Bearing-Only Vision SLAM with Distinguishable Image Features

Patric Jensfelt[1], Danica Kragic[1] and John Folkesson[2]
[1]Centre for Autonomous Systems, Royal Institute of Technology
[2] Dept of Mechanical Engineering, Massachusetts Institute of Technology
[1]Sweden, [2]USA

1. Introduction

One of the key competences for autonomous mobile robots is the ability to build a map of the environment using natural landmarks and to use it for localization (Thrun et al., 1998, Castellanos et al, 1999, Dissanayake et al, 2001, Tardos et al. 2002, Thrun et al., 2004). Most successful systems presented so far in the literature have relied on range sensors such as laser scanners and sonar sensors. For large scale, complex environments with natural landmarks the problem of SLAM is still an open research problem. Recently, the use of vision as the only exteroceptive sensor has become one of the most active areas of research in SLAM (Davison, 2003, Folkesson et al., 2005, Goncavles et al., 2005, Sim et al., 2005, Newman & Ho., 2005).

In this chapter, we present a SLAM system that builds maps with point landmarks using a single camera. We deal with a set of open research issues such as how to identify and extract stable and well-localized landmarks and how to match them robustly to perform accurate reconstruction and loop closing. All of these issues are central to success, especially when an estimator such as the Extended Kalman Filter (EKF) is used. Robust matching is required for most recursive formulations of SLAM where decisions are final. Even for methods that allow the data associations to change over time, e.g. (Folkesson & Christensen, 2004, Frese & Schröder 2006), reliable matching is very important.

One of the big disadvantages with the laser scanner is that it is a very expensive sensor. Cameras, on the other hand, are relatively cheap. Another aspect of using cameras for SLAM is the much greater richness of the sensor information as compared to that from, for example, a range sensor. Using a camera it is possible to recognize features based on their appearance. This provides the means for dealing with one of the most difficult problems in SLAM, namely data association.

The main contributions of this work are i) a method for the initialisation of visual landmarks for SLAM, ii) a robust and precise feature detector, iii) the management of the measurement to make on-line estimation possible, and iv) the demonstration of how this framework can facilitate real-time SLAM even with an EKF based implementation.

2. Related Work

Working with a single camera, the measurements will be of bearing only type. This means that a single observation of a landmark is not enough to estimate its full pose since the depth is unknown. This problem is typically addressed by combining the observations from multiple views as in the structure-from-motion (SFM) approaches in computer vision. The biggest difference between SLAM and SFM is that SFM considers mostly batch processing while SLAM typically requires on-line, real-time performance.

The fact that the full pose of a landmark cannot be estimated from a single observation leads to one of the most important problems that has to be addressed in bearing only SLAM; landmark initialisation. Several approaches have been presented in the literature. In (Davison, 2003) a particle filter was used to represent the unknown initial depth of features. The drawback of the approach is that the initial distribution of particles has to cover all possible depth values for a landmark, which makes it difficult to use when the number of detected features is large. A similar approach has been presented in (Dissanayake et al., 2005) where the initial state is approximated using a Gaussian Sum Filter for which the computational load grows exponentially with number of landmarks. The work in (Lemarie et al. 2005) proposes an approximation with additive growth. It uses a weighted Gaussian sum approximation for the depth estimate of uninitialised landmarks. Gaussians in the sum are deleted when they no longer are supported by subsequent observations. When a single Gaussian remains, the landmark is initialised given that a few other conditions are fulfilled.

Another, more practical problem associated with landmark initialisation comes from the limited field of view of a normal perspective camera in combination with the robot typically moving along the optical axis as pointed out in (Goncavles et al., 2005). To cope with the reconstruction problem, a stereo-based SLAM method was presented in (Sim et al., 2005) where Difference-of-Gaussians (DoG) is used to detect distinctive features which are then matched using SIFT descriptors. An important problem mentioned is that their particle filter based approach is inappropriate for large-scale and textured environments. One of the contributions of our work is that we deal with this problem by identifying only a few high quality features in the scene to perform SLAM.

Another problem mentioned in (Sim et al., 2005) is related to the time-consuming feature matching. We address this by using a KD-tree to make our matching process very fast. The visual feature detector used in our work is the Harris corner detector across different scales represented by a Laplacian pyramid, similar to what is suggested in (Mikolajczyk & Schmid 2003). For feature matching, we use a modified SIFT descriptor in combination with KD-trees.

Working in indoor environments means that the floor is typically flat and the SLAM problem can be simplified by assuming that the robot is constrained to a plane. However, there are many repetitive features stemming from, for example, right angle corners. A single SIFT descriptor is not discriminative enough in an image to solve the data association problem. To address this, ``chunks'' of SIFT points were used to represent landmarks in an outdoor environment in (Luke et al., 2005). This was motivated by the success that SIFT has had in recognition applications where the object/scene was represented as a set of SIFT points. In our approach, the position of a landmark is defined by a series of SIFT points representing different views of the landmark. Each such point is accompanied with a chunk of descriptors that make the matching/recognition of landmarks more robust. Our experimental evaluation shows also that our approach performs successful matching even

with a narrow field of view, which was mentioned as a problem in (Goncavles et al., 2005, Sim et al., 2005).

Yet another problem in SLAM is loop closing, that is the ability to detect when the robot comes back to a position it has been to previously and thereby closing a loop. (Newman & Ho, 2005) argue for using laser for the geometric mapping but to rely on visual input to solve the loop-closing problem. The message is that robustness is best achieved if the same mechanism is not used for the mapping and the loop closing detection. In (Newman & Ho, 2005) visually salient, so called ``maximally stable extremal regions'' or MSERs, encoded using SIFT descriptors, are used to detect when the robot is revisiting an area. In (Gutmann & Konolige, 1999) scan matching is used to detect when loops are closed. We show in this chapter that our framework also can be used for loop closing detection.

In the remainder of this chapter we will make a distinction between recognition and location features. A single location feature will be associated with several recognition features. The recognition features' descriptors then give robustness to the match between the location features in the map and the features in the current image. The key idea is to use a few high quality features to define the location of landmarks and then use the other features for recognition. This contributes to a low complexity (few location features) while maintaining highly robust matching (many recognition features).

3. Feature Description

The SIFT descriptor (Lowe, 1999) has been used frequently in both computer vision and various robot vision applications. It has been shown in (Mikolajczyk & Schmid 2003) to be the most robust descriptor regarding scale and illumination changes. The original version of the SIFT descriptor uses feature points determined by the peaks of a series of Difference of Gaussians (DoG) on varying scales. In our system, peaks are found using Harris-Laplace features, (Mikolajczyk & Schmid 2001) since they respond to regions of high curvature, instead of blob-like image structures obtained by series of DoG. This leads to features accurately localized spatially, which is essential when features are used for reconstruction and localization, instead of just recognition.

In a sparse, indoor environment many of the detected features originate from corner features. The original SIFT descriptor assigns canonical orientation at the peak of smoothed gradient histograms. This means that similar corners but with a significant rotation difference can have similar descriptors. This may potentially lead to many false matches. For example, the four corners of the waste bin in Figure 2. may all match if rotated. Therefore, we use a rotationally 'variant' SIFT descriptor where we avoid the canonical orientation at the peak of smoothed gradient histogram and leave the gradient histogram as it is.

4. Landmark Selection and Initialisation

Landmark initialisation is a key issue in bearing only vision SLAM. To determine which image features that are worth turning into landmarks, we match the features across N frames. Features that are successfully matched over enough frames become candidates for landmarks in the map. Such a matching buffer also allows us to calculate an estimate of the 3D position of the corresponding landmark by multi view triangulation. The SLAM process is fed measurements from the output side of the frame buffer, which means that the

measurements are delayed N frames with respect to the input side of the buffer. Figure 1. illustrates this idea.

Figure 1. A buffer of N image frames is used for matching, selection & triangulation
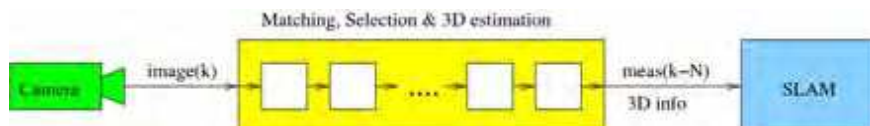

Figure 2. Many structures in indoor environments look similar even when rotated

The benefit of this is that the SLAM process can be fed with few and high quality landmarks. In addition, since an estimate of the 3D position of landmarks can be supplied with the first measurement of a landmark, the landmarks can immediately be fully initialised in the SLAM process. This allows immediate linearisation without the need to apply multiple hypotheses (Lemarie et al., 2005) or particle filtering (Davison, 2003) techniques to estimate the depth. It is important to point out that the approximate 3D position found from the buffer of frames is only used for initialising the point landmark at the correct depth with respect to the camera at the first observation. The uncertainty in depth is still assumed to be very high, as problems with incorporating information twice would otherwise occur. Comparing to a multiple hypothesis approach, it is like knowing which of the multiple hypotheses about the depth is correct right away which saves computations. Having the correct depth allows us, as said before, to reduce the linearisation errors that would results from having a completely wrong estimate of the depth.

Assuming that the delay caused by the length of the buffer is not too large, it is possible to make a quite accurate estimate of the current robot pose by using dead reckoning information to predict forward from the pose estimated by the SLAM process. For typical values of N, the addition to the robot position error caused by the dead reckoning is small and we believe that the benefits of being able to initialise landmarks using bearing-only information and perform feature quality checks are more significant. The predication

forward in time is done in each iteration from the latest pose estimated by SLAM. This way there is no accumulation of dead reckoning errors other than over the short distances corresponding to the size of the buffer.

In addition to requiring that features can be tracked over more than a certain predefined number of frames, we require that the image positions of the feature allow good triangulation and that the resulting 3D point is stable over time in the image. Requiring that the feature can be tracked over several frames removes noise and moving targets that could otherwise severely damage the estimation process. Good triangulations rule out features that have a high triangulation uncertainty, typically because of small baseline or having bearings near the direction of motion. The third requirement removes features that lack sharp positions in all images due to parallax or a lack of a strong maximum in scale space. Difference in scales of the images can also cause apparent motion of features, such as for example a corner of a non-textured object.

We have used a fixed value for N, i.e. the length of the buffer, in our tests. The values between 10 and 50 have been tested. A buffer with all frames acquired from the same camera pose would be of little use for triangulation. Therefore, a new frame is added to the buffer when the camera has moved enough since the last added frame. This way, it is likely that there is enough baseline for estimating the location. The value of N depends very much on the motion of the robot/camera and the camera parameters. For a narrow field of view, camera mounted in the direction of motion of the robot as in our case the effective baseline will be quite small. An omnidirectional camera would offer one way to deal with the small field of view. Another idea is to actively control the direction of the camera as in (Vidal-Calleja et al., 2006).

5. Feature Tracking

The buffer with data from the past N frames does not contain the whole images, but rather the feature points that have been extracted in each frame. An even higher reduction of space could be achieved by using an indexing scheme as in (Nistér & Stewénius, 2006). The feature points are tracked over consecutive frames. To estimate if two feature points match, we use the distance between the descriptors, i.e. between the 128-dimensional vectors associated with the SIFT descriptors. On the left hand side of Figure 3. the organization of the frame memory is shown. Notice the lists that store the associations between the points for the different frames in the buffer. Ideally, each association list corresponds to one landmark in the world and denotes how the projection of this landmark moves in the image as the robot moves.

The SIFT descriptor is invariant to changes in scale and view angle but only up to a certain degree. The change between two consecutive observations in the buffer is however typically quite small and makes tracking possible. The different descriptors in the list correspond to different viewpoints of the same landmark.

As was previously described, the buffer is used to sort the good from the bad landmarks. The output from the frame memory is a small selection of all the features points in the oldest frame. These points are the ones that are judged to be the best with respect to the criteria mentioned earlier. Some of these points correspond to observations of already existing landmarks and some to the first observation of a new landmark. For each new landmark observation, an estimate of the 3D position is obtained by triangulating the points in the corresponding association list. The number of points that are used as observations in each

frame is typically only a small fraction of all points in that frame. This helps reducing the complexity. The time to perform the tracking over frames has constant complexity assuming that the number of features in each frame is bounded.

Only using the similarity of the point descriptor for tracking has two problems. First, it requires that all points in the image are tested for similarity which is computationally expensive and second, it can lead to false matches in cases where there are similar structures in multiple places in the image. To address these issues we predict the approximate image location for the old point features in the new frame using odometry and optical flow estimates. The predicted image location allows us to narrow the search region for each feature match and thus increase efficiency. Notice that the buffer allows us to predict feature points observed not only in the very last frame but also further back. This increases the robustness in the tracking, as some feature points are not present in every frame.
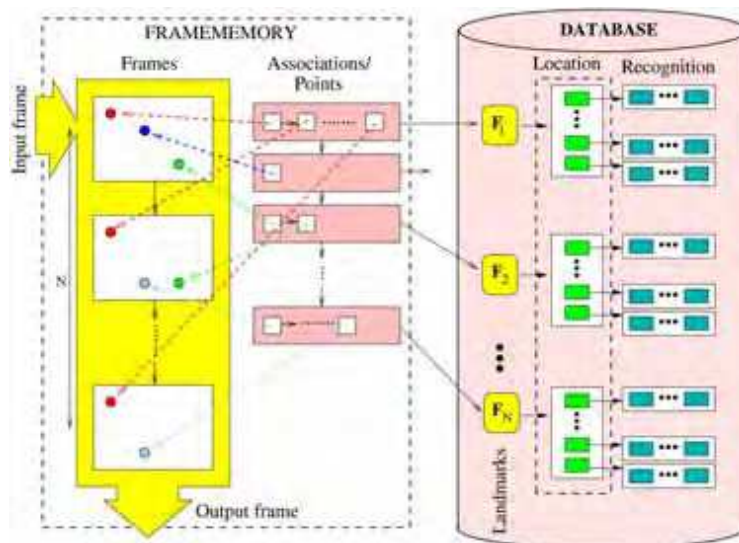

Figure 3. A schematic view of the frame memory and the database

Feature points in a new frame that do not match any of the old feature points with their predicted image locations are matched to a database of initialised landmarks. This allows the system to deal with loop closing situations, i.e. the case where the robot re-visits an area it has been to before. Landmarks are added to this database at the same time, as the first observation is output from the frame memory.

6. Landmark Re-Detection and Loop Closing

The database serves a purpose not only for true loop closing situations but also when the robot turns abruptly. Landmarks not in the field of view will eventually leave the frame memory. When the robot turns the camera back to this region it is important that new landmarks are not created but rather that matches are found to the already existing landmarks. As discussed in the previous section, landmarks appear different from different viewpoints. To handle this, the database stores a number of descriptors for each landmark, corresponding to its appearance from different viewpoints. The different descriptors for a

landmark in the database are provided by the frame memory. For every new observation of a landmark the descriptor is compared to the existing ones and used to augment the descriptor list if it is different enough.

The SIFT point descriptors are not globally unique (see Figure 2. again) and thus matching a single observation to a landmark is doomed to cause false matches in a realistic indoor environment. However, using large number of SIFT descriptors has proven to give robust matching results in object recognition applications. This is why we store, along with the landmark descriptor associated with the location of the landmark, the rest of the descriptors extracted from the same frame and use these for verification. We refer to the rest of the feature points in a frame as recognition features to distinguish them from the location feature associated with the location of the landmark.

The structure of the database is shown on the right hand side in Figure 3. Each landmark $F_1, F_2, ..., F_N$ has a set of location descriptors shown in the dashed box. A KD-tree representation and a Best-Bin-First (Beis & Lowe, 1997) search allow for real-time matching between new image feature descriptors and those in the database. Each location descriptor has a set of recognition descriptors shown to the right.

When we match to the database, we first look for a match between a single descriptor in the new frame and the location descriptors of the landmarks (dashed box Figure 3.). As a second step, we match all descriptors in the new frame to the recognition descriptors associated with candidate location descriptors for verification. As a final test, we require that the displacement in image coordinates for the two location features (new frame and database) is consistent with the transformation between the two frames estimated from the matched recognition descriptors (new frame and database). This assures that it is not just two similar structures in the same scene but that they are at the same position as well. Currently, the calculation is simplified by checking the 2D image point displacement. This final confirmation eliminates matches that are close in the environment and thus share recognition descriptors such as would be the case with the glass windows in Figure 2.

7. SLAM

The previous sections have explained how we track features between frames to be able to determine which make good landmarks and how these are added to, represented in and matched to the database. In our current system, we use an EKF base implementation of SLAM. It is however important to point out that the output from the frame memory could be used as input to any number of different SLAM algorithms. It is possible to use normal EKF despite its limitation regarding complexity since most features extracted from the frames have been discarded by the matching and quality assessment process in the frame memory. Even though hundreds of features are extracted in each frame only a fraction of these are used for estimation. We are also able to supply the approximate 3D location of new landmark so that no special arrangement for this has to be added in the SLAM algorithm. This also makes the plug-n-play of SLAM algorithm easier.

We use the same implementation for SLAM that was used in (Folkesson et al, 2005). This is part of the freely available CURE/toolbox software package. In (Folkesson et al, 2005) it was used for vision SLAM with a camera pointing up in the ceiling.

To summarize, the division is such that the SLAM process is responsible for estimating the location of a landmark and the database for its appearance.

8. Experimental Evaluation

Figure 4. The PowerBot platform with the Canon VC-C4 camera

The camera used in the experimental evaluation is a Canon VC-C4 camera mounted in the front on a PowerBot platform from MobileRobotics Inc (see Figure 4.). The experimental robot platform has a differential drive base with two rear caster wheels. The camera was tilted upward slightly to reduce the amount of floor visible in the image. The field of view of the camera is about 45 degrees in the horizontal plane and 35 in the vertical plane. This is a relatively small field of view. In addition, the optical axis is aligned with the direction of motion of the platform so that it can be used for other navigation tasks. The combination of a small field of view and motion predominantly along the optical axis makes it hard to generate large baselines for triangulation.

The experimental evaluation will show how we are able to build a map of the environment with few but high quality landmarks and how detection of loop closing is performed.

The setting for the experiment is an area around an atrium that consists of loops of varying sizes. We let the robot drive 3 laps following approximately, but not exactly, the same path. Each lap is about 30m long. The trajectory along with the resulting map is shown in Figure 5. The landmarks are shown as small squares. Overlayed on the vision based map is a map built using a laser scanner (the lines). This second map is provided as a reference for the reader only. The laser scanner was not used at all in the vision experiments. Figure 6. shows the situation when the robot closes the loop for the first time. The lines protruding from the camera point out the points that are matched. Figure 7. shows one of the first acquired images along with the image in which the two matches shown in Figure 6. were found just as the loop is closed for the first time.

There are a number of important observations that can be made. First, there are much fewer landmarks than typically seen in maps built using point landmarks and vision, see e.g. (Sim et al., 2005, Se et al., 2002). We can also see that the landmarks are well localized as they fall closely to the walls. Notice that some of the landmarks are found on lamps hanging from the ceiling and that the area in the upper left corner of Figure 6. is quite cluttered. It is a student study area and it has structures at many different depths. A photo of this area is shown in

Figure 8. The line picked up by the laser scanner is the lower part of the bench where people sit and not the wall behind it. This explains why many of the points in this area do not fall on the laser-based line. Some of the spread of the point can also be explained by the small baseline. The depth error is inversely proportional to the baseline (Hartley & Zisserman, 2000).
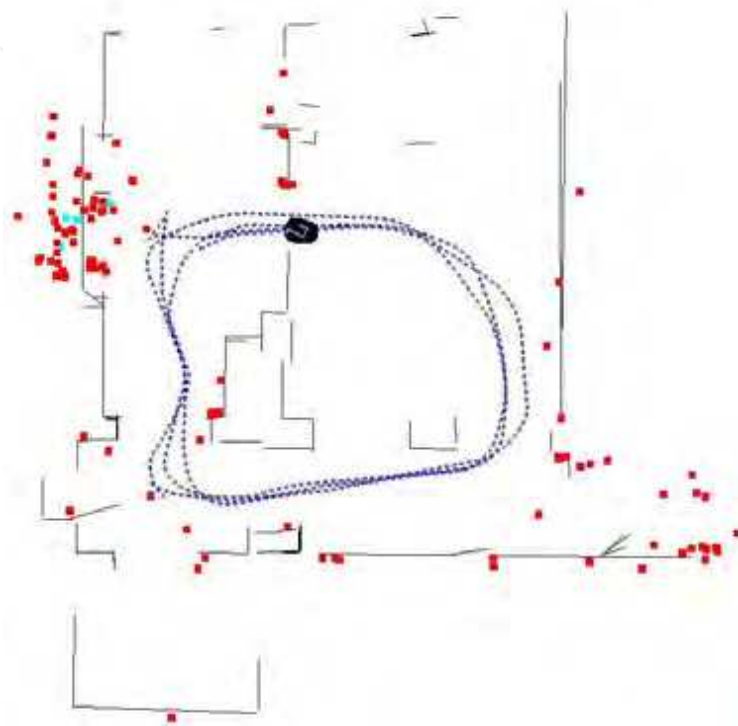

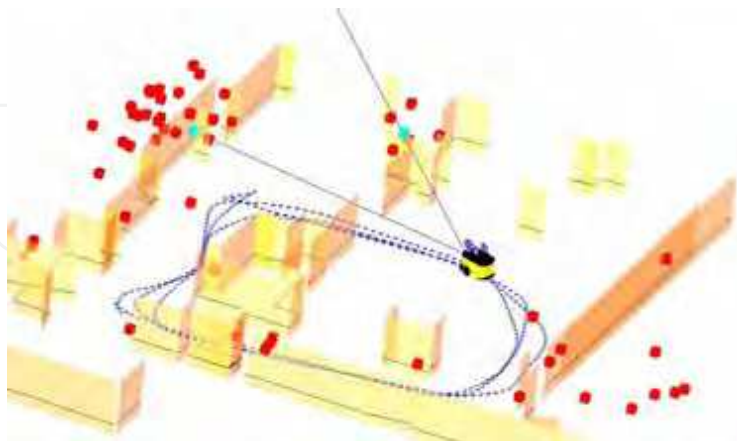Figure 5. The landmark map with the trajectory and reference laser based map

Figure 6. Situation when the first loop is closed. Lines show matched points

Another observation that can be made is that the final map contained 113 landmarks and that most of these were added to the map during the first loop (98). This indicates that landmarks were matched to the database rather than to be added to the map. Had this not been the case one would have expected to see roughly 3 times the number of landmarks.

As many as half of the features in each frame typically do not match any of the old features in the frame memory and are thus matched to the database. A typical landmark in the database has around 10 descriptors acquired from different viewing angles. The matching to the database uses the KD-tree in the first step that makes this first step fast. This often results only in a few possible matching candidates.

Figure 7. One of the matched points in the first loop detection (compare to Figure 6)

Figure 8. Cluttered area in upper right corner of Figure 5

In the experiments, an image resolution of 320x240 was used and images were grabbed at 10Hz. Images were added to the frame buffer when the camera had moved more than 3cm and/or turned 1 degree. The entire experimental sequence contained 2611 images, out of which roughly half were processed. The total time for the experiment was 8min 40s and the processing time was 7min and 7s on a 1.8GHz laptop. This shows that it can operate under real-time conditions

9. Conclusions and Future Work

For enabling the autonomy of robotic systems, we have to equip them with the ability to build a map of the environment using natural landmarks and to be able to use it for localization purposes. Most of the robotic systems capable of SLAM presented so far in the literature have relied on range sensors such as laser scanners and sonar sensors. For large scale, complex environments with natural landmarks the problem of SLAM is still an open research problem. More recently, the use of cameras and machine vision as the only exteroceptive sensor has become one of the most active areas of research in SLAM.

The main contributions presented in this chapter are the feature selection and matching mechanisms that allow for real-time performance even with an EKF implementation for SLAM. One of the key insights is to use few, well localized, high quality landmarks to acquire good 3D position estimates and then use the power of the many in the matching process by including all features in a frame for the verification. Another contribution is our use of a rotationally variant feature descriptor to better deal with the symmetries that are often present in indoor environments. An experimental evaluation was presented on data collected in a real indoor environment. Comparing the landmarks in the map built using vision with a map built using a laser scanner showed that the landmarks were accurately positioned.

As part of the future research we plan to investigate how the estimation process can be improved by using active control of the pan-tilt degrees of freedom of the camera on the robot. By such coupling, the baseline can actively be made larger to improve triangulation/estimation results. It would also allow the system to use good landmarks, otherwise not in the field of view, to improve the localization accuracy and thus the map quality.

10. References

Beis, J.S.; Lowe, D.G. (1997). Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1000-1006.

Castellanos, J.A.; Tardós, J.D. (1999). *Mobile Robot Localization and Map Building: A Multisensor Fusion Approach,* Kluwer Academic Publishers.

Davison A.J. (2003). Real-time simultaneous localisation and mapping with a single camera. *Proceedings of the International Conference on Computer vision (ICCV).*

Dissanayake, G.; Newman, P.; Clark, S.; Durrant-Whyte, H.F.; Corba, M. (2001). A solution to the slam building problem. *IEEE Transactions on Robotics and Automation*, 17, 3, 229-141

Folkesson, J.; Christensen, H.I. (2004). Graphical SLAM - A Self-Correcting Map, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*

Folkesson, J.; Jensfelt, P.; Christensen; H.I. (2005) Vision slam in the measurement subspace. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA).*

Frese, U.; Schröder, L. (2006) Closing a Million-Landmarks Loop, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*

Goncavles, L.; di Bernardo, E.; Benson, D.; Svedman, M.; Ostrovski, J.; Karlsson, N.; Pirjanian, P. (2005) A visual fron-end for simultaneous localization and mapping. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.pp. 44—49.

Gutmann, J.; Konolige, K. (1999) Incremental mapping of large cyclic environments. *Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 318-325

Hartley, R.; Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521623049

Kwok, N.M.; Dissanayake, G.; Ha, Q.P. (2005) Bearing only SLAM using a SPRT based gaussian sum filter. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.

Lemaire, T.; Lacroix, S.; Sola, J. (2005) A practical 3D bearing-only SLAM algorithm, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp 2757-2762

Lowe, D.G. (1999) Object recognition from local scale-invariant features. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1150-57

Luke, R.H.; Keller, J.M.; Skubic, M.; Senger, S. (2005) Acquiring and maintaining abstract landmark chunks for cognitive robot navigation. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*

Mikolajczyk, K.; Schmid, C. (2001) Indexing based on scale invariant interest points. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. pp. 525-531

Mikolajczyk, K.; Schmid, C. (2003) A performance evaluation of local descriptors. *Proceedings of IEEE Conference on Computer Visiona and Pattern Recognition (CVPR)*, pp. 257-263.

Newman, P.; Ho, K. (2005) SLAM-loop closing with visually salient features, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. pp. 644-651.

Nistér, D.; Stewénius, H. (2006) Scalable Recognition with a Vocabulary Tree, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*

Se, S.; Lowe, D.G.; Little, J. (2002) Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *Journal of Robotics Research*, 21, 8, 735-758

Sim, R.; Elinas, P.; Griffin, M.; Little, J. (2005) Vision-based slam using the rao-blackwellised particle filter, *Proceedings of the Workshop on Reasoning with Uncertainty in Robotics (IJCAI)*

Tardós, J.D.; Neira, J; Newman, P.M.; Leonard, J.J. (2002) Robust mapping and localization in indoor environments using sonar data, *Journal of Robotics Research, 4*

Thrun, S.; Fox, D.; Burgard, W. (1998) A probalistic approach to concurrent mapping and localization for mobile robots, *Autonomous Robots*, 5, 253-271

Thrun, S.; Liu, Y.; Koller, D.; Ng, A.; Ghahramani, Z.; Durrant-White, H. (2004) SLAM with sparse extended information filters, *Journal of Robotics Research, 23, 8*, 690—717

Vidall-Calleja, T.; Davison, A.J.; Andrade-Cetto, J.; Murray, D.W. (2006) Active Control for Single Camera SLAM, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1930-36

Vision Systems: Applications

Edited by Goro Obinata and Ashish Dutta

Computer Vision is the most important key in developing autonomous navigation systems for interaction with the environment. It also leads us to marvel at the functioning of our own vision system. In this book we have collected the latest applications of vision research from around the world. It contains both the conventional research areas like mobile robot navigation and map building, and more recent applications such as, micro vision, etc.The fist seven chapters contain the newer applications of vision like micro vision, grasping using vision, behavior based perception, inspection of railways and humanitarian demining. The later chapters deal with applications of vision in mobile robot navigation, camera calibration, object detection in vision search, map building, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Patric Jensfelt, Danica Kragic and John Folkesson (2007). Bearing-Only Vision SLAM with Distinguishable Image Features, Vision Systems: Applications, Goro Obinata and Ashish Dutta (Ed.), ISBN: 978-3-902613-01-1, InTech, Available from: http://www.intechopen.com/books/vision_systems_applications/bearing-only_vision_slam_with_distinguishable_image_features

INTECH
open science | open minds