

A Forward Model of Optic Flow for Detecting External Forces

L. Georgopoulos, G. M. Hayes, G. D. Konidaris

Robot positioning is an important function of autonomous intelligent robots. However, the application of external forces to a robot can disrupt its normal operation and cause localisation errors. We present a novel approach for detecting external disturbances based on optic flow without the use of egomotion information. Even though this research moderately validates the efficacy of the model we argue that its application is plausible to a large number of robotic systems.

I. INTRODUCTION

AMONG various subsystems implemented in intelligent autonomous robots one would distinguish self positioning as one of critical importance. An external force applied to the system may cause such a disturbance that the robot may be unexpectedly displaced. Hence, the localization subsystem may fail to update its parameters appropriately, possibly resulting in failure to localise. In this paper, we present a framework capable of detecting an event in real time using a monocular video sensor and without the use of egomotion information.

This research has been motivated mainly from various robot competitions. There, the robots crashing into each other or nearby walls, or when picked up by a referee, often experience failures due to unmodelled external forces. This sometimes results in inaccurate internal maps of the environment. Consequently, a detection of such events would allow the robot to take corrective action.

Such effects may be detected by incorporating expensive, power consuming and usually heavy equipment such as gyroscopes or acceleration sensors which limit the overall applicability of the system. However, we demonstrate that such events are detectable by using a camera which is inexpensive, portable, light and ordinary in robotic applications.

Our approach has been founded on the following assumption. Given an external force, such that it may cause a disturbance capable of deceiving an arbitrary localization scheme, we should observe a notable discontinuity onto the motion of the system. Hence, it

should be detectable by measuring some quantity of a video sequence recorded from a video sensor mounted onto the robot.

Our method can be summarized in four sections. First, the from a recorded video sequence during a robot's motion the corresponding optic flow fields are computed. Any of the known algorithms for optic flow computation [6, 10, 11, 12, 14, 15, 16, 17] may be incorporated. Second, assuming of local spatiotemporal smoothness of motion we predict the next optic flow field. Third, we utilize a similarity measure to compare the actual and predicted field. Finally, we perform classification with regard to the similarity distribution generated by previous observations of undisturbed motion.

Hence, the model can be analysed in the following components: Estimation of optic flow, Prediction, Comparison and Classification. These are presented in detail below. Additionally, experimental results are provided based on offline analysis of appropriate video sequences. The usability, limitations and possible extensions of this work are examined in the discussion section.

II. OPTIC FLOW

A. Overview

We have chosen a video sensor in order to maximize probable areas of the model's applications. An appropriate notion to describe motion in a video sequence is optic flow. The latter has been employed extensively in numerous applications [1, 3, 4, 5, 8].

Optic flow can be identified as the apparent motion of brightness points on the two dimensional plane. This represents the planar projection of the field of view along two succeeding recorded fragments of motion [7, 11, 12]. When brightness motion is coherent with actual motion then we say that optic flow corresponds to the velocity field, which is the actual planar projection of points in the observed volume. However, that is implausible, mainly due to the Aperture Problem [2]. The latter states our inability to compute the optic flow along the isobrightness contour. Hence, we should consider optic flow as the lower limit approximation of the velocity field [17]. Recovering egomotion information from either the velocity field or optic flow is ill-posed [6]. Thus, specific constraints are usually set, such as spatiotemporal smoothness, irradiance smoothness, image brightness constancy, and Lambertian surfaces [16]. These are usually violated in the general case and specifically in real world applications. Commonly, problems arise when dealing with shadows, highlights, surface translucency

Manuscript received March 15, 2007. This work was supported in part by the Institute of Perception, Action and Behaviour within the School of Informatics of the University of Edinburgh, and the Student Awards Agency for Scotland.

L. Georgopoulos is with EPFL, IC, ISC, LANOS, Nonlinear Systems Laboratory, Station 14, CH-1015, Lausanne, Switzerland (phone: +41 216934684, e-mail: Leonidas.georgopoulos@epfl.ch).

G. M. Hayes is at the Institute of Perception Action and Behaviour, University of Edinburgh, Edinburgh, Scotland, The King's Buildings, Mayfield Road, JCMB-2107C, EH9 3JZ, United Kingdom. (e-mail: gmh@inf.ed.ac.uk, phone: +44 (0) 131 651 3440).

G. D. Konidaris is at the Autonomous Learning Laboratory in the Department of Computer Science at the University of Massachusetts, Computer Science Building, 140 Governors drive, Amherst, MA 01003-9264, USA (e-mail: gdk@cs.umass.edu).

and occlusions, moving objects, transparencies and reflective surfaces [16]. However, the principal effects in the optic flow field are mainly due to egomotion. A disturbance in egomotion, caused by an external force, would probably be propagated as a spatiotemporal discontinuity of optic flow. On this ground we expect an external force to be detected in optic flow.

B. Optic Flow Algorithms

As mentioned, there exist a large number of algorithms that attempt to overcome the difficulties mentioned above. These are distinguished by the constraints imposed, data conservation, smoothness, spatial coherence, modelling assumption, and robustness [14], [16]. Moreover, we may divide them into five major categories based on the way optic flow is computed: feature selection or affine matching, gradient computation, correlation, energy and phase.

Feature selection techniques attempt to detect objects or features among sequential frames and measure the disparity. In gradient computation algorithms [6], [14], [17], the main concept is to estimate the spatiotemporal derivatives of the irradiance among frames. A different approach is correlation. There, the key concept is to correlate parts between frames. In energy based methods a Fourier transformation is performed on the spatiotemporal energy on a plane in frequency space [14]. In phase based algorithms spatiotemporal band pass filters are applied to extract binary edge maps to track the edges [16]. Finally, another approach is to augment any of the above estimators in a robust statistical estimator by incorporating a minimization of a parametric model [17].

C. Selection of the appropriate Optic Flow algorithm

Obviously, from the variety of algorithms summarized above there exists a subgroup that best fits the needs of the presented framework. Specifically, our choice has been based on both theoretical grounds and experiments performed on simplified artificial scenery. In order to preserve the validity of our initial assumption we have set two prerequisites. First, that the algorithm has low computational cost, since we aim at real time applications. Second, that computation along the field preserves spatial features and therefore avoids over-smoothed optic flow field. The latter allows a larger divergence to be detected.

An experimental comparison was performed between the ‘‘Second Order Approach’’, Barron-Beauchemin [14], the ‘‘Robust Gradient Estimator’’, M.J. Black [16] and the ‘‘Laplacian Pyramid’’, P. Anandan [15]. We have compared the three algorithms in regard to speed, flexibility and accuracy. Our goal was to minimize time spent to compute Optic Flow, allow for parameter flexibility which would allow to better fine tune the model to the environment and maximize accuracy in the computation of the flow fields.

The experiments were carried out on two sequences, the well known Yosemite sequence [18] and an artificial two dimensional scene of a square moving in frame with

standard speed. In both occasions the velocity field was known. The three algorithms were optimized in regard to their parameters and the best results were compared with each other. Extensive results are provided in [20].

All three algorithms demonstrated sufficiently low, average magnitude and angular error in areas of adequate texture. However, in areas of contradictory brightness transition, error measures reached maximal values. Among the three the ‘‘Robust Gradient Estimator’’ demonstrated consistent performance in all tests and had always minimal errors. The latter was also the fastest and above all its implementation permitted numerous parameters to fine tune its performance. Moreover, these were much more than pre-processing tasks such as smoothing. In contrast, they affected the evaluation of the spatiotemporal brightness change across images.

In conclusion, we observe that it is feasible to make an approximation of the velocity field from just the optic flow field. Even though that would be far from completely recovering the velocity field in the general case it is feasible to have a measure of egomotion by measuring optic flow. Below, we demonstrate that this may provide detection of an external disturbance.

III. DIRECT OPTIC FLOW PREDICTION

We wish to predict the next optic flow frame from just the current frame. At this point we shall make the following assumption. There exists such a fragment of time in which the motion of the system is normal and smooth. Also, this can be expanded in the following manner. There exists such a fragment of time in which the projected planar motion of the observed velocity volume appears to be normal and smooth. Thus, we may now associate the current field with the next one in a linear fashion.

Specifically, this can be done by associating the next field directly with the previous by assuming local spatiotemporal consistency. Hereon, let us denote the current optic flow field with O , next and previous fields as O_+ and O_- respectively, and the estimate as \hat{O} . An initial approach to estimate the next optic flow would be linearly.

$$\hat{O}_+ = O + N(0, \sigma) \quad (1)$$

Let us assume of spatial local consistency of optic flow in a small neighbourhood around point r_0 . Consequently, the Taylor expansion of the optic flow field around that point r_0 would be:

$$O(r) = O(r_0) + \nabla O(r_0)(r - r_0) + Order^2(r_0) \quad (2)$$

This estimates the optic flow in the neighbourhood of r_0 . Now, let us consider that expansion in the temporal sense. Then, the optic flow is a function of time and space as well. Hence, a Taylor expansion in the temporal neighbourhood around a moment t_0 would be:

$$O(r, t) = O(r, t_0) + \frac{\partial O(r, t_0)}{\partial t} (t - t_0) + \text{Order}^2(r, t) \quad (3)$$

Hence, it is possible to approximate the next optic flow field by knowing its temporal derivative. Now, for a discrete sampling sequence of appropriate frame rate we write:

$$\hat{O}_+ = O + \delta O_- \quad (4)$$

Where δO_- is the spatiotemporal change in the optic flow between the current and previous frame. This can be found from the temporal derivative of optic flow which can be approximated as the difference among subsequent optic flows. By implementing noise in this formula we finally obtain:

$$\hat{O}_+ = O + \varepsilon \delta O_- + N(0, \sigma) \quad (5)$$

Where $N(0, \sigma)$ represents noise in measurement and ε is a small value to ensure temporal locality.

IV. SIMILARITY

The resulting frame from the optic flow algorithm is usually erroneous in two aspects. There is usually a systematic divergence from the correct optic flow field and a random aspect of local noisy estimates. Additionally, these errors are specific to the type of motion and the environment.

The prediction should be compared with the actual frame and then the outcome would be classified as a product of normal motion or a disturbance. In this context we need a divergence measure of the density distributions of pixel velocities. Hence, by comparing the estimated flow with the true flow, we have means of signalling an event.

Thereafter, it is possible to approximate the distribution of divergence and provide a probability level of belief on the current comparison, the next optic flow frame.

Among the possible measures of similarity we distinguish: Euclidean Distance, Cross Entropy, Kullback-Leibler Divergence and Jensen-Shannon Divergence. We have identified that the measure appropriate is *KL*-divergence. Consider two distributions h and g in space $C \in \mathfrak{R}$. Then the *KL*-divergence is defined in the discrete case:

$$KL(h \parallel g) = \sum_{i \in C} h_i \log \left(\frac{h_i}{g_i} \right) \quad (6)$$

The latter is always positive due to the fact that it is always greater than or equal to the cross entropy. In fact expanding the *KL*-divergence results in:

$$KL(h \parallel g) = H(h \parallel g) - H(h) \quad (7)$$

But, the cross entropy between two arbitrary distributions $H(h \parallel g)$ is always greater than the entropy H of the distribution.

KL-divergence does not in fact determine the divergence of distributions but provides a comparative

measure of disorder. Hence, it is a measure of cross entropy disparity. When applying *KL*-divergence we perform a comparison on the deviation of cross entropy from entropy. This allows inference onto the magnitude of divergence between entropy functions of distributions.

The key notion is that during normal motion systematic errors in the transformation from velocity field to optic flow and finally to the predicted frame would be propagated along with the entropy of the optic flow field. Hence, the divergence of the prediction and the actual next field should be constant plus added some noise. In case of an event we should measure a similarity outside the distribution of similarities of normal motion, due to the increased disorder in the field of optic flow.

V. CLASSIFICATION

We have selected a simple classification scheme to provide basic signalling for an event. We assume that when the system moves unaffected then the divergence should be near constant. Thus, a Gaussian distribution describes appropriately the distribution of observations. Hence, the probability of a new observation $P(K_k | K_{k-\lambda:k-1})$ belonging to the distribution can be evaluated from:

$$P(K_k | K_{k-\lambda:k-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{K_{k-\lambda}}^{K_{k-1}} \exp(-(u - \mu)^2 / 2\sigma^2) du \quad (8)$$

Where, K_k is the *KL*-divergence corresponding to optic flow O_k from the previous optic flow O_{k-1} . The parameters mean μ and variance σ , are estimated using observations $k-1$ to $k-\lambda$. By selecting an appropriate threshold we the new observation can be classified. Likewise, we employed a Maxwell distribution which in some situations yielded better results. A detailed discussion can be found in [20].

VI. THE FORWARD MODEL OF OPTIC FLOW

We have laid out the essential components of the model and have set out a number of assumptions. First, we have assumed that there exists a fragment of time in the motion of the robot for which we can assume normality and smoothness. Secondly, we assume that the disturbance caused from an external force applied onto our system would be such that the first constraint would be violated.

The first constraint ensures that there will be local constancy of similarity disparity between actual and predicted frame. Also, it allows for prediction using a spatiotemporal Taylor expansion of the optic flow field. The second constraint ensures that any other motion is disturbance which will cause significant variation in the predicted-actual frames similarities.

The general framework may be summarized below:

1. From the video sequence extract the image frames $I_i, i = k-2, k-1, k, k+1$

2. Obtain the frames of optic flow for $O_i, i = k - 1, k, k + 1$
3. Compute the corresponding optic acceleration from

$$\delta O_k \approx \left(\frac{\partial O}{\partial t} \right) = \lim_{\delta t \rightarrow 0} \frac{O_k - O_{k-1}}{\delta t} \quad (9)$$

4. Predict the next field

$$\hat{O}_{k+1} = O_k + \delta O_k + N(0, \sigma) \quad (10)$$

5. Estimate KL -divergence for the predicted \hat{O}_{k+1} and the observed O_{k+1} fields.
6. Calculate the probability $P(K_k | K_{k-\lambda:k-1})$ as defined in (8), where λ is the number of frames since the previous disturbance (sampling span), and update the distribution parameters.

During the first three steps we capture the necessary image frames, compute the optic flow fields and the optic acceleration along them. In step four the predicted frame is computed for frame k . Thereafter, it is straightforward to compute KL -divergence and the probability of the observation belonging to the distribution of similarities. We have assumed a Gaussian probability distribution of similarities and used a classification threshold to perform classification. Other classifier schemes could also be used but this was sufficient for our purposes.

VII. EXPERIMENTS

We performed experiments using a Koala™ wheeled robot, which would move in a straight-line manner without displaying any significant deviation from the expected route. The latter, could perform a variety of manoeuvres such as accelerating smoothly or rapidly, or turning on the spot, or even following a certain pre-programmed path. Hence, it was possible using a mounted camera to record video sequences of certain motion as we thought fitting. The experiments were carried out within the lab where we had setup an artificial arena of boxes wrapped with newspaper, in order to provide sufficient texture. We shot at two frame rates of 12fps and 25fps and two focal lengths of 36mm and 78mm in controlled lighting environment. We include the results for two experiments.

The first motion, which we refer to as a γ -motion type, consisted of an acceleration phase, traveling with constant speed and a sudden stop at 1.7 seconds; then on the spot right turn and then left turn, traveling forward for about 4.3 seconds and finally halt. Additionally, we wanted to display the ability of the model to actually detect when it is being picked up. A different robot was incorporated, which had significantly less stability during its motion along with random vibrations from the suspension. The focal length was 35mm and we performed analysis at 12fps image sequences. Results are displayed for assumed Gaussian and Maxwell similarity density distributions only for the case of the first robot γ -motion. Complete

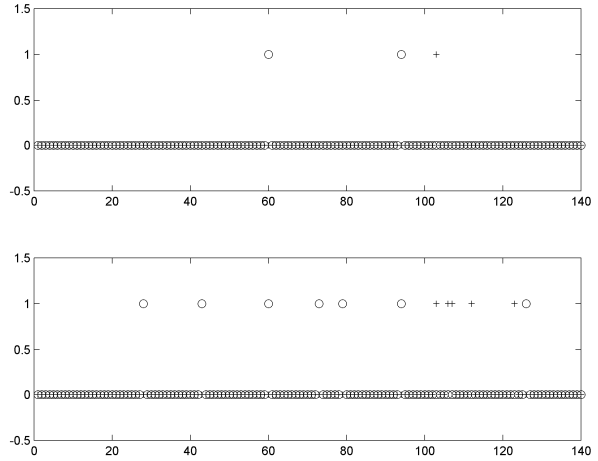


Fig 1: Signaling output of picking up a robot

Top: Gaussian. Bottom: Maxwell. **Angular** component is noted with circle (o) and the **range** with a cross (+).

results are included in [20], which demonstrate the efficacy of our approach.

VIII. EXPERIMENTAL RESULTS

We have included two figures displaying results from the experiments described in the previous section. Looking at Fig. 1 we present the signalling output of the model when a robot is picked up. The two graphs display performance for an assumed Gaussian distribution of similarity density (top) and Maxwell (bottom). Angular and range components are noted with circle and cross respectively. The first graph demonstrates the model's accuracy in detecting only the event of first grabbing the robot and then picking up the robot at frames 60 and 100-120 respectively. Excessive signalling is displayed in the second graph which demonstrates the sensitivity to small variations when altering the classification scheme. These disturbances were generated from the suspension system. We observe that the range component has successfully detected major disturbances whereas the angular is more sensitive to small vibrations. On Fig. 2 KL -divergence is displayed during motion of type γ . The top four graphs demonstrate the divergence of the angular component of the optic flow field. The rest are for the range component. The figures on the left column are for analysis performed at 25fps and on the right at 12fps.

Looking at any of the figures we identify four major disturbances, at about 30, 70, 130 and 260, for 25fps figures. These are better distinguished at lower frame rate when looking on the right side figures (15, 35, 65 and 130 respectively). These may be identified as the four points disturbances during the motion of the robot, sudden stop, turn right, turn left, travel for a while and then stop. Also, one can distinguish the relative nature of the magnitude of disturbance. Particularly at frame 30 (25fps), a sudden stop is obvious. The latter presents in most graphs a significantly large value when compared with the previous

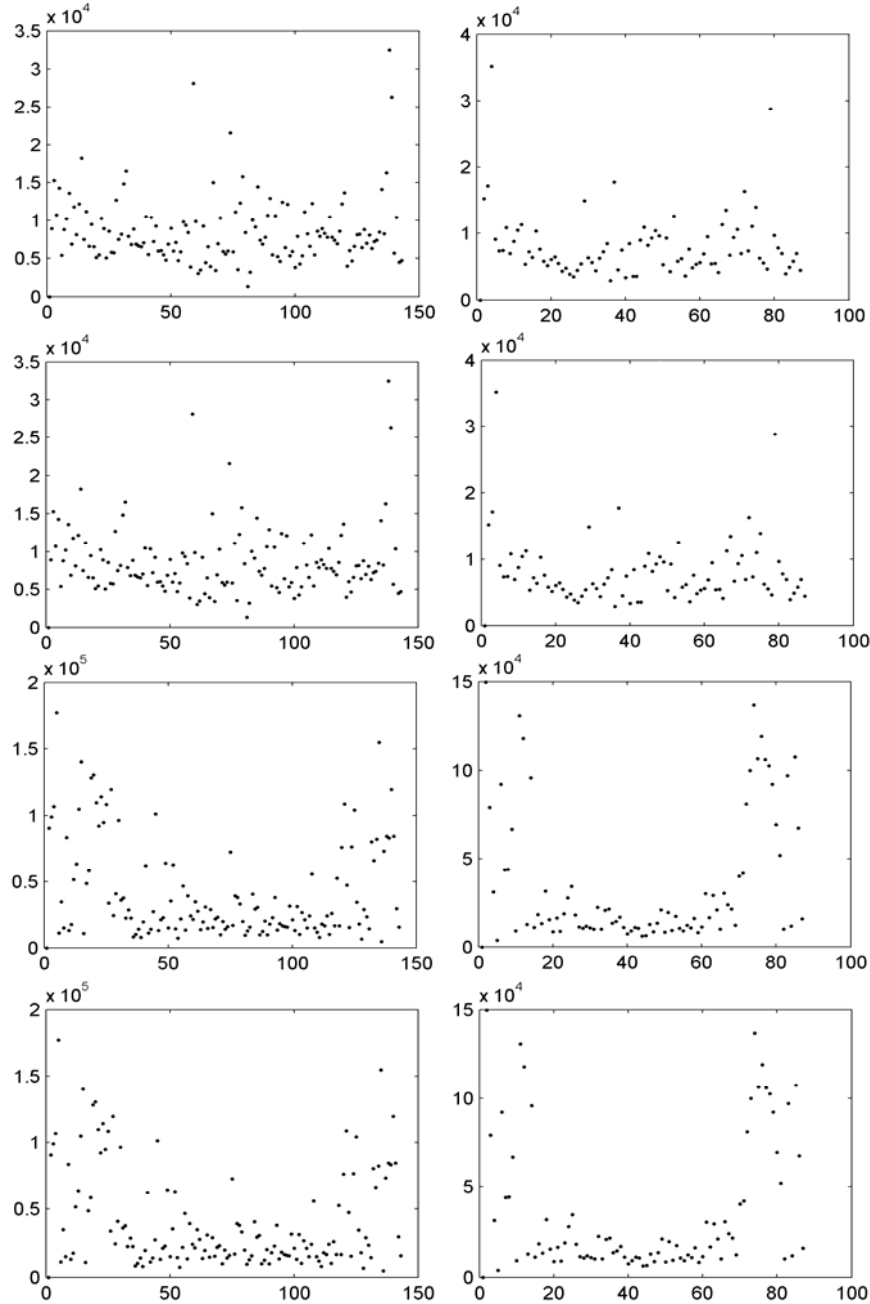


Fig. 2: KL Divergence, γ -motion type

Scales are: horizontal axis: Corresponding frame number, vertical axis: KL Divergence.

From left to right and down, frames per second and focal ratio are noted:

1. angular at 25fps with 35mm, 2. angular at 12fps with 35mm, 3. angular at 25fps with 78mm, 4. angular at 12fps with 78mm, 5. range at 25fps with 35mm, 6. range at 12fps with 35mm, 7. range at 25fps with 78mm 8. range at 12fps with 78mm

observations of the KL-divergence.

IX. DISCUSSION AND CONCLUSIONS

The possibility of a robot being displaced by an external force is rather significant given the increasingly complex areas of application in modern robotics. Specifically, such situations have been observed in various robot competitions where such disturbances often lead to failure of the localization algorithm. We approach the problem by making a prediction, based on a Forward Model of Optic

Flow, and comparing that with the actual optic flow field.

The experiments, firstly, demonstrate the ability to detect an event such as picking up a robot while it moves forward. Secondly, we present that the effects of the disturbance are propagated onto the entire optic flow field and can be detected through the *KL*-divergence. Additionally, further experimentation [20] has shown that detection is feasible in a dynamic environment as well.

Notably, we have demonstrated that prediction of optic flow based on a noisy and ambiguous field is plausible.



Figure 3: γ -motion type images from the sequence

From left to right and bottom selected images from the sequence are displayed. Frame positions (25fps): 10, 40, 75, 95, 105, 115, 130, 200, 255

We have displayed that the detection of external forces can be founded on two related notions, discontinuity and incoherency. The exact prediction of the next optic flow frame is rather insignificant. Our objective is to predict the output of the optic flow algorithm for the next frame given the current. However, the presented algorithm is based on local temporal approximation which unfortunately includes the effects of an external force. If the force is applied smoothly enough then its effects would go undetectable. Nevertheless, that is a function of time and therefore it is possible to select such a frame rate that the disturbance will be rapid and thus detectable. Evidently, further work is needed regarding this matter. Finally, implementing an advanced classification algorithm would render the algorithm robust in performing detection for various motions.

Concluding, we have made negligible assumptions about the environment and minimal regarding the egomotion. This makes the model applicable in a wider area of robotic systems. However, further work is necessitated regarding motion types of higher complexity. Moreover, extensions may be considered regarding classification and model parameter initialization. Even though, we have used a simple classification scheme, we have demonstrated that the algorithm can provide detection of external forces.

ACKNOWLEDGMENTS

We kindly thank Dr. M.J. Black and Dr. S.S. Beauchemin for providing us with source code. This research was partly funded by the Students' Award Agency for Scotland. This publication was kindly endorsed and supported by LANOS and EPFL.

REFERENCES

- [1] T. Morris, D. Britch, "Object-based intra-frame wavelet video coding", Proc. of the 2nd International Symposium, 2001, 599 - 603.
- [2] Adiv G., "Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field", IEEE Trans. Pattern Anal. Mach. Intell. 11, 1989, 477-489.
- [3] R. Berthilsson, A. Heyden, and G. Sparr, "Recursive structure and motion from image sequences using shape and depth spaces", Computer Vision Pattern Recognition, 1997, 444-449.

- [4] T. J. Broida, R. Chellappa, "Estimating the kinematics and structure of a rigid object from a sequence of monocular images", IEEE Trans. Pattern Anal. Mach. Intell. 13, 1991, 497-513.
- [5] C. Colombo, "Time to Collision from First-Order Spherical Image Motion", Robotics & Autonomous Systems, vol. 31, pp. 5-15, 2000.
- [6] B. Horn and B.G. Schunck, "Determining Optical Flow", Artificial Intelligence, vol. 17, nos. 1-3, pp. 185-203, Aug. 1981.
- [7] K. Rangachar, C. Ramesh, "Computer Vision: Principles", chapter 6: "Dynamic Vision", IEEE Computer Society Press, 1991
- [8] J.T. Todd. "Theoretical and biological limitations on the visual perception of three-dimensional structure from motion", In High Level Motion Processing: Computational, Neurobiological and Psychophysical Perspectives, 1998.
- [9] J.A. Perrone, L.S. Stone, "A Model of Self-Motion Estimation within Primate Extrastriate Visual Cortex", Vision Research, vol. 34, no. 21, pp. 2917-2938, 1994.
- [10] A. Singh, "Optic Flow Computation: A Unified Perspective", Los Alamitos, IEEE Computer Society Press, 1991.
- [11] E. Trucco, A. Verri, "Introductory Techniques for 3-D Computer Vision", Prentice Hall, pp. 177-218, 1998
- [12] B. K. P. Horn, B. Schunck, "Determining Optical Flow", Artificial Intelligence, 17 (1981), pp. 185-203
- [13] J. L. Barron, D. J. Fleet, S. S. Beauchemin. "Performance of optical flow techniques". Int. Journal of Comp. Vision, 12(1):43--77, 1994.
- [14] A. Giachetti, V. Torre. , "Refinement of optical flow estimation and detection of motion edges", In ECCV96, pages II:151--160, 1996.
- [15] P. Anandan. "Measuring Visual Motion from Image Sequences", PhD thesis, CS Department, University of Massachusetts, 1987.
- [16] M.J. Black. "Robust Incremental Optical Flow", PhD thesis, Yale University, Department of Computer Science, 1992.
- [17] M. Proesmans, L. Van Gool, E. Pauwels, A. Oosterlinck, "Determination of optical flow and its discontinuities using non-linear diffusion", Computer vision, ECCV '94, Lecture Notes in Computer Science, Vol. 801, Springer, Berlin, 295-304, 1994.
- [18] Produced by Lynn Quam at SRI, University of Western Ontario: ftp://ftp.csd.uwo.ca/pub/vision/testdata/YOSEMITE_DATA/.
- [19] S.S.Beauchemin, <http://www.csd.uwo.ca/faculty/beau/SOFTWARE>
- [20] L. Georgopoulos, "Developing a Forward Model of Optic Flow for Detecting External Forces", MSc Thesis, University of Edinburgh, School of Informatics, 2005, available online at: <http://www.inf.ed.ac.uk/publications/thesis/msc.html> .