# Extracting Minimalistic Corridor Geometry from Low-Resolution Images

Yinxiao Li⋆, Vidya N. Murali⋆, and Stanley T. Birchfield

Department of Electrical and Computer Engineering,
Clemson University, USA
{yinxial,vmurali,stb}@clemson.edu

**Abstract.** We propose a minimalistic corridor representation consisting of the orientation line (center) and the wall-floor boundaries (lateral limit). The representation is extracted from low-resolution images using a novel combination of information theoretic measures and gradient cues. Our study investigates the impact of image resolution upon the accuracy of extracting such a geometry, showing that accurate centerline and wall-floor boundaries can be estimated even in texture-poor environments with images as small as $16 \times 12$. In a database of 7 unique corridor sequences for orientation measurements, less than 2% additional error was observed as the resolution of the image decreased by 99%. One of the advantages of working at such resolutions is that the algorithm operates at hundreds of frames per second, or equivalently requires only a small percentage of the CPU.

**Keywords:** Low-Resolution, Robot Navigation, Geometry Estimation.

## 1 Introduction

Psychological studies have shown that while driving or walking human beings tend to focus their eye gaze on the direction of the goal and also along tangent points in roadways/hallways to steer toward. Land and Tatler in their classic paper [5] speak of the tendency of race car drivers to steer along the direction of the goal while allowing the bend points (tangent points) to hold their gaze intermittently to judge the future steering angle. In indoor environments, this phenomenon loosely translates into pedestrians having a tendency to look near wall-floor boundaries when nearing a corner.

While working on visual sensing for robot navigation/exploration, the question of what resolution is sufficient for basic navigation tasks is an inherent question. Psychological studies have shown that the human visual system does not require high-resolution images to ascertain information about the environment for basic navigation. The "selective degradation hypothesis", developed by Leibowitz [7], states that some visual abilities such as vehicle steering and speed control remain relatively easy despite loss in visual acuity and color vision. Torralba *et al.* [13,14] in their recent work have presented convincing psychovisual evidence that $32 \times 24$ bit images are sufficient for humans to successfully performs basic scene classification, object segmentation, and

---

⋆ Indicates equal contribution.

| $320 \times 240$ | $160 \times 120$ | $64 \times 48$ | $32 \times 24$ | $16 \times 12$ |

**Fig. 1.** A typical corridor image at five different resolutions. Even at $32 \times 24$ resolution, it is easy to see the structure of the corridor. For display purposes, the downsampled images were upsampled to ensure that they are all at the same size.

identification. The work of Tovar *et al.* [15] and O'Kane and LaValle [11] focuses on identifying simple and generic solutions to basic robot exploration/navigations tasks, with emphasis on minimalistic representations.

Motivated by this idea, we describe a system to develop a minimalistic structural representation of an indoor corridor for basic navigation tasks by a mobile robot. We also show that the recovered critical steering information does not degrade significantly with reduction in resolution. The corridor is represented by three lines: the center of the corridor, the left wall-floor boundary, and the right wall-floor boundary. While in theory the center line is redundant (it is simply the vertical line passing through the intersection of the two wall-floor boundary lines), it is helpful to keep it distinct, particularly when the robot is not looking straight down the corridor. To detect this representation, we combine and extend two previous approaches. In previous work, Murali and Birchfield presented the use of ceiling lights for determining the center of the corridor using low-resolution techniques [10]. We extend this work by using other metrics like maximum entropy and maximum symmetry to estimate the center of the corridor when ceiling lights are not visible. Li and Birchfield [8] deveoped a method for extracting wall-floor boundaries for indoor environments that is much more computationally efficient than previous approaches to perform geometric reasoning on images [4,6]. We extend this work by detecting wall-floor boundaries of typical indoor office environments using low-resolution images. Fig. 1 shows a typical indoor corridor image varying from high resolution to low resolution, from which it can be seen that structural information in the scene is visually discernible even at very low resolutions.

## 2   Orientation Line Estimation

We model the structure of a corridor by three lines in the image. A vertical line indicates the orientation line, or centerline, of the corridor, which passes through the vanishing point. The wall-floor boundary is then captured by two diagonal lines that meet at the same point on the orientation line. (As mentioned above, the vertical line is redundant in theory but helps in the case when both diagonal lines are not visible.) Our approach consists of two steps: First we estimate the orientation line in the image by combining multiple cues, then we estimate the wall-floor boundary.

In this section we describe the orientation line estimation. Our approach is adapted from the work done by Murali and Birchfield [10], which uses the median of bright pixels (ceiling lights), maximum entropy, and maximum symmetry measures in the

corridor image to determine its center and therefore the orientation. This approach has several advantages over existing techniques: It is simple, computationally efficient, and yields good results even for low-resolution images.

### 2.1 Median of Bright Pixels

The ceiling lights, which are usually symmetric with respect to the main corridor axis, provide an important cue. When lights are not in the center of the corridor, we can use the $k$-means algorithm to overcome this difficulty, where $k = 2$. The median horizontal position of the brighter of the two regions is calculated, yielding an estimate of the center position. In order to overcome specular reflections of the walls, we use Ullman's formula for local contrast [16]. The horizontal coordinate is transformed to an angle by applying the same scalar factor using the equation $f_l(I) = \psi(\text{med}\{x : (x, y) \in \mathcal{R}_{bright}\})$, where $I$ is the image, $\mathcal{R}_{bright}$ is the set of bright pixels, med is the median, and $\psi = \alpha x$ converts from pixels to degrees, where the factor $\alpha$ is determined empirically. Sample results of orientation line estimation using bright pixels are shown in Fig. 2.
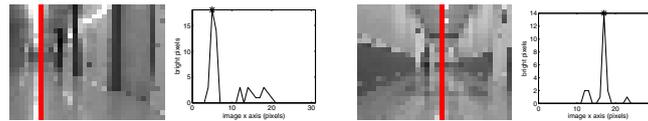


**Fig. 2.** Variation of bright pixels and corresponding orientation estimate in corridor images with resolution of $32 \times 24$

### 2.2 Maximum Entropy

Empirically, entropy is maximum when the camera is pointing down the corridor [10]. The reason for this perhaps surprising result is that such an orientation causes scene surfaces from a variety of depths to be visible, yielding an increase of image information at this orientation. A similar observation has been noted by other researchers in the context of using omnidirectional images [2,3]. We divide the image into overlapping vertical slices and computing the graylevel entropy of the image pixels in each slice. The maximum entropy along the horizontal axis is then used to estimate the orientation. Sample results of orientation line estimation using entropy are shown in Fig. 3.
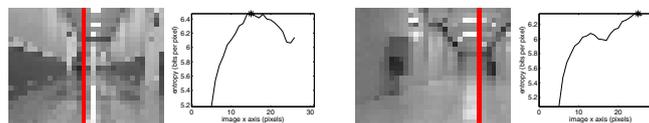


**Fig. 3.** Variation of entropy and corresponding orientation estimate in corridor images with resolution of $32 \times 24$

### 2.3   Symmetry by Mutual Information

Another important feature of corridors is symmetry. One easy way to find symmetry is to compare the two regions using mutual information by calculating entropy. As with entropy, for each horizontal coordinate $x$ a column of pixels $\mathcal{C}(x)$ is considered. The column is divided in half along its vertical center into two columns $\mathcal{C}_L(x)$ and $\mathcal{C}_R(x)$. The normalized graylevel histograms of these two regions are used as the two probability mass functions (PMFs), and the mutual information between the two functions is computed:

$$MI(x) = \sum_{v \in \mathcal{V}} \sum_{w \in \mathcal{V}} p(v, w) \log \frac{p(v, w)}{p_L(v)p_R(w)}, \tag{1}$$

where $p(v, w)$ is the joint PMF of the intensities, and $p_L(v)$ and $p_R(w)$ are the PMFs computed separately of the intensities of the two sides. As before, the orientation estimate is given by $f_s(I) = \psi(\arg\max_x MI(x))$. Sample results of orientation line estimation using mutual information are shown in Fig. 4.
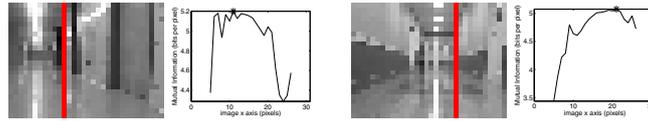


**Fig. 4.** Variation of symmetry and corresponding orientation estimate in corridor images with a resolution of $32 \times 24$

### 2.4   Combining the Metrics

We combine the estimates as a weighted average: $f(I) = \alpha_l(I)f_l(I) + \alpha_h(I)f_h(I) + \alpha_s(I)f_s(I)$. Because of the reliability of the bright pixels, we set $\alpha_l = 0.8$, $\alpha_h = \alpha_s = 0.1$. An example result obtained for different resolutions is shown in Fig. 5.
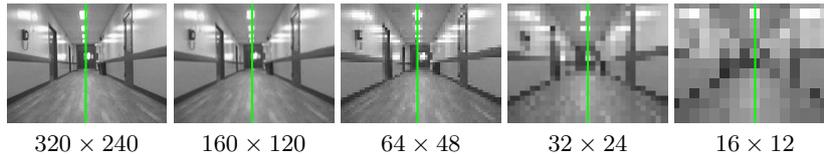


| $320 \times 240$ | $160 \times 120$ | $64 \times 48$ | $32 \times 24$ | $16 \times 12$ |

**Fig. 5.** The orientation line estimate (vertical green line) for the images shown in Fig. 1. The results remain essentially unchanged from the original resolution down to a resolution of $16 \times 12$

## 3   Wall-Floor Boundary

We use the idea of the floor segmentation method introduced by Li and Birchfield [8] which has been shown to be fairly robust to reflections on the floor. For the seven different resolutions, we compute the minimum acceptance length of the horizontal line segments $l_h$ as $l_h = \log \eta d$, where $d = \sqrt{w^2 + h^2}$ is the length of the diagonal of the image, $w$ and $h$ are the width and height of the image, respectively, $\eta = 5$ is a scaling factor, and log is the natural logarithm.

According to the floor segmentation method [8], there are three different scores (structure score, homogeneous score, and bottom score) that contribute to the final wall-floor boundary detection. When applying the method to different resolutions, we noticed the structure score always shows the best accuracy, while the bottom score always fails when decreasing the resolution. Therefore, we adapt the weights for the three scores according to the resolution so that $\Phi_{total}(\ell_h)$ is relatively high for line segments near the wall-floor boundary. At the same time, when combining with the orientation line, we compute the intersection of the orientation line and the wall-floor boundary, which is considered as the vanishing point. Then we apply the line-fitting algorithm to both half wall-floor boundaries separated by the vanishing point. Using the slopes and the computed vanishing point, it is easy to find the two terminal points on the image border. Finally, we connect the vanishing point, two terminal points, as well as the orientation line and obtain the structure of the corridor. The sample results are shown in Fig. 10 and the second row of Fig. 6.
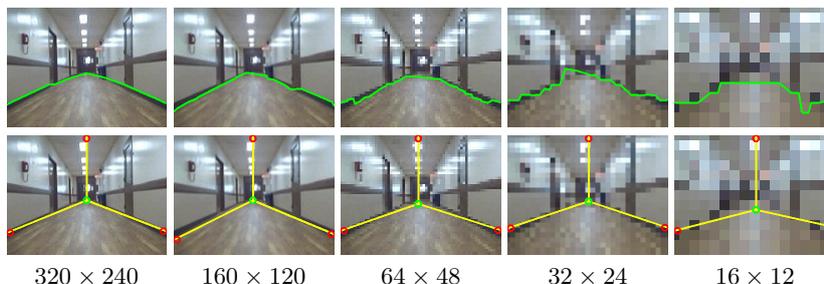


| $320 \times 240$ | $160 \times 120$ | $64 \times 48$ | $32 \times 24$ | $16 \times 12$ |

**Fig. 6.** TOP: The wall-floor boundary found by the algorithm described in [8] for the different resolution images of Fig. 1. The accuracy degrades slightly until the resolution of $32 \times 24$, after which the errors become more pronounced. BOTTOM: The three-line model estimate of the corridor found by combining the orientation line with the wall-floor boundary, on the same images. As before, the structure of the corridor remains intact even in the resolution of $32 \times 24$, with only slight errors visible in $16 \times 12$.

## 4 Experimental Results

For orientation, we collected data for 4 different buildings, 8 unique corridors (1 training + 7 for testing). For every unique corridor, at equally spaced intervals along the corridor (4.5m), we rotated the robot from $-20°$ to $+20°$ and collected corresponding odometry (heading), laser readings (span of $-90°$ to $+90°$) and images. We ran the entropy detector, light detector, and symmetry detector on the images and compared with ground truth (odometry and/or laser). Since a linear relationship exists between the detected pixel location corresponding to the center of the corridor and the robot orientation as explained in previous sections, we use either the estimate $f_l$ or $(f_h + f_s)/2$.

For wall-floor boundary and corridor reconstruction, we collected data for 11 distinct corridors in 6 different buildings. We drove the robot three times (middle, left, right separated by 0.5 m) along each corridor and collected images along with their
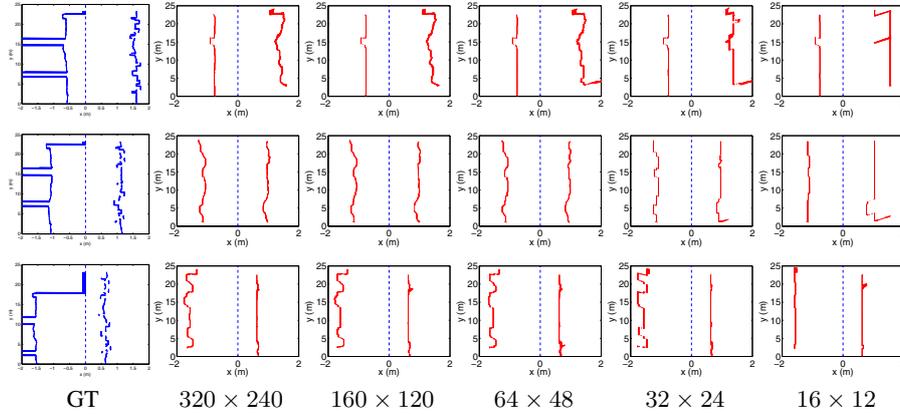
| GT | $320 \times 240$ | $160 \times 120$ | $64 \times 48$ | $32 \times 24$ | $16 \times 12$ |

**Fig. 7.** Corridor structure reconstruction from the wall-floor boundary, displayed as a top-down view. The first column in blue shows the ground truth location of the walls (Cartesian conversion of polar laser readings), and the next 5 columns in red show the reconstruction results from the wall-floor boundaries on different resolution images. Each row represents a different run of the robot in the same corridor, with the robot placed at a different lateral position in the corridor for each run.

corresponding laser readings ($-90°$ to $+90°$ sweep). The position of the orientation line with respect to the wall-floor boundaries gives the lateral position in the corridor. The distance between the two end-points in the wall-floor boundary yields the width of the corridor (in pixels). We use a homography obtained during a calibration process to transform to world coordinates. Several examples of the recovered corridor structure are shown in Fig. 7. The robot's position was determined using an image to top-down view calibration procedure utilizing a homography between a square pattern on the floor of the corridor and the image of its four corners. Ground truth was provided by laser readings, which were converted from polar to Cartesian coordinates to yield a top-down measurement of the corridor for every image in the sequence. Using the detected wall-floor boundary and applying a homography to get the top-down structure of the corridor and the lateral position, we achieved the reconstruction shown in Fig. 7. The Normalized Root Mean Square Error (NRMSE) between ground truth (laser) and predicted values of estimated corridor width and lateral position was calculated for each of the 7 resolutions considered ($320 \times 240$, $160 \times 120$, $80 \times 60$, $64 \times 48$, $32 \times 24$, $16 \times 12$ and $8 \times 6$), for three trials in a corridor. The results are shown in Fig. 8.

The parameters for a linear fit between the location of the orientation line and predicted orientation were estimated by using one of the corridors as a training set. Using the trained parameters, the orientations for all the other data for the remaining 7 test corridors ($\theta$) were predicted from the mean pixel locations using the above equations. The Normalized Root Mean Square Error between ground truth (laser) and predicted values of heading was calculated for each of the 7 resolutions considered. The results are shown in Fig. 9.

In Fig. 10 we show a variety of corridors in which our results are successful, including one without ceiling lights, where the result is based only on maximum symmetry
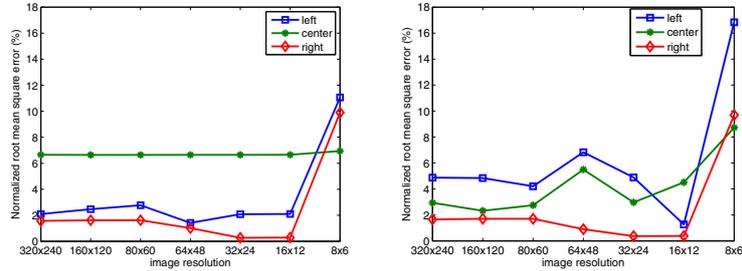
**Fig. 8.** LEFT: NRMSE for estimating the lateral position of the robot for three runs in a single corridor. The structure was accurately captured in all three cases. RIGHT: Mean NRMSE for the estimation of the corridor width. There is not much difference in estimation error rates across the different resolutions, and in fact the error drops in some cases for $32 \times 24$ and $16 \times 12$ sizes due to the removal of noise and artifacts by downsampling.
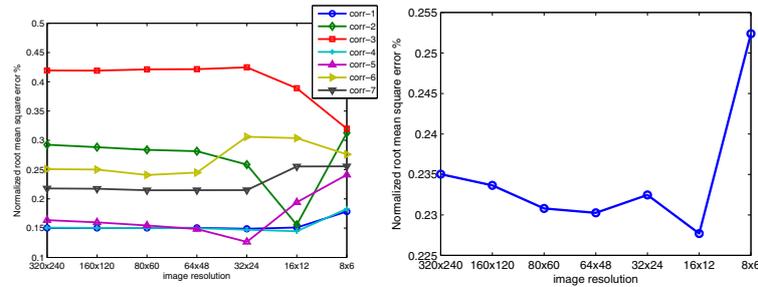


**Fig. 9.** Minimalistic geometric information is obtained by the algorithm at very low image resolutions. LEFT: NRMSE for all 7 test corridors. RIGHT: Mean NRMSE for the corridors. The orientation estimation error remains relatively stable across different image resolutions. In fact, the error drops for a few corridors at $32 \times 24$ and $16 \times 12$, primarily due to the fact that downsampling removes artifacts such as reflections and other noise on the walls and floor.

and maximum entropy.[1] In addition, empirically we found that when approaching the end of a corridor, the angle between the two wall-floor boundary lines increases toward $180°$, thus providing some indication of the distance to the end of the corridor. Fig. 11 shows normalized error in continuous frames with resolutions of $320 \times 240$ and $32 \times 24$. We also found that entropy decreases sharply when the robot reaches the end of the corridor [10]. Therefore, these could be used to detect the end of a corridor.

Estimating the pose of the robot or the orientation of the robot in a typical indoor corridor is one of the necessary tasks for robot exploration/navigation. While many authors have approached this problem by estimating vanishing points in a corridor [1,9,12], we have discovered that the approach of clustering detected lines performs poorly in low-resolution and textureless images because lines are not easily detected in such images. A more recent approach by Kong *et al.* [4] approaches the problem similarly but uses

---

[1] See http://www.ces.clemson.edu/~stb/research/
minimalistic_corridor for videos of the results.

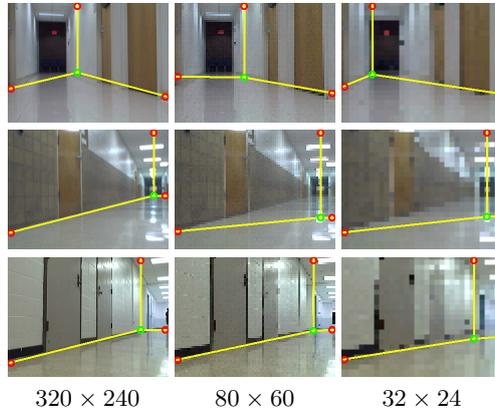$$320 \times 240 \qquad 80 \times 60 \qquad 32 \times 24$$

**Fig. 10.** Additional results for other corridors, including one without ceiling lights
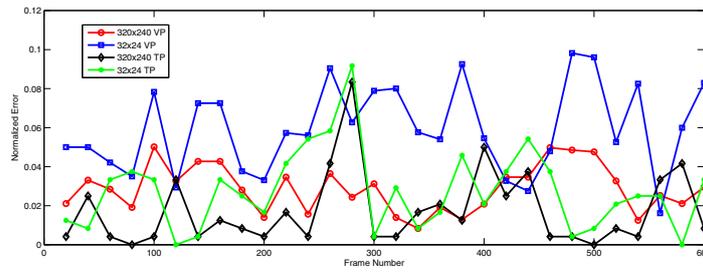


**Fig. 11.** Normalized error of vanishing point (VP) and terminal points (TP) in Sirrine Hall video sequence compared with ground truth. Results do not change significantly when 99% of the pixels are discarded.
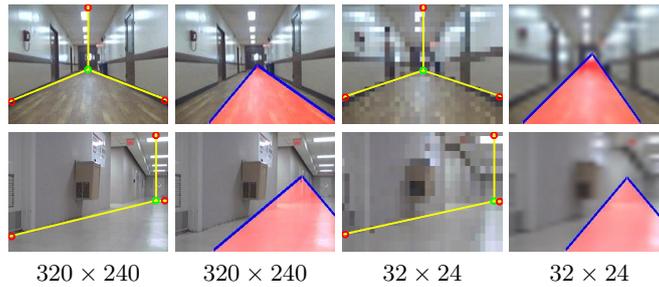


$$320 \times 240 \qquad 320 \times 240 \qquad 32 \times 24 \qquad 32 \times 24$$

**Fig. 12.** Comparison between our results (three yellow lines) and those of Kong *et al.* [4] (pink region). Our algorithm achieves more accurate estimation of both the orientation line and the wall-floor boundary in indoor scenes, particularly at low resolutions.

texture orientation rather than explicit line detection. In their approach, Gabor filters yield texture estimates, and an adaptive voting scheme allows pixels to decide the confidence of an orientation. Not only is their approach much more computationally intensive than ours, but with indoor low-resolution images the results are significantly less accurate. See Fig. 12 for some examples.

## 5    Conclusion

We have proposed an algorithm to extract a minimalistic geometric representation of a typical indoor corridor environment using low resolution images. Motivated by the "selective degradation hypothesis", our approach exploits the redundancy of image information in order to extract useful information for mobile robotic tasks with minimal processing. Our algorithm combines two ideas: extracting the wall-floor boundary by combining intensity edges and specular reflection removal, and centerline estimation using a combination of information theoretic cues and bright pixel estimation. Previous approaches for these problems have been extended and modified to facilitate low resolution processing. The proposed algorithm was tested on images from several different corridors, showing that the accuracy of the estimation of the orientation line or corridor geometry changed very little even when more than 99% of the original information was discarded by downsampling the image to an extremely low resolution. Our approach can be seen as an exploration into identifying how much information is needed for basic mobile robot tasks such as corridor exploration and navigation. By reducing the resolution required for these basic tasks, the CPU time is freed for other tasks that potentially require higher resolutions and more involved processing.

## References

1. Bayramoglu, E., Andersen, N., Kjolstad Poulsen, N., Andersen, J., Ravn, O.: Mobile robot navigation in a corridor using visual odometry. In: International Conference on Advanced Robotics, pp. 1–6 (June 2009)
2. Bonev, B., Cazorla, M., Escolano, F.: Robot navigation behaviors based on omnidirectional vision and information theory. Journal of Physical Agents 1(1), 27–35 (2007)
3. Escolano, F., Bonev, B., Suau, P., Aguilar, W., Frauel, Y., Saez, J., Cazorla, M.: Contextual visual localization: cascaded submap classification, optimized saliency detection, and fast view matching. In: IEEE International Conference on Intelligent Robots and Systems (2007)
4. Kong, H., Audibert, J.Y., Ponce, J.: General road detection from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition (2009)
5. Land, M.F., Tatler, B.W.: Steering with the head: The visual strategy of a racing driver. Current Biology 11(15), 1215–1220 (2001)
6. Lee, D., Hebert, M., Kanade, Y.: Geometric reasoning for single image structure recovery. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR (2009)

7. Leibowitz, H.W., Rodemer, C.S., Dichgans, J.: The independence of dynamic spatial orientation from luminance and refractive error. Perception & Psychophysics 25(2), 75–79 (1979)

8. Li, Y., Birchfield, S.T.: Image-based segmentation of indoor corridor floors for a mobile robot. In: IROS (2010)

9. McLean, G., Kotturi, D.: Vanishing point detection by line clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 1090–1095 (1995)

10. Murali, V.N., Birchfield, S.T.: Autonomous exploration using rapid perception of low-resolution image information. Autonomous Robots 32(2), 115–128 (2012)

11. O'Kane, J.M., LaValle, S.M.: Almost-sensorless localization. In: Proc. IEEE International Conference on Robotics and Automation (2005)

12. Segvic, S., Ribaric, S.: Determining the absolute orientation in a corridor using projective geometry and active vision. IEEE Transactions on Industrial Electronics 48(3), 696–710 (2001)

13. Torralba, A.: How many pixels make an image? Visual Neuroscience 26(01), 123–131 (2009)

14. Torralba, A., Fergus, R., Freeman, W.T.: 80 million tiny images: A large data set for non-parametric object and scene recognition. IEEE TPAMI 30(11), 1958–1970 (2008)

15. Tovar, B., Guilamo, L., Lavalle, S.M.: Gap navigation trees: Minimal representation for visibility-based tasks. In: Proceedings of the Workshop on the Algorithmic Foundations of Robotics, pp. 11–26 (2004)

16. Ullman, S.: On visual detection of light sources. Biological Cybernetics 21, 205–212 (1976)