Speed and Texture: An Empirical Study on Optical-Flow Accuracy in ADAS Scenarios

Naveen Onkarappa and Angel Domingo Sappa, *Senior Member, IEEE*

Abstract—Increasing mobility in everyday life has led to the concern for the safety of automotives and human life. Computer vision has become a valuable tool for developing driver assistance applications that target such a concern. Many such vision-based assisting systems rely on motion estimation, where optical flow has shown its potential. A variational formulation of optical flow that achieves a dense flow field involves a data term and regularization terms. Depending on the image sequence, the regularization has to appropriately be weighted for better accuracy of the flow field. Because a vehicle can be driven in different kinds of environments, roads, and speeds, optical-flow estimation has to be accurately computed in all such scenarios. In this paper, we first present the polar representation of optical flow, which is quite suitable for driving scenarios due to the possibility that it offers to independently update regularization factors in different directional components. Then, we study the influence of vehicle speed and scene texture on optical-flow accuracy. Furthermore, we analyze the relationships of these specific characteristics on a driving scenario (vehicle speed and road texture) with the regularization weights in optical flow for better accuracy. As required by the work in this paper, we have generated several synthetic sequences along with ground-truth flow fields.

Index Terms—Advanced driver assistance systems (ADASs), optical flow, regularization parameters, road texture, vehicle speed.

I. INTRODUCTION

THE developments in computer vision and computing systems have drawn the interest of the automotive industry to make use of them toward advanced driver assistance systems (ADASs). ADASs include lane departure warning, collision avoidance, parking assistance, and autonomous navigation (e.g., see [1] and [2]). These systems involve tasks such as egomotion estimation, moving-object detection, and 3-D reconstruction. One of the well-known tools for estimating motion that can be used in many of the aforementioned tasks is the optical flow. Optical flow is a displacement vector field of patterns between two images. In a driving scenario, optical flow is estimated between successive video frames captured by

a camera that is mounted on a vehicle. The seminal methods of estimating optical flow were proposed in 1981 in [3] and [4]. The literature shows that there have been several attempts to improve optical-flow accuracy, with increased interest in recent years, particularly on variational approaches that typically involve data and regularization terms. The balancing between the regularization and data terms has to be tuned to get better flow fields. Almost all the state-of-the-art approaches empirically select this weight for a fixed set of image sets used for evaluation.

In the ADAS domain, it can happen that the vehicle is driven in different environments (e.g., urban, highway, and countryside) [5] with different speeds and different road textures, making it difficult to achieve the same optical-flow accuracy all over the vehicle's trajectory; in turn, it reduces the confidence and effectiveness of ADAS applications. It is very important to adjust the regularization weight based on the environment where the vehicle is being driven. This motivates us in this paper to analyze the effect of some specific properties of the driving environment on the optical-flow accuracy. There are many factors that affect the flow accuracy, such as illumination, occlusions, specularity, texture, structure, and large displacements. In particular, in this paper, we study the influence of onboard vision system speed and also the road texture on optical-flow accuracy.

As motivated by the natural way of representing a vector in terms of polar coordinates, it is also demonstrated that this representation exhibits statistical independence on image sequences of ADAS scenarios in this paper. The polar-represented optical-flow estimation [6] involves the following two regularization terms: 1) orientation and 2) magnitude. This formulation gives the advantage of independently tuning each term, unlike in Cartesian-represented optical-flow estimation. Fig. 1 shows image frames of different speeds and textures and the estimated flow fields based on [6]. The error values [average angular error (AAE) and average endpoint error (EPE)] for the same flow fields are given in Table I, where S1 corresponds to the sequence with the lowest speed, whereas S4 corresponds to the sequence with the highest speed. On the other hand, T1 corresponds to the lowest texture contrast, and T3 corresponds to the highest texture contrast. An analysis of errors for a fixed set of regularization weights and different speeds and textures in both Fig. 1 and Table I reveals the importance of regularization weights for an accurate flow-field estimation. In this paper, we analyze the variation in accuracy of the optical flow by varying the weights of regularization on several sequences of different speeds and road textures. First, the analysis of the influence of just speed is performed. Second, different textures

S1, T1, $frame1$ S1, T1, $frame2$

S1, T3, $frame1$ S1, T3, $frame2$

S1, T1, $(\alpha_\theta, \alpha_m) = (10, 10)$ S1, T1, $(\alpha_\theta, \alpha_m) = (120, 120)$

S1, T3, $(\alpha_\theta, \alpha_m) = (10, 10)$ S1, T3, $(\alpha_\theta, \alpha_m) = (120, 120)$

S4, T3, $(\alpha_\theta, \alpha_m) = (10, 10)$ S4, T3, $(\alpha_\theta, \alpha_m) = (120, 120)$
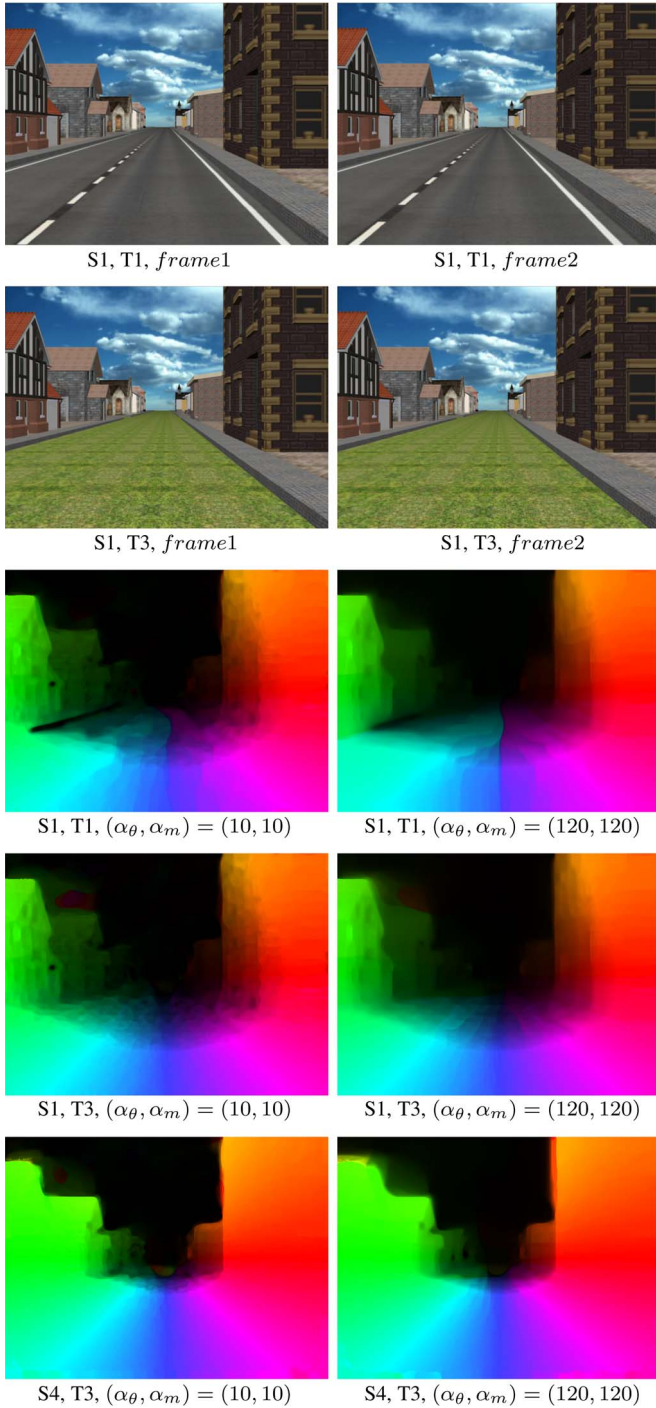
Fig. 1. Image frames of different textures and speeds, and computed optical flows for different regularization weights.

TABLE I
AAEs AND EPEs FOR FIXED REGULARIZATION WEIGHTS FOR
SEQUENCES OF DIFFERENT TEXTURES AND SPEEDS
(FLOW FIELDS ARE SHOWN IN FIG. 1)

Sequence	(AAE, EPE) for $(\alpha_\theta, \alpha_m) = (10, 10)$	(AAE, EPE) for $(\alpha_\theta, \alpha_m) = (120, 120)$
S1,T1	(2.3974, 0.1006)	(3.4395, 0.1537)
S1,T3	(1.4916, 0.0657)	(1.5065, 0.0628)
S4,T3	(1.0999, 0.2210)	(1.1156, 0.1736)

urban scenario for the required cases are rendered using 3-D models that were generated with the graphic editor Maya[1]; the corresponding ground-truth flow fields are also generated using a ray-tracing technique.

In summary, the contributions of this paper are listed as follows.

1) The statistical independence of polar representation is exploited on ADAS scenarios.
2) The dependency of regularization weights (both for magnitude and orientation) are analyzed for different speeds of the onboard vehicle camera, for different road textures, and for different combinations of both speed and texture together.
3) Several synthetic sequences of driving scenarios for different speeds and different road textures are generated with the corresponding ground-truth flow fields.

This paper is organized as follows. The next section presents the related work. Then, Section III provides a brief comparative study of the use of polar representation with respect to the Cartesian in the context of ADAS applications. Next, Section IV presents the polar optical-flow formulation used in this paper. The texture measures needed to evaluate the different scenarios are presented in Section V, whereas the framework used to generate scenes is detailed in Section VI. Experimental results, discussions, and conclusions are given in Sections VII–IX, respectively.

II. RELATED WORK

Optical-flow techniques can be classified as global and local approaches. Global approaches produce dense flow fields using variational energy minimization, whereas local approaches produce sparse flow fields using a least squares criterion over small neighborhoods. During the last three decades, several approaches on optical-flow estimation have been proposed [7] to improve accuracy and efficiency. Hence, there is a need to evaluate this large amount of contributions. Performance evaluation of different methods on complex data sets have been presented in [8]–[11]. Recently, Baker *et al.* [12] have proposed benchmarking sequences with ground-truth data and a methodology for evaluation.

The denseness of global optical-flow approaches make them useful in many applications. Typically, global optical-flow estimation (e.g., see [13] and [14]) that is formulated as a variational energy minimization consists of a data term that matches some properties between images and a smoothness term, also called a regularization term, that makes the problem

are analyzed. Finally, the analysis that combines both speed and textural properties is done.

This empirical analysis requires the following image sequences: 1) to analyze the influence of speed, having sequences of different speeds with the same geometrical structure and texture is needed and 2) to analyze the influence of texture, having sequences with the same geometrical scene structure but with a different texture is needed. It is impossible to have such real-life scenarios and also the corresponding ground-truth optical flow. In this paper, several synthetic sequences of an

[1]www.autodesk.com/maya

well posed. Attempts have been made to improve on data terms, regularization, and energy minimization. Improvements in data terms are made by robust penalizing functions [15] and by adopting higher order terms [16]. Developments in regularization terms are focused on preserving motion discontinuities [17] and using the temporal coherence [18]. A method that combines the advantages of both local and global approaches is proposed in [16]. Recently, Sun *et al.* [19] have explored concepts such as preprocessing, coarse-to-fine warping, graduated nonconvexity, interpolation, derivatives, robustness of penalty functions, and median filtering, and then, their influence on optical-flow accuracy is revealed. Using the best of the explored concepts and weighted nonlocal median filtering, an improved model is proposed in [19]. A multiframe optical-flow estimation technique based on the temporal coherence of flow vectors across image frames is proposed in [20]. Motivated by the natural representation of a vector and the statistical independence of the polar coordinates to the Cartesian coordinates, recently, an optical-flow estimation approach based on polar representation has been proposed in [6]. A top-performing method that intelligently preserves small and large motion in a coarse-to-fine approach is proposed in [21]. In summary, there has been increased interest on optical-flow approaches in the last few years, which can be appreciated on the number of publications and released code [11].

We can notice that, in almost all optical-flow methods, the weights for regularization are empirically chosen. There are very few attempts in this direction to automatically select such parameters. Krajsek *et al.* [22] present a Bayesian model that automatically weighs different data terms and a regularization term. This model that estimates optical flow and several parameters together is very complex to minimize. Recently, Zimmer *et al.* [23] has proposed to automatically select the regularization weight based on the *optimal prediction principle*. In their work, the optimal regularization weight is obtained as the one that can produce a flow field with which the next frame in a sequence is best predicted. Inherently, this approach involves a brute-force method to select the optimal weight based on the average data constancy error, and hence, it is computationally expensive. On the other hand, there is an attempt [24] to use several different optical-flow methods for a sequence by selecting the best suitable method per pair of frames or per pixel.

In a preliminary work [25], we present a study on optical-flow accuracy for different speeds of the vehicle. In that work, the size of the video sequence is very small, and the frames considered for speed analysis involve a different geometrical scene structure. In this paper, in particular, the study on speed is improved by adding longer sequences. In addition, both the study of texture and the combined study on speed and texture have been performed on sequences that include complex scenarios.

III. Polar Versus Cartesian Representation of Flow Vectors

The most commonly used representation in optical-flow estimation is the Cartesian coordinate system. However, represent-
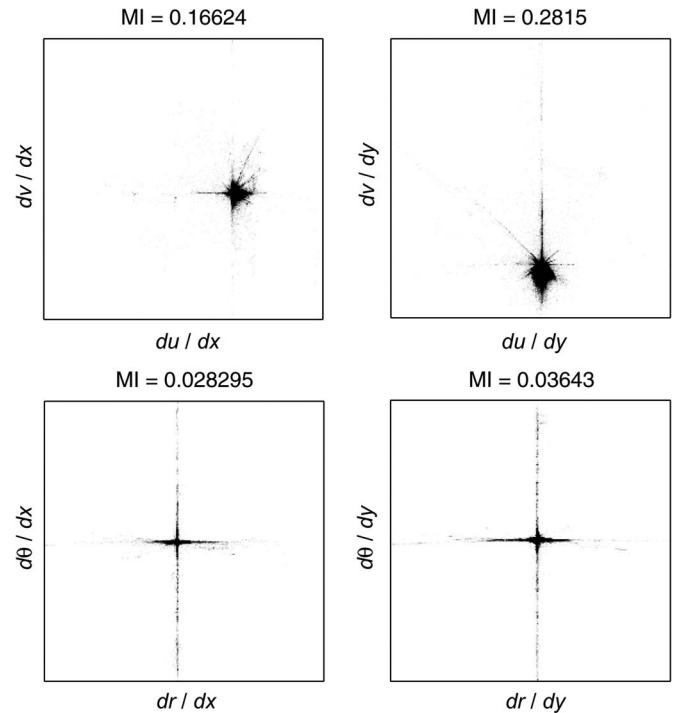

Fig. 2. Joint histograms of flow derivatives in the Cartesian and polar coordinates of an estimated flow field in a synthetic sequence of an urban road scenario. On top of each plot, the MI value is depicted.

ing a vector in terms of its magnitude and orientation is a natural way that is referred to as polar representation. As presented in [6], the analysis of spatial derivatives distribution of a flow field represented in polar shows a significant statistical difference among its components compared to the components of a Cartesian representation. Furthermore, the polar components show higher statistical independence compared to the Cartesian components when the mutual information (MI) between the derivatives of flow components in the respective representations are analyzed, as shown in [6] and [26].

A similar analysis is shown in Fig. 2. This analysis is performed on the estimated optical-flow field from a pair of images in an urban driving scenario (shown in Fig. 3, left column). Fig. 2 shows the joint histograms of flow derivatives in both the Cartesian and polar coordinate systems. The MI between the coordinate components that were computed using these joint histograms are depicted on top of each plot in Fig. 2. The lower the values of MI, the higher the statistical independence. As shown in Fig. 2, the representation of flow field in polar is more independent than the Cartesian system. A similar analysis on the ground-truth flow field between the same pair of images has shown zero MI (in the cases in Fig. 2, *bottom left* and *bottom right*), for polar coordinates. For the Cartesian coordinates (in the cases in Fig. 2, *top left* and *top right*), the MI values are 0.27082 and 0.50335, respectively, when the ground-truth flow field is considered. This shows that, in the ideal case of translational motion, polar coordinates are mutually exclusive (totally independent).

A polar representation of flow vectors for optical-flow estimation is proposed in [6], and its implications are studied. It is shown that polar-represented optical flow performs almost
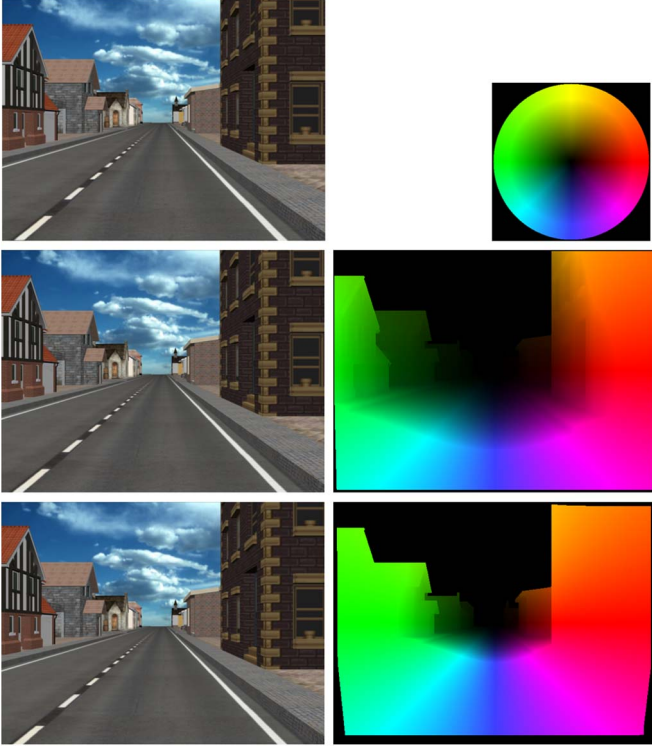
Fig. 3. Images from sequences of different speeds. *Top left*: First frame common for all sequences. *Top right*: Color map used to show the flow fields. *Left column*: Second frame from the sequences of different speeds in increasing order (second and third rows). *Right column*: Ground-truth flow fields between the respective first and second frames.

similar to the state-of-the-art Cartesian coordinates represented optical-flow estimation on traditional image data sets. Furthermore, it is shown that, for specular and fluid flow image sets, polar representation adds the advantage by independently allowing regularization in either coordinate component. In the vehicle-driving scenario, the majority of the motion is translation. The expected flow field in such a scenario is diverging, and the variation in magnitude is higher compared to the variation in orientation. In such a motion scenario, the polar optical flow becomes convenient. This paper exploits the possibility of independent tuning of regularization terms.

IV. Overview of the Polar Optical Flow

A typical variational formulation of the optical-flow energy function using the Cartesian representation looks like

$$E(u,v) = \int\int_{\Omega} \left\{ \underbrace{(I(x+u, y+v, t+1) - I(x,y,t))}_{Data\,Term} \right.$$

$$\left. + \alpha \underbrace{\left(|\nabla u_1|^2 + |\nabla u_2|^2\right)}_{Regularization} \right\} dx\,dy, \quad (1)$$

which contains a data and a regularization term. Here, $I(x,y,t)$ is the pixel intensity value at (x,y) at time t, α is the regularization weight, and (u,v) is the flow-field vector to be

estimated using Euler–Lagrange equations [14] or alternative methods [13].

This section presents a brief description of the polar optical-flow formulation proposed in [6]. According to that work, the flow vector at a pixel (x,y) can be represented in terms of polar coordinates as

$$flow(x,y) = (m(x,y), \theta(x,y)) \quad (2)$$

where m is the magnitude, and θ is the orientation at (x,y). The energy formulation using the polar representation allows us to separate the regularization terms as follows:

$$E(\theta(x,y), m(x,y))$$

$$= \int\int_{\Omega} \left\{ \psi\left(I(x+m\cos\theta, y+m\sin\theta, t+1) - I(x,y,t)\right) \right.$$

$$\left. + \alpha_\theta \psi_\theta\left(\rho_\theta(\theta)\right) + \alpha_m \psi_m\left(\rho_m(m)\right) \right\} dx\,dy, \quad (3)$$

where ψ is a robust penalty function for the data term, and ψ_θ and ψ_m are robust penalty functions, respectively, for the orientation and magnitude components' regularization (see [6] for more details). Similarly, α_θ and α_m are regularization weights, and ρ_θ and ρ_m are differential operators (in a simpler case, the first derivative). All these ρ_*, ψ_*, and α_* can be varied, depending on the image sequences or application of interest.

To avoid the difficulty of m being negative, the following equivalence relation is defined over values of m and θ:

$$(m,\theta) \sim \begin{cases} (m,\theta) & \text{if } m > 0 \\ (-m, \theta+\pi) & \text{if } m < 0. \end{cases} \quad (4)$$

Due to the periodic nature of θ, the orientation is expressed in terms of two parameters as

$$s(x,y) = \sin\theta(x,y)$$
$$c(x,y) = \cos\theta(x,y) \quad (5)$$

where the constraint $s^2 + c^2 = 1$ is called the *coherence constraint*, which ensures proper representation of orientation.

Using the Lagrange multiplier λ and assuming that it as a predetermined parameter, the energy function can be formulated to minimize three parameters (c, s, m) as

$$E(c,s,m)$$

$$= \int\int_{\Omega} \left\{ \lambda(s^2 + c^2 - 1)^2 \right.$$

$$+ \psi\left(I(x+mc, y+ms, t+1) - I(x,y,t)\right)$$

$$\left. + \alpha_\theta \psi_\theta\left(\rho_\theta(c), \rho_\theta(s)\right) + \alpha_m \psi_m\left(\rho_m(m)\right) \right\} dx\,dy, \quad (6)$$

where λ is a pixelwise predetermined parameter that is updated every iteration as $\lambda = e^{(s^2+c^2-1)^2}$ using the previous iteration values of c and s. Equation (6) can be minimized using Euler–Lagrange equations.

V. Texture Measures

To study the effect of texture on optical-flow accuracy, it is necessary to quantify the texture property. There are several ways of measuring the texture content of a given sequence [27]; in this paper, three of the most widely used statistical texture metrics, i.e., *contrast*, *correlation*, and *homogeneity*, are considered. These metric values are computed over a co-occurrence matrix of gray values of images [28] and are correlated with the optical-flow error measures. The texture metrics computed over the co-occurrence matrix, which is also called normalized gray-level co-occurrence matrix (GLCM) of an image, are defined as

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}; \quad |i-j| = n \tag{7}$$

$$Correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{8}$$

$$Homogeneity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i-j|} \tag{9}$$

where $p(i,j)$ is the (i,j)th entry in the normalized GLCM, N_g is the number of distinct gray levels in the quantized image, and 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 $p_y(j) = \sum_{i=1}^{N_g} p(i,j)$.

VI. Synthetic Sequence Generation Framework

To analyze the influence of speed on optical-flow accuracy, we need to have image sequences of the same scene, but with the onboard vision system moving with different speeds on exactly the same trajectory. Similarly, to analyze the impact of texture, we need image sequences of the same scene (i.e., surrounding scene structure) but with just different textures. In reality, it is impossible to have such scenarios and to generate ground-truth optical flow. Although it is possible to have such sequences in a controlled laboratory environment, there does not exist any sensor to generate ground-truth flow fields. The only way to have such scenarios is to build virtual 3-D models and use them in the aforementioned setup (speed and texture). In fact, there are several data sets available with ground-truth in the literature (e.g., [12], [29], and [30]). In [12], different synthetic and real data sets for general motion scenarios with the ground truth are presented, whereas in [29], many synthetic sequences with the ground truth for ADAS scenarios are presented. The work presented in [30] provides synthetic sequences of varied complexities that were created using the open movie Sintel as a benchmark in optical-flow research. Recently, [31] has proposed a data set (KITTI) of real sequences with ground-truth flow in ADAS scenarios. Although KITTI has real sequences of ADAS scenarios, it does not contain sequences for studying the influence of speed/texture. In summary, none of the available data sets are suitable for the study proposed in this paper. Mac Aodha *et al.* in [24] presented
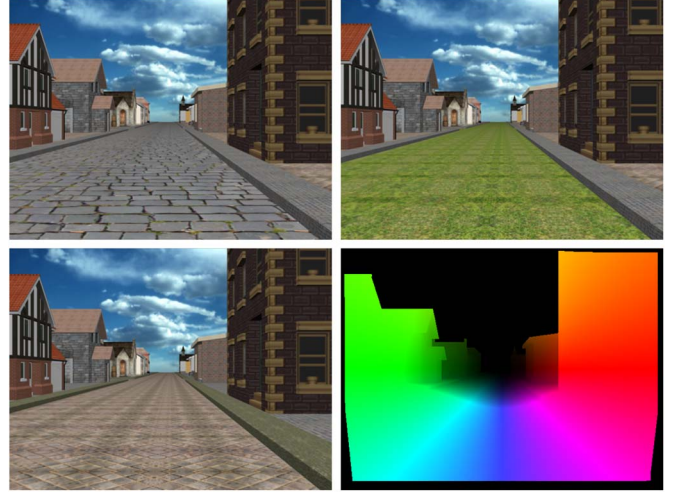

Fig. 4. *Top left, top right*, and *bottom left*: Frames with different texture from different sequences. *Bottom right*: Ground-truth flow field for all the pairs of images on the *top left, top right*, and *bottom left*. All of them have the same scene geometry and same speed, but with different textures.

a framework for generating synthetic data sets, which is used here as detailed in the following.

In this paper, we build a synthetic 3-D urban scenario that consists of a straight road and buildings around with appropriate texture; it is developed using Maya. A camera that is assumed to be fixed in a vehicle moves along the road in the model, and images are rendered for different speeds of the vehicle along the road. Fig. 3 shows generated synthetic frames and the corresponding ground-truth flow fields. On the top left is the first image, which is common to all the sequences of different speeds. The color map on the top right is used to represent the flow fields in this paper. On the second row to the left is the second image in the sequence, and on the second row to the right is the ground-truth flow between the images on the top left and the second row to the left. On the bottom left is the second image of another sequence of higher speed. On the bottom right is the ground-truth flow between images on the top left and bottom left. Ground-truth optical-flow values are computed using a ray-tracing technique directly over the 3-D synthetic models, considering the camera position in consecutive frames [24]. The maximum displacement in lower speed sequences is 8.31 pixels, and it is 33.67 pixels in higher speed sequences.

In ADAS scenarios, the road surface covers a major part in the images taken through a vehicle's camera. The flow vectors computed from this surface are more reliable, because there could be inaccuracies due to occlusions and specularities in the roadside structures. To analyze texture influence, in particular, we are changing only the texture of the road surface. Hence, for a given speed, several sequences with different road textures are rendered. For example, Fig. 4 shows the images with three different textures on the road surface. On the *bottom right* is the ground-truth flow field for all the three image pairs, i.e., *top left, top right*, and *bottom left*, with their corresponding next ones. Note that, in this case, the same ground truth is valid for all the sequences with different textures, because the scene geometry is exactly the same in all of them. All the flow fields in Figs. 3 and 4 are diverging, because the vehicle is moving straight on
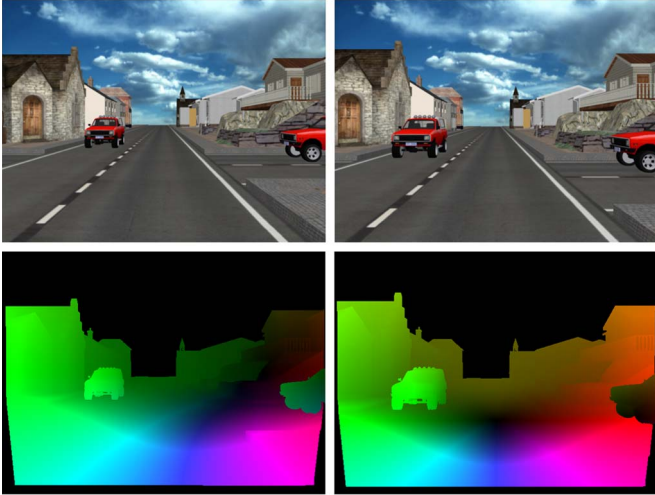
Fig. 5. *Top*: Two different image frames from a sequence with independently moving vehicles and different egomotion. *Bottom*: Ground-truth flow fields between the top frames and to their next ones in the sequence.

Fig. 6. RoIs used to calculate the error measures. *Left*: Speed analysis. *Right*: Texture, and speed together with texture analysis.

a road. In addition to these sequence sets, which are simple in motion, we have created another set of complex sequences for different speeds and textures similar to the previous set of sequences. The new complex sequences contain two moving vehicles: one vehicle moves along the road and comes toward the onboard camera vehicle, and the other vehicle comes from a cross road toward the onboard camera vehicle. The new sequences also contain changes in yaw and pitch angles during the vehicle's trajectory. The yaw is $0.25°$ to the left/right, and the pitch is $0.25°$ to the up/down. Two of the image frames from this new sequence and the ground-truth optical flows with their next image frames are shown in Fig. 5. All the rendered images are of a resolution of 480×640 pixels, and the camera focal length is 35 mm. All this data set (i.e., rendered frames for different video sequences and corresponding ground-truth flows) is available through our website.[2]

VII. EXPERIMENTAL ANALYSIS

This section presents the empirical study of the optical-flow accuracy of scenes, where the following conditions hold: 1) the camera moves at different speeds; 2) the texture of the scene changes; and 3) both speed and texture changes are considered together. First, we perform the study for all these three cases on a set of simple sequences, where there is no complexity, and the vehicle's camera moves straight on a road with different speeds and with different road textures. Such a simple sequence enables us to easily analyze the influence of speed and textures. Then, we also present the study of the influence of speed and texture together with another set of sequences that has complex egomotion.

A. Analysis for Speed

Following the framework presented in Section VI, we have generated four sequences of different speeds with an incremental translation of 0.25, 0.5, 0.75, and 1 cm along the optical
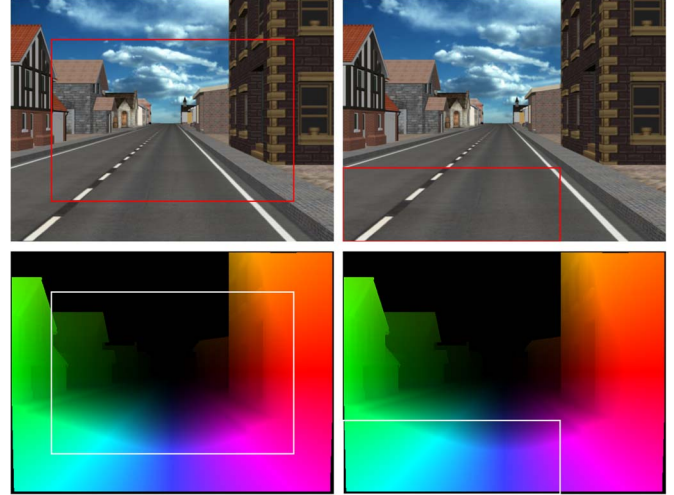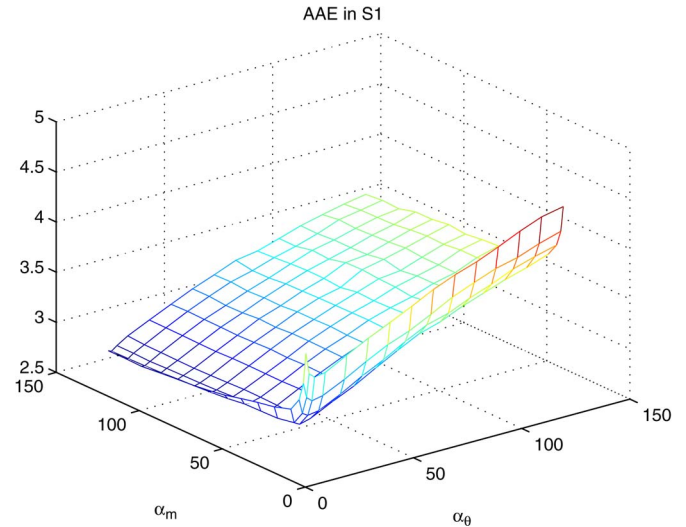
Fig. 7. Three-dimensional plot of the AAEs from S1 for varying α_θ and α_m values.

axis of the vehicle camera in a Maya model with a working unit as centimeter. Let us call these sequences S1, S2, S3, and S4 in increasing order of speed. The ground-truth optical flows for these sequences are also generated. The scene and texture of all these sequences are shown in Fig. 3. The first aim is to study optical-flow accuracy for the change in speed and to find its relationship with respect to the regularization parameters in the optical-flow formulation. We use the polar optical flow presented in Section IV, because this formulation provides the possibility of separately tuning different regularization parameters, which is an attractive feature in the ADAS domain. Furthermore, it involves two regularization terms that allow an independent study of their influence. Initially, experimentation is performed to find the optimal range for the regularization weights. It consists of computing optical flow on a pair of images from one of the sequences for a wide range of weights of both regularization terms. Based on this experiment, it is determined that the following range of values for experimentation is sufficient: 1, 2.5, 5, 10, 20, 30, ..., 120.
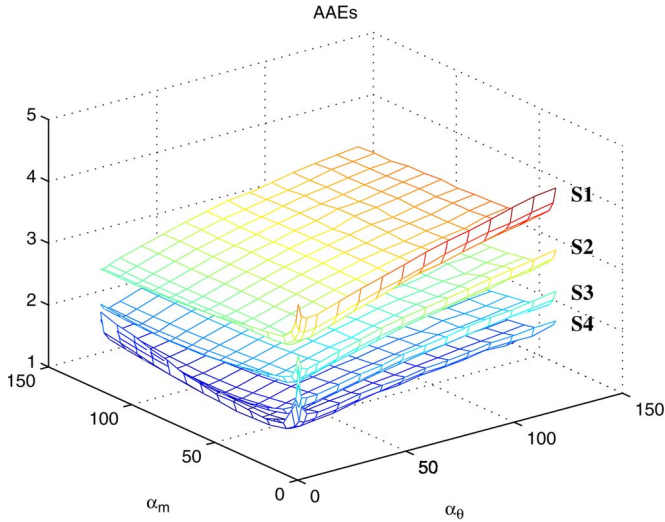
AAEs

Fig. 8.   Three-dimensional plot of the AAEs of all sequences for varying α_θ and α_m values.

Further, for analysis of the influence of speed, it is not good to have an equal number of frames in all of the sequences of different speeds. Because each sequence has a different displacement per frame, having an equal number of frames in all of them will result in having the vehicle camera move different distances and ending up processing different sequences. Because there is a different scene geometry with different buildings in the 3-D model, the nth frame in S1 will have a different scene geometry from the nth frame in S2, S3, and S4. Because the scene geometry also affects optical-flow accuracy, in this experiment, we have generated sequences of different speeds, but the vehicle camera travels a constant distance in all of them along the camera axis of the 3-D model, hence generating varying numbers of frames in different sequences. This way, all the sequences cover exactly the same geometric scene, but with a different number of frames. Therefore, we have 40 frames in S1, 20 frames in S2, 13 frames in S3, and 10 frames in S4. The average of error measures of all the frames in a sequence are considered for analysis. We have considered both AAE and EPE for analysis. All the errors in this analysis are computed over a region of interest (RoI) of size 320×480 at the center of the flow field. The considered RoI is shown in Fig. 6, *left*.

Fig. 7 shows a 3-D representation of the AAE for sequence S1 for varying values of two regularization weights α_θ and α_m. The 3-D error representations of AAEs from all the four sequences are shown in Fig. 8. The minimum AAEs and the corresponding regularization weights for all the sequences are given in Table II. Observing the meshes in Fig. 8 and by analyzing the minimum AAE values in Table II, we can conclude that the error in the sequence of lower speed is always higher than the error in the sequence of higher speed at almost all combinations of regularization weights. The values of α_θ and α_m in Table II reveal that α_θ is constant around 2.5 and 5, where α_m values decrease as the speed increases. It can be inferred that, overall, the AAE decreases with the increase in speed of the vehicle, α_θ has to slightly be increased, and α_m should be tuned with the change in speed of the vehicle.

TABLE II
REGULARIZATION PARAMETER VALUES THAT PRODUCE THE
LOWEST AAEs IN EACH OF THE SEQUENCES

Sequence	α_θ	α_m	AAE
S1	2.5	120	2.9261
S2	2.5	90	2.2481
S3	5	60	1.7618
S4	5	50	1.5695

EPE in S1
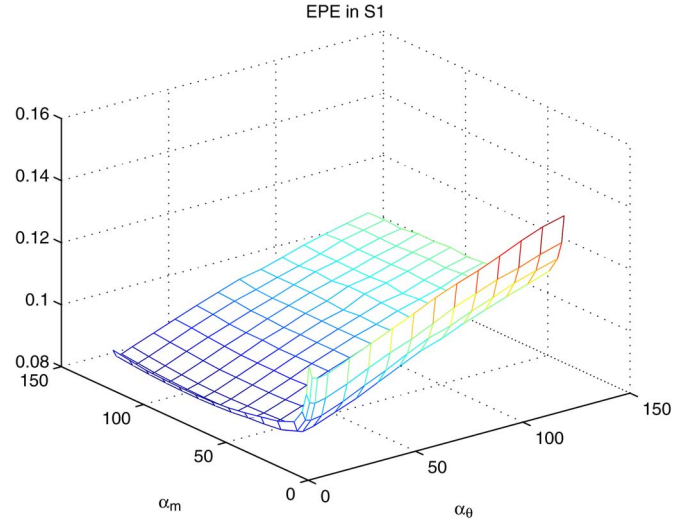
Fig. 9.   Three-dimensional plot of the EPEs from S1 for varying α_θ and α_m values.
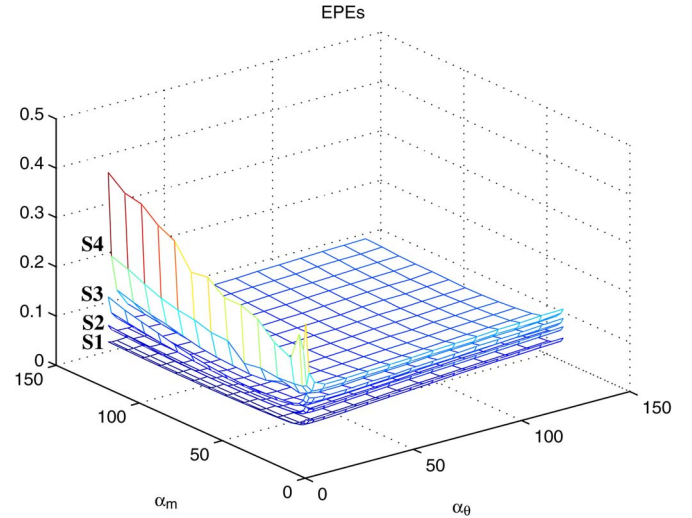
EPEs

Fig. 10.   Three-dimensional plot of the EPEs of all sequences for varying α_θ and α_m values.

A similar analysis is also done using the EPE. Fig. 9 shows the 3-D representation of the EPE for S1 for all combinations of two regularization weights. The 3-D representations of EPEs of all four sequences are depicted in Fig. 10. The minimum EPEs for all four sequences with corresponding regularization weights are shown in Table III. It is observed in the error maps in Fig. 10 and Table III that the EPE in a lower speed sequence is lower than in a sequence of higher speed for any combination of both regularization weights. In Table III, α_θ increases from a smaller value as the speed increases, whereas

TABLE III
REGULARIZATION PARAMETER VALUES THAT PRODUCE THE
LOWEST EPEs IN EACH OF THE SEQUENCES

Sequence	α_θ	α_m	EPE
S1	2.5	70	0.0887
S2	5	60	0.1063
S3	10	60	0.1163
S4	20	60	0.1338

TABLE IV
TEXTURE METRICS FOR THE DIFFERENT SEQUENCES

Sequence	*Contrast*	*Correlation*	*Homogeneity*
T1	0.0488	0.9386	0.9765
T2	0.0850	0.8611	0.9575
T3	0.1255	0.7168	0.9372

α_m keeps constant at around value 60. From the point of view of EPE, α_m has to be kept constant at a higher value, and α_θ should be tuned according to the change in speed of the vehicle. One interesting conclusion from this first study is that, depending on the required accuracy (AAE or EPE, i.e., angular or magnitudinal) needed for a given application, different tuning of regularization parameters has to be applied. Furthermore, it is clear that there is a relationship between this parameter tuning and the current speed of the vehicle.

B. Analysis for Texture

The aim of the work in this section is to analyze the influence of road texture on optical-flow accuracy and to identify the way of adjusting the regularization weights for better results. We have generated several sequences with different road textures, and some of the images of these sequences are shown in Fig. 4. The study in this section is performed considering three sequences with the increasing value of texture *contrast*. Hereinafter, they are referred to as T1, T2, and T3. These sequences are of the same speed as S1, but with different road textures. The texture metrics are computed over a small RoI of size 146×430 on the road surface. This RoI is shown in Fig. 6, *right*. Again, in this section, the polar-represented optical flow described in Section IV is used. The optical flow is computed on all image pairs from these sequences, which were obtained by assuming that the onboard vision system travels at the same speed. The average error values of all the flow fields in the same small RoI (where texture metrics were calculated) in a sequence are computed. We consider both the AAE and EPE for analysis. Table IV gives the texture metrics for the sequences. Figs. 11 and 12 show 3-D representations of AAEs and EPEs, respectively, for three sequences of different textures T1, T2, and T3. The minimum AAEs and EPEs with the corresponding regularization weights are shown in Tables V and VI, respectively. By observing Figs. 11 and 12 and Tables V and VI, it can easily be confirmed that both the AAE and EPE measures decrease with the increase in texture contrast. The regularization weights in Tables V and VI reveal that both values should increase with the increase in texture *contrast* for better results. Similarly, these results can be correlated with other textural properties such as *correlation* and *homogeneity* in Table IV.
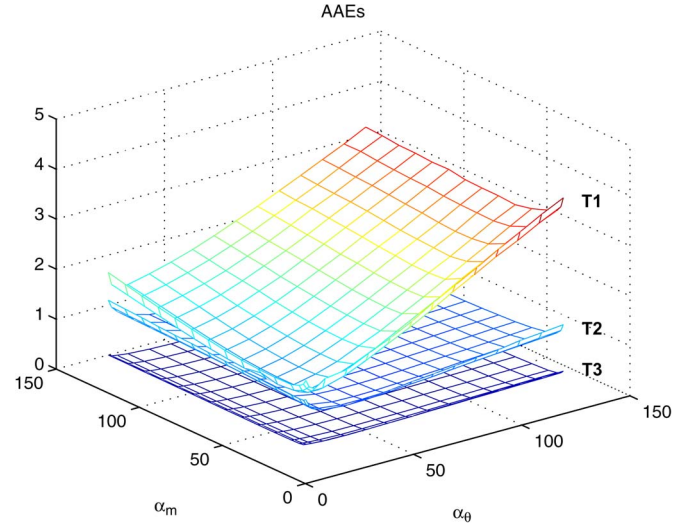

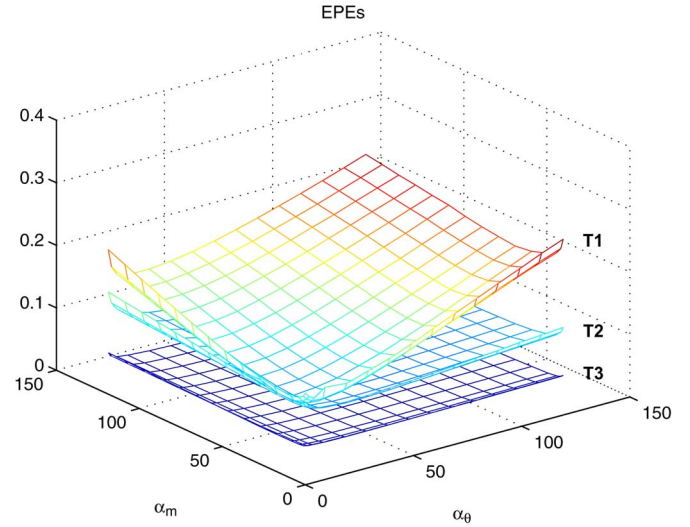Fig. 11. Three-dimensional plot of the AAEs from three different textured sequences for varying α_θ and α_m values.

Fig. 12. Three-dimensional plot of the EPEs from three different textured sequences for varying α_θ and α_m values.

TABLE V
REGULARIZATION PARAMETER VALUES WITH THE LOWEST AAEs

Sequence	α_θ	α_m	AAE
T1	10	20	1.5564
T2	40	90	0.7237
T3	60	120	0.6057

TABLE VI
REGULARIZATION PARAMETER VALUES WITH THE LOWEST EPEs

Sequence	α_θ	α_m	EPE
T1	10	20	0.1095
T2	60	80	0.0548
T3	80	120	0.0466

C. Analysis for Both Speed and Texture

Furthermore, we have performed experiments to analyze the influences of both speed and texture together. We use 12 different sequences of four different speeds and three different textures. Optical flow is estimated on all the frames in these
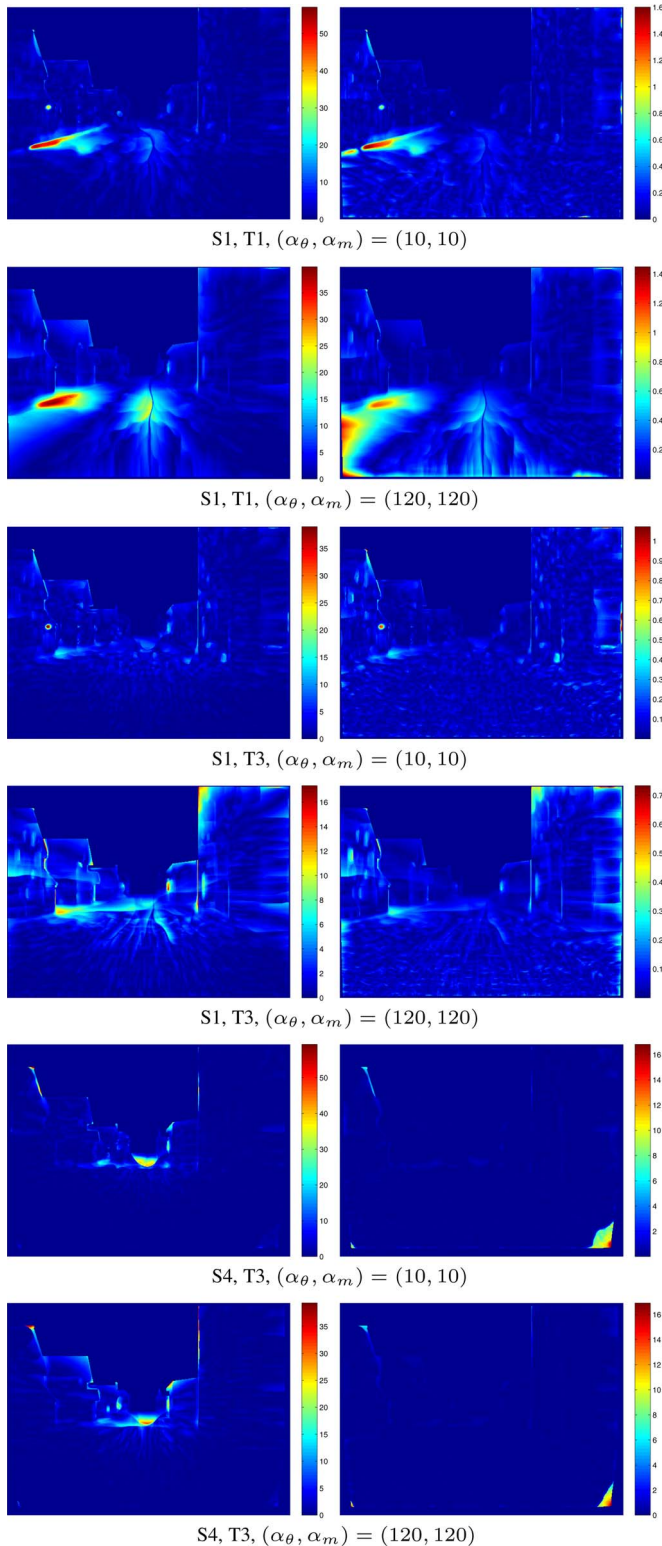
S1, T1, $(\alpha_\theta, \alpha_m) = (10, 10)$

S1, T1, $(\alpha_\theta, \alpha_m) = (120, 120)$

S1, T3, $(\alpha_\theta, \alpha_m) = (10, 10)$

S1, T3, $(\alpha_\theta, \alpha_m) = (120, 120)$

S4, T3, $(\alpha_\theta, \alpha_m) = (10, 10)$

S4, T3, $(\alpha_\theta, \alpha_m) = (120, 120)$

Fig. 13. Error images for the same image pairs shown in Fig. 1 and Table I. *Left*: AAEs. *Right*: EPEs.

TABLE VII
MINIMUM AAEs AND THEIR CORRESPONDING
REGULARIZATION WEIGHTS $(\alpha_\theta, \alpha_m)$

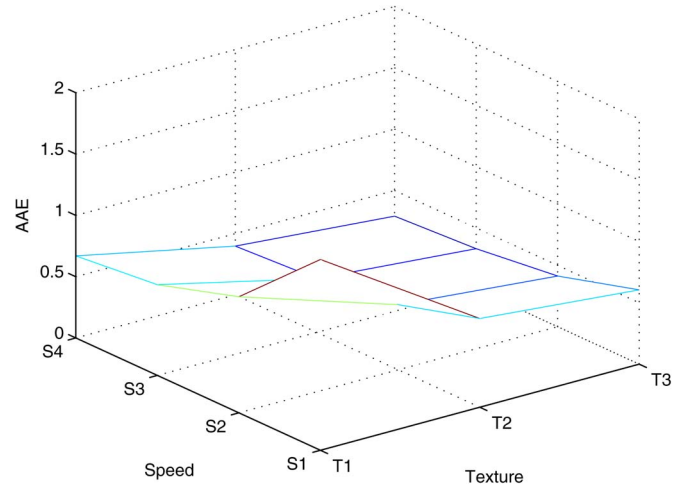| Sequence	T1	T2	T3
S1	1.5564 (10,20)	0.7237 (40,90)	0.6057 (60,120)
S2	0.9472 (10,10)	0.5307 (50,60)	0.4126 (80,70)
S3	0.7402 (20,10)	0.4382 (70,50)	0.3305 (100,70)
S4	0.6694 (30,5)	0.3984 (80,20)	0.2906 (110,20)

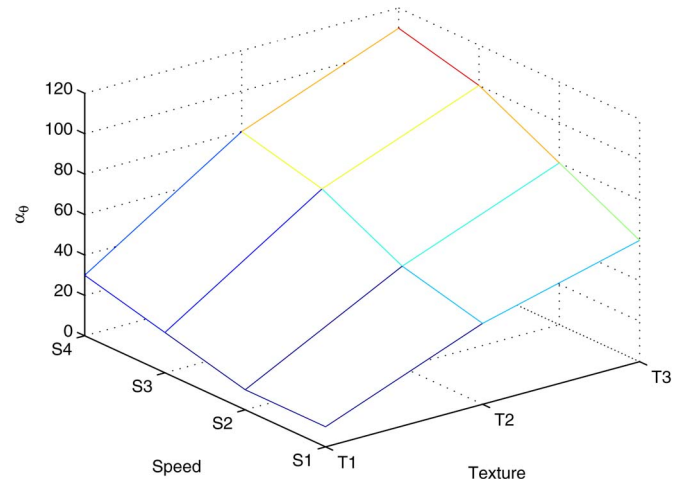Fig. 14. Three-dimensional plot of the minimum AAE for all sequences with different speeds and textures for varying α_θ and α_m values.

Fig. 15. Three-dimensional plot of α_θ that corresponds to the minimum AAE for all sequences with different speeds and textures.

sequences, and errors are computed. The error for a particular sequence is the average of errors in all the flow fields in that sequence. Error heat maps are shown in Fig. 13 for the same flow fields shown in Fig. 1. The errors are calculated on a small RoI of size 146×430 on the road surface, which is the same as in the previous subsection. Table VII shows the minimum AAEs

for 12 different sequences of different speeds and textures. The regularization weights that correspond to the minimum errors are also mentioned in brackets in Table VII. Three-dimensional plots of the minimum AAEs and the corresponding α_θ's and α_m's are shown in Figs. 14–16, respectively. In Table VII and Fig. 14, we can notice that the AAE reduces with the increase in texture $contrast$, as well as with the increase in speed. With respect to the AAE, Fig. 15 indicates that α_θ has to be kept small and slightly increase when the speed increases for a sequence of lower texture, but it has to be higher and to increase
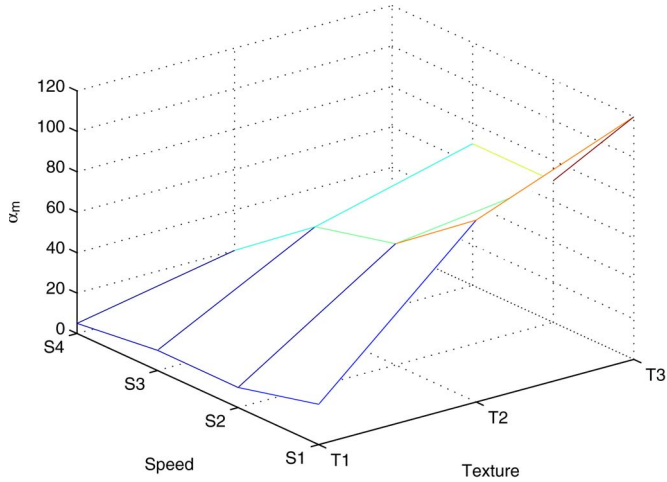
Fig. 16. Three-dimensional plot of α_m that corresponds to the minimum AAE for all sequences with different speeds and textures.

TABLE VIII
MINIMUM EPEs AND THEIR CORRESPONDING
REGULARIZATION WEIGHTS $(\alpha_\theta, \alpha_m)$

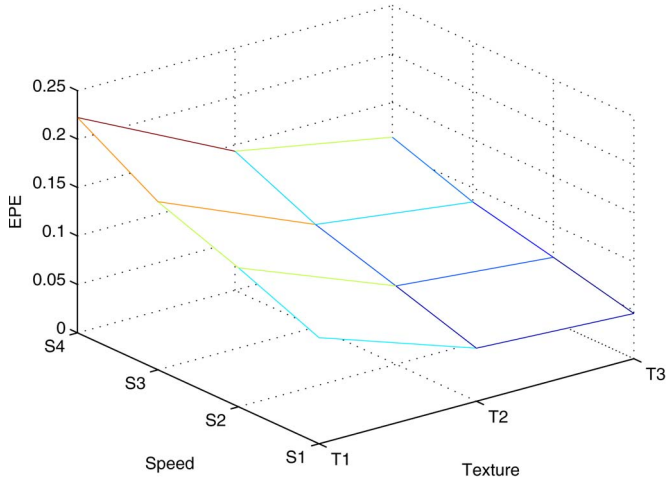| Sequence	T1	T2	T3
S1	0.1095 (10,20)	0.0548 (60,80)	0.0466 (80,120)
S2	0.1434 (20,10)	0.0809 (90,70)	0.0664 (120,120)
S3	0.1735 (20,10)	0.1061 (110,60)	0.0854 (120,80)
S4	0.2223 (60,10)	0.1435 (120,10)	0.1141 (120,50)

Fig. 17. Three-dimensional plot of the minimum EPE for all sequences with different speeds and textures for varying α_θ and α_m values.

when the speed increases for a sequence of higher texture. In conclusion, α_θ has to increase with the increase in speed and texture. The 3-D representation in Fig. 16 indicates that α_m has to decrease with the increase in speed and to increase with the increase in texture.

A similar experiment on all the 12 sequences is performed considering the EPEs. Table VIII shows the minimum EPEs and the corresponding regularization weights in brackets. Figs. 17–19 are the 3-D representations of the minimum EPEs, α_θ's, and α_m's, respectively. In Table VIII and Fig. 17, we
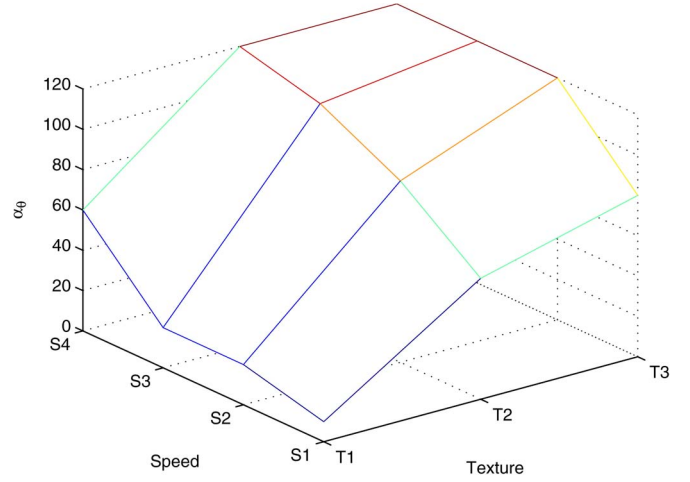

Fig. 18. Three-dimensional plot of α_θ that corresponds to the minimum EPE for all sequences with different speeds and textures.
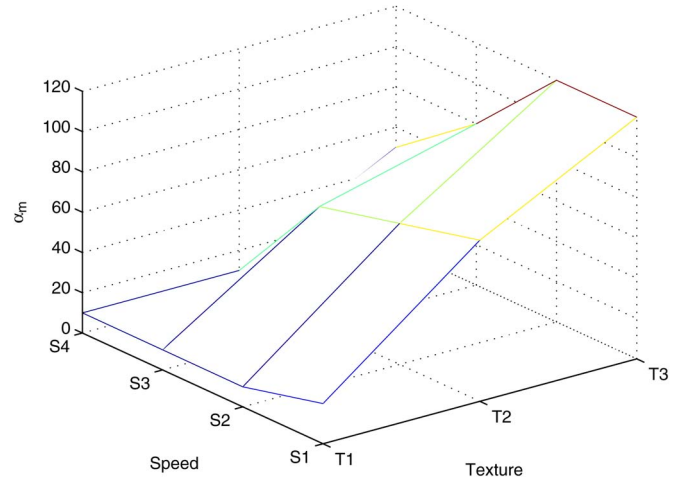

Fig. 19. Three-dimensional plot of α_m that corresponds to the minimum EPE for all sequences with different speeds and textures.

can observe that the EPE reduces with the increase in texture contrast but increases with the increase in speed. Fig. 18 shows that α_θ has to increase with the increase in speed and with the increase in texture contrast, whereas the 3-D representation in Fig. 19 indicates that α_m has to increase with the increase in texture but has to decrease with the increase in speed, except for lower textured sequences (e.g., the sequence with texture T1 and with speeds S1, S2, S3, and S4).

Furthermore, to ensure the conclusions based on Tables II and VII about the decrease in AAE with the increase in speed and the conclusion that the AAE will decrease with the increase in texture contrast, we analyzed the AAEs of all 12 sequences, keeping α_θ and α_m constant at different values. Table IX shows the AAEs in all sequences for fixed α_θ and α_m values at 40. The AAEs in this table confirm our conclusion that the AAE decreases with the increase in speed and texture. Comparing the error values in Tables VII and IX, it is clear that tuning of regularization weights is very important in getting accurate optical flow. Similarly, Table X shows the EPEs for a fixed α_θ and α_m value of 40 for the sequences. This also reaffirms that the EPE increases with the increase in speed but decreases

TABLE IX
AAEs FOR FIXED REGULARIZATION WEIGHTS: $\alpha_\theta = 40$ AND $\alpha_m = 40$

Sequence	T1	T2	T3
S1	1.9815	0.7376	0.6282
S2	1.0912	0.5336	0.4154
S3	0.8105	0.4447	0.3371
S4	0.7250	0.4116	0.2979

TABLE X
EPEs FOR FIXED REGULARIZATION WEIGHTS: $\alpha_\theta = 40$ AND $\alpha_m = 40$

Sequence	T1	T2	T3
S1	0.1333	0.0559	0.0494
S2	0.1649	0.0828	0.0696
S3	0.1988	0.1112	0.0915
S4	0.3049	0.1672	0.1255

TABLE XI
MINIMUM AAEs AND THEIR CORRESPONDING
REGULARIZATION WEIGHTS $(\alpha_\theta, \alpha_m)$

Sequence	T1	T2	T3
S1	1.6947 (5,10)	0.7859 (30,60)	0.7030 (40,60)
S2	1.2964 (10,10)	0.6670 (40,50)	0.5370 (40,60)
S3	1.2213 (30,20)	0.5392 (60,50)	0.4328 (70,60)
S4	0.9361 (40,20)	0.4582 (80,20)	0.3469 (110,60)

TABLE XII
MINIMUM EPEs AND THEIR CORRESPONDING
REGULARIZATION WEIGHTS $(\alpha_\theta, \alpha_m)$

Sequence	T1	T2	T3
S1	0.1388 (10,30)	0.0589 (60,80)	0.0520 (70,110)
S2	0.1644 (20,20)	0.0830 (80,70)	0.0683 (100,100)
S3	0.2240 (40,20)	0.1035 (90,50)	0.0824 (120,90)
S4	0.2529 (60,20)	0.1271 (110,30)	0.1000 (120,50)

with the increase in texture and comparing values based on Tables VIII and X reveals that tuning of regularization weights is needed.

Finally, a similar analysis as in the previous one has been performed, but by adding complexity to the motion. The new sequences involve large changes in (*yaw* and *pitch*) angles. All these sequences are ten frames long. The optical flow is computed for varying regularization weights, and the errors are computed on a small RoI (the one shown in Fig. 6, *right*). The minimum AAE and EPE are shown in Tables XI and XII, respectively, for all the 12 different sequences in this complex set. All these sequences have the same degree of egomotion, but the onboard camera moves at different speeds and on different textures. Here, we can observe almost the same trends in error values and regularization weights as in the previous study. α_θ has to be increased when the speed and texture contrast increase for both the AAE and EPE, whereas α_m has to be increased when the texture contrast increases for both the AAE and EPE. Because sequences have egomotion, changes in α_m do not affect much in the AAE; it is almost constant with the increase in speed for the AAE, and it has to slightly decrease with the increase in speed for the EPE.

For the completeness of our study, we have added few independently moving vehicles in the scene and performed similar analysis. In this particular case, the RoI corresponds to the one shown in Fig. 6, *left*. As expected, independently moving objects are another source of errors that cannot be tackled merely by tuning regularization weights, although it could improve. The relative motion of onboard camera and moving vehicles causes different groups of flow vectors. Errors due to occlusions and due to moving objects present in the scene are not under our control by tuning regularization weights. Overall, we can tune on the things that are related with static and under our control (such as speed and texture), but not on the behavior of dynamic moving vehicles present in the given scenario.

VIII. DISCUSSION

Although it is out of the scope of this paper the question on how we can tune the regularization parameters could arise. In general, based on the aforementioned study, we can say that having the best set of parameters would depend on the current

scenario; however, for a given set of regularization parameters (independently, whether it is the best set), we can adapt its values according to the speed and texture using the information presented in the previous section. We perceive that this analysis is just a tip on how to proceed and a rigorous study and validation should be performed to define a rule to adapt regularization parameters to particular characteristics of a sequence. A much deeper study is required to conclude adaptation rules in the case of combinations of several characteristics of any sequence. The study in this paper is a starting point in that direction.

Although an increase in speed of the vehicle camera can be compensated for by increasing the camera cycle time, which is by a higher number of frames per second (FPS), the increase in the number of frames leads to higher computation burden. A related work [32] in that direction proposes to change the resolution of images for varied FPS in a variational hierarchical framework. Moreover, the maximum number of FPS of a camera is limited by its hardware. In this paper, the influence of vehicle speed for a fixed cycle time of a camera is considered.

IX. CONCLUSION

This paper has shown that the polar representation of flow vectors is very convenient in ADAS scenarios due to its freedom of differently weighing regularization terms and has further used polar-represented optical-flow estimation for the analysis. The analysis of optical-flow accuracy to specific characteristics of a driving scenario, e.g., vehicle speed and road texture, is performed. It is concluded that there is a need to tune the regularization parameters, depending on the needed accuracy (angular or magnitudinal) and for varying speeds and textural properties of the road. This paper has also presented a framework and generated synthetic video sequences along with ground-truth flow fields for different scenarios of speeds and road textures.

REFERENCES

[1] A. Jazayeri, H. Cai, J. Y. Zheng, and M. Tuceryan, "Vehicle detection and tracking in car video based on motion model," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 583–595, Jun. 2011.

[2] S. Cherng, C.-Y. Fang, C.-P. Chen, and S.-W. Chen, "Critical motion detection of nearby moving vehicles in a vision-based driver-assistance system," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 70–82, Mar. 2009.

[3] B. K. P. Horn and B. G. Schunk, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1–3, pp. 185–203, Aug. 1981.

[4] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision (DARPA)," in *Proc. DARPA Image Understanding Workshop*, Apr. 1981, pp. 121–130.

[5] I. Tang and T. P. Breckon, "Automatic road environment classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 476–484, Jun. 2011.

[6] Y. Adato, T. Zickler, and O. Ben-Shahar, "A polar representation of motion and implications for optical flow," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Colorado Springs, CO, USA, Jun. 2011, pp. 1145–1152.

[7] A. Bruhn, "Variational optic flow computation: Accurate modeling and efficient numerics," Ph.D. dissertation, Dept. Math. Comput. Sci., Saarland Univ., Saarbrücken, Germany, 2006.

[8] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 43–77, Feb. 1994.

[9] B. Galvin, B. Mccane, K. Novins, D. Mason, and S. Mills, "Recovering motion fields: An evaluation of eight optical flow algorithms," in *Proc. British Mach. Vis. Conf.*, Southampton, U.K., 1998, pp. 195–204.

[10] M. Otte and H.-H. Nagel, "Estimation of optical flow based on higher order spatiotemporal derivatives in interlaced and noninterlaced image sequences," *Artif. Intell.*, vol. 78, no. 1/2, pp. 5–43, Oct. 1995.

[11] [Online]. Available: http://vision.middlebury.edu/flow/

[12] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 1–31, Mar. 2011.

[13] A. Wedel, T. Pock, C. Zach, D. Cremers, and H. Bischof, "An improved algorithm for TV-L1 optical flow," in *Proc. Dagstuhl Motion Workshop*, Dagstuhl Castle, Germany, Sep. 2008, pp. 23–45.

[14] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High-accuracy optical flow estimation based on a theory for warping," in *Proc. Eur. Conf. Comput. Vis.*, Prague, Czech Republic, May 2004, vol. 3024, pp. 25–36.

[15] M. Black and P. Anandan, "Robust dynamic motion estimation over time," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Maui, HI, USA, Jun. 1991, pp. 296–302.

[16] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods," *Int. J. Comput. Vis.*, vol. 61, no. 3, pp. 211–231, Feb. 2005.

[17] H.-H. Nagel and W. Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 5, pp. 565–593, Sep. 1986.

[18] J. Weickert and C. Schnörr, "Variational optic flow computation with a spatiotemporal smoothness constraint," *J. Math. Imaging Vis.*, vol. 14, no. 3, pp. 245–255, May 2001.

[19] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 2432–2439.

[20] S. Volz, A. Bruhn, L. Valgaerts, and H. Zimmer, "Modeling temporal coherence for optical flow," in *Proc. IEEE Int. Conf. Comput. Vis.*, Barcelona, Spain, Nov. 2011, pp. 1116–1123.

[21] L. Xu, J. Jia, and Y. Matsushita, "Motion detail preserving optical flow estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1744–1757, Sep. 2012.

[22] K. Krajsek and R. Mester, "Bayesian model selection for optical flow estimation," in *Proc. DAGM Symp.*, Heidelberg, Germany, Sep. 2007, pp. 142–151.

[23] H. Zimmer, A. Bruhn, and J. Weickert, "Optic flow in harmony," *Int. J. Comput. Vis.*, vol. 93, no. 3, pp. 368–388, Jul. 2011.

[24] O. Mac Aodha, G. J. Brostow, and M. Pollefeys, "Segmenting video into classes of algorithm—Suitability," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 1054–1061.

[25] N. Onkarappa and A. Sappa, "An empirical study on optical flow accuracy depending on vehicle speed," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2012, pp. 1138–1143.

[26] S. Roth and M. J. Black, "On the spatial statistics of optical flow," in *Proc. IEEE Int. Conf. Comput. Vis.*, Beijing, China, Oct. 2005, pp. 42–49.

[27] P. Wu, Y. M. Ro, C. S. Won, and Y. Choi, "Texture descriptors in MPEG-7," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, Warsaw, Poland, Sep. 2001, pp. 21–28.

[28] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[29] T. Vaudrey, C. Rabe, R. Klette, and J. Milburn, "Differences between stereo and motion behavior on synthetic and real-world stereo sequences," in *Proc. Image Vis. Comput.*, Christchurch, New Zealand, Nov. 2008, pp. 1–6.

[30] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A naturalistic open source movie for optical flow evaluation," in *Proc. Eur. Conf. Comput. Vision*, A. Fitzgibbon, P. Perona, Y. Sato, and C. Schmid, Eds., Oct. 2012, pp. 611–625, Springer-Verlag.

[31] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI Vision Benchmark Suite," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Providence, RI, USA, Jun. 2012, pp. 3354–3361.

[32] Y. Kameda, A. Imiya, and T. Sakai, "Hierarchical properties of multiresolution optical flow computation," in *Proc. ECCV Workshops Demo.*, 2012, vol. 7584, pp. 576–585, Springer Berlin Heidelberg.

Naveen Onkarappa received the B.Sc. degree in computer science from Kuvempu University, Shimoga, India, in 1999 and the M.Sc. degree in computer science and the M.Sc.Tech. degree (by research) in computer science and technology from the University of Mysore, Mysore, India, in 2001 and 2007, respectively. He is currently a Ph.D. student, under the supervision of Dr. Angel D. Sappa, at the Computer Vision Center, Autonomous University of Barcelona, Barcelona, Spain.

From 2001 to 2005, he was a Guest Lecturer with Kuvempu University and the University of Mysore. From 2007 to 2009 he was with HCL Technologies, Bangalore, India. He is a member of the Advanced Driver Assistance Systems Group, Computer Vision Center, Autonomous University of Barcelona. His research interests include optical-flow estimation and its applications to driver assistance systems.

Angel Domingo Sappa (S'94–M'00–SM'12) received the B.Eng. degree in electromechanical engineering from the National University of La Pampa, General Pico, Argentina, in 1995 and the Ph.D. degree in industrial engineering from the Polytechnic University of Catalonia, Barcelona, Spain, in 1999.

In 2003, after holding research positions in France, the U.K., and Greece, he joined the Computer Vision Center, Autonomous University of Barcelona, Barcelona, Spain, where he is currently a Senior Researcher. He is a member of the Advanced Driver Assistance Systems Group, Computer Vision Center, Autonomous University of Barcelona. His research interests include 2-D and 3-D image processing, with emphasis on stereoimage processing and analysis, 3-D modeling, dense optical-flow estimation, and visual SLAM for driving assistance.