Monocular Obstacle Avoidance for Blind People using Probabilistic Focus of Expansion Estimation

Sebastian Stabinger Antonio Rodríguez-Sánchez Justus Piater
University of Innsbruck
Sebastian.Stabinger@uibk.ac.at

Abstract

Visually impaired people have a much higher chance of head injuries in daily life because of obstacles that cannot be reliably detected using conventional aids. We present part of a solution to this problem, using only one head mounted camera and optical flow techniques. As part of the system, a novel method to estimate the focus of expansion is presented, which also provides a metric for the quality of the estimate. The final result is a real time capable software system, which can detect obstacles at eye level.

2016 IEEE Winter Conference on Applications of Computer Vision (WACV)
http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7477608

1. Introduction

Visually impaired people usually do very well in avoiding general obstacles using a white cane or a guide dog, but there are obstacles that are hard to detect using those aids. Especially overhanging structures like tree branches, signs, opened truck doors, etc. can not or only unreliably be detected. Such obstacles might be a nuisance in general, but can be a real danger at head level since even collisions at walking speed can lead to severe injuries. A survey by Manduchi & Kurniawan [18] shows that more than 50% of blind people report head-level accidents happening more often than once a year. We therefore wanted to research possibilities to warn users of such obstacles.

Since we did not want to rely on any specialized hardware, only one head-mounted camera is used. Wearing a camera on the head is clearly advantageous for detecting obstacles at head height, but since the head can be moved independently of the walking direction, we cannot assume that obstacles will be in the center of the image. Additionally, a camera being carried by a walking person also captures large amounts of unstructured motion. The system therefore has to be resilient in high noise environments and also detect the true heading direction of the user.

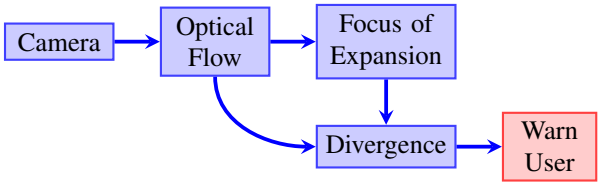


Figure 1: Overview of the presented collision detection system

Our system is constructed as follows: A camera, carried on the head, delivers a video stream. From this stream, a dense optical flow field is calculated using off-the-shelve methods. We then calculate the probability distribution of the focus of expansion (FOE) from this flow field with a novel method. Under some restrictions, the FOE coincides with the heading direction of the camera. Least squares fitting of an affine motion model is used to estimate the divergence in the heading direction, which can be used to warn the user of possible collisions. Figure 1 shows an overview of the system. The proposed method is real time capable. We tested our system in cooperation with a blind person.

2. Related Work

Starting with the influential work of Gibson [8] on the perception of motion, there has been extensive work on optical flow analysis from human subjects while walking (Warren & Kurtz [40]; Bardy *et al.* [1]; Hanada & Ejima [9]; Li *et al.* [16]; to name a few). Of special interest is the seminal paper from Warren [39] which presents research studying visual locomotion over the past 40 years. Psychophysics and Neurophysioloy have been the source for some human motion models (*e.g.*, Prazdny [27]; Lappe & Rauschecker [15]; Hildreth *et al.* [11]; Langer & Mann [14]; Chessa *et al.* [6]). In Computer Vision most of the effort has been invested in cameras mounted on a vehicle (*e.g.* Souhila *et al.* [34]), or a robot (*e.g.* McQuirk *et al.* [21];

Sahin & Guadiano *et al*. [31]). Such methods cannot be directly transferred to a camera carried by a human being, since the motion in this case is much more irregular than on a vehicle or robot.

A considerable amount of research has been going into developing obstacle avoidance aids for the visually impaired, but only a small proportion of these systems is using computer vision for this task. Most commonly used (*e.g*. by Shin & Lim [33]) are ultrasound sensors. Such systems suffer from interference of ambient noises and mirror effects, and it is often necessary to wear quite bulky equipment. In addition, cameras as sensors can be used for many additional tasks besides obstacle avoidance (*e.g*. navigation, text recognition, ...). This would lower the amount of equipment that has to be carried by a blind person.

The existing computer vision based systems for obstacle avoidance can mainly be separated using two properties: Whether a monocular or stereo camera was used, and how the camera is worn. *E.g*. Alberto Rodríguez *et al*. presented in [30] and [29] a system using a stereo camera worn on the chest. Martínez & Ruiz [19] developed a system to explicitly warn from overhanging obstacles using a stereo camera worn on the shoulder. Pradeep *et al*. [26] developed a system using a head mounted stereo camera. The system by Tapu *et al*. [35] uses a smartphone (*i.e*. monocular camera) fixed to the chest. Peng *et al*. [25] also presented a smartphone based system which is hand–held and where the user has to maintain a 45°angle between the smartphone and the floor. Chen *et al*. [5] also presented a system using a stereo camera, but did not do actual experiments with the camera worn by a person. Molton *et al*. [23] described a method for obstacle avoidance using four cameras worn on the upper body for detecting obstacles on the ground. Of great interest is the Kalman filter approach to predict camera motion about half a second into the future which was developed further by Molton and Brady [22]. Recently, Liyanage & Perera [17] proposed an optical-flow based obstacle avoidance system for visually impaired people, though the test scenario was a simulated environment that neglected the self-motion of a human.

To our knowledge we are the first to investigate a head mounted monocular camera for obstacle detection.

A lot of research has been done on extracting the focus of expansion from image sequences. Thomasi & Shi [37] solved the direction of heading through analyzing image deformation, thus eliminating the effects of rotation at the input. Prazdny [28] is optimizing over the variance of intersections of multiple optical flow vectors. Burger & Bhanu [4] is one of the few methods which, similar to the method presented in this paper, delivers an error measure for the focus of expansion. Negahdaripour & Horn [24] presented a very simple method based on counting vector directions. The FOE can also be estimated using a least squares formulation as presented by Tistarelli *et al*. [36]. Hatsopoulos & Warren [10] apply a linear two-layer neural network to the problem. Sazbon *et al*. [32] are using matched filters to detect the focus of expansion. Kuiaski *et al*. [13] are sampling from intersections between vector pairs and estimate the FOE using k-means.

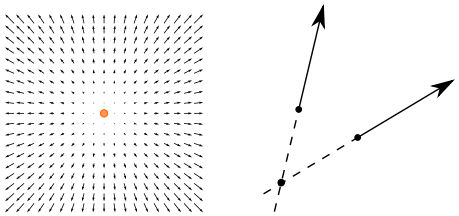


Figure 2: Purely translational optical flow field, with its focus of expansion indicated, and triangulation of the FOE from two vectors.

3. Methods

We will now present the proposed system in detail. As shown in Figure 1, we first approximate the motion field by extracting the optical flow. From this, we can compute the focus of expansion and divergence, which will be evaluated to warn the user.

To calculate the dense optical flow we used the method presented by Farnebäck [7] in the implementation provided by the OpenCV library [3]. The optical flow is calculated between two images which are separated by 5 frames in the video stream to get a bigger amount of motion between the two frames.

3.1. Determining the Focus Of Expansion

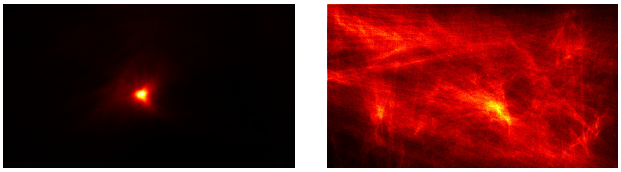
The *focus of expansion (FOE)* of an optical flow field is defined as the point from which the majority of flow vectors point away (left side of Figure 2). The FOE coincides, under some restrictions, with the heading direction of the camera. With an ideal optical flow field, the FOE can be obtained by simple triangulation (right side of Figure 2). In practice we cannot obtain a noise free optical flow. Therefore, the triangulation method is too simple for practical situations.

Still, in some sense such a point of intersection increases the likelihood of the FOE being in that position. It could therefore be assumed that calculating the intersection points of many pairs of vectors and selecting the point with the highest density of intersections would give a reasonable estimate of the FOE. Using this assumption we propose Algorithm 1 to estimate the focus of expansion.

Algorithm 1 The proposed algorithm for estimating the focus of expansion

procedure SAMPLEFOE($\mathbf{F}$)
  $intersections \leftarrow nil$
  for $i \leftarrow 1, samples$ **do**
    $\mathbf{p}, \mathbf{q} \leftarrow$SELECTPOSITIONS($\mathbf{F}.width, \mathbf{F}.height$)
    $\mathbf{r}, \mathbf{s} \leftarrow \mathbf{F}(\mathbf{p}), \mathbf{F}(\mathbf{q})$
    $ray_r \leftarrow$ RAY($\mathbf{p}, -\mathbf{r}$)
    $ray_s \leftarrow$ RAY($\mathbf{q}, -\mathbf{s}$)
    $\mathbf{i} \leftarrow$ INTERSECT(ray_r, ray_s)
    if $\mathbf{i} \neq nil$ **then**
      $intersections \leftarrow$ APPEND($intersections, \mathbf{i}$)
    end if
  end for
  FOE $\leftarrow$ select point with highest density of intersections
  return FOE
end procedure

▷ get the focus of expansion from a flow field $\mathbf{F}$
▷ Initialize list of intersections
▷ Select two positions
▷ Get flow vectors at positions $\mathbf{p}$ and $\mathbf{q}$
▷ Get rays in opposite direction of flow vectors
▷ Get point of ray intersection
▷ $\mathbf{i}$ = point of intersection or nil if no intersection



(a) Low noise (b) High noise

Figure 3: Visualization of intersection histograms to determine the focus of expansion at different noise levels.

3.1.1 FOE Probability Distribution

A big advantage of our proposed method is the fact that we not only obtain a single estimate for the focus of expansion, but a full probability distribution over the whole image by interpreting a higher density of intersections as a higher probability for the FOE to be at this position. This can be seen by visualizing the density of intersections using a heat map (Figure 3).

Using the method from Algorithm 1 it is easy to determine the reliability of the FOE estimate. If the votes are spread over a large area the estimate will probably be unreliable. If the votes are tightly clustered the estimate is probably reliable. In practice we calculate a measure which we call *inlier proportion* which is the amount of intersections in a region around the FOE in relation to all calculated intersections.

3.2. Relation between Divergence and Time to Contact

After the FOE has been determined, we know at which position in the image collisions should be detected. Now, the time to contact has to be calculated in that direction. We use the inverse of the divergence as a measure for the time to contact. The mathematical reasoning behind this will be discussed in this section. We will also empirically evaluate this relationship in Section 4.2.

Let x, y be a system of Cartesian coordinates in a two dimensional euclidean space and $\mathbf{i}$, $\mathbf{j}$ the two basis vectors. A two dimensional vector field is then defined as $\mathbf{F} = U\mathbf{i} + V\mathbf{j}$. The divergence of $\mathbf{F}$ is then defined as follows:

$$\operatorname{div} \mathbf{F} = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} \tag{1}$$

Translated to the discrete case of an optical flow field, this can be approximated for a certain position (x_1, y_1) by

$$\operatorname{div} \mathbf{F} = \frac{u_{x_2} - u_{x_1}}{\Delta_x} + \frac{v_{y_2} - v_{y_1}}{\Delta_y} \tag{2}$$

where $x_2 = x_1 + \Delta_x$ and $y_2 = y_1 + \Delta_y$ and u_{x_1}, u_{x_2} are the horizontal components of the flow field at the horizontal positions x_1 and x_2 and v_{y_1}, v_{y_2} are the vertical components of the flow field at vertical positions y_1 and y_2.

Since the FOE is the point towards which we are moving we will not see any optical flow component at the FOE. Assuming we are choosing x_1, y_1 to be the FOE, then $u_{x_1} = v_{y_1} = 0$. So Equation 2 transforms to:

$$\operatorname{div} \mathbf{F} = \frac{u_{x_2}}{\Delta_x} + \frac{v_{y_2}}{\Delta_y} \tag{3}$$

The time to contact τ can be calculated (as presented by Tresilian [38]) using the optical flow components along the x-axis, where d_x is the distance of position x from the focus of expansion by:

$$\tau_x = \frac{d_x}{u_x} \tag{4}$$

By setting $d_x = \Delta_x$ we get:

$$\tau_x = \frac{\Delta_x}{u_{x_2}} \tag{5}$$

Calculating the time to contact using the flow along the y-axis can be done analogously to Equation 4 and Equation 5.

We know that the time to contact has to be a value that is independent of the axis of the flow field we are using in the calculation and therefore:

$$\tau = \tau_x = \tau_y = \frac{\Delta_x}{u_{x_2}} = \frac{\Delta_y}{v_{y_2}} \tag{6}$$

Using Equation 5 and Equation 6, we can reformulate Equation 3 as follows:

$$\operatorname{div} \mathbf{F} = \frac{1}{\tau} + \frac{1}{\tau} = \frac{2}{\tau} \tag{7}$$

and therefore

$$\frac{1}{\operatorname{div} \mathbf{F}} \propto \tau \tag{8}$$

So the inverse of the divergence, at the focus of expansion, is proportional to the time to contact in direction of the FOE. This formal result will be empirically validated in Section 4.2.

3.3. Estimating Divergence Using an Affine Motion Model

If we assume a smooth optical flow field that can be locally approximated by a first-order Taylor expansion, it can be described by a so called *affine optical flow field* (Equation 9) as explained by Jähne *et al.* [12][p 383]. As long as the patch of optical flow that we want to describe is small enough, the smoothness constraint is usually fulfilled.

$$\begin{bmatrix} u_{xy} \\ v_{xy} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_4 \end{bmatrix} + \begin{bmatrix} c_2 & c_3 \\ c_5 & c_6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{9}$$

u_{xy} and v_{xy} are the two components of the flow vector at position x and y. An affine optical flow field is completely defined by the six parameters c_1 to c_6. c_1 and c_4 correspond to the two translation components. $c_3 - c_5$ corresponds to the rotation of the flow field. $c_3 + c_5$ and $c_2 - c_6$ correspond to deformations in the two axis. The divergence of an affine optical flow is given by $c_2 + c_6$. We again refer to the work of Jähne *et al*. [12][p 383ff.] for further information.

If a vector field is given, as it is in our case, the parameters u_{xy}, v_{xy}, x, and y from Equation 9 are known. Since we have two equations, but search for six unknowns we have an underconstrained system. Fortunately, we know that the same affine model is valid (*i.e*. has the same parameters c_1 to c_6 for Equation 9) for the whole patch used to calculate the divergence. We can therefore use multiple instances of Equation 9 at different positions x and y of the patch to construct an overdetermined system. We then solve this system using least squares, and get an estimate of the divergence using $c_2 + c_6$.

3.4. Warning the User

After the divergence has been calculated, a threshold can be used to inform the user about a possible collision.

Unfortunately, the time series of the divergences is noisy. Especially since rapid motions of the camera can produce very high or low divergence peaks which are counterproductive for a reliable and stable system.

As a solution, we discard divergence measures for frames for which the sampling density of the FOE estimation is below a certain threshold. This generally indicates a frame with too much unstructured motion to be used reliably for calculating the divergence. We empirically determined a threshold of 0.02 on the inlier proportion which we use as a measure for the sampling density. As an additional precaution we only accept frames as good if more than three consecutive frames have an inlier proportion above the threshold. As will be shown in the evaluation, the possibility to filter properties using the inlier proportion is one of the most interesting consequences of the approach presented in Algorithm 1.

4. Results

The system was tested on synthetic as well as real life image sequences. Figure 4 shows a single frame from all the sequences that were used.

The synthetic sequences (S1 to S5) were created using Blender by the Blender Online Community [2]. In addition to the videos, we also generated the ground truth heading direction of the camera. The scenes contain different amounts of textures in an attempt to simulate circumstances in which the calculation of the optical flow is more or less difficult.

The real world sequences (R1 to R4) were created using a GoPro Hero Black 3 which was strapped to the head. The sequences were recorded with 60 frames per second and were used without any pre processing (*i.e*. no image stabilization was done).

The sequences B1 and B2 were recorded by a blind person to have recordings with head movements representative for the target users of the system. We used the same setup as for sequences R1 to R4. The videos including the FOE ground truth (where available) can be found at `https://goo.gl/5kUAzX`.

4.1. Estimating Heading Direction

Heading direction estimation was first evaluated on synthetic image sequences. This allowed us to use the ground truth heading direction for comparison. We compared our method to the widely used *counting method* which is a method used to estimate the focus of expansion. The counting method was presented by Negahdaripour & Horn [24] and used for example by Souhila *et al*. [34], Liyanage & Perera [17], and McCarthy & Barnes [20]. The counting

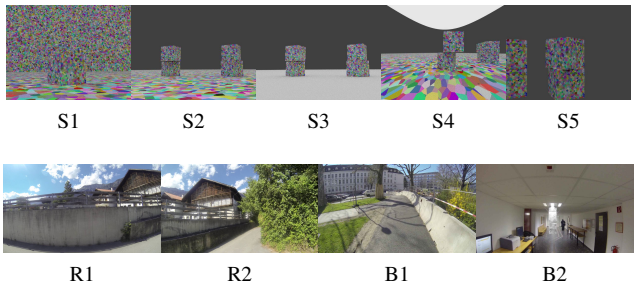


Figure 4: Single frames from image sequences used for testing.

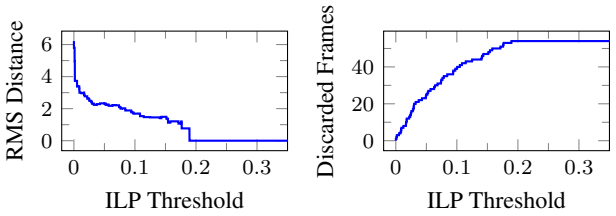
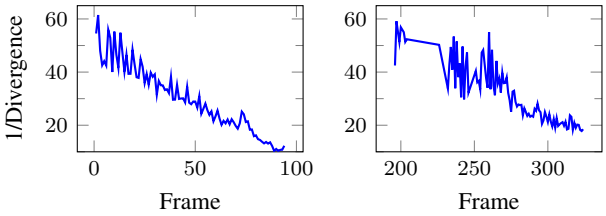


Figure 5: The Root-Mean-Square distance and the number of discarded frames in relation to the the inlier proportion threshold used to filter frames for sequence S1.

Table 1: Root-Mean-Square distance for different image sequences and estimation methods for the focus of expansion

Sequence	Counting	Sampling	filtered Sampling (discarded frames)
S1	26.36	8.57	2.07 (9.25%)
S2	14.58	6.81	2.27 (9.25%)
S3	85.33	31.14	7.20 (35.19%)
S4	104.84	20.04	18.46 (8.6%)
S5	6.74	2.62	2.10 (9.6%)



(a) Synthetic sequence S1 (b) Real sequence R1

Figure 6: Inverse of the divergence of a 60×60 pixel patch in the center of the image for a synthetic and real image sequence. Divergence measurements for low sample density frames were discarded.

method is based on the fact that a majority of the vectors point away from the FOE.

To compare the accuracy of the different methods, we are using the *Root-Mean-Square distance* (Equation 10) as an error measure, where $dist(H, F_i)$ is the euclidean distance in pixels between the true heading direction H and the estimated focus of expansion for frame i, F_i. N is the total number of frames.

$$E = \sqrt{\frac{\sum_{i=0}^{N} \text{dist}(H, F_i)^2}{N}} \tag{10}$$

Additionally we are comparing against a filtered version of the sampling algorithm. The filtering is done by discarding frames for which the inlier proportion is below 0.02.

The influence of the threshold on the Root-Mean-Square distance and the number of discarded frames can be seen in Figure 5 for sequence S1.

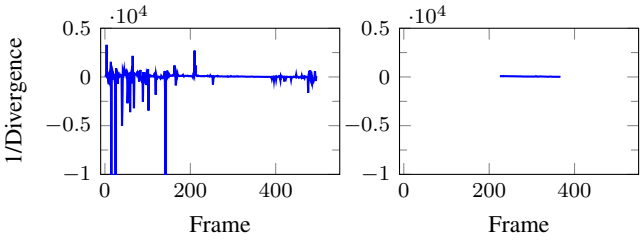
Our sampling method has big advantages compared to the counting method, and filtering improves the reliability even further (Table 1). Heading direction estimation was also indirectly tested for real image sequences during the evaluation of the whole system in Section 4.3.

4.2. Collision detection

We first tested whether the proportionality between time to collision and the inverse of the divergence actually holds in practice. To test this assumption we used image sequences in which the camera moves towards an obstacle with a constant speed and tried to observe a linear decrease of the inverse of the divergence over time. The method from Section 3.3 was used to calculate the divergence.

Figure 6a shows that the inverse of the divergence could be approximated quite well by a linear slope. As previously stated, the camera is moving towards the obstacle with constant speed. Therefore the time to contact will also decrease in a linear fashion. This indicates a linear relationship between the inverse of the divergence and the time to collision as theoretically derived in Section 3.2. Figure 6b, calculated for a real sequence, indicates that the person started slowing down before colliding with the obstacle since the slope gets flatter at the end. This can also be observed by looking at the video. In addition, as would be expected, the relationship is not as clear as for the synthetic image sequence because of unstructured motion during walking, but it still suggests that the proportionality of the time to contact and the inverse of the divergence also holds in practice.

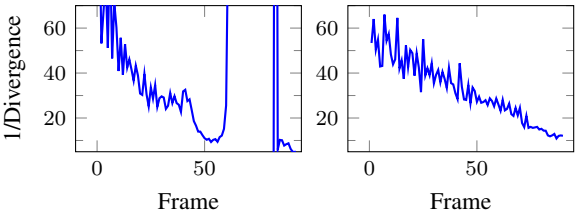
We also analyzed the effect of filtering on the divergence. This was done by discarding frames for which the inlier proportion was above an empirically determined threshold of 0.02. Figure 7 shows the influence of filtering, which eliminates a lot of the noise in the divergence measurement and only leaves the measurements that correspond to a possible



(a) Without filtering

(b) Filtered by sampling density (cf. Figure 9a)

Figure 7: Influence of filtering using the sampling density of the FOE estimate on the measured divergence for video R2.



(a) 1/Divergence in the image center

(b) 1/Divergence at the estimated FOE

Figure 8: Visualization of the importance of estimating the heading direction.

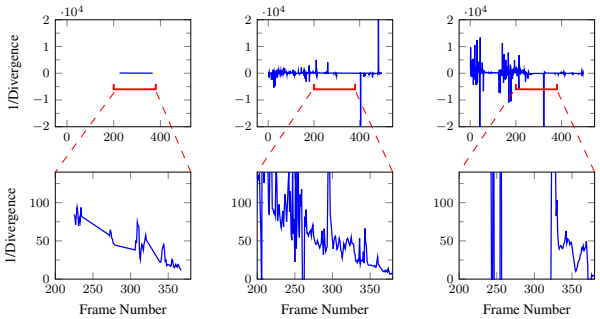
collision.

4.3. Evaluating FOE and divergence together

The complete system, combining heading direction estimation and calculating the time to collision, was tested on synthetic and real image sequences. The ultimate test is how well the divergence, in the estimated heading direction, corresponds to the true time to collision. We also compare against a system which calculates the divergence for the center of the image, as well as against the focus of expansion estimated by the counting method.

Figure 8 shows a comparison between the divergence in the image center (Figure 8a) and the one calculated at the focus of expansion (Figure 8b) for image sequence S5. In sequence S5 the camera moves towards a point on the left margin of the image. The divergence at the FOE reliably corresponds to the time to contact, while the divergence in the center does not, so it is clear that the heading direction is important in accurately detecting possible collisions.

Image sequence R2 has a person walking towards a bush, while looking to the left, with a camera fixed to the head. This results in a heading direction to the right side of the image. In addition, the sequence contains a lot of unstructured motion, which interferes heavily with the calculation



(a) Sampling

(b) Counting

(c) Image center

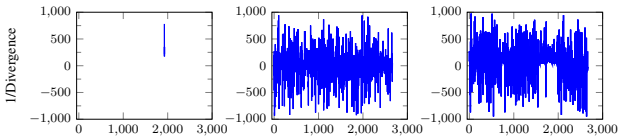
Figure 9: Comparing results of different methods on image sequence R2.

of the divergence. If the sampling density of the FOE estimation is used as a quality measure for the current optical flow field, we get a relatively noise free divergence (Figure 7).

Figure 9 shows a comparison between three different methods. Figure 9a shows the inverse of the divergence at the focus of expansion, calculated using the novel sampling approach presented in Algorithm 1 and filtered using the sampling density of the FOE estimation. Figure 9b shows the inverse divergence at the FOE, calculated using the counting method. Since the counting method does not give any measure of reliability we cannot use it to filter the divergences. Figure 9c shows the inverse of the divergence in the center of the image. Filtering is also not possible in this case.

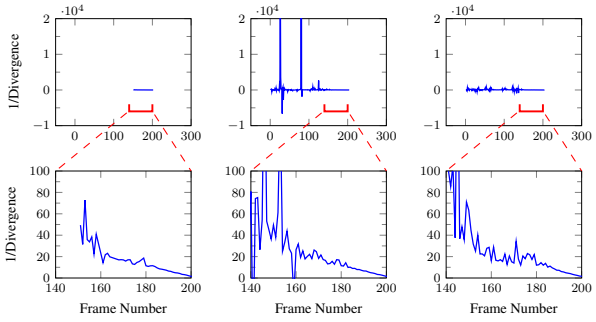
The sampling approach with filtering only returns divergence measures for the region where a collision is actually imminent. Looking at the same frames using the counting method, the falling of the inverse of the divergence can be observed, but the data contains much more noise. Looking at all the frames it would be difficult to find the frames corresponding to the collision. Calculating the divergence in the image center introduces even more noise and the approaching obstacle can only be detected very briefly before the collision actually happens, since it is only detected once the obstacle is big enough to get into the center of the image.

We also tested our system on two video sequences recorded together with a blind person. The first one (B1) was recorded to test the system on a sequence with no obstacles and check whether the system would erroneously warn the user. The blind person was taking a short walk for this with a camera mounted on the head. Our sampling method ignored the whole sequence (except for five frames) since there was no danger of a collision during the whole sequence. The five frames which were considered reliable all had a high inverse of the divergence and therefore did not warn of a collision. The other two methods produce very



(a) Sampling (b) Counting (c) Image center

Figure 10: Comparing results of different methods on image sequence B1.



(a) Sampling (b) Counting (c) Image center

Figure 11: Comparing results of different methods on image sequence B2.

noisy divergence measures (Figure 10).

We also recorded a sequence together with the blind person simulating a collision using the same setup and parameters (B2). Figure 11 shows that the system is able to detect the obstacle. In this special case the methods using the counting method or simply the center of the image also give good results, since the image sequence only consists of walking straight towards an obstacle, but as previously observed, the other two methods also report highly variable divergences in times where no obstacle is present.

4.4. Performance

We evaluated the performance of the system on an *Intel i7-3770K 3.5GHz Quad Core with Hyper Threading* using SIMD instructions and a parallel implementation. Using the OpenCV implementation of the optical flow method presented by Farnebäck [7], the whole system needs a *little under 40 ms* to process one 320×280 frame, using a divergence patch size of 30×30, and 20,000 samples for the FOE sampling method. This leads to about *25 frames processed per second*. Excluding the optical flow calculation the whole method takes on average 1.82 ms per frame which shows that the optical flow calculation need most of the computational resources.

Using the counting method (not including calculating the optical flow) takes about 7.2 ms per frame. It should be noted though that we did not optimize the counting method.

5. Discussion

As the results show, the presented method for estimating the FOE generally gives better results than the widely used counting method. The biggest advantage is in our opinion the possibility to quantify the quality of the FOE estimate. We used this property to filter other measurements taken from the optical flow field (*i.e.* the measures of the divergence). In an obstacle avoidance system, such a quality measure can be used to not warn the user in cases the system cannot detect obstacles reliably. Alternatively, the system could inform the person that it is currently not working reliably so the user can be especially careful. In essence this leads to less false warnings for the user. We can think of other uses this quality measure could have:

6. Conclusion

We presented a novel method of estimating the focus of expansion in a probabilistic way. We further showed that the metric, determining the reliability of the FOE estimate, can be interpreted as a general measure of the quality of an optical flow field. We used this insight to eliminate frames with a low confidence on the FOE estimate from further processing. We were able to greatly improve on the noise level of the measured divergence and through this the time to collision. This noise reduction is especially important for systems where the camera is carried by a person since it generally introduces a lot of unstructured motion into the image sequences.

7. Appendix — Implementation

Algorithm 1 can be implemented in a very efficient manner which we will present here.

We are only interested in the heading direction with pixel accuracy. In addition we can discard points of intersection which are outside the image. Thus we can use a 2D histogram to tally up the number of intersections per pixel directly instead of managing a list of intersection points. This gives us a memory requirement independent of the number of samples and ensures constant time execution of all following computations. This histogram can be computed concurrently since all intersection points are independent of each other. Figure 3 shows two visualizations of such histograms under low and high noise conditions.

The region of highest density of intersections and therefore the most probable position of the FOE is determined by summing up all bins of the histogram within a small window which slides over all positions. The window position that gives the highest sum is assumed as being the highest density region and therefore the location of the FOE.

We are using the so called *inlier proportion* as a metric for the quality of our measure. The inlier proportion is the proportion of the number of intersections within the highest

density window determined in the previous paragraph and the number of all generated samples. If the inlier proportion gets bigger, more of the samples are concentrated in a small area of the histogram (*i.e*. the variance is smaller).

References

[1] B. G. Bardy, W. H. Warren, and B. A. Kay. The role of central and peripheral vision in postural control duringwalking. *Perception & psychophysics*, 61(7):1356–1368, 1999. 1

[2] Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation, Blender Institute, Amsterdam, 2014. 4

[3] G. Bradski. Opencv library. *Dr. Dobb's Journal of Software Tools*, 2000. 2

[4] W. Burger and B. Bhanu. On computing 'fuzzy' focus of expansion for autonomous navigation. In *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on*, pages 563–568. IEEE, 1989. 2

[5] L. Chen, B.-l. Guo, and W. Sun. Obstacle detection system for visually impaired people based on stereo vision. In *Genetic and Evolutionary Computing (ICGEC), 2010 Fourth International Conference on*, pages 723–726. IEEE, 2010. 2

[6] M. Chessa, F. Solari, and S. P. Sabatini. Adjustable linear models for optic flow based obstacle avoidance. *Computer Vision and Image Understanding*, 117(6):603–619, 2013. 1

[7] G. Farnebäck. Two-frame motion estimation based on polynomial expansion. In *Image Analysis*, pages 363–370. Springer, 2003. 2, 7

[8] J. J. Gibson. The visual perception of objective motion and subjective movement. *Psychological Review*, 61(5):304, 1954. 1

[9] M. Hanada and Y. Ejima. Effects of roll and pitch components in retinal flow on heading judgement. *Vision research*, 40(14):1827–1838, 2000. 1

[10] N. G. Hatsopoulos and W. H. Warren. Visual navigation with a neural network. *Neural Networks*, 4(3):303–317, 1991. 2

[11] E. C. Hildreth, J. M. Beusmans, E. R. Boer, and C. S. Royden. From vision to action: experiments and models of steering control during driving. *Journal of Experimental Psychology: Human Perception and Performance*, 26(3):1106, 2000. 1

[12] B. Jähne, H. Haussecker, and P. Geissler. *Handbook of computer vision and applications*. Academic Press, 2000. 4

[13] J. R. Kuiaski, A. E. Lazzaretti, and H. V. Neto. Focus of expansion estimation for motion segmentation from a single camera. *VII Workshop de Visão Computacional*, pages 272–277, 2011. 2

[14] M. S. Langer and R. Mann. Optical snow. *International Journal of Computer Vision*, 55(1):55–71, 2003. 1

[15] M. Lappe and J. P. Rauschecker. A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Computation*, 5(3):374–391, 1993. 1

[16] L. Li, J. Chen, and X. Peng. Influence of visual path information on human heading perception during rotation. *Journal of Vision*, 9(3):29, 2009. 1

[17] D. Liyanage and M. Perera. Optical flow based obstacle avoidance for the visually impaired. In *Business Engineering and Industrial Applications Colloquium (BEIAC), 2012 IEEE*, pages 284–289. IEEE, 2012. 2, 4

[18] R. Manduchi and S. Kurniawan. Mobility-related accidents experienced by people with visual impairment. *Research and Practice in Visual Impairment and Blindness*, 4(2):44–54, 2011. 1

[19] J. M. S. Martinez and F. E. Ruiz. Stereo-based aerial obstacle detection for the visually impaired. In *Workshop on Computer Vision Applications for the Visually Impaired*, 2008. 2

[20] C. McCarthy and N. Barnes. A unified strategy for landing and docking using spherical flow divergence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):1024–1031, 2012. 4

[21] I. S. McQuirk, B. K. Horn, H.-S. Lee, and J. L. Wyatt Jr. Estimating the focus of expansion in analog vlsi. *International journal of computer vision*, 28(3):261–277, 1998. 1

[22] N. Molton and M. Brady. Modelling the motion of a sensor attached to a walking person. *Robotics and Autonomous Systems*, 34(4):203–221, 2001. 2

[23] N. Molton, S. Se, J. Brady, D. Lee, and P. Probert. A stereo vision-based aid for the visually impaired. *Image and vision computing*, 16(4):251–263, 1998. 2

[24] S. Negahdaripour and B. K. Horn. A direct method for locating the focus of expansion. *Computer Vision, Graphics, and Image Processing*, 46(3):303–326, 1989. 2, 4

[25] E. Peng, P. Peursum, L. Li, and S. Venkatesh. A smartphone-based obstacle sensor for the visually impaired. In *Ubiquitous Intelligence and Computing*, pages 590–604. Springer, 2010. 2

[26] V. Pradeep, G. Medioni, and J. Weiland. Robot vision for the visually impaired. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 15–22. IEEE, 2010. 2

[27] K. Prazdny. Egomotion and relative depth map from optical flow. *Biological cybernetics*, 36(2):87–102, 1980. 1

[28] K. Prazdny. Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer. In *1981 Technical Symposium East*, pages 199–206. International Society for Optics and Photonics, 1981. 2

[29] A. Rodríguez, L. M. Bergasa, P. F. Alcantarilla, J. Yebes, and A. Cela. Obstacle avoidance system for assisting visually impaired people. In *Proceedings of the IEEE Intelligent Vehicles Symposium Workshops, Madrid, Spain*, volume 35, page 16, 2012. 2

[30] A. Rodríguez, J. J. Yebes, P. F. Alcantarilla, L. M. Bergasa, J. Almazán, and A. Cela. Assisting the visually impaired: obstacle detection and warning system by acoustic feedback. *Sensors*, 12(12):17476–17496, 2012. 2

[31] E. Sahin and P. Guadiano. Visual looming as a range sensor for mobile robots. *CAS/CNS Technical Report Series*, 1998. 2

[32] D. Sazbon, H. Rotstein, and E. Rivlin. Finding the focus of expansion and estimating range using optical flow images and a matched filter. *Machine Vision and Applications*, 15(4):229–236, 2004. 2

[33] B.-S. Shin and C.-S. Lim. Obstacle detection and avoidance system for visually impaired people. In *Haptic and Audio Interaction Design*, pages 78–85. Springer, 2007. 2

[34] K. Souhila, O. Djekoune, D. Djebrouni, and D. Meriche. On the use of optical flow in robot navigation. In *Signal Processing and Communications, 2007. ICSPC 2007. IEEE International Conference on*, pages 1287–1290. IEEE, 2007. 1, 4

[35] R. Tapu, B. Mocanu, A. Bursuc, and T. Zaharia. A smartphone-based obstacle detection and classification system for assisting visually impaired people. In *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pages 444–451. IEEE, 2013. 2

[36] M. Tistarelli, E. Grosso, and G. Sandini. Dynamic stereo in visual navigation. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 186–193. IEEE, 1991. 2

[37] C. Tomasi and J. Shi. Image deformations are better than optical flow. *Mathematical and computer modelling*, 24(5):165–175, 1996. 2

[38] J. R. Tresilian. Perceptual information for the timing of interceptive action. *Perception*, 19(2):223–239, 1990. 3

[39] W. H. Warren. Visually controlled locomotion: 40 years later. *Ecological Psychology*, 10(3-4):177–219, 1998. 1

[40] W. H. Warren and K. J. Kurtz. The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, 51(5):443–454, 1992. 1