

Using Symmetrical Regions of Interest to Improve Visual SLAM

Gert Kootstra and Lambert R. B. Schomaker

Artificial Intelligence

University of Groningen, The Netherlands

g.kootstra@ai.rug.nl, l.schomaker@ai.rug.nl

Abstract—Simultaneous Localization and Mapping (SLAM) based on visual information is a challenging problem. One of the main problems with visual SLAM is to find good quality landmarks, that can be detected despite noise and small changes in viewpoint. Many approaches use SIFT interest points as visual landmarks. The problem with the SIFT interest points detector, however, is that it results in a large number of points, of which many are not stable across observations. We propose the use of local symmetry to find regions of interest instead. Symmetry is a stimulus that occurs frequently in everyday environments where our robots operate in, making it useful for SLAM. Furthermore, symmetrical forms are inherently redundant, and can therefore be more robustly detected. By using regions instead of points-of-interest, the landmarks are more stable. To test the performance of our model, we recorded a SLAM database with a mobile robot, and annotated the database by manually adding ground-truth positions. The results show that symmetrical regions-of-interest are less susceptible to noise, are more stable, and above all, result in better SLAM performance.

I. INTRODUCTION

Being able to make a representation of the environment and to use this representation to estimate the current position is one of the most fundamental skills of an autonomous robot. For that reason, Simultaneous Localization and Mapping (SLAM) has been an important field of research [1]. With the use of precise range sensors, like laser-range scanners, SLAM systems currently achieve good quality maps of the environment, and accurate localization. A more challenging problem is visual SLAM, using cameras instead of range sensors. The advantage of using a camera is that it is a cheap and light-weight sensor. Furthermore, a camera is a passive sensor that, unlike most range sensors, does not emit energy in the environment. Moreover, a camera provides a rich source of information that opens possibilities for more advanced representations of the environment.

One of the challenges in visual SLAM is to find high-quality visual landmarks to represent the environment [2]. A good landmark needs to be reliably detectable from multiple viewpoints in the presence of noise. In the current work, we focus on this problem, and propose a model for reliable landmark detection.

A common approach for the selection of landmarks is to detect interest points in the camera image [3]. To reliably detect a landmark, it needs to be robust, that is, it needs to be detected despite noise and changing light conditions. Moreover the landmark needs to be stably detected over a sequence of observations. And finally, it needs to be detectable when the robot revisits a location.

Many current approaches to visual SLAM detect interest points based on contrast features, for instance using the Scale-Invariant Feature Transform (SIFT) [4], [5], Speeded-Up Robust Features (SURF) [6], [7] or Harris corners [8]. SIFT has been proven to be one of the best performing interest point detectors for SLAM [3], as well as for object recognition [9]. For that reason, we compare our method to SIFT interest point detector. However, in [10], it was shown that using local symmetry instead of contrast, results in more robust interest points. Moreover, it was shown that the performance of the SLAM system using local symmetry for landmarks selection outperformed the system using SIFT.

In the presented research, we propose the use of local symmetry to find regions-of-interest instead of interest points. The problem with interest points is that a large number of points are found in the image, but many of them are unstable. This results in a high computational load, and reduces the quality of the map of the environment. Although the use of symmetry reduces the number of interest points, and improves the stability, the large number of unstable points remains a problem. We hypothesize that using regions-of-interest will result in fewer and more stable landmarks, since larger areas contribute to the regions-of-interest, providing more evidence and making the method less susceptible to noise. Similar findings were done in [11] and [12].

The motivation to use symmetry in robotic vision is inspired by human vision. It has been shown that predictions by local symmetry models correlate well with human eye fixations [13]. Especially early fixations are on highly symmetrical parts of the visual field [14]. Furthermore, humans can rapidly detect symmetry, particularly when patterns have multiple axes of symmetry [15], and they can recall and discriminate symmetrical forms better than non-symmetrical forms [16]. Since it is apparently advantageous for humans to pay attention to locally symmetrical parts of the visual field, it might be beneficial for robots to do so as well. Moreover, symmetry is highly abundant in our everyday environments. This can be exploited for robotic SLAM by using local-symmetry detection for landmark selection.

Although contrast features have received most attention in computer vision research (e.g., [4], [17]), symmetry is successfully used in a number of studies. Marola [18], for instance, used symmetry for detection and localization of objects in planar images, and found that symmetry is insensitive to noise. Symmetry has also been used as a feature to control the gaze of an active vision system [19], and to

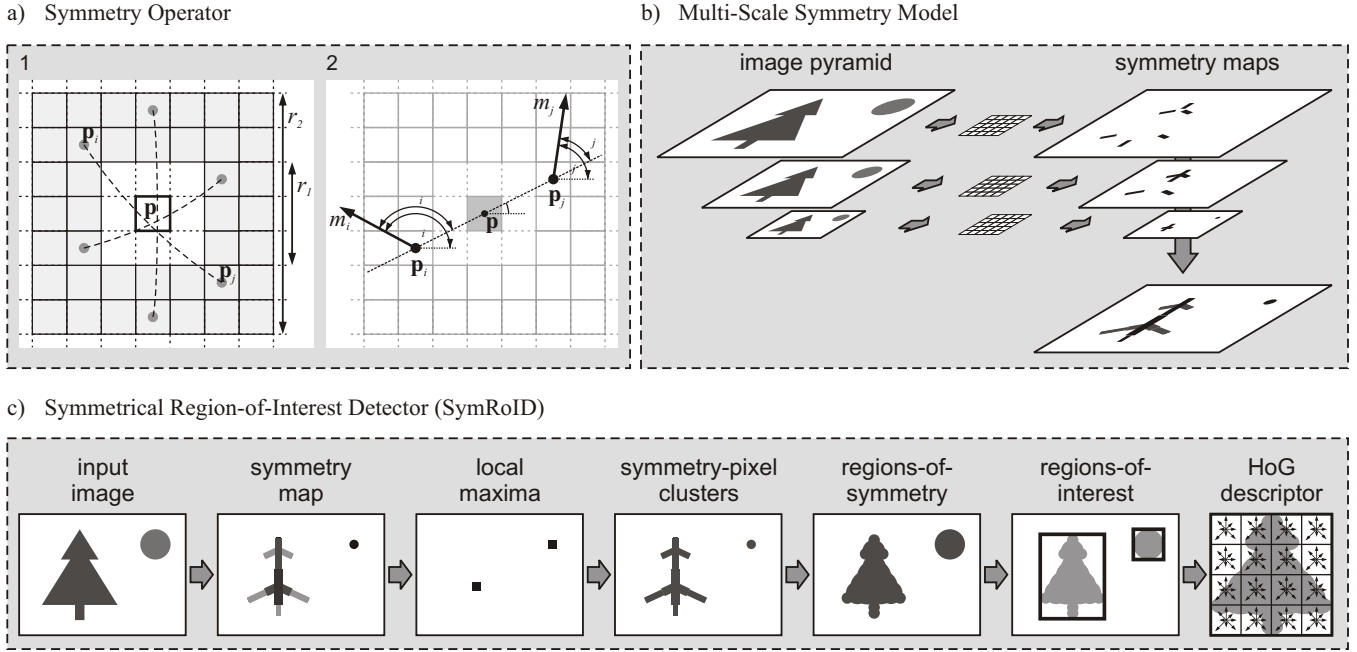


Fig. 1. The Symmetrical Region-of-Interest Detector. The symmetry operator (a) is at the basis of the detector. All pixel pairs that lie in the symmetry kernel (the gray marked area) contribute to the symmetry value at position \mathbf{p} (a1). The symmetry contribution of a pixel pair is determined by comparing the gradients of both pixels (a2). In the multi-scale symmetry model (b), an image pyramid is constructed. The symmetry operator is applied to all images in the pyramid. The resulting symmetry maps are rescaled, and summed up to result in the multi-scale symmetry map. The complete SymRoID model (c) first calculates the symmetry map for the input image. In this map, the local maxima are found. These local maxima serve as the seeds for the region-growing algorithm that clusters all symmetry pixels. Next, the most contributing radius of each symmetry pixel in a cluster is found and marked with a circle in the regions-of-symmetry map. The regions-of-interest are finally determined by the bounding box of the marked regions. Subsequently, the regions are described using histograms-of-gradients (HoGs).

direct the attention of a robot [20]. Furthermore, a context-free symmetry operator has been proposed for the detection of facial features [21].

In this paper we show that the detecting regions-of-interest based on symmetry results in landmarks that are more robust and stable than interest points detected using contrast features. Moreover, we test the performance of our landmark selection method on a SLAM database that we annotated with ground-truth positions. The results show that our model results in better SLAM performance.

II. SYMMETRICAL REGION-OF-INTEREST DETECTOR

Our symmetrical region-of-interest detector, SymRoID, is an extension of the symmetry model that we proposed in [13], [14]. The model is based on the symmetry operator of Reisfeld *et al.* [22]. In this section we first describe the symmetry operator. We then discuss the multi-scale symmetry model. Next we present the SymRoID model, and we end with the region-of-interest descriptor.

A. The symmetry operator

For every position, $\mathbf{p} = (x, y)$, in the image, a symmetry kernel is applied that calculates the amount of symmetry by comparing the intensity gradients of the surrounding pixels. Pixel pairs in the neighborhood contribute to the symmetry value. A pixel pair consists of pixel \mathbf{p}_i and \mathbf{p}_j , so that $\mathbf{p} = (\mathbf{p}_i + \mathbf{p}_j)/2$ (see fig. 1a). The contribution of the pair is

calculated by comparing the intensity gradient \vec{g}_i at \mathbf{p}_i and gradient \vec{g}_j at \mathbf{p}_j according to:

$$s(i, j) = d(i, j, \sigma) \cdot c(i, j) \cdot m_i \cdot m_j \quad (1)$$

where m_i is the magnitude of the gradient, and $d(i, j, \sigma)$ is a Gaussian weighting function on the distance between p_i and p_j with a standard deviation of σ . The multiplication with the gradient magnitudes assures that only strong edges contribute. The symmetry measurement is:

$$c(i, j) = (1 - \cos(\gamma_i + \gamma_j)) \cdot (1 - \cos(\gamma_i - \gamma_j)) \quad (2)$$

where $\gamma_i = \theta_i - \alpha$ is the angle between the orientation of gradient, θ_i , and the angle, α , of the line between \mathbf{p}_i and \mathbf{p}_j (see fig. 1a). The first term in (2) has a maximum value when $\gamma_i + \gamma_j = \pi$, which is true for gradients that are mirror symmetric with respect to \mathbf{p} . Using only this term would result in high value for points that lie on a straight edge. Since we are not interested in edge detection, but in finding the centre of symmetrical patterns, the second term demotes pixel pairs with similar gradient orientations.

The symmetry value at position \mathbf{p} is calculated by summing up the contributions of all pixel pairs in the neighborhood. This neighborhood is defined by an inner and an outer square centered around \mathbf{p} . The size of the sides of the squares are respectively r_1 and r_2 (see fig. 1a). All pixels that lie inside the outer square, but outside the inner square

are considered. $\Gamma(\mathbf{p})$ is the set of contributing pairs. In our experiments we used $r_1 = 5$ and $r_2 = 17$. Smaller values of r_1 result in too small symmetry patterns, and larger values of r_2 are too computationally expensive, and make the operator view the image with too much detail. The total symmetry value at $\mathbf{p} = (x, y)$ is then:

$$S_k(x, y) = \sum_{(i, j) \in \Gamma(\mathbf{p})} s(i, j) \quad (3)$$

where S_k is the symmetry map at scale k . The different scales are discussed in the next section.

B. The multi-scale symmetry model

A region-of-interest detector for visual SLAM needs to be able to detect structures of various sizes since the appearance of landmarks changes drastically when the robot moves around in the environment. Although the symmetry operator can detect symmetry within the neighborhood radius, it cannot detect patterns on larger scales. Increasing the radius is not a good idea due to the quadratic complexity of the operator. Moreover, at larger radii, the operator takes into account too much detail, making the operator more susceptible to noise. Therefore, we propose a multi-scale symmetry model, similar to that used in [13], [14].

In fig. 1b, the multi-scale symmetry model is depicted. The scale space consists of an image pyramid that is built by progressively applying a Gaussian filter to the image, followed by a downscaling of the image by a factor of two, where scale zero is the image in its original resolution. Secondly, the symmetry operator is applied to all images in the pyramid, resulting in a pyramid of symmetry maps. Finally, the symmetry maps at the different scales are resized to the size of the first scale, and then summed up to result in the overall symmetry map:

$$S(x, y) = \bigoplus_{k=\kappa_1}^{\kappa_2} S_k(x, y) \quad (4)$$

where κ_1 is the first, and κ_2 is the last scale. The operator \bigoplus rescales all maps to the first scale, and subsequently sums the values of the different scales.

Since we are interested in all symmetrical regions in our robotic system, we do not apply the normalization that we used in [13], [14], because that promotes symmetry maps with only one dominant salient point.

C. The SymRoID model

A simplified flow chart of the complete SymRoID model is given in fig. 1c. It consists of a number of steps:

- 1) The symmetry map is calculated by the multi-scale symmetry model as described earlier.
- 2) Local maxima. A pixel (x_m, y_m) is a local maximum if it has the highest value in its 3×3 pixels neighborhood, and its symmetry value $S(x_m, y_m) \geq \tau$, where we used $\tau = 0.5$ in our experiments.
- 3) The local maxima are seeds for a region-growing algorithm. The flood-fill algorithm that we applied,

takes a local maximum, and grows the area to add all neighboring pixels that have a symmetry value of $S(x, y) \geq \lambda \cdot S(x_m, y_m)$, where the threshold is a ratio, λ , of the symmetry value of the local maximum. Connecting regions are merged. The region growing results in clusters of symmetry pixels. In our experiments, we used $\lambda = 0.5$.

- 4) The symmetry-pixel clusters contain the pixels that are the centra of symmetry. Since we are interested in symmetrical regions-of-interest, the complete symmetrical pattern that contributed to these symmetry centra need to be found. To do so, we store, for each pixel, the radius that contributed most to its symmetry value. If \mathbf{p}_i and \mathbf{p}_j form the pixel pair with the highest symmetry contribution $s_{\max}(i, j)$, then $r_s = \|\mathbf{p}_i - \mathbf{p}_j\|/2$ is the maximally contributing symmetry radius. A circle with center \mathbf{p} and radius r_s is then marked in the regions-of-symmetry map.
- 5) Finally, the regions-of-interest are determined by taking the bounding box of the different regions in the regions-of-symmetry map. The regions-of-interest can overlap.

Some examples of regions-of-interest found by SymRoID can be found in fig. 3. It shows two pairs of subsequent images from the SLAM database.

D. The region-of-interest descriptor

We represent the regions-of-interest with a histograms-of-gradients (HoGs) descriptor, similar to the SIFT descriptor [4]. A region is first resampled to a 16×16 pixels descriptor window. This window is then divided in 16 squares (see fig. 1c). For each square, a histogram-of-gradients is calculated from the intensity gradients of the 4×4 pixels that are in the square. Such a histogram contains 8 bins, for the different gradient orientations, i.e., $[0, \frac{1}{4}\pi), [\frac{1}{4}\pi, \frac{2}{4}\pi)$, etc.

The values of the 8 bins in each of the 16 histograms form the 128-dimensional region-of-interest descriptor. Like in [4], the descriptor is normalized to achieve invariance to changes in intensity, resulting in a vector of unit length.

Since the descriptor window adapts to the size of the region-of-interest, and the size of the region-of-interest itself is determined by the observed symmetrical pattern, the SymRoID model is scale invariant. This makes it possible to detect a landmark from different distances. Moreover, the descriptor is relatively invariant to small shift changes and affine transformations, as discussed in [4]. Unlike the standard SIFT descriptor [4], we did not add rotational invariance, since our robot drives on flat surfaces, and will therefore not encounter rotational transformations of the stimulus.

III. THE VISUAL SLAM SYSTEM

To ensure that only stable landmarks are added to the map of the environment, we use a visual buffer that tests the stability of the regions-of-interest. When a landmark passes the buffer, it is added to the SLAM system.

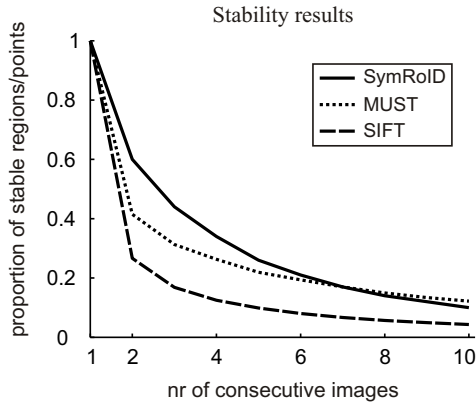


Fig. 2. Stability of the regions-of-interest and interest points. The graphs show the proportion of regions (for SymRoID) or points (for MUST and SIFT) that are stably found in all observed images.

A. The visual buffer

Fig. 2 shows the stability of the symmetrical regions-of-interest compared to the stability of interest points obtained by MUST [10] and SIFT [4]. To test this stability, multiple sequences of images recorded by our robot are taken (see section IV for the experimental setup). A region or point in the last image of the sequence is considered stable if it is matched in all the previous images of the sequence. Two regions or points match when the Euclidean distance between the two descriptors, \mathbf{d}_i and \mathbf{d}_j , is $\|\mathbf{d}_i - \mathbf{d}_j\| < 0.6$.

Although the stability of the symmetrical regions is considerably higher than that of the MUST and SIFT interest points, it remains a problem that many of the regions are not matched over the full sequence of images. To make sure that only stable landmarks are added to the map of the environment, we use a visual buffer to test the stability of the regions-of-interest, similar to [10], [11].

The visual buffer contains the N most recent camera images. The regions in the current image are compared to those in the $N - 1$ previous images. A region, i , passes the buffer if it is matched in at least M of the previous images. Two regions, i and j , match when the descriptors, \mathbf{d}_i and \mathbf{d}_j , are sufficiently similar. This is true when the Euclidean distance is below the threshold τ_1 ,

$$\|\mathbf{d}_i - \mathbf{d}_j\| < \tau_1, \quad (5)$$

and when the *best-to-next-best ratio* is smaller than δ_1 ,

$$\|\mathbf{d}_i - \mathbf{d}_j\| / \|\mathbf{d}_i - \mathbf{d}_l\| < \delta_1, \quad (6)$$

where \mathbf{d}_l is the descriptor of the second most-similar region in the previous image. This ratio ensures uniqueness. In our experiments we obtained good results with $N = 7$, $M = 5$, $\tau_1 = 0.6$, and $\delta_1 = 0.8$.

Finally, the position of the landmark in the environment is obtained by comparing the different observations of the region. Estimates of the position are made by triangulation using the bearings of the observations and the displacement of the robot, and by inferring depth information from the

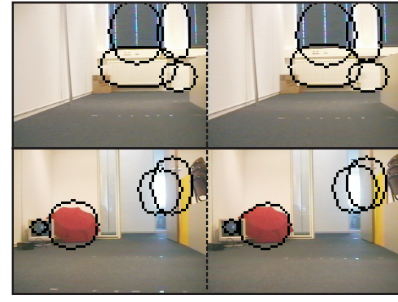


Fig. 3. Examples of regions-of-interest found in the images. Both pairs contain two sequential images.

change in area of the regions-of-interest and the displacement of the robot. This results in a set of K estimations:

$$P = \{\mathbf{p}_k | \mathbf{p}_k = \langle r_k, \theta_k \rangle \wedge 1 \leq k \leq K\}, \quad (7)$$

where r_k and θ_k are respectively the range and bearing of the estimation. The position of the landmark is then determined by the mean of P , and its uncertainty by its covariance matrix.

B. The SLAM system

We use a standard implementation of the Extended Kalman Filter (EKF) as basis of the SLAM system [23]. Our method and results, however, are also valid for other SLAM approaches. We will not discuss the full EKF method, rather only the incorporation of the landmark observations.

A landmark i with descriptor \mathbf{d}_i that results from the buffer is classified as either a new landmark, or a previously observed landmark that is already in the map. It concerns a previously observed landmark if the landmark in the database with the most similar descriptor, \mathbf{d}_j , fulfills three criteria:

- 1) Similarity in descriptors:

$$\|\mathbf{d}_i - \mathbf{d}_j\| < \tau_2 \quad (8)$$

- 2) A small best-to-next-best ratio:

$$\|\mathbf{d}_i - \mathbf{d}_j\| / \|\mathbf{d}_i - \mathbf{d}_l\| < \delta_2 \quad (9)$$

where \mathbf{d}_l is the second most similar descriptor in the database. This is to only match unique landmarks.

- 3) A small distance in the EKF map, measured by the Mahalanobis distance:

$$\sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{S}_j^{-1} (\mathbf{x}_i - \mathbf{x}_j)} < \eta \quad (10)$$

where \mathbf{S}_j is the uncertainty covariance matrix, discussed in the next paragraph.

We obtained good results with $\tau_2 = 0.5$, $\delta_2 = 0.5$, and $\eta = 5.0$ in our experiments.

The landmark is classified as new only if none of the three criteria is fulfilled. For a new landmark, the state matrix and covariance matrix are augmented using the observation, \mathbf{z}_i , and the uncertainty covariance matrix, \mathbf{S}_i , where \mathbf{z}_i is set to mean(P), and \mathbf{S}_i is determined using the uncertainty of the observation, cov(P), and the uncertainty of the robot's position in the EKF. When a landmark is matched with an existing landmark in the database, \mathbf{z}_i and \mathbf{S}_i are used to update the EKF.

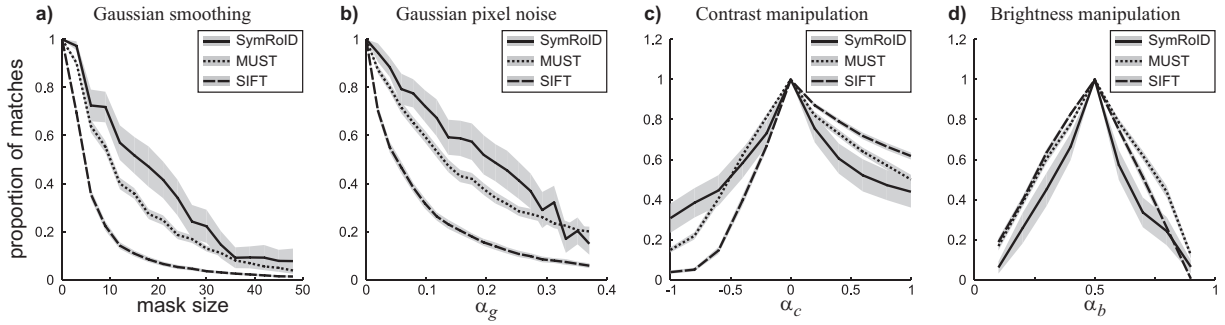


Fig. 4. Robustness to noise and changing light conditions. The lines give the mean proportion of matched regions-of-interest or interest points between the distorted and the original images of 57 runs. The gray areas around the lines depict the 95% confidence intervals on the mean. Note that these intervals for MUST and SIFT are sometimes small, and therefore hardly visible.

IV. EXPERIMENTS AND RESULTS

A. Experimental setup

To test the stability of the symmetrical regions-of-interest, and to test the SLAM performance using SymRoID, we recorded a SLAM database with a Pioneer II DX robot, equipped with a Sony D31 camera. The database contains the camera images and the odometric information of ten different runs, in which the robot drove four laps through an office environment. Each lap was approximately 35 meters, and the robot drove at an average speed of 0.3 m/s. Camera images of 320×240 were stored at 5Hz. At intervals of one meter, the true location of the robot was logged by hand. This enabled us to quantify the SLAM performance.

In the experiments, we tested the performance of SymRoID, and compared it to MUST [10] and SIFT [4].

B. Robustness

In [12] a benchmark for testing the robustness of detectors is presented. Unfortunately, it is not suited for testing landmark selection methods for SLAM in indoor environments since it contains outdoor scenes. We therefore took images from one of the runs in our database with intervals of 3 meters to represent the complete environment. To test the noise robustness of SymRoID, we smoothed the image with a Gaussian kernel and added Gaussian noise to the pixels. We furthermore manipulated the contrast and brightness of the images to test the robustness to changing light conditions.¹

The robustness is measured by the proportion of matching regions between the original and the manipulated images. Two regions match when (5) and (6) are met, where $\tau_1 = 0.6$ and $\delta_1 = 0.75$. Additionally, the distance between the two regions in the image should be less than 3 pixels.

¹The functions used for the manipulations are: (1) *Gaussian pixel noise*: $I'(x, y) = I(x, y) + X(\alpha_g)$, where $I(x, y) \in [0, 1]$ is the intensity of pixel (x, y) , and $X(\alpha_g)$ is a random sample from the normal distribution $N(0, \alpha_g^2)$. (2) *Gaussian smoothing*: $I' = I * G_s$, where G_s is a Gaussian mask of size $s \times s$, with a standard deviation of $\sigma = s/6$. (3) *Contrast manipulation*: $I'(x, y) = I(x, y) + \alpha_c (I(x, y) - \bar{I}_{x,y})$, where $\bar{I}_{x,y}$ is the local average in a neighborhood of 21×21 pixels around pixel (x, y) . The contrast for $\alpha_c > 0$, and decreases otherwise. (4) *Brightness manipulation*: $I'(x, y) = I(x, y)^{\log \alpha_b / \log 0.5}$. For $\alpha > 0.5$, the pixels are brightened, for $\alpha < 0.5$, the pixels are darkened.

The results in fig. 4 show the robustness results. The lines give the mean performance over the 57 runs, with the 95% confidence intervals on the mean given by the gray areas. The symmetry models are significantly less affected by the two types of noise than SIFT (fig. 4a and b). Moreover, using regions-of-interest instead of interest points gives a significant improvement. For the contrast manipulation, SymRoID performs significantly better when there is low contrast. With enhances contrast, on the other hand, SIFT performs better (fig. 4c). Using regions-of-interest results in worse performance for the brightness manipulation. MUST scores best when brightness is enhanced (fig. 4d).

C. SLAM performance

To test the SLAM performance, we calculated the Euclidean distance between the estimated position of the robot by the EKF and the ground-truth position. The estimation error is the average distance in the last of the four lap that the robot drove through the environment.

The parameters for the buffer and for the matching of regions with the landmark ($\tau_1, \tau_2, \delta_1, \delta_2$) are optimized for all three landmark selection methods. The overall best settings for SymRoID are given in section III.

Fig. 5 shows the estimation error. The bars give the mean over the 10 runs, and the error bars depict the 95% confidence intervals. The horizontal dashed line and the horizontal gray bar show the mean and 95% confidence interval of the odometry error. It can be appreciated that the use of symmetry by SymRoID and MUST gives significantly better SLAM performance. Moreover, the use of symmetrical regions-of-interest significantly outperforms the use of interest points. This is true for both the best settings per runs and the overall best settings.

Also computationally, SymRoID outperforms the other models. SIFT selects on average 120 interest points, MUST 40, but SymRoID selects only half a dozen regions per image, thereby greatly reducing the computations in the buffer. Moreover, both SIFT and MUST take in the order of a second to calculate interest points from an image, while SymRoID finds regions about four times as fast. This improvement is due to the coarse scale space used by SymRoID, in contrast with the detailed scale space used by SIFT and MUST.

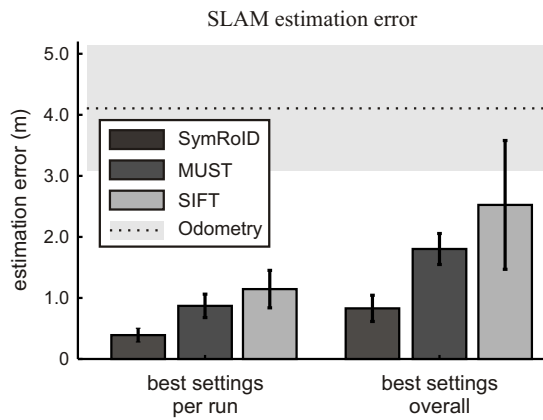


Fig. 5. The slam performance. The bars give the mean estimation error in the last lap through the environment, after a map has been established in the previous three laps. The error bars are the 95% confidence intervals on the mean. The first group of bars give the results where the best buffer and matching settings per run are used, the second group shows the performance for the overall best settings. The horizontal dashed line gives the mean error when only the odometry information is used. The height of the horizontal gray bar represents the 95% interval.

V. DISCUSSION

We presented a symmetrical region-of-interest detector, SymRoID. The model selects regions-of-interest in the image using local symmetry. The model is used to select landmarks for a visual SLAM system. We compared the stability, robustness, and SLAM performance of the model with that of MUST, a symmetrical interest-point detector [10], and SIFT, an interest-point detector based on contrast features [4]. We showed that the use of symmetry improves the stability and robustness to noise, and yields significantly better SLAM performance. Moreover, the use of regions-of-interest outperforms the use of interest points. SymRoID also is more robust to decreased contrast in the image. However, for enhanced contrast and brightness manipulation, the model scores worse than the others.

The improved robustness to noise of the symmetry model can be explained by the fact that symmetrical regions are intrinsically redundant. Furthermore, by using regions, instead of points, more evidence of the existence of symmetry is gathered, making the model more robust. Robustness to noise is an important property, since robots usually operate under noise conditions.

The worse robustness to changing light conditions requires more future research, since it is an important property when a robot needs to operate in an environment over extended periods. Since the contrast and brightness manipulations mainly effect the gradient magnitudes, we believe that the robustness of the model can be increased by adjusting the role of m_i and n_j in 1.

Concluding, the use of symmetry for landmark selection has shown to improve SLAM performance.

REFERENCES

- [1] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, Massachusetts: The MIT Press, 2005.
- [2] M. Giovanni, "Attentional landmarks and active gaze control for visual slam," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1054–1065, 2008.
- [3] O. M. Mozos, A. Gil, M. Ballesta, and O. Reinoso, "Interest point detectors for visual slam," in *Lecture Notes in Computer Science*, D. Borrajo, L. Castillo, and J. M. Corchado, Eds. Springer-Verlag, 2008, vol. 4788, pp. 170–179.
- [4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [5] S. Se, D. G. Lowe, and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *International Journal of Robotics Research*, vol. 21, no. 8, pp. 735–758, 2002.
- [6] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," in *ninth European Conference on Computer Vision (ECCV)*, Graz, Austria, 2006.
- [7] A. C. Murillo, J. J. Guerrero, and C. Sagues, "Surf features for efficient robot localization with omnidirectional images," in *IEEE International Conference on Robotics and Automation (ICRA)*, Rome, Italy, 2007, pp. 3901–3907.
- [8] A. J. Davison and D. Murray, "Simultaneous localization and map-building using active vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 865–880, 2002.
- [9] P. Moreels and P. Perona, "Evaluation of features detectors and descriptors based on 3d objects," *International Journal of Computer Vision*, vol. 73, no. 3, pp. 263–284, 2007.
- [10] G. Kootstra, S. de Jong, and L. R. B. Schomaker, "Using local symmetry to select landmarks for visual slam," in *International Conference on Vision Systems (ICVS)*, Liège, Belgium, 2009.
- [11] S. Frintrop and P. Jensfelt, "Attentional landmarks and active gaze control for visual slam," *IEEE Transactions on Robotics*, vol. 24, no. 5, 2008.
- [12] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1/2.
- [13] G. Kootstra, A. Nederveen, and B. de Boer, "Paying attention to symmetry," in *British Machine Vision Conference (BMVC2008)*, M. Everingham, C. Needham, and R. Fraile, Eds., Leeds, UK, 2008, pp. 1115–1125.
- [14] G. Kootstra and L. R. B. Schomaker, "Prediction of human eye fixations using symmetry," in *Cognitive Science Conference (CogSci)*, Amsterdam, Netherlands, 2009.
- [15] S. E. Palmer and K. Hemenway, "Orientation and symmetry: Effects of multiple, rotational, and near symmetries," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 4, no. 4, pp. 691–702, 1978.
- [16] F. L. Royer, "Detection of symmetry," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 7, no. 6, pp. 1186–1210, 1981.
- [17] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [18] G. Marola, "Using symmetry for detecting and locating objects in a picture," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 179–195, 1989.
- [19] G. Backer, B. Mertsching, and M. Bollmann, "Data- and model-driven gaze control for an active-vision system," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1415–1429, 2001.
- [20] G. Sela and M. D. Levine, "Real-time attention for robotic vision," *Real-Time Imaging*, vol. 3, pp. 173–194, 1997.
- [21] D. Reifeld and Y. Yeshurun, "Preprocessing of face images: Detection of features and pose normalization," *Computer Vision and Image Understanding*, vol. 71, no. 3, pp. 413–430, 1998.
- [22] D. Reifeld, H. Wolfson, and Y. Yeshurun, "Context free attentional operators: The generalized symmetry transform," *International Journal of Computer Vision*, vol. 14, pp. 119–130, 1995.
- [23] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part i," *Simultaneous Localization and Mapping: Part I*, vol. 13, no. 2, pp. 99–108, 2006.