

Torchlight Navigation

Michael Felsberg
Linköping University
mfe@isy.liu.se

Fredrik Larsson
Linköping University
larsson@isy.liu.se

Wang Han
Nanyang Technological
University
hw@ntu.edu.sg

Anders Ynnerman
Linköping University
andyn@itn.liu.se

Thomas B. Schön
Linköping University
schon@isy.liu.se

Abstract—A common computer vision task is navigation and mapping. Many indoor navigation tasks require depth knowledge of flat, unstructured surfaces (walls, floor, ceiling). With passive illumination only, this is an ill-posed problem. Inspired by small children using a torchlight, we use a spotlight for active illumination. Using our torchlight approach, depth and orientation estimation of unstructured, flat surfaces boils down to estimation of ellipse parameters. The extraction of ellipses is very robust and requires little computational effort.

Keywords—torchlight, pose estimation, active illumination, plane estimation, ellipses;

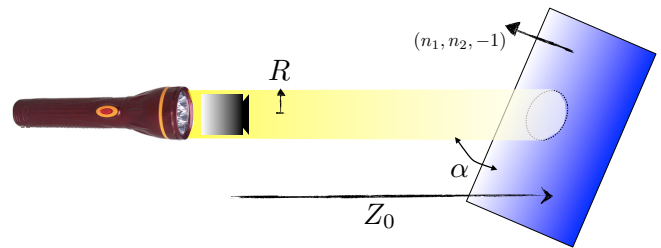
I. INTRODUCTION

Controlled illumination for computer vision is a well known technique for solving hard vision problems or achieving high accuracy. Examples include estimation of depth maps using structured light [1], range cameras using sheets of light [2], shape from shading [3], and BDRF estimation [4]. Humans also use active illumination for analysing the depth-structure of a scene, e.g. small children using a torchlight (flashlight in AE). However, to the best of our knowledge, simple torchlights have not been used for computer vision so far.

A common computer vision task is navigation and mapping. Many indoor navigation tasks require depth knowledge of flat, unstructured surfaces (walls, floor, ceiling). With passive illumination only, this is an ill-posed problem, but using our torchlight approach, it becomes straightforward. Since a camera is a projective sensor, the illumination source must either be located at some distance from the camera, or the emitted light must be bundled by a mirror, displacing the virtual locus of the light source.

A potential field of application is robot navigation in (partly) collapsed buildings, where no accurate maps are available, no or bad illumination forces the robot to carry along its own light source, and the floor, walls, and ceiling might be covered by dust. Algorithms must be robust under these circumstances. Processing must be simple and fast, since the computations must be performed onboard and resolution might be poor. Our torchlight approach fulfills these requirements.

Active illumination as used in the literature, mostly makes use of points and lines (laser, grids, etc), because this is easy to analyse in geometric terms. Due to the small spatial



$$Z_0 = \frac{Rfa}{b^2}$$

$$\alpha = \tan^{-1} \sqrt{n_1^2 + n_2^2} \quad \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = -\frac{f}{b^2} \mathbf{z}_0$$

Figure 1. The relative pose between the camera and a planar surface can be obtained by fitting an ellipse to the projection of the light beam. The final equations needed for the estimation can be seen below the illustration, see section 2 for nomenclature.

support of these light-patterns, they are deemed to be brittle when it comes to rough surfaces or occlusions. Using a light-beam from a torchlight is probably more robust and the resulting pattern on a flat surface (a filled ellipse) is easy to analyse.

We propose a simple but robust method for estimating 3D plane parameters from a single perspective view of the light-beam reflection, see Fig. I. The algorithm consists of three steps: boundary extraction, ellipse fitting, and plane parameter computation. The method is tested in a setting with low-cost equipment, consisting of a rechargeable spotlight and a laptop webcam.

The paper is structured as follows: In the second section, we give the formulation of ellipses in terms of Fourier descriptors, derive the geometry of the projected light beam and the estimation equations for the plane parameters, and describe the experimental setup. The results are documented in the third section and the paper is concluded with a discussion of the results.

II. METHODS

Since Fourier descriptors [5] are complex valued and in order to simplify the algebraic expressions, all in-plane coordinates are represented using complex numbers.

A. Ellipses and Fourier Descriptors

Let \mathbb{C} denote the set of complex numbers. Then an ellipse $\mathcal{E} \subset \mathbb{C}$ is defined by two foci $f_1 \in \mathbb{C}$ and $f_2 \in \mathbb{C}$, such that all ellipse points $z \in \mathcal{E}$ fulfill the property

$$|z - f_1| + |z - f_2| = 2a, \quad (1)$$

for some $a \in \mathbb{R}$. The distance between the foci is $2c$.

If we rotate the ellipse such that its principal axes are aligned with the coordinate system (this is achieved by multiplying with $w = (f_2 - f_1)/2c$), we obtain

$$\frac{\operatorname{Re}\{w(z - z_0)\}^2}{a^2} + \frac{\operatorname{Im}\{w(z - z_0)\}^2}{b^2} = 1, \quad (2)$$

where $b = \sqrt{a^2 - c^2}$ and $z_0 = (f_1 + f_2)/2$ is the center of the ellipse.

An ellipse can be parametrized by one angle using two complex exponentials

$$\mathcal{E} = \{z | z = \bar{w} \frac{a+b}{2} e^{i\phi} + w \frac{a-b}{2} e^{-i\phi} + z_0\}, \quad (3)$$

since

$$\frac{a+b}{2} e^{i\phi} + \frac{a-b}{2} e^{-i\phi} = a \cos \phi + ib \sin \phi. \quad (4)$$

That means that the Fourier descriptors of an elliptic contour are all zero except for the frequencies $\{-1, 0, 1\}$ where the DC part is given by the center z_0 . If the extracted contour deviate from an ideal ellipse, restricting the Fourier descriptors to these frequencies corresponds to minimizing the quadratic error of the contour according to Bessel's inequality and Parseval's theorem [6].

B. Projected Cylinder Sections

From a geometric point of view, there is no difference between a stereo rig and a rig carrying a camera and a torchlight - the geometry is the same. One implication is that estimates becomes more accurate with increasing baseline. We achieve this by bundling the light with a mirror and thus displacing the effective center of the light source.

We assume a parabolic mirror in the torchlight, which results in a collinear lightbeam [7]. The collinear light beam is modeled as a light cylinder of radius R which shares its axis with the optical axis of a perspective camera. The coordinate system is placed in the optical center with Z being the optical axis. The focal length is denoted by f .

The light cylinder is hence given as

$$L(X, Y, Z) = \begin{cases} 1 & X^2 + Y^2 \leq R^2 \\ 0 & X^2 + Y^2 > R^2 \end{cases}. \quad (5)$$

The light is reflected by a plane \mathcal{P} at distance Z_0 parametrized over (X, Y) as

$$Z(X, Y) = n_1 X + n_2 Y + Z_0, \quad (6)$$

such that $(n_1, n_2, -1)$ is normal to the plane. This vector and Z_0 fully parametrize the plane in 3D space. We assume that the reflectance does not decay with the angle to the normal vector.

The image of the reflected light is obtained by computing the projection (note the complex parameterization of the image plane $z = x + iy$)

$$z = \frac{(X + iY)f}{Z(X, Y)} = \frac{(X + iY)f}{\operatorname{Re}\{\bar{n}(X + iY)\} + Z_0}, \quad (7)$$

where $n = n_1 + in_2$. It is not trivial to see that (7) is an ellipse if (X, Y) are points of a circle. Using homogeneous coordinates, one can derive the corresponding conic [8], p.59

$$C = \begin{pmatrix} \frac{Z_0^2}{R^2} - n_1^2 & -n_1 n_2 & f n_1 \\ -n_1 n_2 & \frac{Z_0^2}{R^2} - n_2^2 & f n_2 \\ f n_1 & f n_2 & -f^2 \end{pmatrix}. \quad (8)$$

Which is easy to verify by

$$\begin{pmatrix} \operatorname{Re}\{z\} & \operatorname{Im}\{z\} & 1 \end{pmatrix} C \begin{pmatrix} \operatorname{Re}\{z\} \\ \operatorname{Im}\{z\} \\ 1 \end{pmatrix} = 0, \quad (9)$$

for points on the cylinder, such that $X^2 + Y^2 = R^2$.

From C we compute the ellipse parameters according to [9] that are required for identifying (3). For C describing an ellipse, we require $Z_0^2 > R^2 |n|^2$. We get the center of the ellipse (DC Fourier descriptor) as

$$z_0 = \frac{fn}{|n|^2 - \frac{Z_0^2}{R^2}} \quad (10)$$

The orientation of the major axis is given by the angle of z_0 : $\bar{w} = z_0 / |z_0| = n / |n|$. The coefficients a and b in (2) are given as

$$a = \frac{fZ_0}{\left(\frac{Z_0^2}{R^2} - |n|^2\right)R}, \quad (11)$$

$$b = \frac{f}{\sqrt{\frac{Z_0^2}{R^2} - |n|^2}}. \quad (12)$$

This can be verified by setting $X + iY = R \exp(i\theta)$ in (7) and plugging z , z_0 , w , a , and b into (2).

C. Estimation of Plane Parameters

Ellipse parameters can be estimated in many different ways, three of which are

- Fourier descriptors (see Section II-A)
- Using four points [10]
- Using statistical moments (covariance matrix) [9]

The Fourier descriptor method results in a least-squares solution as the covariance method does, but in addition to the latter, a measure of the deviation from the ellipse model is obtained. The quotient of the ellipse energy by the total descriptor energy determines the confidence of a correct estimate.

In all three cases, the first step is to extract the parameters a, b and w . Let z_k denote the Fourier descriptor with frequency $k \in \mathbb{Z}$. For the special geometry with identical axis, w is obtained by the angle of z_0 . Next, we obtain from (3)

$$a = w(z_1 + z_{-1}) , \quad (13)$$

$$b = w(z_1 - z_{-1}) . \quad (14)$$

Using the four-point method, a, b and w are usually given directly. The second step is to compute Z_0 and n . From (11) and (12) we obtain

$$\frac{fa}{b^2} = \frac{Z_0}{R} , \quad (15)$$

such that the depth estimate is given as

$$Z_0 = \frac{Rfa}{b^2} . \quad (16)$$

The angle between the plane normal and the optical axis is given as

$$\alpha = \tan^{-1}(|n|) , \quad (17)$$

where we estimate the steepness $|n|$ by plugging (12) into (10), giving $z_0 = -nb^2/f$, and thus

$$n = -\frac{f}{b^2} z_0 . \quad (18)$$

The signal to noise ratio for measuring the distance (16) becomes better for smaller f and R . On the other hand, (18) can be rewritten as

$$n = -\frac{z_0 Z_0}{Ra} , \quad (19)$$

which means that the signal to noise ratio for measuring the steepness becomes better for larger R and smaller Z_0 . In total, this means that for a certain light beam radius R , a small focal length and small distances should be used in order to obtain good results.

D. Experimental Setup

In our experiments, we use an off-the-shelf rechargeable spotlight, see Fig. 2. It would be desirable to have a light with a parabolic mirror to generate collinear light and avoid spherical aberration [7], but the conical mirror in our light still produces acceptable results. The radius R is measured to be 60 mm.

A more severe problem that occurs is the anisotropy of the resulting light-beam. This is presumably caused by the shape of the filament in the halogen lamp. This anisotropy



Figure 2. Experimental setup: rechargeable spotlight and a webcam glued to the center of its frontglass.

is modeled by varying the radius with the angle θ . The light beam is not assumed to be circular, but elliptic. Hence

$$R(\theta) = R(\cos(\theta - \theta_0) + \alpha \sin(\theta - \theta_0)) . \quad (20)$$

The angle θ_0 is the measured orientation of the filament. The eccentricity given in terms of the aspect ratio factor α is calibrated from an image sequence.

As a camera, we use a standard laptop webcam with a resolution of 640×480 pixels. This camera is glued to the front glass of the spotlight, see Fig. 2. We calibrate this camera using the OpenCV calibration tool [11]. The resulting focal lengths are 557.8 respectively 554.1 pixels and the center is at 314.8 respectively 235.8, in x and y direction respectively.

We placed the torchlight in front of a blue poster wall at distances of 25 cm, 30 cm, 35 cm, 40 cm, and 50 cm, see Fig. 6. The poster wall is rotated around the vertical axis using a manual turntable with angles from 0° to 70° in eight steps. However, no knowledge of this restriction has been used in the general estimation equation (18). For each distance and angle, at least 90 images were taken.

Videos showing experiments can be found at <http://www.cvl.isy.liu.se/research/torchlight/> .

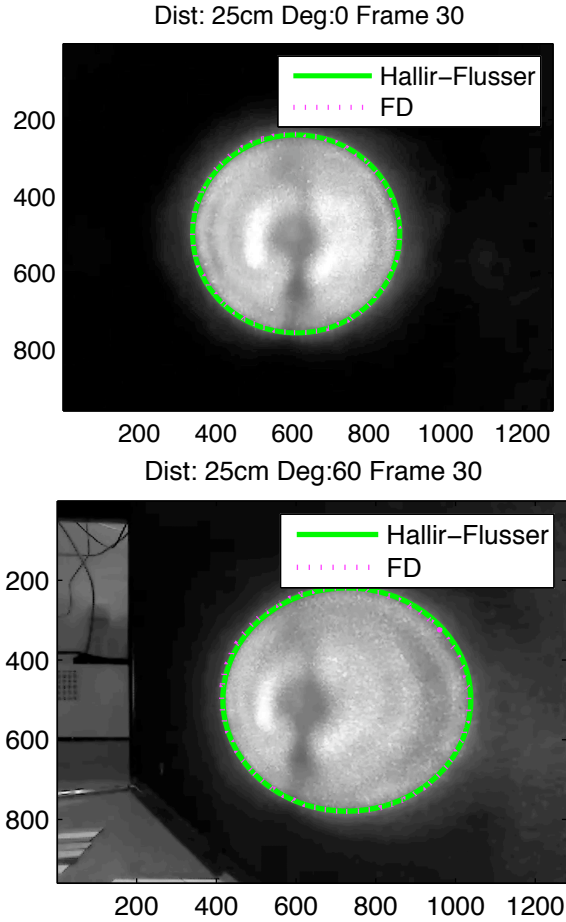


Figure 3. Two examples of camera views (25 cm and 0° respectively 60°) with the fitted ellipses using Halir-Flusser and Fourier descriptors.

III. RESULTS

From the taken images (see Fig. 3 for two examples), we extract the ellipse parameters using the Halir-Flusser method (HF) and Fourier descriptors (FD). We removed the first ten and last ten estimates, since these were often influenced by manually starting and stopping the recording (shaking poster wall).

From the ellipse parameters, we extracted the distance estimates. Fig. 4 shows box plots of the achieved results for the HF method (top) and the FD case (bottom). Accuracy is decreasing with increasing distance in both cases, but the HF-based estimates seem to be biased, presumably by the contour being placed at integer positions. Fig. 5 shows the corresponding results for the angular error of the estimated rotation of the poster wall. Again, accuracy is decreasing with increasing distance.

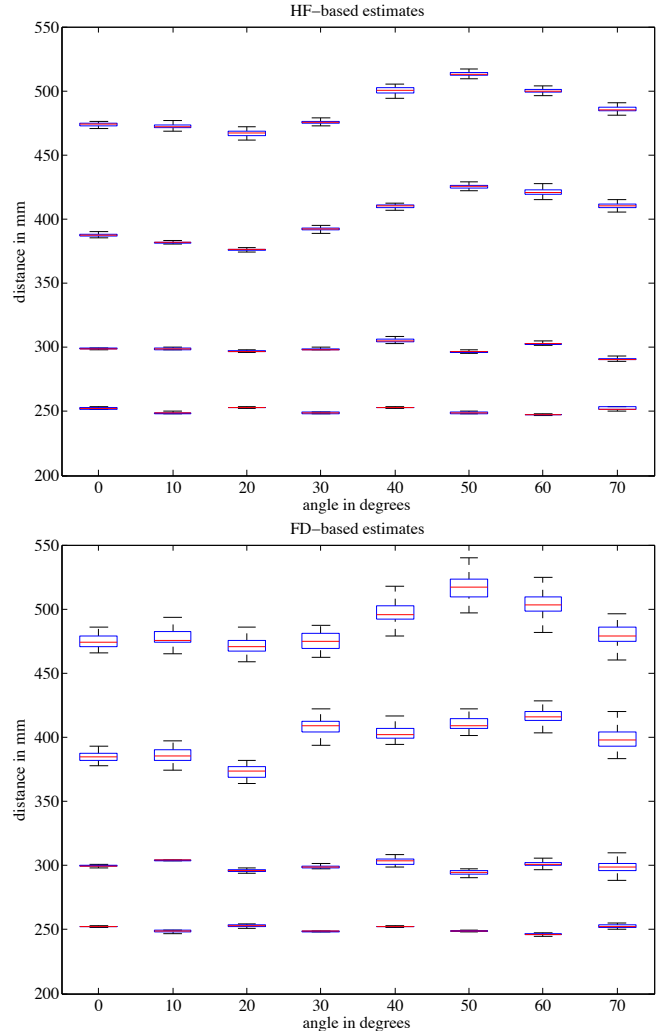


Figure 4. Box plots (median, 25% and 75% quantile, inlier bounds) of distance estimates.

IV. CONCLUSION

We have shown that distance and plane orientation can be estimated using the torchlight approach. In order to solve for ambiguities, increase accuracy, and remove outliers in the estimates, the estimates should be computed using a sequential (in time) filtering approach. This is also required to build maps. A system fusing torchlight estimates to form maps will be topic of future work.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n $^\circ$ 215078 DIPLECS. The two last authors would like to thank the Strategic Research Center MOVIII, funded by the Swedish Foundation for Strategic Research, SSF and CADICS, a Linnaeus center funded by the Swedish Research Council.

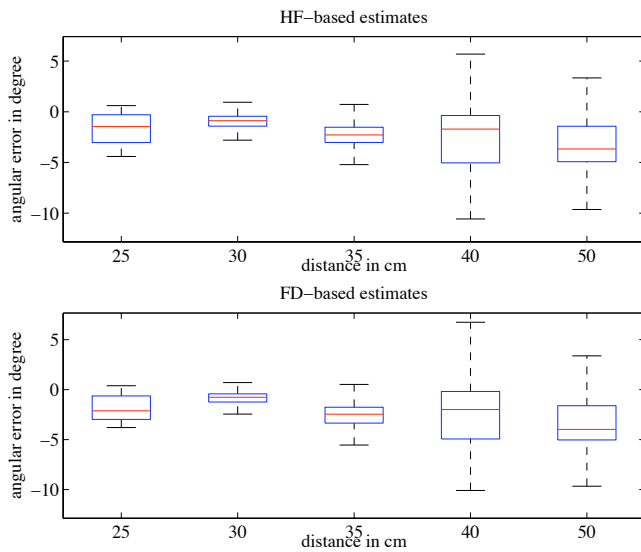


Figure 5. Box plots (median, 25% and 75% quantile, inlier bounds) of angular error.



Figure 6. The experimental setup.

REFERENCES

- [1] D. Scharstein and R. Szelisky, "High-accuracy stereo depth maps using structured light," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003)*, vol. 1, 2003, pp. 195–202.
- [2] M. Johannesson, A. Åström, and P.-E. Danielsson, "An image sensor for sheet-of-light range imaging," in *Proceedings of IAPR Workshop on Machine Vision Applications*, 1992, pp. 43–46.
- [3] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape from shading: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 690–706, 1999.
- [4] G. A. Atkinson and E. R. Hancock, "Two-dimensional brdf estimation from polarisation," *Computer Vision and Image Understanding*, vol. 111, no. 2, pp. 126 – 141, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/B6WCX-4PT1SGW-1/2/4f93736235500bcc7a23e18e7fc7a0ff>
- [5] G. H. Granlund, "Fourier preprocessing for hand print character recognition," *IEEE Trans. on Computers*, vol. C-21, no. 2, pp. 195–201, 1972.
- [6] R. N. Bracewell, *The Fourier transform and its applications*. McGraw Hill, 1986.
- [7] R. Fitzpatrick, "Spherical mirrors," <http://farside.ph.utexas.edu/teaching/316/lectures/node136.html>, 2007-07-14.
- [8] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [9] P.-E. Forssén and A. Moe, "View matching with blob features," *Image and Vision Computing, Canadian Robotic Vision Special Issue*, vol. 27, no. 1-2, pp. 99–107, 2009.
- [10] R. Halíř and J. Flusser, "Numerically stable direct least squares fitting of ellipses," in *Proc. 6th International Conference in Central Europe on Computer Graphics and Visualization. WSCG '98*, 1998, pp. 125–132.
- [11] OpenCV, "Camera calibration and 3D reconstruction," http://opencv.willowgarage.com/documentation/python/camera_calibration_and_3d_reconstruction.html, accessed 2010-01-20.