Optical Flow in Onboard Applications: A Study on the Relationship Between Accuracy and Scene Texture

Naveen Onkarappa, Sujay M. Veerabhadrappa and Angel D. Sappa

Abstract Optical flow has got a major role in making advanced driver assistance systems (ADAS) a reality. ADAS applications are expected to perform efficiently in all kinds of environments, those are highly probable, that one can drive the vehicle in different kinds of roads, times and seasons. In this work, we study the relationship of optical flow with different roads, that is by analyzing optical flow accuracy on different road textures. Texture measures such as *contrast*, *correlation* and *homogeneity* are evaluated for this purpose. Further, the relation of regularization weight to the flow accuracy in the presence of different textures is also analyzed. Additionally, we present a framework to generate synthetic sequences of different textures in ADAS scenarios with ground-truth optical flow.

Keywords Optical flow accuracy · Texture metrics · Ground-truth optical flow

N. Onkarappa (✉) · A. D. Sappa
Computer Vision Center, Edifici O, Campus UAB 08193 Bellaterra, Barcelona, Spain
e-mail: naveen@cvc.uab.es
URL: www.cvc.uab.es

A. D. Sappa
e-mail: asappa@cvc.uab.es

S. M. Veerabhadrappa
Department of Electrical and Electronics, PES Institute of Technology and Management, Shivamogga 577204, India
e-mail: sujay.veerabhadrappa@gmail.com

Mohan S. and S. S. Kumar (eds.), *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012)*, Lecture Notes in Electrical Engineering 221, DOI: 10.1007/978-81-322-0997-3_24,

1 Introduction

Computer vision has got many applications in human life safety, assistance and comfort. Among them, the safety of automotives and people has got relevant importance in the present world. With the advance in computing performance, computer vision plays a major role in making these assistance and safety applications a reality. The information about the visual motion is very important for ADAS applications such as egomotion, moving object detection, autonomous navigation etc. (e.g., [1, 2]). A well known visual motion estimation instrument is optical flow. Optical flow is the apparent displacement vector field between two image frames. The seminal work on optical flow dates back to 1981. Horn and Schunck proposed a variational formulation to estimate dense optical flow in [3]. At the same time, Lucas and Kanade [4] proposed another approach that computes sparse flow field. In general, optical flow methods can be classified as global and local. Global methods produce dense, whereas local methods give sparse flow fields. A huge amount of work has been proposed during the last three decades; interesting surveys can be found in [5] and [6]. An empirical evaluation of optical flow methods on complex image sequences is presented in [5]. In [7], Galvin et al. evaluate eight different optical flow algorithms. Later evaluation of optical flow algorithms with benchmarking suite of image sequences and tools are proposed by McCane et al. in [8].

The research on optical flow has been getting lots of interests in recent years.[1] Most of the approaches are variational methods [3, 9, 10], which produce dense flow fields. These works concentrate on robust edge-preserving regularization (e.g., [11–13]), and sophisticated data terms [14]. The developments in respective parts of optical flow estimation are discussed in detail in [6]. Recently, in [15], the concepts such as pre-processing, coarse-to-fine warping, graduated non-convexity, interpolation, derivatives, robustness of penalty functions, median filtering are explored and the best formulation out of variants of all these is discussed.

A typical variational method involves a data term and a regularization term that makes the problem well posed. As per our knowledge almost all methods fix the regularization weight empirically. Even though there is a large amount of work on optical flow, there are no considerable efforts to adapt the regularization weight based on some features of the given sequence. In [13], an automatic selection approach based on optimal prediction principle is presented. It predicts the regularization weight based on the computed flow field and data constancy error. But it also involves a brute-force method of selecting the weight empirically. In ADAS domain, the vehicle can be driven in any kind of environment such as urban, highway and countryside, and at different times in a day and in different seasons. Hence, it is impossible to compute accurate optical flow in all the scenarios with a fixed regularization weight. Note that for an ADAS involving visual motion perception, it is very important to estimate accurate optical flow in all such scenarios.

[1] http://vision.middlebury.edu/flow/

In that direction, it is clear that it is necessary to adapt the motion perception algorithms for the occurring scenario directly using some features of the image frames being captured at times. Different scenarios can be regarded as different textures and different structural scenes. In the current work, we study the effect of different textures on optical flow in ADAS domain. Specifically, the change in accuracy of flow field with different regularization weights in the presence of different textures is analyzed; trying to find correlations between accuracy, texture information, and regularization weights.

For the work planned above, it is required to have datasets of image sequences with ground-truth optical flow. Few sequences with ground-truth optical flow have been proposed in [16, 17]. But none of them are suitable for our intended study. It is difficult to obtain ground-truth optical flow for real-world scenarios unless it is done in controlled environments. Particularly for the work in this paper, it is not possible to have or create a real scenario, neither possible to generate dense ground-truth flow field. Therefore, we present a framework to generate synthetic image sequences where the same geometrical scene is used but with different textures; furthermore this framework allows to compute the ground-truths.

The paper is organized as follows. An overview of the basic variational optical flow estimation is presented in Sect. 2. Section 3 describes the texture measures. The framework for generating the sequences and ground-truth is described in Sect. 4. Section 5 presents the experimental analysis. Finally, the work is concluded in Sect. 6.

2 Optical Flow Overview

The classical variational method of Horn and Schunck [3] assumes the constancy of brightness, which is also called *optical flow constraint* (OFC). The OFC can be formulated as: $I_1(\boldsymbol{x}+\boldsymbol{u}) - I_0(\boldsymbol{x}) = 0$, where I_0 and I_1 are two images, $\boldsymbol{x} = (x_1, x_2)$ is the pixel location within the image space $\Omega \subseteq \mathbf{R^2}$; $\boldsymbol{u} = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}))$ is the two-dimensional flow vector. Linearizing the above equation using first-order Taylor expansion we get OFC as: $(I_{x_1}u_1 + I_{x_2}u_2 + I_t)^2 = 0$, where subscripts denote the partial derivatives. Using only local intensity constraints do not provide enough information to infer meaningful flow fields, make the problem ill-posed. In particular, optical flow computation suffers from two problems: first, no information is available in un-textured regions. Second, one can only compute the normal flow, i.e., the motion perpendicular to the edges. This problem is generally known as the *aperture problem*. In order to solve this problem it is clear that some kind of regularization is needed. The Horn and Schunk [3] method overcomes this by assuming the resulting flow field globally smooth all over the image, that can be realized as penalizing large flow gradients. Combining OFC and regularization in a single variational framework and squaring both constraints yields the following energy function:

$$E(\mathbf{u}) = \int_{\Omega} \{ \underbrace{(I_{x_1}u_1 + I_{x_2}u_2 + I_t)^2}_{Data\ Term} + \alpha \underbrace{(|\nabla u_1|^2 + |\nabla u_2|^2)}_{Regularization} \} \, d\mathbf{x}, \tag{1}$$

where α is the regularization weight. This energy function can be minimized by solving the corresponding Euler-Lagrange equations. Another well known approach to minimize variational energies is by dual formulation [18].

3 Texture Measures

It is necessary to quantify texture properties to study the influence of different textures on the accuracy of optical flow. These texture properties are used to correlate with optical flow accuracy. In this direction, we use basic texture measures such as *contrast*, *correlation* and *homogeneity*. These textural measurements are computed over Gray-Level Co-occurrence Matrix (GLCM) [19] of an image. A GLCM is a matrix that is defined over an image as the distribution of co-occurring values at a given offset. A co-occurrence matrix P is defined over an image I of size $m_1 \times m_2$, parameterized by an offset $(\Delta x, \Delta y)$ as:

$$P_{\Delta x,\Delta y}(i,j) = \sum_{k=1}^{m_1} \sum_{l=1}^{m_2} \begin{cases} 1, & \text{if } I(k,l) = i \text{ and } I(k+\Delta x, l+\Delta y) = j. \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The most widely used, computed on the texture metrics on the normalized GLCM, are the following:

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \}; \ |i-j| = n \tag{3}$$

$$Correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{4}$$

$$Homogeneity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1+|i-j|} \tag{5}$$

where $p(i,j)$ is the (i,j)th entry in normalized GLCM; N_g is the number of distinct gray levels in the quantized image; μ_x, μ_y, σ_x, and σ_y are the means and standard deviations of p_x and p_y: $p_x(i) = \sum_{j=1}^{N_g} p(i,j)$ and $p_y(j) = \sum_{i=1}^{N_g} p(i,j)$.

4 Framework for Synthetic Dataset

The specific goal of this work is to study the effect of different road textures on the accuracy of optical flow computation in the ADAS scenario and to unveil their relationship. For that, it is required to have images with the same structural scene but with different textures. It is impossible to find such a scenario in the real world. Moreover, it is not possible to obtain ground-truth optical flow from such real-world scenarios unless it is done in sophisticated controlled laboratory environments. So, in the current work, we present a framework to generate synthetic image sequences. The framework can generate image sequences with exactly similar scene structures and with exactly the same acquisition condition (i.e., vehicle speed, camera pose, et.,), but with different textures. There are some works to generate synthetic datasets with ground-truth data in the literature (e.g., [16, 17, 20] and [21]). In [16], authors provide ground-truth information for several synthetic as well as real sequences, whereas in [17], synthetic sequences for ADAS are provided. Later Aodha et al. present a framework for synthetic dataset generation in [20]. Following the framework proposed in [20] a set of synthetic sequences for different speeds of the vehicle are presented in [21]. The datasets provided in these previous works can be used to compare different optical flow techniques, but do not provide sequences with different textures. For the intended study, we generate our own dataset and make it available for the whole community for further research through our website.[2] We use a framework similar to [20] and construct a 3D urban model, which contains road, sky and several buildings using Maya.[3] A camera is made to move in the model along the road, mimicking the camera fit inside a moving vehicle. The image frames are rendered for the camera movement timeline. Similarly, several different image sequences are generated by changing road textures in the model. Also, the ground-truth flow vectors are generated for the camera movement using ray-tracing technique on 3D Maya model. In ADAS, the road surface covers major area in the visibility of the vehicle camera and the structures in the sides of the road vary a lot depending on the environment (urban, highway, hilly etc.). So the flow vectors on the road surface are generally preferred in ADAS applications since they are more reliable and

[2] www.cvc.uab.es/adas

[3] www.autodesk.com/maya

cover a large part of the image. Hence, using the framework described in this section we generate several sequences of the same structural scene with different road textures. The generated images and ground-truth flow fields are of resolution 640×480. Figure 1 shows images from sequences of different textures, but having the same structural scene. The image in Fig. 1 (*bottom-left*) shows the ground-truth flow field. The ground-truth flow field is the same for all the sequences shown there, as the ray-tracing technique used gives the same displacement information based on the structural geometry of the scene irrespective of the texture.

5 Experimental Analysis

In order to make the analysis and conclusion precise and easier, we have considered three sequences of different textures. The textures are different at least on the road surface, but the structures of the scene are exactly the same in all the three sequences. Let us refer to these sequences as T1, T2 and T3. These sequences are selected in the increasing order of the textural property, *contrast*. For further experimentation, we have selected several pairs of images from the sequences at specific locations and present the results as average values of all these selected pairs. Hereinafter, the results obtained from these sets of image pairs are referred to as the results obtained from the corresponding sequences.

Now the task is to find the relation between the accuracy of optical flow, textural property of images and the regularization weight involved in optical flow estimation. For computing the optical flow, we use the method proposed in [15], that involves weighted median filtering. First, optical flow is computed on the sets of image pairs from different sequences. Then, errors (both average angular error (AAE) and average end-point error (AEPE)) are computed between the estimated flow fields and the ground-truth flow fields. The ground-truth flow fields for all the three sequences are the same.

To study the effect of regularization weight on the accuracy, first we empirically analyzed optical flow error on different pairs of images for a wide range of values of regularization weights and determined that the range from 1 to 22 at intervals of 1 or 2 would be fine for further analysis. Then, we estimated optical flow for values of regularization weight within the most suitable range of values (i.e., 1–22). Figure 2 (*left*) shows the curves of AAEs corresponding to sequences T1, T2 and T3 for different values of regularization weight. The minimum AAE for each sequence is marked as $*$ on the curve. Similarly, AEPE error curves are shown in Fig. 2 (*right*).

Further, textural measures explained in Sect. 3 are computed on one of the images of each sequence. Since road is the major and reliable surface in ADAS scenarios, we compute texture metrics on a Region Of Interest (ROI) on the road

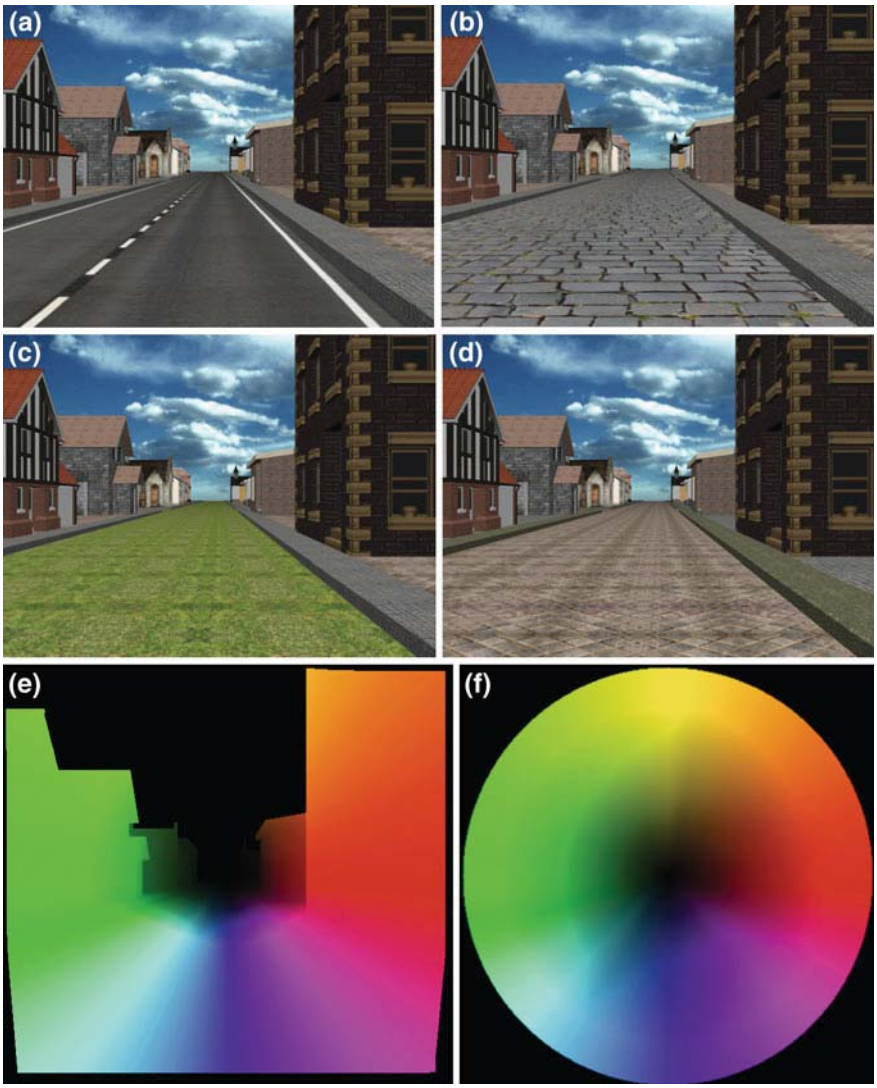

Fig. 1 *Top-left* consecutive frames of a sequence; *top-right* and *middle* frames of different texture; *bottom-left* ground-truth flow field from images in *top-left*; *bottom-right* colormap used to show the flow field

surface. The texture metrics, minimum AAEs, minimum AEPEs, and the regularization weights corresponding to the minimum errors are depicted in Table 1.

The curves in Fig. 2 and texture metrics in Table 1 indicate that image sequences with higher texture *contrast* produce smaller AAE and AEPE,

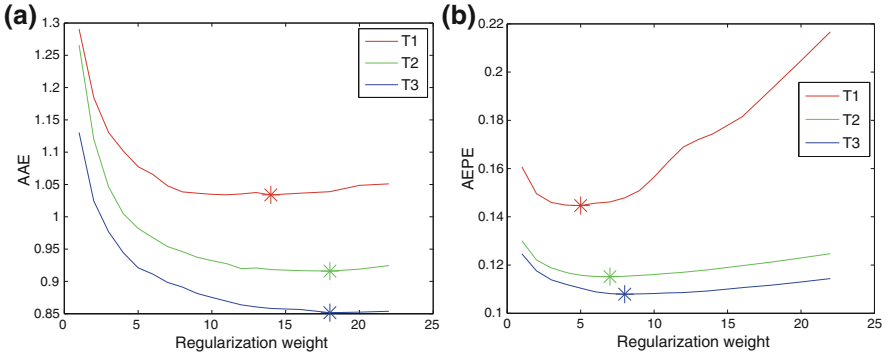

Fig. 2 *Left* AAEs for three sequences for different regularization weights; *right* similarly AEPEs

Table 1 Texture metrics, minimum AAEs, minimum AEPEs, regularization wights for the minimum errors for T1, T2 and T3

Sequence	T1	T2	T3
min AAE	1.0339	0.9159	0.8515
α of min AAE	14	18	18
Contrast	0.0488	0.0850	0.1255
Correlation	0.9386	0.8611	0.7168
Homogeneity	0.9765	0.9575	0.9372
min AEPE	0.1447	0.1152	0.1079
α of min AEPE	5	7	8

independently of the regularization factor α. We can see in Fig. 2, both (*left*) and (*right*), that curve of T1 is above T2, and T2 is above T3. In Table 1, T3 has higher *contrast* than T2, and T2 has higher *contrast* than T1. Similarly, the sequence with lower values of *correlation* or *homogeneity* produce smaller errors. One can see in Fig. 2 (*right*), the AEPE of T1 increases drastically with the increase in α. So one should be careful in tuning the α parameter when dealing with sequences of low texture *contrast*. Another interesting conclusion that can be drawn from this study is that, for higher texture *contrast* sequences, the minimum error is obtained by increasing the regularization factor α. This can be observed in Fig. 2 marked by $*$ and also in the Table 1, the α values of minimum errors. Although in the plots only three curves have been depicted, this conclusion has been extracted from a larger set up using nine sequences of different textures. Figure 3 presents four additional ROIs with the textures from the road surface used to validate the results from our studies. In these four cases, similar behavior like the ones presented in Fig. 2 and Table 1 were obtained.

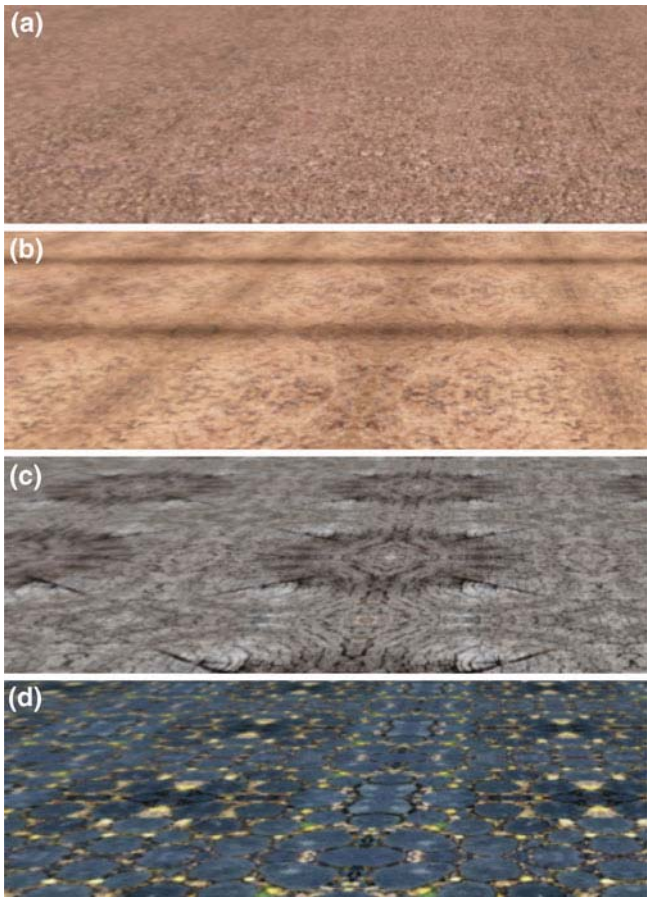

Fig. 3 ROIs of road surface from different sequences with different textures

6 Conclusions

In this work, the behavior of optical flow accuracy for different textural properties and their dependency on the regularization factor in ADAS scenarios is studied. It is evident that the sequence with higher textural *contrast* gives more accurate flow estimation and a more special care should be taken while tuning the regularization factor for sequences with low contrast more than those with higher contrast. It is

also observed that, increasing the regularization factor gives better results with increase in textural *contrast* of sequences. A framework to generate synthetic sequences of different textures with the ground-truth optical flow is also presented along with the dataset generated. The presented framework motivates deeper study of the relationship between the optical flow accuracy and scene texture using more sophisticated texture metrics.

Acknowledgments This work has been partially supported by the Spanish Government under Research Program Consolider Ingenio 2010: MIPRCV (CSD2007-00018) and Project TIN2011-25606. Naveen Onkarappa is supported by FI grant of AGAUR, Catalan Government. The authors would like to thank Oisin Mac Aodha for providing the Python code for raytracing with Maya.

References

1. Giachetti A, Campani M, Torre V (February 1998) The use of optical flow for road navigation. Robot Autom IEEE Trans 14(1):34–48
2. Onkarappa N, Sappa AD (2010) On-board monocular vision system pose estimation through a dense optical flow. In: 7th international conference on image analysis and recognition, vol 1, pp 230–239
3. Horn BKP, Schunk BG (1981) Determining optical flow. Artif Intell 17:185–203
4. Lucas BD, Kanade T (1981) An iterative image registration technique with an application to stereo vision (DARPA). In: DARPA image understanding workshop, pp 121–130
5. Barron JL, Fleet DJ, Beauchemin SS (1994) Performance of optical flow techniques. Intern J Comput Vis 12(1):43–77
6. Bruhn A (2006) Variational optic flow computation: accurate modelling and efficient numerics. PhD thesis, Department of Mathematics and Computer Science, Saarland University, Saarbrücken
7. Galvin B, Mccane B, Novins K, Mason D, Mills S (1998) Recovering motion fields: an evaluation of eight optical flow algorithms. In: British machine vision conference, pp 195–204
8. McCane B, Novins K, Crannitch D, Galvin B (October 2001) On benchmarking optical flow. Comput Vis Image Underst 84(1):126–143
9. Brox T, Bruhn A, Papenberg N, Weickert J (2004) High accuracy optical flow estimation based on a theory for warping. In: European conference on computer vision, vol 3024 of LNCS. Springer, pp 25–36
10. Wedel A, Pock T, Zach C, Cremers D, Bischof H (2008) An improved algorithm for TV-L1 optical flow. In: Dagstuhl motion workshop, Dagstuhl Castle, pp 23–45
11. Weickert J, Schnörr C (December 2001) A theoretical framework for convex regularizers in pde-based computation of image motion. Intern J Comput Vis 45(3):245–264
12. Wedel A, Cremers D, Pock T, Bischof H (2009) Structure- and motion-adaptive regularization for high accuracy optic flow. In: International IEEE conference of computer vision, Kyoto, pp 1663–1668
13. Zimmer H, Bruhn A, Weickert J (2011) Optic flow in harmony. Intern J Comput Vis 93(3):368–388
14. Steinbruecker F, Pock T, Cremers D (2009) Advanced data terms for variational optic flow estimation. In: Modeling vision and visualization workshop, Braunschweig, pp 155–164
15. Sun D, Roth S, Black MJ (June 2010) Secrets of optical flow estimation and their principles. In: Conference IEEE on computer vision and pattern recognition, San Francisco, pp 2432–2439

16. Baker S, Scharstein D, Lewis JP, Roth S, Black MJ, Szeliski R (2007) A database and evaluation methodology for optical flow. In: IEEE international conference of computer vision, pp 1–8
17. Vaudrey T, Rabe C, Klette R, Milburn J (2008) Differences between stereo and motion behaviour on synthetic and real-world stereo sequences. In: Image and vision computing New Zealand, Christchurch, pp 1–6
18. Zach C, Pock T, Bischof H (September 2007) A duality based approach for realtime TV-L^1 optical flow. In: Symposium DAGM on pattern recognition, Heidelberg, pp 214–223
19. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621
20. Mac Aodha O, Brostow GJ, Pollefeys M (2010) Segmenting video into classes of algorithm-suitability. In: IEEE conference on computer vision and pattern recognition, San Francisco, pp 1054–1061
21. Onkarappa N, Sappa A (2012) An empirical study on optical flow accuracy depending on vehicle speed. In: IEEE intelligent vehicles symposium, pp 1138–1143