Robust Brightness Description for Computing Optical Flow

M. Kharbat, N. Aouf, A. Tsourdos and B. White

Autonomous Systems Group
Department of Informatics and Sensors
Cranfield University, Shrivenham, SN6 8LA

Paper 246

Abstract

Most optical flow algorithms are based on the assumption of brightness constancy of individual pixels while moving on the image plane. Although being attractive analytically, this assumption is often violated under non-ideal visual conditions resulting in poor flow estimates. This paper presents an approach to support the validity of the assumption under such conditions. The method describes the grey-level of each pixel by the content of its neighbourhood using a geometric moment rather than its individual intensity function value. Then, the description of each pixel is normalised and made insensitive to fluctuations of intensity. As a result, the optical flow algorithm becomes much more reliable and robust against visual phenomena like varying illumination, specular reflections and shadows. The proposed approach is applied to a regression method and comprehensive results on synthetic and real data are reported.

1 Introduction

The computation of optical flow is a fundamental task in the analysis of image sequences. The objective is to obtain measurements which describe the 2D pixel motion in images, known as the motion field. This motion is basically the projection of the relative 3D motion, between the camera and the objects' surfaces in 3D space, on the image plane. However, the information which can be obtained by optical flow analysis is limited to what is known as *the apparent motion* [8], that is the optical flow is an approximation of the motion field depending on the observed motion of points in the image.

Optical flow has a broad spectrum of applications in image processing such as estimating the relative objects-observer motion, recovering surfaces' structure, change detection and objects tracking. Most applications demand high accuracy of optical flow measurements to present a close approximation of the true 2D motion field and consequently its corresponding 3D values.

A large number of algorithms for computing optical flow have been proposed, a number of which showed promising results and were subjected to more extensive research (e.g. [2, 13, 1]).

BMVC 2008 doi:10.5244/C.22.46

Successive evaluations and comparisons showed that most optical flow computation algorithms can work only on a specific form of images taken under nearly ideal visual conditions. Such conditions are characterised spatially like constant brightness and textureness of the image, or temporally like close sampling (small motion) and noise-free scenes. This restriction is mainly related to assumptions made when algorithms were formalised to facilitate a solution. As a result, a new generation of optical flow algorithms has been needed in order to employ optical flow outside of the research laboratory.

In this paper, we focus on computing accurate optical flow of image sequences under non-ideal visual conditions including varying illumination effects, shadows and specular reflections. Most optical flow methods rely on a data conservation constraint in the form of brightness constancy assumption. It is assumed that the intensity values of the object's individual pixels stay constant while moving on the image plane. While this assumption can hold approximately within indoor conditions, it is often violated in outdoor environments producing poor estimates of image motion.

Varying illumination, specular reflection, and shadows can all be described as complete of partial change in image contrast. Non-uniform illumination is often considered to be global, occurring throughout the entire image due to change of the source of illumination [4]. Specularities and shadows affect the image in local regions resulting in very high or low saturation of image intensity, respectively.

Several methods have been proposed to deal with the problem. They can roughly be classified into two main groups. The first one has investigated more realistic assumptions and formulated constraints which model specific expected visual conditions (e.g. [4, 7, 11]). These methods perform well when dealing with specific phenomena for which they were designed in the first place. However, they have limited capabilities otherwise. The second group has worked on techniques to detect and explain the violations that exist. Their goal has been to develop methods that perform well even when violations are present (e.g. [4, 19]). Such approaches often search for the prominent motion in each sub-region of the image assuming that the minor motion is coarse error. Apparently, this scheme can only adapt to limited and small violations.

This research investigates the observation of the general grey-level description of a local regions rather than the brightness of its individual pixels. This is achieved by using normalised geometric moments to describe the intensity distribution of each pixel's local neighbourhood. These descriptors have high immunity against fluctuations of intensity values over the image. The principle of conservation of such descriptions is used to obtain a robust and accurate optical flow.

Few researchers have investigated in this direction [6, 20]. In [6], it was suggested to use a set of rotation-invariant Zernike moments as invariant features. Then a set of linear equations that solves for optical flow is formed over each small local region based on the principle of conservation of those invariant features. This work was extended in [20] by integrating an improved mathematical expression of the brightness change. Although the concept showed promising results, it came at a very high computational expense.

The following section provides a quick review of optical flow and geometric moments. It also illustrates the principle of conservation of normalised geometric moments and applies it to the optical flow problem. Experiments on synthetic and real data sets are presented in Section (3) along with their error analysis. Conclusions are presented in Section (4).

2 Robust computation of optical flow

In this section, a robust optical flow computation technique is presented in a working environment of varying illumination, specular reflections and shadows. The proposed approach is based on moments invariant which are insensitive to the absolute intensity of the studied image. We start by a description of the optical flow problem and an introduction of image moments followed by the proposed methodology.

2.1 Optical flow

Most current techniques for computing optical flow are based on the data conservation constraint as an outcome of the image brightness constancy assumption. This constraint works as a model linking between the brightness variation in the image and its corresponding optical flow. This is derived from the idea that the brightness intensity function of small regions remains constant although their position might be changing. This implicitly states that the image irradiance (light at each point of the image plane) is always proportional to the scene radiance (light emitted by each point of the lighting source) [17], and assumes that the proportionality factor is the same across the entire image plane. The brightness constancy assumption can be given as

$$I(x,y,t) = I(x+u\delta t, y+v\delta t, t+\delta t) \tag{1}$$

where $I(x,y,t)$ is the image intensity function and (u,v) is the horizontal and vertical image velocity at a pixel (x,y), respectively. The equation simply states that the intensity value of a pixel (x,y) at time t is the same as the value in a later image at a location offset by the optical flow.

There are several formulations of the data conservation constraint. Gradient-based formulations have shown to be the most popular due to their relative high accuracy and low complexity. This kind of technique deals with dense measurements that is at each image pixel, and relates the spatial and temporal derivatives of the image brightness function at each point to the optical flow vector (u,v). Expanding Equation (1) using Taylor series and ignoring the terms higher than first order gives

$$\frac{\partial I}{\partial x} \cdot u + \frac{\partial I}{\partial y} \cdot v + \frac{\partial I}{\partial t} = 0 \tag{2}$$

If restricted to information at one pixel in the image, the solution of Equation (2) is partially constrained. The typical way of handling this is to add another constraint in the form of a spatial coherence assumption to sufficiently constrain the solution. An implicit way of imposing this constraint is to combine the data conservation information from a local region assuming that the flow in that region is constant or modeled as a parametric function of the image coordinate [12, 3], and solving for the flow as a regression problem. This method has proved to give accurate and fast results when compared with other approaches under standard conditions [2]. It is given that,

$$\sum_{\Re} \frac{\partial I}{\partial x} \cdot \mathbf{u}(\mathbf{a}) + \frac{\partial I}{\partial y} \cdot \mathbf{v}(\mathbf{a}) + \frac{\partial I}{\partial t} = 0 \tag{3}$$

where $\mathbf{u}(\mathbf{a})$ is a model of the flow within a region, $\mathbf{a}$ holds the parameters of the model, and $\Re$ is a local region in the image [4].

2.2 Image moments

Image moments has been an active research field in image processing since it was introduced to the computer vision community in 1962 [9]. It is a major approach in pattern recognition and can be applied to a wide range of applications such as pose estimation, image coding and reconstruction. When applied to an image, moments describe its content with respect to its axis. They are designed to capture geometric information about the image and to extract properties that have analogies in statistics and mechanics.

A general definition of the moment functions Φ_{pq} of order $(p+q)$, of an $N \times N$ image region $I(x,y)$ can be given as

$$\Phi_{pq} = \sum_{x=1}^{N} \sum_{y=1}^{N} \Psi_{pq}(x,y) I(x,y) dxdy \tag{4}$$

where $\Psi_{pq}(x,y)$ is a function of the local coordinates (x,y) of the image region and is known as the moment weighting kernel or the basis set. The indices p,q usually denote the degree of the coordinates x,y respectively, as defined in Ψ.

Geometric image moments are the simplest among moment functions, with the kernel function defined as a product of the pixel coordinates. Geometric moments are also sometimes referred to as Cartesian moments, or regular moments. If the basis set is x^p, y^p, then the $(p+q)^{th}$ order two-dimensional geometric moments are denoted by m_{pq}, and can be expressed as

$$m_{pq} = \sum_{x=1}^{N} \sum_{y=1}^{N} x^p y^q f(x,y) dxdy \tag{5}$$

where $p,q = 0,1,2,3....$ Equation (6) has the form of the projection of the image region $I(x,y)$ onto the monomial $x^p y^p$.

The geometric moments of different orders represent different spatial characteristics of the image intensity distribution. A set of moments can thus form a global shape descriptor of an image. For example, the moment of order zero represents the total intensity of an image. The first order moment provides the intensity moment about the x- or y-axis. The second-order moments are a measure of the variance of the image intensity distribution about the origin. Higher orders characterise different features.

When image moments were introduced to pattern recognition tasks, geometric moments were further developed to have a unique set of shape descriptors which are invariant to one or more of image transformation such as translation, rotation and scale variation [15] or image distortion such as blurring, varying intensity, high contrast and noises (e.g. [10, 18]), so they can be used to identify the elements of the image regardless of their size, orientation, distance of the camera, and the quality of the image. The development and review of different invariants is beyond the scope of this paper; interested readers are suggested to refer to [5] for a comprehensive tutorial. In Section (2.3), we illustrate the development of a moment invariant to non-uniform illumination and intensity change in general and then employ it for computing optical flow.

2.3 Proposed methodology

Various kinds of radiometric degradation are introduced to the image during the acquisition process, mainly due to environmental conditions and also by factors related to the imaging system such as imaging geometry, lens aberration, and random sensor errors. Analysis and interpretation of this kind of images is a key problem in many applications. Moments-based approaches appear to show promise in providing a solution. The basic idea here is to describe the grey-level of a pixel by the content of its neighbourhood using a geometric moment. Moreover, normalising this description grants it immunity against fluctuations of intensity values over the whole neighbourhood.

Equation (6) defines the geometric moment for an image. It is also possible to compute the same moments over a small window around each pixel. These expressions, known as local moments, are associated with image features in a pixel neighbourhood. For a region $\Re$ of size $(2N+1\times 2N+1)$, with the pixel (x,y) as its centre, the corresponding $(p+q)^{th}$ order local moment at (x,y) on $\Re$ is defined as [14]

$$m_{pq}(x,y) = \sum_{i=-N}^{N}\sum_{j=-N}^{N} i^p j^q I(x-i,y-j)dxdy \tag{6}$$

The previous equation can be seen as a neighbourhood operation that can be interpreted as a convolution of the image with a mask (kernel) [14] which can be generated according to the region size. If the same image region is acquired again in different lighting conditions, the intensity function of the new image $I'(x,y)$ can be approximated to be equal to the initial intensity function $I(x,y)$ multiplied by a factor c [10]. That is,

$$I'(x,y) \approx c\cdot I(x,y) \tag{7}$$

The new intensity function will result in moments that differ by the same factor c compared to the moments that are obtained by the initial intensity function; i.e.

$$m'_{pq} \approx c\cdot m_{pq} \tag{8}$$

It is evident that moments can be independent of the factor c by dividing them by any non-zero moment. This normalisation results in descriptors of image regions insensitive to imaging conditions during the acquisition process.

This outcome is employed to solve for the optical flow computation. Every pixel in the image is a centre of a neighbourhood $\Re$ which is approximately conserved in frame sequence. Therefore, similar to the original analysis based on the brightness constancy assumption, the data conservation constraint becomes

$$M(x,y,t) = M(x+u\delta t, y+v\delta t, t+\delta t) \tag{9}$$

where M is the normalised moment value of the local image region surrounding the pixel (x,y).

Expanding the previous equation using Taylor series and ignoring the terms higher than the first order gives

$$\frac{\partial M}{\partial x}\cdot u + \frac{\partial M}{\partial y}\cdot v + \frac{\partial M}{\partial t} = 0 \tag{10}$$

A solution can be found if there are two normalised moments computed for each pixel. However, such an approach is not stable numerically because of the approximations made when calculating the temporal and spatial derivatives. Alternatively, the optical flow can be solved as a regression problem, over a small neighbourhood, in which a constant flow is assumed. The overdetermined system of linear equations can be written as

$$A\mathbf{u} = b \tag{11}$$

where **u**=[u,v], and
the i^{th} row of A is $[\frac{dM^i}{dx}\ \frac{dM^i}{dy}]$, where i denotes the order of the normalised moment.

The least-squares solution for the computation of optical flow can then be given as

$$\mathbf{u} = [A^T A]^{-1} A^T b \tag{12}$$

The next section provides more details about implementing the method along with experiments and error analysis.

3 Experimental results

We have tested the performance of our technique by conducting experiments on synthetic and real sequences. Experiments were performed over regions of different sizes using the second order normalised moment (divided by the first order moment). Local moments were computed by convoluting the images masks of sizes similar to the studied regions.

All experiments were held with no pre-smoothing. Derivatives were obtained by averaging the four first differences taken over a $2 \times 2 \times 2$ cube in the image sequence. Quantitative results of the synthetic sequences were compared to those of the regular regression method. The flows were filtered out removing unreliable estimation values. Those were identified according to the eigenvalues (λ_1, λ_2) of $A^T A$ in Equation (12), which depends on the spatial derivatives of the moments. If $\lambda_1 + \lambda_2 > \tau$, the flow vectors are computed from Equation (12) where τ is a pre-set threshold value. Although multiple normalised moments can be used in describing one pixel, we found that one moment can capture the details of the neighbourhood reasonably accurately and there is no need to compute further descriptors.

3.1 Error analysis

In order to quantify the accuracy of our technique, we compute relative error measures between the correct flow $\mathbf{u_c} = [u_c, v_c]$ and the estimated one $\mathbf{u_e} = [u_e, v_e]$ in magnitude and direction [16]. We report their mean and standard deviation in addition to a robustness measure similar to the one used in [1].

The first error measure describes the relative error in velocity magnitude (E_r) while the second error measure, the directional error (E_d), captures the deviation from the correct flow and thus how accurately the correct direction has been recovered:

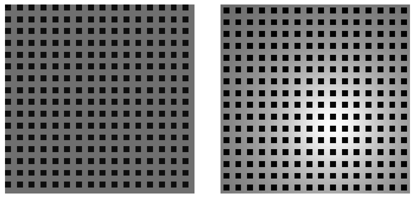
$$E_r = \frac{|(\|\mathbf{u_c}\| - \|\mathbf{u_e}\|)|}{\|\mathbf{u_c}\|} \cdot 100[\%] \qquad E_d = \arccos\left(\frac{\mathbf{u_c} \cdot \mathbf{u_e}}{|\mathbf{u_c}||\mathbf{u_e}|}\right)[^\circ]$$

We also report a robustness measure RX denoting the percentage of pixels that have an error measure above X. We report $R15\%$ and $R7.5^\circ$ for the magnitude and directional measures, respectively.

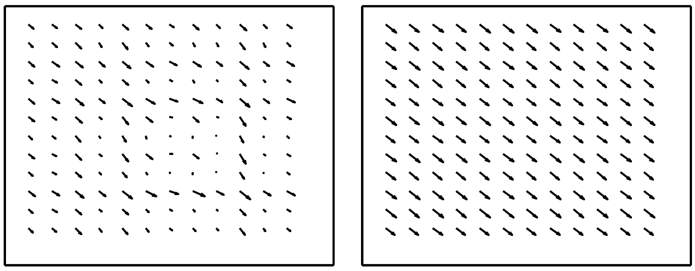
3.2 Experiments

3.2.1 A synthetic sequence

We synthesised a 64×64 image sequence to highlight the case when the lighting condition of a frame is different than the one before it. Figure (1) shows two frames of a square-pattern sequence translating with a velocity $(1, 1)$. The second frame (b) is affected by an Omni-type lighting. We compute optical flow first using a regression method with the regular brightness description (c) and then using the proposed methodology (d). The result illustrates the improved outcome. To compare the two approaches quantitatively magnitude and directional errors are worked out for both techniques in addition to two robustness measures and listed in Table (1).



(a) (b)



(c) (d)

Figure 1: Synthetic sequence. (a) Frame 1 (b) Frame 2 (c) Flow estimates using the regular brightness description (d) Flow estimates using the proposed methodology.

(a) Squares pattern - Original

	Error [%]	S.D.	Error [°]	S.D.	R15%	R7.5°
Existing	11.23	11.69	4.17	9.20	27.50	3.41
Proposed	7.26	6.58	3.29	8.09	5.05	4.39

(b) Squares pattern - Illumination affected

	Error [%]	S.D.	Error [°]	S.D.	R15%	R7.5°
Existing	34.33	22.42	10.81	12.88	73.78	49.97
Proposed	8.41	7.90	5.84	9.40	7.91	14.18

Table 1: Error analysis for the synthetic sequence.

3.2.2 A real sequence

An illumination-affected real sequence was used to validate the proposed method qualitatively. Figure (2) show frames of the Rubik sequence, widely used for testing optical flow algorithms, where a cube is rotating counterclockwise on a turntable. The frame in (a) shows the first original frame of the sequence while Figures (b–d) correspond to the second frame subject to an increasing lighting effect that is altering the contrast of the image.

A sub-sampled version of the estimated optical flow is shown in Figures (e–h). (e) and (f) were obtained based on the original brightness description for frames (b) and (c), respectively. On the other hand, (g) and (h) were obtained using the proposed methodology for the same frames. The results demonstrates the robustness of the proposed approach in sequences influenced by un-modelled variation of intensity.

4 Discussion and conclusions

When computing optical flow, the brightness constancy assumption is well-known to be fragile and easily violated in environments of non-ideal visual conditions. In this paper, we present a method in which each pixel is described by the grey-level of its neighbourhood using geometric moments. The description is then normalised and made insensitive to the absolute intensity value of each neighbourhood.

The paper proposes to replace the usual intensity function with a moment-based one as an initial stage to tackling challenges of computing optical flow. Our experiments on synthetic and real sequences illustrated the advantages of the proposed method. As found out from Table (1), the proposed approach performs better even under ideal visual conditions. In the experiment on the real sequence, it is apparent that the quality of the flow estimated by the proposed method is highly robust and compares favourably to techniques with the original brightness assumption.

The proposed approach is very competitive in terms of execution time as well. The additional step of calculating normalised moments by convoluting the image with a mask adds very little computation expense so that the overall computation required for both approaches is almost the same. The time requirement for the proposed method is compared against that of the existing method in Table (2)[1].

Future research includes developing techniques to tackle violations of optical flow underlying assumptions other than brightness constancy. Methods to detect outliers are of high interest to explain particular properties of the scene like motion boundaries, lack of texture, occlusions, etc. Modelling of various visual conditions can still be useful to get even more accurate flow estimates especially when specific visual distortions are foreseen.

Sequence	Region size [pixel]	Proposed - time [s]	Existing - time [s]
Synthetic	3	0.06	0.06
Rubik	7	0.98	0.96

Table 2: Time requirement for the proposed approach.

[1]MATLAB implementation on an Intel Core2 Duo 1.86GHz machine with 1GB RAM

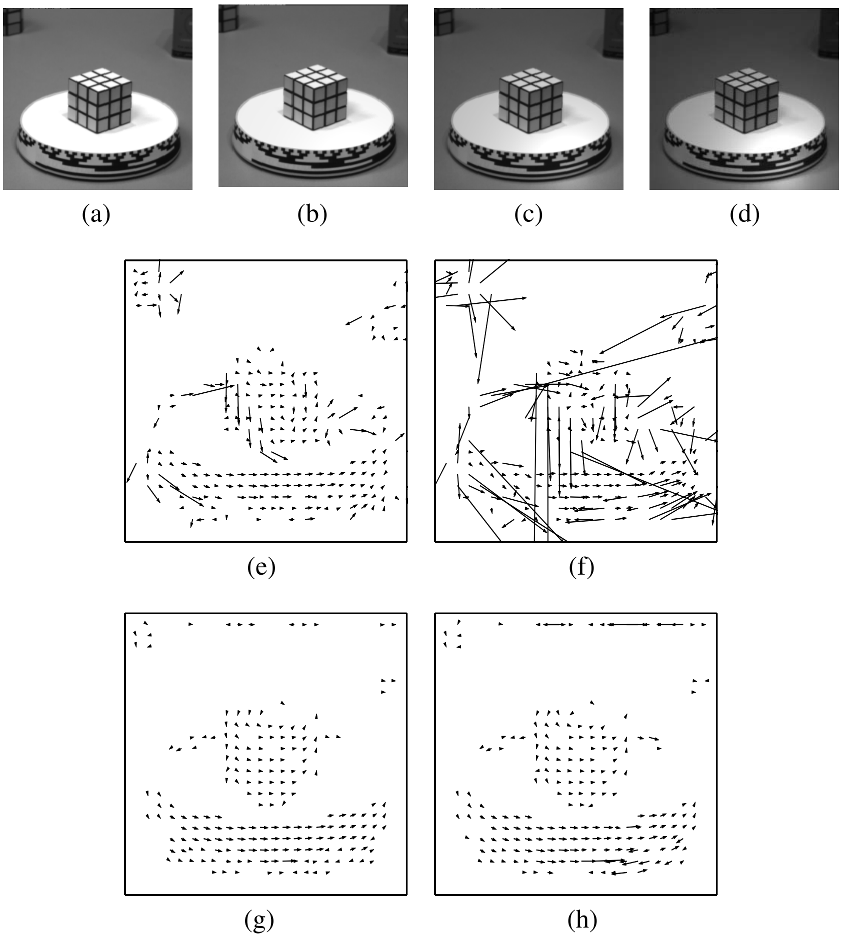


Figure 2: Rubik sequence. (a) Original frame 1. (b-d) Frame 2 affected by increasing illumination effect. (e-f) Flow estimates using a regression method based on the regular brightness description for frames (b-c), respectively. (g-h) Flow estimates using the proposed method for frames (b-c), respectively.

References

[1] Simon Baker, Stefan Roth, Daniel Scharstein, Michael J. Black, J.P. Lewis, and Richard Szeliski. A database and evaluation methodology for optical flow. *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8, 14-21 Oct. 2007.

[2] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[3] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. *ECCV*, 92:237–252.

[4] M.J. Black, D.J. Fleet, and Y. Yacoob. Robustly estimating changes in image appearance. 78(1):8–31, April 2000.

[5] Jan Flusser, Barbara Zitova, and Timas Suk. Moments and moment invariant in image analysis. Technical report, Accessed online on the 25th of Feb. 2008 http://staff.utia.cas.cz/zitova/tutorial/icip07.html, 2007.

[6] S. Ghosal and R. Mehrotra. Robust optical-flow estimation using semi-invariant local features. 30(2):229–237, February 1997.

[7] H.W. Haussecker and D.J. Fleet. Computing optical flow with physical models of brightness variation. 23(6):661–673, June 2001.

[8] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. In *Shape recovery*, pages 389–407. Jones and Bartlett Publishers, Inc., 1992. 136790.

[9] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *Information Theory, IEEE Transactions on*, 8(2):179–187, 1962.

[10] T.M. Hupkens and J. de Clippeleir. Noise and intensity invariant moments. 16:371–376, 1995.

[11] Yeon-Ho Kim, Aleix M. Martinez, and Avi C. Kak. Robust motion estimation under varying illumination. *Image and Vision Computing*, 23(4):365–375, 2005.

[12] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI81*, pages 674–679, 1981.

[13] B. McCane, K. Novins, D. Crannitch, and B. Galvin. On benchmarking optical flow. *Computer Vision and Image Understanding*, 84:126–143, 2001. doi:10.1006/cviu.2001.0930.

[14] R Mukundan and K R Ramakrishnan. *Moment Functions in Image Analysis: Theory and Application*. World Scientific, 1998.

[15] C. Ruberto and A. Morgera. A comparison of 2-d moment-based description techniques. pages 212–219, 2005.

[16] H. Spies, B. Jahne, and J.L. Barron. Range flow estimation. 85(3):209–231, March 2002.

[17] Emanuele Trucco and Alessandro Verri. *Introductory techniques for 3-D computer vision*. Prentice Hall, 1998.

[18] L.J. Van Gool, T. Moons, and D. Ungureanu. Affine/photometric invariants for planar intensity patterns. pages I:642–651, 1996.

[19] M. Ye and R.M. Haralick. Two-stage robust optical flow estimation. pages II: 623–628, 2000.

[20] Qiang Zhou, D. Chelberg, Yujing Zeng, and J. Starzyk. Robust optical flow estimation using invariant feature. *Proceedings of the 33rd Southeastern Symposium on System Theory*, pages 263–267, Mar 2001.