A Local Approach for Robust Optical Flow Estimation under Varying Illumination

Yeon-Ho Kim[1], Aleix M. Martínez[2] and Avi C. Kak[1]

[1]School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN 47907, USA
{yeonho, kak}@ecn.purdue.edu

[2]Department of Electrical and Computer Engineering
The Ohio State University, Columbus, OH 43210, USA
aleix@ee.eng.ohio-state.edu

Abstract

The problem of motion estimation, in general, is made difficult by large illumination variations and by motion discontinuities. In recent papers, we and others have proposed global approaches to deal with both problems simultaneously within the regularization framework. A major drawback of such global methods is that several regularization parameters responsible for the integration of the illumination and motion components need to be determined in advance. This has reduced the applicability of global methods. In this paper, a parameter-free local approach, which solves a linear regression problem using a simple parametric model, is presented. To achieve robustness for the linear regression problem, we introduce a modified version of the least median of squares algorithm. We show quantitative error comparisons between the results obtained by our local approach and those produced by several global methods. Our results show that our local method is comparable to the best results obtained by the global approaches yet does not require any manual selection of parameters.

1 Introduction

The estimation of motion in images is a basic task in computer vision with many interesting applications such as segmentation, structure from motion, robot navigation and object recognition. A primarily goal in the field is to estimate the scene or object motion as precisely as possible. A classical way to compute the scene or object motion from a sequence of images is the optical flow technique, introduced by Horn and Schunck [5].

Most optical flow estimation techniques are based on either the brightness constancy or the motion smoothness constraint or both [2, 6]. However, in realistic scenarios, such constraints almost never hold. For example, a pixel can change its brightness value because an object moves (translates or rotates) to another part of the scene with different lighting or because the illumination of the scene (globally or locally) changes in time.

One way to deal with large illumination variations is to replace the assumption of constancy of pixel brightness value with a more realistic model. Gennert and Negahdaripour [4] were among the first to propose a method robust to illumination changes. Their algorithm assumes that the brightness at time $t + \delta t$ is related to the brightness at time t through a set of parameters that can be estimated from the image sequence.

The other basic assumption, that of motion smoothness where one assumes neighboring pixels to belong to the same object, is obviously violated when an object moves in front of another at a different velocity. This problem can be solved by recovering the structure of the motion as indicated by the majority of the pixels in the vicinity of a motion boundary while eliminating those pixels that are inconsistent with the majority. The latter pixels are referred to as "outliers". The optical flow estimation problem with outliers falls into a general robust parameter estimation problem. Many different techniques have been proposed to estimate accurate optical flow in the presence of multiple motions by adapting a robust estimation technique that is less sensitive to the outliers than standard estimation methods [1, 3, 6, 11].

Most of these robust estimation techniques for optical flow calculation can be classified into two different groups. The first group corresponds to the global methods where estimation is made robust by minimizing a regularization function for all the pixels of the image. For example, Black and Anandan [3] use an M-estimator instead of the quadratic estimator as the regularization function to be minimized. The second group includes the local approaches, which addresses robustness by solving a set of over-determined linear equations for pixels in a local area of the image. Examples are the algorithms of [1] and [11] where the authors use the least median of squares (LMS) method instead of the standard least squares (LS) method to solve the regression problem. The LMS method has a high breakdown point of 0.5, which is regarded as a practical limit in robust statistics, and has been broadly used in computer vision [15].

In our previous work, we presented a formulation that simultaneously addressed the two problems – the problem caused by large illumination variations and the problem caused by motion discontinuities – in the form of a global method [6]. To do this, we reworked Gennert and Negahdaripour's [4] formulation within the robust statistical framework of Black and Anandan [3]. This allowed us to be simultaneously robust to both motion discontinuities and illumination changes. Other global approaches that integrate a robust estimation framework and a varying illumination model can be found in [14, 16].

However the global methods described in the preceding paragraph have many parameters that need to be carefully tuned in order to obtain desirable results. This drawback has lead us to the local approach described in this paper. For this, we solve Negahdaripour and Yu's [9] over-determined system of equations with a technique similar to that originally proposed by Rousseeuw and Leroy [13].

The rest of this paper is organized as follows. In Section 2, we summarize our global algorithm. Section 3 presents our new local approach. Section 4 details on the experimental results. We conclude in Section 5.

2 A Global Approach for Robust Optical Flow Estimation under Varying Illumination

To estimate optical flow robustly, with regard to the problem caused by large motion discontinuities and the problem caused by large illumination variations, we have previously integrated Gennert and Negahdaripour's varying illumination model [4] with Black and Anandan's robust estimation framework [3] in a form of global approaches [6]. In our previous integrated method, optical flow is estimated by minimizing the following regularization function over all the pixels of the image

$$E = \int \int \left\{ \begin{array}{c} \rho(I_x u + I_y v + I_t - mI - c, \sigma_d) \\ + \lambda_s [\rho(\nabla u, \sigma_s) + \rho(\nabla v, \sigma_s)] \\ + \lambda_m \rho(\nabla m, \sigma_m) \\ + \lambda_c \rho(\nabla c, \sigma_c) \end{array} \right\} dx\,dy, \tag{1}$$

where $I_x u + I_y v + I_t - mI - c$ is the error term for the brightness constraint used, I_x, I_y and I_t are the spatial and temporal partial derivatives of the image, $u = \frac{\partial x}{\partial t}$ and $v = \frac{\partial y}{\partial t}$ are the motion parameters (i.e., optical flow), $m = \frac{\partial M}{\partial t}$ and $c = \frac{\partial C}{\partial t}$ are the radiometric parameters in which M and C are a multiplicative and an additive term for varying illumination effect used in Gennert and Negahdaripour model.

To regularize the brightness constraint, smoothness constraints are used for motion and radiometric parameters in the form of minimizing the gradient of parameters, u, v, m, and c. For these constraints, λ_s, λ_m, and λ_c are used to determine the relative importance of each term. To cope with outliers mostly at the motion boundaries, we use the Lorentzian function defined as $\rho(x, \sigma) = \log(1 + \frac{1}{2}(\frac{x}{\sigma})^2)$, where σ controls the weight to be given to the variable x. Using the same technique as in [3], we can find the global minimum of the objective function E, by first choosing large σ values which create a convex approximation for E, and then decrease σ values to find a more accurate minimum. This process is repeated until we converge to a solution. Implementation details are in [6].

Though we could present superior results to the other methods that deal with the problem of motion discontinuities and the problem of varying illumination separately, we could not free ourselves from the cumbersome task of parameter tuning that is an unpleasant but inherent drawback of any global method. To get the best result, one needs to experimentally determine the values for each of the λs. To avoid this shortcoming of global methods, we have turned our efforts in the design of the local method presented below.

3 A Local Approach for Robust Optical Flow Estimation under Varying Illumination

In contrast to the global method mentioned above, which needs manual parameter tuning, the local method presented in this section does not. This can be achieved thanks to the assumption that motion or varying illumination effect for all pixels in a small local region of images can be restricted to a simple parametric model. The parameters of such a model can be estimated from a set of over-determined linear equations by means of the classical least squares algorithm [8, 9].

Unfortunately, the least squares approach is known to be very sensitive to outliers, as for example to those produced by motion discontinuities. The goal is thus to define a robust algorithm able to estimate the correct value for each of the parameters even when outliers are present. Among many different local robust parametric estimation techniques previously used in computer vision [15], the least median of squares (LMS), the iteratively re-weighted least squares (IRLS), as well as voting-based approaches have shown success [1, 7, 10, 11]. Since the computing time of the voting technique increases exponentially with the number of parameters to be estimated and the IRLS method relies on a good initial guess, we will focus on the LMS algorithm. The general algorithm of the least median of squares method has been well studied; see, for example, [13].

We use the LMS algorithm to *locally* solve an optical flow model which *integrates both the radiometric and motion parameters*. In this paper we show derivations for the following model [4]:

$$(u,v,m,c) = \arg\min_{(u,v,m,c)} median_{x,y\in R}\left(I_x u + I_y v + I_t - Im - c\right)^2,\qquad(2)$$

where R is the local region. Since there is no explicit closed-form solution for this equation, we need to define a way to find an approximation. To this end, we use a modified method based on the standard LMS algorithm of Rousseeuw and Leroy [13].

Formally, we rewrite (2), as

$$\hat{\theta} = \arg\min_{\theta} median_i\, r_i(\theta)^2,\qquad(3)$$

where $\theta = (u,v,m,c)^T$. The indices $i = 1,2,...,n$ index the image point $(x,y)_i$ in the region of interest R and the residual error $r_i(\theta) = (I_x, I_y, -I, -1)\cdot\theta + I_t$ at the coordinate $(x,y)_i$. Following the standard LMS method, p (i.e., 4 in our case) number of pixels are randomly chosen from R and a temporary solution $\tilde{\theta}$ is calculated from a system of p linear equations (*we modify this part and the modification is explained at the end of this section*). Using the temporary solution $\tilde{\theta}$, the squared residuals $r_i(\tilde{\theta})^2$ are calculated for all the other pixels in R and then the median $M = median_i r_i(\tilde{\theta})^2$ is searched. After repeating this calculation of median M_j for $j = 1,2,...,s$ sub-samples, we choose a set of parameters $\hat{\theta}$ corresponding to the minimum median $\hat{M}$ among $M_1, M_2,...,M_s$. To get a global minimum median, the number of sub-samples s is $_nC_p$ (i.e., the number of all possible combinations of p linear equations from n pixels in the region R). Obviously the computation cost of such a naive approach is too high. For this reason, a smaller number of sub-samples are sought – for which the probability of at least one of the s sub-samples consisting of p inlier data is to be close to 1. For large n/p, the probability can be expressed by $1 - (1 - (1 - \varepsilon)^p)^s$ where ε is the fraction of outliers in the region of interest. Moreover, to improve the statistical efficiency of the crude LMS method, the final solution is sought by employing the weighted LS method whose weights are defined to reject outliers as follows.

The initial scale estimate σ^0 is calculated first based on the minimum median $\hat{M}$ and some correction factors, i.e.

$$\sigma^0 = 1.4826(1 + 5/(n-p))\sqrt{\hat{M}},\qquad(4)$$

where the factor $1/\Phi^{-1}(0.75) = 1.4826$ is selected to guarantee that $median_i|z_i|/\Phi^{-1}(0.75)$ is a consistent estimator of σ when the z_i's are distributed as normal distribution of

$N(0, \sigma^2)$ and $\Phi(x)$ is the standard normal cdf for x. The multiplication with the factor $1 + 5/(n - p)$ is introduced empirically for more feasible estimations in the case of small samples [13]. The initial scale estimate σ^0 is then used to determine an initial weight w_i for each pixel at (x_i, y_i):

$$w_i = \begin{cases} 1 & \text{if } |r_i/\sigma^0| \leq 2.5 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The bound is chosen to be 2.5 based on the assumption that there are few residuals larger than 2.5σ for the normally distributed data [13]. The final robust scale estimate $\hat{\sigma}$ is then calculated using data in which outliers are rejected by the initial weights as

$$\hat{\sigma} = \sqrt{(\sum_{i=1}^{n} w_i r_i^2) / (\sum_{i=1}^{n} w_i - p)} \,. \tag{6}$$

The final weights are computed by substituting σ^0 by $\hat{\sigma}$ in (5) and by using the weighted LS method

$$\theta = \arg\min_{\theta} \sum_{i=1}^{n} w_i r_i^2 \,. \tag{7}$$

Note that this algorithm searches for a good approximate solution based on the assumption that a majority (more than 50%) of pixels represents a coherent motion correctly. To support this assumption, the accuracy of the brightness constraint equation is very important and it relies on the spacial and temporal derivatives of the image intensity. The partial image derivatives can be calculated accurately by convolving the derivative of 3D-Gaussian function with a large number of images [1]. However due to many practical reasons, optical flow estimation using only two image frames has usually been preferred. In such a case, to diminish the effect of crude image derivatives, greater care needs to be taken in selecting good samples, for example choosing p number of pixels that are close together or have high gradient magnitudes [11]. In any case, if the observations are largely contaminated by noise such as large illumination changes, each temporary solution $\tilde{\theta}$ from only p samples will be scattered around the correct one and the probability of the temporary solution being the correct one will be very low. To obtain a good solutions from nosy data, we randomly choose a region smaller than the original local area [11], but use *all the pixels in the region* instead of taking p pixels only. Then, we solve a set of over-determined equations using the least squares method.

4 Experimental Results

We will first show optical flow results on synthetic and real data for which the ground-truth is known. By grouping the results into two groups (global and local method), we will be able to compare general performance of optical flow estimations between these two paradigms. In particular, we want to show *the sensitivity of the global methods to the weighting parameters*. We will also show *the improvement obtained by using integrated methods* by comparing their results with those obtained by using separated methods. We also include the results that show *the superiority of our modified LMS method* by comparing its results with the ones from the standard LMS originally proposed by Rousseeuw and Leroy.

The global methods for which the results are shown here include the method of Horn and Schunck [5] (G1); the method of Black and Anandan [3] (G2); the method of Gennert and Negahdaripour [4] (G3); and our previous integrated algorithm [6] (G4). The methods of (G1) and (G2) are based on the brightness constancy constraint but the methods of (G3) and (G4) are based on a varying illumination model. The methods of (G1) and (G3) use the quadratic estimator whereas the methods of (G2) and (G4) use robust estimators. For the local methods, we show results with the method of Lucas and Kanade [8] that is based on the brightness constancy constraint and the LS method (L1); the method based on the brightness constancy constraint and the standard LMS method (L2-1)[1]; the method based on the brightness constancy constraint and our modified LMS method proposed in this paper (L2-2); the method of Negahdaripour and Yu [9] that is based on the varying illumination model and the LS method (L3); the integrated method that is based on the varying illumination model and the standard LMS method (L4-1), and the integrated method based on the varying illumination model and the modified LMS method proposed in this paper (L4-2).

We will then show the results on two real image sequences for which the illumination is not set constant. In the first example, two image frames are taken from an image sequence where a soda can moves in front of a stationary background. The second example consists of a person lifting his arm and rotating his head simultaneously. Both image sequences incorporate a real illumination change caused by a moving light source which is located at the right hand-side of the camera. The reason we do not show comparative results on these images is because the ground truth is not known. However it is instructive to examine, albeit in a qualitative sense only, the results obtained with distinct algorithms.

To make a fair comparison for different techniques shown here, we apply several minor modifications to the original implementations. For instance, we do calculation of image derivatives using two image frames as described in [5] for all the methods used in this paper. To include large estimation errors in the comparison, we do not limit the maximum magnitude of the optical flow to be estimated in any methods. We do not use any confidence measurements that do not contribute to the optical flow computation itself. For all the local methods, we use the same size and same weighting values in all of the windows.

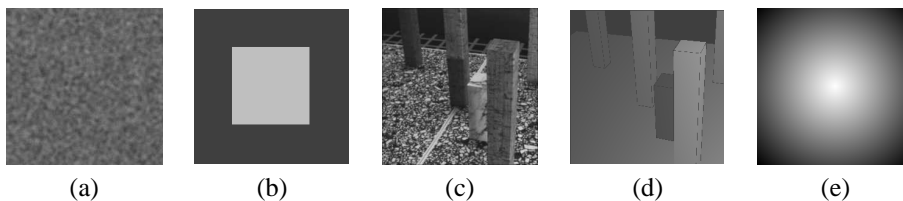| (a) | (b) | (c) | (d) | (e) |

Figure 1: For the Random-dot sequence, (a) first image frame, (b) correct flow. For the Otte sequence, (c) first image frame, (b) its correct flow. (e) Multiplicative varying illumination effect.

Figure 1(a-b) show the first image frame and the horizontal component of the correct optical flow between the first and second image frames of the synthetic random-dot image

---

[1]This method is exactly same as the method of Bab-Hadiashar and Suter [1] except that we use image derivatives calculated from only two image frames as described in [5].

sequence. The second image frame for this image sequence is created by translating the pixels on the square at the center of the first image frame by one pixel to the right and by one pixel downward and translating the pixels on the background image by one pixel to the left and by one pixel upward. To simulate illumination change, the intensity of pixel in the second image frame is multiplied by a factor that varies linearly from the center of the image to the corner of the image radially (1.25 at the center and 0.75 at the corner, as shown in Figure 1(e)), an offset (of 10) is then added to the result.

Figure 1(c) is an image frame from the original "Otte" sequence [12] but the size of image is reduced to half to avoid comparison for erroneous large motions. Figure 1 (d) is a gray scale image representing the horizontal component of the correct optical flow between the first image frame and the second image frame for the Otte sequence. The Otte sequence is taken by a camera moving toward a scene and the objects in the scene are stationary except for a marble block that moves to the left. To simulate varying illumination, the intensity of the pixels at the next image frame is modified by the same factors that were used in the random-dot sequence.

Random-dot sequence				Otte sequence			
Method	Avg.	Std.	Dens.	Method	Avg.	Std.	Dens.
G1(25)	20.16	17.73	100	G1(50)	15.84	12.97	100
G1(1.0)	46.85	35.29	100	G1(1.0)	46.22	29.96	100
G2(0.5)	30.62	31.17	100	G2(2.0)	18.30	9.16	100
G2(1.0)	38.98	40.17	100	G2(1.0)	20.21	19.90	100
G3(0.4,100,100)	5.56	7.66	100	G3(0.3,10,10)	10.89	6.11	100
G3(1,1,1)	52.14	0.90	100	G3(1,1,1)	17.57	11.75	100
G4(0.2,10,10)	3.54	10.90	100	G4(0.4,1,1)	9.88	7.23	100
G4(1,1,1)	49.46	4.38	100	G4(1,1,1)	10.11	6.38	100
L1	23.75	23.66	100	L1	24.51	25.40	100
L2-1	10.68	12.25	63.20	L2-1	20.46	17.37	81.28
L2-2	13.18	12.85	61.57	L2-2	12.23	11.98	61.09
L3	5.52	9.70	100	L3	9.09	8.13	100
L4-1	7.17	11.98	100	L4-1	9.06	7.79	100
L4-2	3.89	8.65	100	L4-2	8.75	7.73	100

G1: Horn and Schunck Method, G2: Black and Anandan Method,

G3: Gennert and Negahdaripour Method, G4: Our global integrated method presented in [6],

L1: Lucas and Kanade Method based on the least squares method,

L2-1: The standard LMS Method,

L2-2: The modified LMS-Method proposed in this paper,

L3: Negahdaripour and Yu Method based on the LS method,

L4-1: The integrated method based on the standard LMS Method,

L4-2: The integrated method based on the modified LMS method proposed in this paper

Table 1: Comparative results on the Random-dot sequence and the Otte sequence

For the quantitative comparison, we calculate errors between the estimated values and the correct values of the optical flow using the error measurement method presented in [2]. There the authors calculate the errors in optical flow by measuring the angle between the 3D normalized versions of the correct optical flow vector and the estimated optical

flow vector. The normalized 3D vector is defined as

$$\vec{v} = \frac{1}{\sqrt{u^2 + v^2 + 1}} (u, v, 1)^T \qquad (8)$$

and the angular error is calculated as $E = \arccos(\vec{v_c}, \vec{v_e})$ where the normalized 3D vector $\vec{v_c}$ corresponds to the correct velocity and $\vec{v_e}$ to the estimated velocity.

In Table 1, the columns are divided into two parts depending on the image sequence used. The first part is for the random-dot sequence and the second part for the Otte sequence. The abbreviations representing the different methods are expanded at the bottom of the table. The first column in each part of the table determines the method applied to calculate the error statistics. "Avg." and "Std." denote the mean and the standard deviation of the error associated with optical flow estimation. "Dens." denotes the number of pixels that produce valid value of optical flow divided by the total number of pixels that are used for the optical flow calculation in the image frame. The number in bracket at the side of G1 and G2 methods denotes the value of weighting parameter for the motion smoothness. The numbers at the side of G3 and G4 methods denote the value of weighting parameter for the motion smoothness, the values of weighting parameters for the smoothness of the multiplicative radiometric term and for the smoothness of the additive radiometric term respectively. The first results of the same global method are obtained with manually tuned parameters. The second results are obtained by un-tuned parameters (i.e., all parameters are set to 1). For all the local methods, 15 by 15 windows are used and for all the LMS based methods, 30 sub-samples are used. We note that the results of our local method are comparable or superior to the result of our global algorithms. Yet the local method does not require any parameter tuning. We can also verify the superiority of our integrated approach to the separated approaches by comparing the results of L4-1 and L4-2 with the results of L2-1 and L2-2. Moreover, the results of L4-2 that uses our modified LMS method proposed in this paper show clear improvement by the results of L4-1 that uses the standard LMS methods.
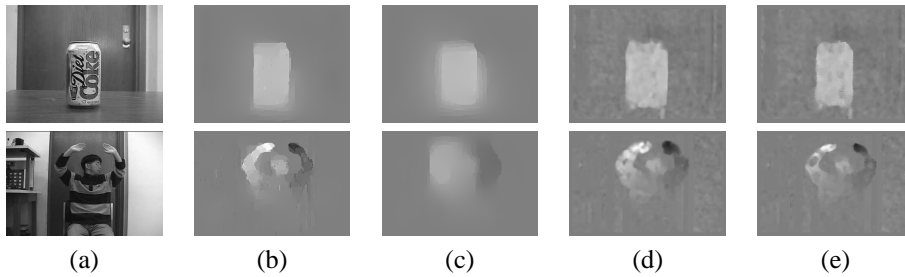

| (a) | (b) | (c) | (d) | (e) |

Figure 2: (a) First image frames. (b) Results obtained by our global method using optimally tuned parameters. (c) Results obtained by our global integrated method using un-tuned parameters. (d) Results obtained by Negahdaripour and Yu method based on the LS method. (e) Result obtained by our local integrated method based on the modified LMS method proposed in this paper.

For the real image case, we show results for two real image sequences taken under a real illumination variation. The first image sequence, Can sequence, shown in the first row of Figure 2 (a) corresponds to a scene of a stationary background and a can that moves

along the plane parallel to the scene in the foreground. The second image sequence, Human sequence, shown in the second row of Figure 2 (a) corresponds to a more complex scene of a stationary background and a human lifting his arms and rotating his head in the foreground. Two image frames are taken from both image sequences and used for the calculation of optical flow. Since no ground truth is available for these sequences, the comparison can only be visual. As we can easily expected from the results of the previous experiments, results of the methods based on the brightness constancy constraint contain large errors and are omitted in Figure 2 to save space. Obviously, the results of the global method are very sensitive to the value of the parameters as expected. Our local method based on the modified LMS method shows clear improvement from the method based on the LS algorithm especially at motion boundaries.

5   Conclusion

In this paper, we have presented a local approach for estimating optical flow on image sequences recorded under conditions of varying illumination and with large motion discontinuities. Our current local approach is to be contrasted with the global method we used in our previous work. In the local approach, we integrate Gennert and Negahdaripour's varying illumination model and our modified LMS method that is based on the standard LMS.

The experimental results reported above, show the superiority of our integrated method to other algorithms based on both global and local methods when illumination changes significantly from one image to the next. We also presented the advantage of our local approach by comparing the accuracy of our algorithm with that obtained using global approaches. We finally showed the improvement obtained with our local approach when employing the modified LMS algorithm defined in this paper. We have shown this using statistical comparisons on synthetic image sequences for which the ground truth was known. A qualitative comparison with two real image sequences was also shown.

6   Acknowledgements

References

[1] A. Bab-Hadiashar and D. Suter, Optic flow calculation using robust statistics, In Proceedings of Computer Vision and Pattern Recognition, pp. 988-993, 1997.

[2] J. L. Barron, D. J. Fleet and S. S. Beauchemin, Performance of optical flow techniques, International Journal of Computer Vision, pp. 34-77, 1994.

[3] M. J. Black and P. Anandan, A framework for the robust estimation of optical flow, In Proceedings of International Conference on Computer Vision, pp. 231-236, 1993.

[4] M. A. Gennert and S. Negahdaripour, Relaxing the brightness constancy assumption in computing optical flow, MIT AI Lab. Memo No.975, 1987.

[5] B. K. P. Horn and B. G. Schunck, Determining optical flow, Aritifical Intelligence 17, pp. 185-203, 1981.

[6] Y.-H. Kim, A. M. Martinez and A. C. Kak, Robust motion estimation under varying illumination, Image and Vision Computing, accepted.

[7] S.-H. Lai and M. Fang, Robust and efficient image alignment with spatially varying illumination models, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2:167-172, 1999.

[8] B. D. Lucas and T. Kanade, An iterative image registration technique with an application to stereovision, Int. Joint Conf. on Artificial Intelligentc, pp. 674-679, 1981.

[9] S. Negahdaripour and C. H. Yu, A generalized brightness change model for computing optical flow, In Proceedings of International Conference on Computer Vision, pp. 2-11, 1993.

[10] P. Nesi, A. Del Bimbo, and D. Ben-Tzvi, A robust algorithm for optical flow estimation, Computer Vision and Image Understanding, 62(1): 59-68, 1995.

[11] E. Ong and M. Spann, Robust optical flow computation based on least-median-of-squares regression, International Journal of Computer Vision, 31(1):51-82, 1999.

[12] M. Otte and H. H. Nagel, Optical flow estimation: advances and comparisons, In Proceedings of European Confernce on Computer Vision, pp. 51-60, 1994.

[13] P. J. Rousseeuw and A. M. Leroy, Robust Regression and Outlier Detection, Wiley, New York, 1987.

[14] D. Shulman and J.-Y. Herve, Regularization of discontinuous flow fields, In Proceedings of IEEE Workshop on Visual Motion, 1989.

[15] Charles V. Stewart, Robust parameter estimation in computer vision, SIAM Reviews, 41(3):513-537, 1999.

[16] P. Treves and J. Konrad, Motion estimation and compensation under varying illumination, In Proceedings of IEEE International Conference on Image Processing, 1:373-377, 1994.