J Math Imaging Vis (2013) 47:48–59
DOI 10.1007/s10851-012-0384-5

# A Novel Space Variant Image Representation

**Naveen Onkarappa · Angel D. Sappa**

Published online: 26 September 2012


**Abstract** Traditionally, in machine vision images are represented using cartesian coordinates with uniform sampling along the axes. On the contrary, biological vision systems represent images using polar coordinates with non-uniform sampling. For various advantages provided by space-variant representations many researchers are interested in space-variant computer vision. In this direction the current work proposes a novel and simple space variant representation of images. The proposed representation is compared with the classical log-polar mapping. The log-polar representation is motivated by biological vision having the characteristic of higher resolution at the fovea and reduced resolution at the periphery. On the contrary to the log-polar, the proposed new representation has higher resolution at the periphery and lower resolution at the fovea. Our proposal is proved to be a better representation in navigational scenarios such as driver assistance systems and robotics. The experimental results involve analysis of optical flow fields computed on both proposed and log-polar representations. Additionally, an egomotion estimation application is also shown as an illustrative example. The experimental analysis comprises results from synthetic as well as real sequences.

**Keywords** Space-variant representation · Log-polar mapping · Onboard vision applications

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

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

## 1 Introduction

Current technology allows to take larger and larger images, at the same time computational power is also increasing, being able to process large amount of information in short period of time. Unfortunately the trade-off between computation time and amount of information to process is still there and needs specific solutions. One way to tackle this challenging problem is by means of the use of space-variant representations, which have been motivated by the mammalian biological vision systems [5, 23], and allows to reduce the amount of information to be processed. The mammalian retina is a space variant sensor with higher sensory neurons at fovea and gradually reducing towards periphery. This leads to high resolution at the center and wide field of view simultaneously without many sensing elements in the periphery. There have been many attempts to design space-variant imaging sensors [27] as well as approaches based on such representations for computer vision applications (e.g., [13, 32]). On the other hand, space variant represented images can be obtained from the conventional rectangular sensors and can have the same advantage of space-variant sensors. The Log-Polar representation (LPR) of images has been originated from such biological vision and has been extensively studied in the computer vision community. Additionally, it has been also exploited in the robotics and active vision communities for pattern recognition [28] and navigation [10] tasks. The LPR has many advantages with respect to the conventional cartesian representation of images [26]; the most important are the reduction in the data and invariance to scale and rotation. The data reduction due to the polar mapping and logarithmic sub-sampling leads to high resolution in the fovea and low resolution in the periphery, which is a desired feature for instance in the active vision community. The importance of loss of information in

LPR due to reduced resolution in the periphery depends on the application of interest.

A review of log-polar imaging is presented in [27] for robotic vision applications such as: visual attention, target tracking and 3D perception. All these applications benefit from the high resolution of the fovea region. Also, there have been attempts to use LPRs for motion analysis (e.g., [6, 10, 29]), mainly based on the estimation of optical flow (e.g., [3, 11, 12, 30]). For instance, [26] presents the advantages of polar and log-polar mapping to the cartesian representation and proposes a technique to estimate time-to-impact using optical flow. Log-polar mapping and its properties are described in [12]; and it proposes the computation of normal optical flow. In [33], a novel optical flow computation approach is proposed. It is based on the concept of variable window and generalized dynamic image model. The variable window adapts its size along the LP space. Also working in the LP space, [11] analyzes the polar deformation and proposes several local optical flow estimation techniques on log-polar plane. Based on LPR, an approach for active monocular fixation is proposed in [34] and an application of docking using rotational-component of log-polar optical flow is proposed in [3]. The optical flow approaches presented in [3, 11, 12, 26] and [33] are all local approaches. However the dense flow fields would be more useful for driver assistance applications such as obstacle detection, egomotion estimation or collision avoidance. In this paper we use a global optical flow technique to obtain dense flow-field on space-variant represented images.

In the particular contexts of robotics and advanced driver assistance systems (ADAS), LPR has attracted the attention of many researchers. In general, in these fields LPRs are obtained using the *vanishing point* (VP) as a center of the log-polar reference system. $VP_{(x,y,z)}$ corresponds to a point at $z \to \infty$ where two parallel lines of a road appear to converge in the image plane. Since LPR results in a high sampling in the fovea region, points in the periphery are undersampled. It should be noted that the periphery of the 2D image (in particular the lateral and bottom parts) corresponds to regions in the 3D space near to the camera reference system, hence are the most important areas for robotics navigation tasks and ADAS applications. Furthermore, points near to the camera are not only useful for detection tasks but also for an accurate calibration; note that the accuracy of 3D data decreases with the depth.

In the current work classical LPR is studied and a new space variant representation scheme called Reverse Log-Polar Representation (RLPR) is proposed. It is intended to overcome the disadvantage of LPR with respect to the sampling in periphery. In addition to the benefit of reduction in information to be processed similar to LPR, RLPR produces higher accurate results in navigational tasks. Unlike LPR which has advantages in pattern recognition applications with rich fovea and sparse periphery, the newly proposed spatial representation has advantages in navigational tasks. These space variant representations are used for optical flow estimation in forward facing motion problems. The performance of both representations is analyzed using dense optical flow. Additionally the accuracy of both space variant representations to estimate the egomotion parameters of a moving camera is evaluated. The paper is organized as follows. Section 2 initially overviews the dense optical flow estimation; Section 3 presents a statistical analysis that support the use of polar representations, which motivates the current work. Then, the proposed space variant representation is introduced in Sect. 4. Experimental results and comparative studies are given in Sect. 5 together with an application to the egomotion estimation problem. Finally, the work is concluded in Sect. 6.

## 2 Optical Flow Overview

Optical flow is defined as a velocity field between an image pair, which transforms one image into the next image in a sequence. Even though many attempts have been made to estimate motion since a long time, concrete formulations were first proposed by Horn and Schunck [16], and Lucas and Kanade [19] in 1981. Broadly, optical flow methods can be classified as local and global methods. The local methods give sparse flow fields, whereas global methods give dense flow fields using variational methods. The first variational optical flow method was proposed by Horn and Schunck [16]. The classical variational method is based on two assumptions: the *brightness constancy assumption* (BCA) and homogeneous *regularization*. BCA is also called as *optical flow constraint*, assumes the grey value of an object remains constant over time. Whereas regularization assumes that the resulting flow field varies smoothly all over the image. The BCA can be formulated as:

$$I_1\big(\boldsymbol{X} + \boldsymbol{V}(\boldsymbol{X})\big) - I_0(\boldsymbol{X}) = 0, \tag{1}$$

where $I_0$ and $I_1$ are two consecutive images of a given sequence; $\boldsymbol{X} = (x, y)$ is a pixel location within the rectangular image domain $\Omega \subseteq \mathbf{R}^2$; $\boldsymbol{V} = (u(X), v(X))$ is the two-dimensional displacement field. Linearising above equation using first-order Taylor expansion we get *optical flow constraint* as:

$$I_x u + I_y v + I_t = 0, \tag{2}$$

where $I_x$, $I_y$ and $I_t$ are the partial derivatives with respect to $x$, $y$ and $t$ respectively. Since optical flow is a highly ill-posed inverse problem, using only local intensity constraints does not provide enough information to infer meaningful flow fields. In particular, optical flow computation suffers

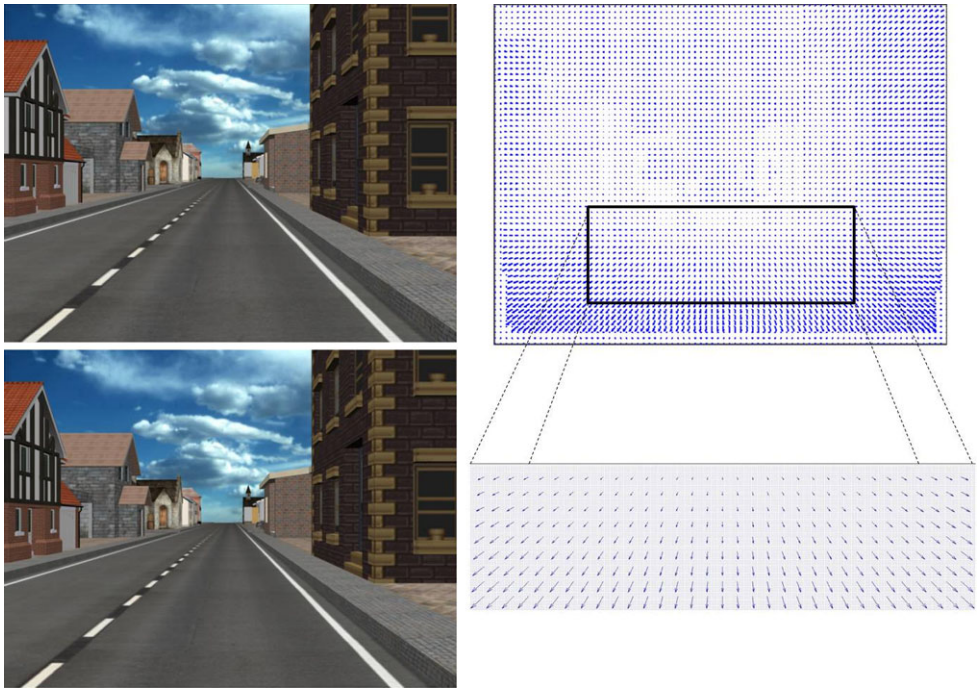

**Fig. 1** Synthetic sequence of an urban scenario: (*top-left*) Image $I_0$ at time $t$; (*bottom-left*) Image $I_1$ at time $t+1$; (*top-right*) optical flow vectors computed from $I_0$ and $I_1$; (*bottom-right*) enlarged region of the flow field

from two problems: first, no information is available in non-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*. It is clear that in order to solve this problem some kind of regularization is needed. The Horn and Schunk [16] method overcomes this by assuming resulting flow field changes smooth all over the image. This can be formulated as penalizing large spatial flow gradients $|\nabla u|$ and $|\nabla v|$. Combining BCA and smoothness assumptions in a single variational framework and squaring both constraints in order to penalize negative and positive derivations in the same way, the following energy functional is obtained.

$$E(\boldsymbol{V}) = \int_{\Omega} \{\underbrace{(I_x u + I_y v + I_t)^2}_{\textit{Data Term}} + \alpha\underbrace{(|\nabla u|^2 + |\nabla v|^2)}_{\textit{Regularization}}\} d\boldsymbol{X}, \tag{3}$$

where $\alpha$ is a regularization parameter.

Variational optical flow energy functions can be minimized in a number of ways. The most used way is to express and solve the set of Euler-Lagrange equations of the energy model. The thesis [7] presents various numerical linear and non-linear equation systems solvers such as: basic Gauss-Seidel method, its variants, advanced methods such as successive overrelaxation (SOR) technique, unidirectional multigrid methods in the form of coarse-to-fine strategies and bidirectional multigrid methods. Another popular way proposed in [35] using a dual formulation is based on iterative alternating steps to solve TV-L1 optical flow energy model.

For the work presented in this paper, we use a recent variational optical flow technique proposed in [24]. It explores the basic formulation and some concepts such as pre-processing, coarse-to-fine warping, graduated non-convexity, interpolation, derivatives, and median filtering. In [24], an improved model underlying median filtering is proposed by using best of the explored concepts in optical flow computation. As an illustration Fig. 1 shows two images and the ground-truth optical flow field between them.

## 3 Image Motion Statistics

Image motion statistics, in particular through the use of optical flow, have been recently studied and exploited (e.g., [22] and [1]). In the current section we follow previous studies using synthetic sequences of urban road scenarios (see Fig. 1) where cartesian and polar representations are used to depict motion information between two consecutive frames.

Figure 2 shows joint histograms of image motion derivatives in cartesian and polar coordinates—the image motion refers to the flow field estimated as indicated in the previous section. These joint histograms are then used for computing the mutual information (MI) between the variables. In probability theory, the mutual information of two random variables is a quantity that measures the mutual dependence of the two variables (see [9, 15]). Formally, it can be defined as:

$$MI(R;S) = \sum_{r\in R}\sum_{s\in S} p(r,s)\log\frac{p(r,s)}{p(r)p(s)}, \tag{4}$$

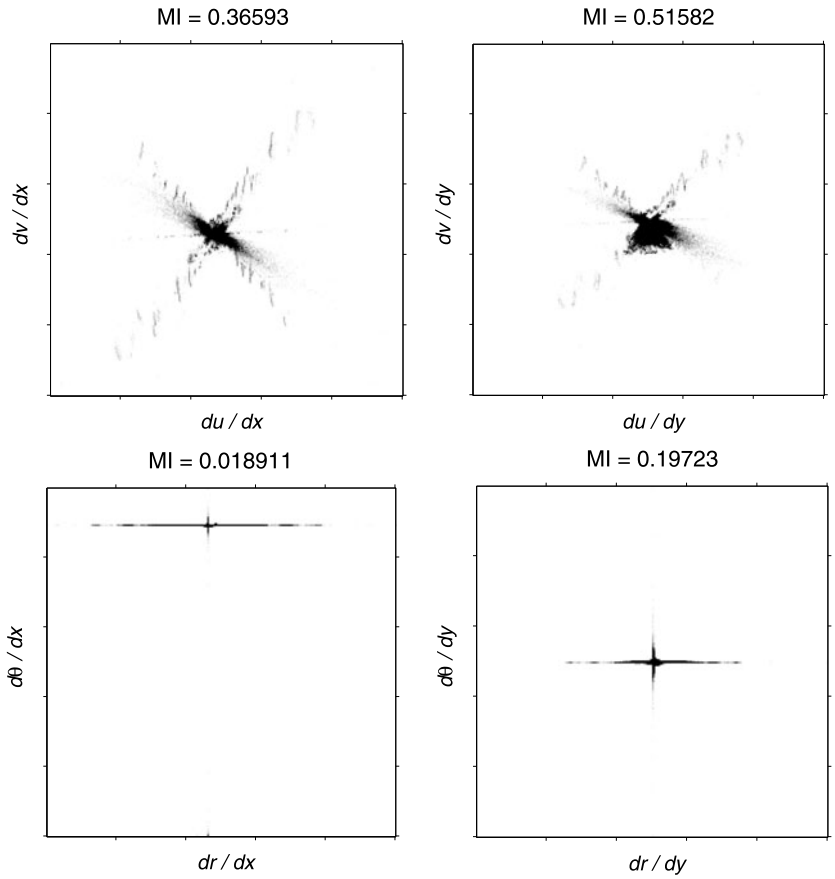


**Fig. 2** Joint histograms of image motion derivatives in cartesian and polar coordinates of synthetic sequences of an urban road scenario. On top of each figure *MI* value is depicted

where $p(r, s)$ is the joint probability distribution function of $R$ and $S$; $p(r)$ and $p(s)$ are the marginal probability distribution functions. The joint probability function $p(r, s)$ is directly obtained from the cell $(r, s)$ in the joint histogram of image motion, while $p(r)$ and $p(s)$ are found by summing along each dimension of that joint histogram. Small values of MI indicate approximate statistical independence. Hence, as shown in Fig. 2 (see MI at the top of every illustration), it can be concluded that it is more appropriate to represent the image motion in a polar coordinate system. A similar conclusion has been reached in [1] and [22].

The work in [1] is based on a polar representation of the flow vectors in optical flow formulation. The results using this formulation are shown to be same as the state of the art results on traditional datasets, and better on specular and fluid flow datasets. The statistics presented in the current section suggest us that in the particular case of forward facing vision systems, such as those used in robotics or driver assistance, motion is better represented in a polar way. In the current work we use polar based image representations; their advantages have been vastly used by the robotics community. The classical log-polar representation is explored and a new space variant representation is proposed as presented in the next section.

## 4 Space-Variant Representations

As presented in the previous section a polar representation offers advantages with respect to the cartesian one (in particular due to the independence of motion coordinates [1]). In the current section polar space-variant image representations are discussed.

A log-polar representation is a polar mapping with logarithmic distance along the radial axis. For a given pixel $(x, y)$, the log-polar $(\rho, \theta)$ are defined as:

$$\begin{aligned} \rho &= \log\bigl(\sqrt{(x - x_0)^2 + (y - y_0)^2}\bigr), \\ \theta &= \arctan\bigl((y - y_0)/(x - x_0)\bigr), \end{aligned} \tag{5}$$

where $(x_0, y_0)$ is the origin of mapping. The current work focuses on the study of a particular scenario of forward facing moving platforms, hence the origin of the reference system corresponds to the vanishing point.

As mentioned in Sect. 1, LPR oversamples the fovea and undersamples the periphery. This leads to the non-preservation of vital information of the periphery useful for forward facing vision applications. The latter motivates us

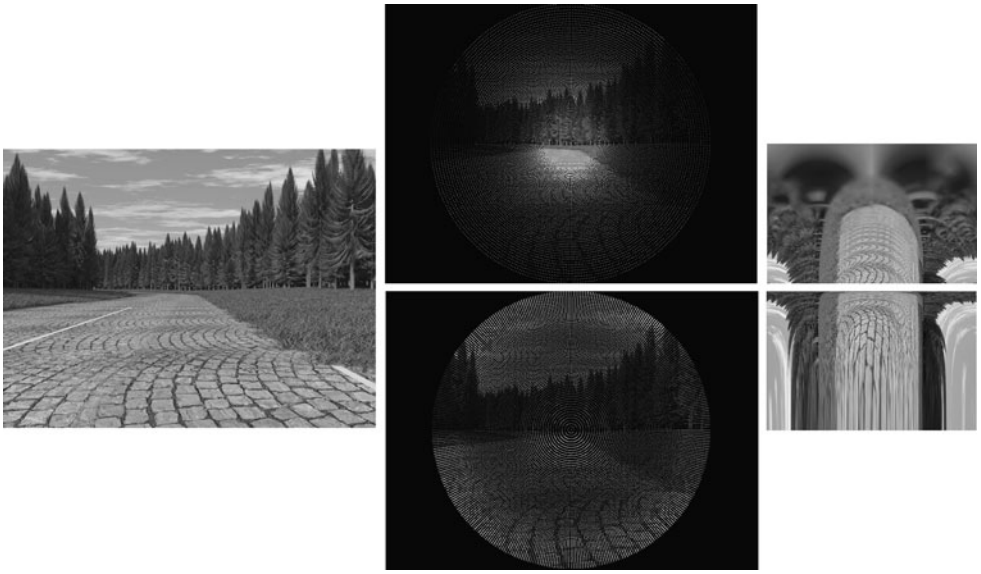

**Fig. 3** (*Left*) A synthetic image. (*Middle-top*) Log-Polar sampling. (*Middle-bottom*) Reverse-Log-Polar sampling. (*Right-top*) Cartesian representation of LP image. (*Right-bottom*) Cartesian representation of RLP image

to propose a more appropriate space variant representation, where a $(x, y)$ pixel is mapped as:

$$\rho = \log\bigl(r_{\max} - \sqrt{(x-x_0)^2 + (y-y_0)^2}\bigr),$$
$$\theta = \arctan\bigl((y-y_0)/(x-x_0)\bigr), \tag{6}$$

where $r_{\max}$ is the radius of the circle around VP; typically $r_{\max}$ is the largest inner circle within the cartesian image boundary around VP. The image contained inside the circular region is to be mapped. The proposed mapping is different from LPR in the sense that logarithmic subsampling is from the periphery towards the center and will be referred to as *Reverse Log-Polar Representation* (RLPR). Figure 3 (*right*) shows LP (*top*) and RLP (*bottom*) representations of the same image Fig. 3 (*left*). In both cases the images are sparsely sampled as depicted in Fig. 3 (*middle*). Since the LP/RLP transformations involve both many-to-one and one-to-many mapping, the LP/RLP images cannot be straight forwardly dense. The dense images presented in Fig. 3 (*right*) are obtained by querying for each $(\rho, \theta)$ to the cartesian and by bi-linear interpolation—horizontal axis is angles $(\theta' s)$ and vertical axis is distances $(\rho' s)$. As can be seen in the grids in Fig. 3 (*middle*), qualitatively, the RLPR sampling better preserves the periphery information, which covers most part of the road at the bottom in the scenario of a moving vehicle. Further experiments in this paper are performed on these images (shown in Fig. 3 (*right*)), represented as rectangular images, but sampled using LP and RLP sampling (Fig. 3 (*middle*)) from the original image shown in Fig. 3 (*left*).

## 5 Experimental Results

As mentioned in Sect. 1, there have been many applications using LP represented images, some of them based on the optical flow estimation in that space. The current work aims to estimate the optical flow on RLP represented images and compare them with results from LPRs. Further these optical flow fields are evaluated by estimating orientations of a moving camera.

In LP/RLP representations of images the origin of mapping should be the vanishing point in the scenario of a forward facing moving vehicle, so that the mapped images better suit the applications. In the current work, vanishing points computed from a RANSAC based approach [20] are used. The images are mapped to LP/RLP as explained in Sect. 4 and then, the optical flow is computed on these LP and RLP rectangular represented images. The bottleneck to compare the flow fields from LP and RLP representations is that the flow field patches at a particular location in both representations correspond to different regions of the image in cartesian with varied resolution. Hence, the framework proposed to perform the comparisons consists in inverse mapping the flow fields back to cartesian and compare them in the cartesian space. Figure 4 shows an image pair in cartesian (*top-left*), their LPRs (*middle-left*) and RLPRs (*bottom-left*) and in the (*top-right*) the ground-truth flow field, the computed flow field of LPR (*middle-right*) and RLPR (*bottom-right*). The color map used to display optical flow is shown in Fig. 4 bottom right corner. Since the image pairs correspond to translation along the camera focal axis, the ground-truth flow field in cartesian looks diverging (see Fig. 4 (*top-right*)), vectors appear originating from a VP toward the boundary in all directions. Hence, the computed flow fields in both LP/RLP representations looks blue in color indicating all the vectors point downwards. In order to compare the flow fields of LP and RLP, these flow fields are mapped back to cartesian. Figure 5 depicts the flow fields of both LP and RLP mapped back to cartesian. These flow

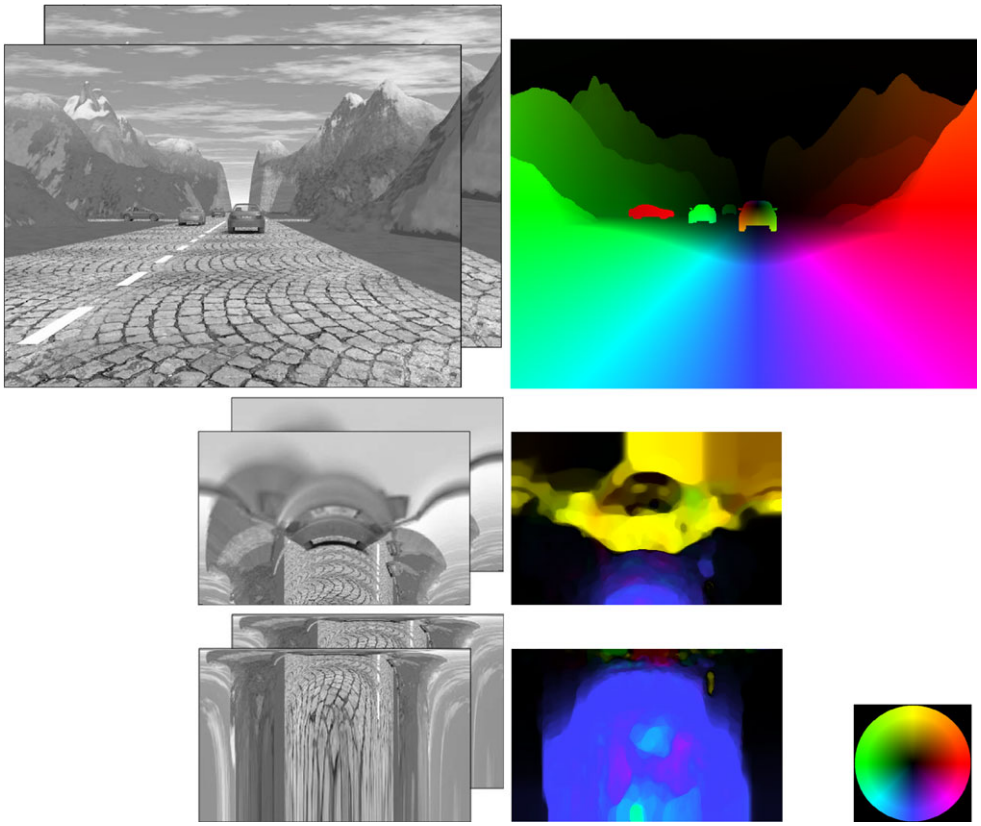

**Fig. 4** (*Top*) Pair of images and flow fields in Cartesian. (*Middle*) LP representations. (*Bottom*) RLP representations and (*bottom-right-corner*) colormap used to display flow fields

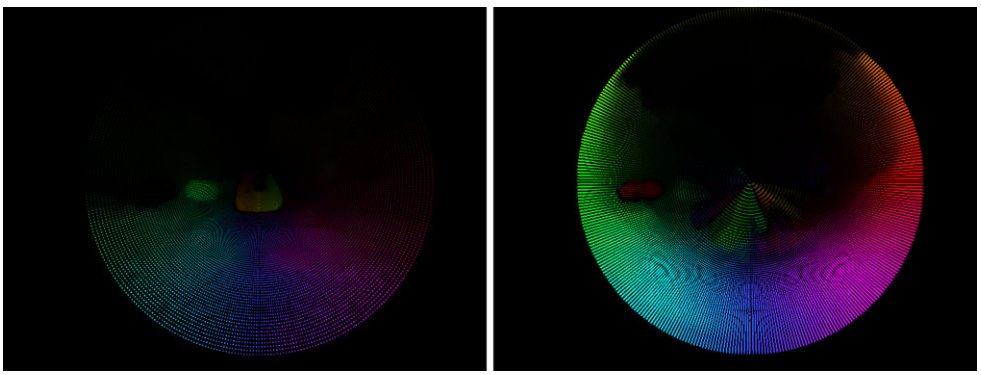

**Fig. 5** Inverse mapped flow fields from (*left*) LP and (*right*) RLP

fields in cartesian are sparse. Hereinafter, the LP and RLP representations of flow fields refer to these mapped back to cartesian.

The well known error measures to compare flow fields are Average Angular Error (AAE) and Average End-Point error (AEP) [2, 4]. The AAE is firstly chosen for the detailed study in the current work. According to [4], the angular error $e$ between two vectors $(u_1, v_1)$ and $(u_2, v_2)$ is given by:

$$e\big((u_1, v_1), (u_2, v_2)\big) = \arccos\left(\frac{u_1u_2 + v_1v_2 + 1}{\sqrt{(u_1^2 + v_1^2 + 1)(u_2^2 + v_2^2 + 1)}}\right). \quad (7)$$

Since the flow fields from LP and RLP representations are sparse and of different resolutions along radial direction, in order to do a fair comparison a common set of pixels (intersection mask) is selected. Figure 6 shows the masks (i.e., locations where the flow vectors exist) of LPR (*left*) and RLPR (*middle*) flow fields and the intersection mask (*right*) that is the common set of positions those have flow values in both representations. Then, AAE between LP and ground-truth flow fields, and between RLP and ground-truth flow fields are computed using the obtained intersection mask. Table 1 shows AAEs of flow fields from ten selected different image pairs from sequence-1 of set-2 of [31]. The images are of resolution $480 \times 640$. They are mapped to LP and RLP representations of resolution $230 \times 360$, placing the vanish-

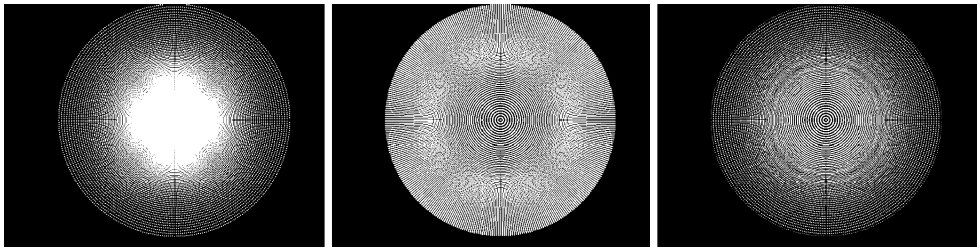

**Fig. 6** (*Left*) LP mask. (*Middle*) RLP mask. (*Right*) Mask from their intersection

**Table 1** AAEs (deg.) for selected ten flow fields from sequences [31] in LPR and RLPR

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Seq. 1 | LPR | 24.3 | 24.3 | 23.9 | 23.9 | 23.9 | 23.8 | 23.6 | 23.7 | 23.5 | 23.4 |
| | RLPR | 20.9 | 19.1 | 19.0 | 19.4 | 18.3 | 18.4 | 18.4 | 18.0 | 17.9 | 18.9 |
| Seq. 2 | LPR | 24.3 | 24.5 | 27.4 | 27.1 | 24.3 | 24.6 | 24.7 | 24.5 | 24.7 | 24.8 |
| | RLPR | 21.6 | 21.8 | 27.2 | 26.6 | 21.6 | 23.3 | 24.2 | 23.8 | 23.8 | 22.0 |

ing point at (230, 340), computed from [20]. Then, optical flow is computed on these images using [24]. The flow fields are mapped back to cartesian and then using the intersection mask shown in Fig. 6 (*right*), the AAEs are computed. The AAEs in Table 1 show that flow fields estimated in RLP representations are more accurate than flow fields from LPRs.

A similar experiment on sequence-2 of set-2 of [31] is performed; results are also presented in Table 1. Vanishing points in sequence-2 varies, but the most of the frames' VP lies in (240, 320); hence the resolution of the mapped images is fixed at $240 \times 360$. Also in the results of sequence-2, the RLPR flow fields are more accurate than LPR flow fields. The difference in AAEs between LP and RLP representations of sequence-2 is smaller than the results of sequence-1 because the displacement between consecutive images in sequence-2 is very high. These large displacements lead to more stretching in RLP represented images and hence increase in errors in flow fields.

In order to analyze the behavior of space-variant representations, we have done experiments for different resolutions in the two sequences presented above (sequence-1 and sequence-2). The experiments consist of testing different radial and angular resolutions. Since it is intuitive to keep the angular resolution equal or higher than 360, otherwise leads to higher loss of information, we have performed experiments for different angular resolutions such as 360, 720 and 1440. From this first test we conclude that increasing the angular resolution does not improve results, it only leads to higher processing. On the contrary, for radial resolution, in addition to the deep study that will be presented in Fig. 7, we tested two values for sequence-1. These values corresponds to the maximum inner circle (230) and minimum bounding circle (378) covering the whole rectangular cartesian image respectively from the VP. Similarly, for sequence-2 we considered 240 and 400 corresponding to inner circle and outer circle. In sequence-1, 230 is the distance from the VP to the nearest boundary. Note that in the case of using the outer circle, a lot of empty space in LP and RLP representations will be included, which is not desired. Values smaller than 230, but covering the same area, would yield to more sparse sampling in both LPR and RLP that could introduce more errors in the estimated optical flow fields. Therefore, as mentioned above, we have considered $230 \times 360$ (radial $\times$ angular) and $240 \times 360$ (radial $\times$ angular) as appropriate resolutions for sequence-1 and sequence-2 respectively. These resolutions are the ones used in all our experiments, which correspond to the regions contained in the largest inner circle centered on the VP.

Further experiments are done to analyze how the error evolves along the space in these representations. Different concentric circular regions of the flow field around the vanishing point, with an increase in the radius of the circle within the flow field boundary, are considered. At each sized circle, the AAE is calculated for the region inside the circle and also for the region outside the circle. This experiment is done for both LPR and RLPR on sequence-1. Since the radial axis for the flow fields in sequence-1 is of length 230, nine circles with increasing radius from 23 till 207, in steps of 23, are considered. Figure 7 (*top-left*) and (*middle-left*) show the AAEs in colormap for the region inside the circle at radius 115 for LPR and RLPR. Figure 7 (*top-right*) and (*middle-right*) show the AAEs in colormap for the region outside the circle at radius 115 for LPR and RLPR respectively. In Fig. 7 (*bottom-left*), solid line indicates AAEs (the average of 10 flow fields' region inside the circle) in LPR with the increase in radius. The AAE increases as the inner area increases. This proves that the flow field near the fovea is more accurate than in the periphery in LPR. The dashed line correspond to AAEs in RLPRs with the increase in radius. In the plot Fig. 7 (*bottom-left*) the AAE of RLPR

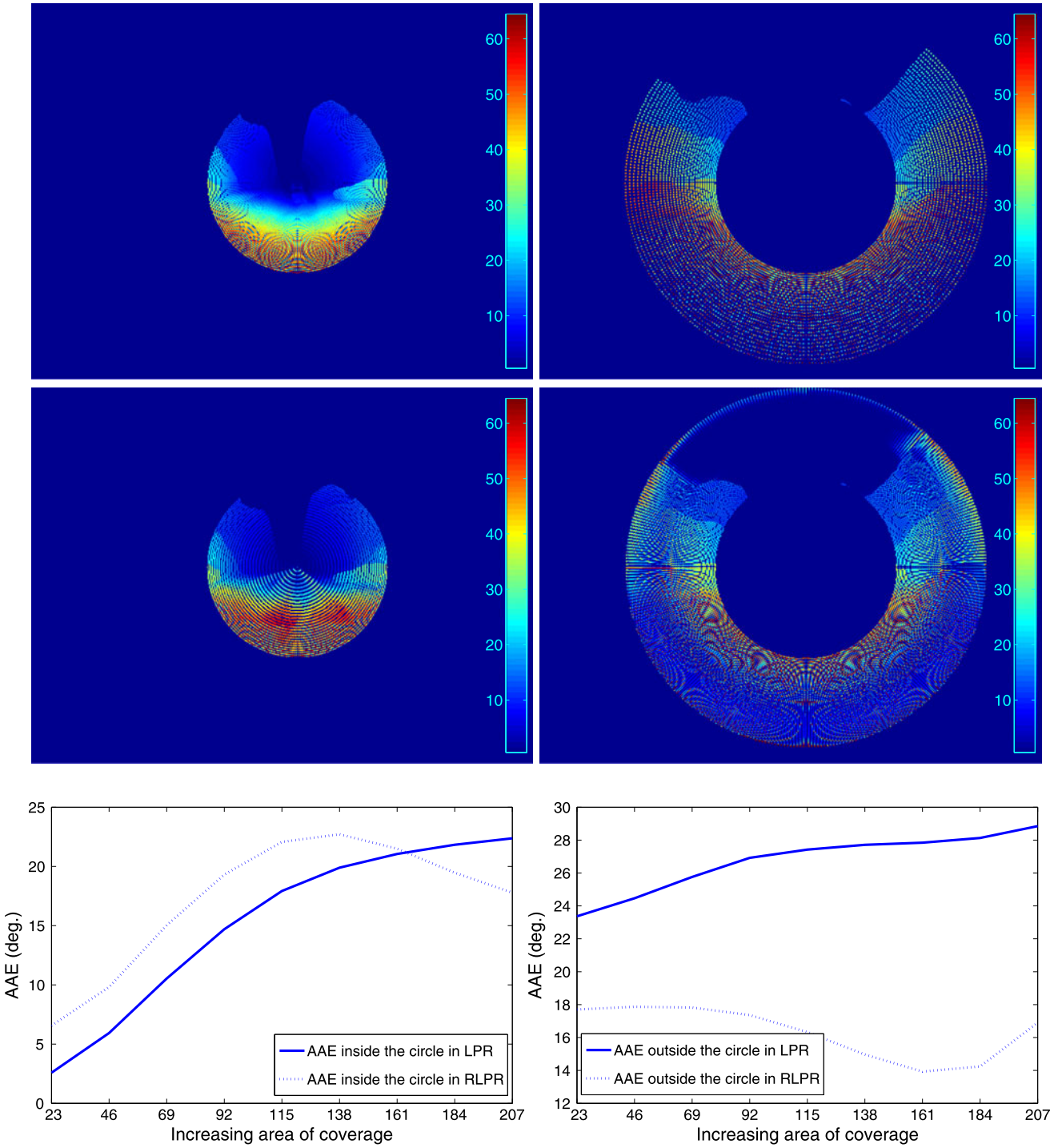


**Fig. 7** Analysis of AAEs over space in LPR and RLPR (values in colormap scale computed from Eq. (7). (*Left*) Region inside circle. (*Right*) Region outside circle

decreases from radius 138 till the boundary. At radius 207, where most of the image area is covered, the AAE of RLPR is lower than the AAE of LPR. This shows RLPR is better at periphery than LPR.

Figure 7 (*bottom-right*) shows the AAEs of LPR and RLPR, outside the circles, with the increase in radii of the circles. The outer area is getting reduced with the increase in radius of the circle. The solid line indicating AAE of LPR increases as the outer area decreases, whereas the dashed line indicating AAE of RLPR decreases as the outer area decreases till the circle with radius 161. Then, it increases due to some artifacts in the extreme periphery of RLPR flow

**Table 2** Mean AAE and mean AEP of synthetic sequences: [31] (Sequence-1 and Sequence-2 in set-2) and [21], and a real sequence [18]

| | | AAE | AEP |
|---|---|---|---|
| Sequence 1 | RLPR | 18.3145 | 1.4043 |
| | LPR | 23.9895 | 1.7151 |
| Sequence 2 | RLPR | 35.9757 | 6.1254 |
| | LPR | 47.41 | 6.4059 |
| Sequence 3 | RLPR | 26.1972 | 3.2783 |
| | LPR | 36.8119 | 3.9058 |
| Real Sequence | RLPR | 36.5783 | 2.1985 |
| | LPR | 63.7911 | 2.6398 |

field. Figure 7 (*middle-right*) shows the artifact, thin band of circular arc on the top, whereas this band is absent in the LPR (*top-right*) flow field. This plot (*bottom-right*) gives the same conclusion obtained from the plot in (*bottom-left*).

RLPR is better than LPR not only by AAE, Table 2 shows the mean AAE and mean AEP of all frames in sequence-1 and sequence-2 of set-2 [31] named as Sequence 1 and Sequence 2 in the table. Sequence-1 contains 99 flow-fields and sequence-2 contains 394 flow fields. These two sequences resembles a countryside scenario with very good road texture. Table 2 also provides results of similar experiments on another synthetic sequence and a real sequence. Sequence 3 in the table is another synthetic sequence of ten image frames [21] of an urban scenario with asphalt road texture, which does not contain as much texture as previous two sequences. The error values for this sequence also show that still RLPR is better in this case. Finally, a real sequence is also used to validate the usefulness of the proposed space variant representation. Figure 8 shows a pair of images from the real sequence used here with an annotated flow field (*top-right*), and the corresponding LP (*bottom-left*) and RLP (*bottom-right*) representations. Results from this real sequence [18] are depicted in the last row of Table 2. This sequence contains 37 image frames of an urban driving scenario. This dataset has annotated ground-truth optical flow. The results for this sequence also reveal the goodness of RLPR. From the experiments on these four sequences, it can be confirmed that RLPR is better than LPR in both AAE and AEP on both synthetic and real sequences.

From the above experiments, it is evident that RLPR gives more accurate optical flow than LPR. In order to show the advantages of the better accuracy in RLPR, both representations are further evaluated by computing camera egomotion [25] on them. The egomotion is the estimation of the motion (rotation and translation) of a camera from the image sequence acquired by it. This is one of the fundamental tasks in many advanced driver assistance and robotic applications. An extensive review of egomotion estimation techniques, from the ADAS field, can be found in [8]. Here we estimate three orientation parameters (*pitch*, *yaw* and *roll*) of the camera egomotion using the optical flow of the space variant representations. This experiment is performed on sequence-2 of set-2 that has different motions of camera. We use the 5-point algorithm [17] to estimate these orientation parameters. For inputting the point sets to 5-point algorithm, we use a RANSAC [14] based approach that randomly select $k$ sets of ten points and compute the rotation angles. In the current work $k$ has been set to 25 samples. Table 3 shows the mean (in radians) and standard-deviation of the errors in *pitch*, *yaw* and *roll* angles obtained using ground-truth, Subsampled-Cartesian, LPR and RLPR flow fields of all frames in the sequence. The mean and standard-deviation

**Fig. 8** (*Top-left*) Pair of images of a real sequence [18]. (*Top-right*) Annotated ground-truth flow field between the pair of consecutive frames in the real sequence. (*Bottom-left*) LP representations and (*Bottom-right*) RLP representations

**Table 3** Mean and standard-deviation of errors in *pitch*, *yaw* and *roll* in radians of sequence-2 in set-2 [31]

| | Pitch | | Yaw | | Roll | |
|---|---|---|---|---|---|---|
| | Mean | Std | Mean | Std | Mean | Std |
| GTOF | 4.08E-07 | 4.34E-07 | 5.06E-07 | 5.90E-07 | 1.15E-04 | 1.66E-04 |
| Sub. cart. | 2.39E-03 | 5.44E-03 | 1.43E-03 | 4.59E-03 | 2.63E-03 | 3.71E-03 |
| LPR | 1.67E-03 | 2.62E-03 | 9.16E-04 | 3.10E-03 | 5.89E-04 | 8.83E-04 |
| RLPR | 1.33E-03 | 1.63E-03 | 7.08E-04 | 1.49E-03 | 6.77E-04 | 8.54E-04 |

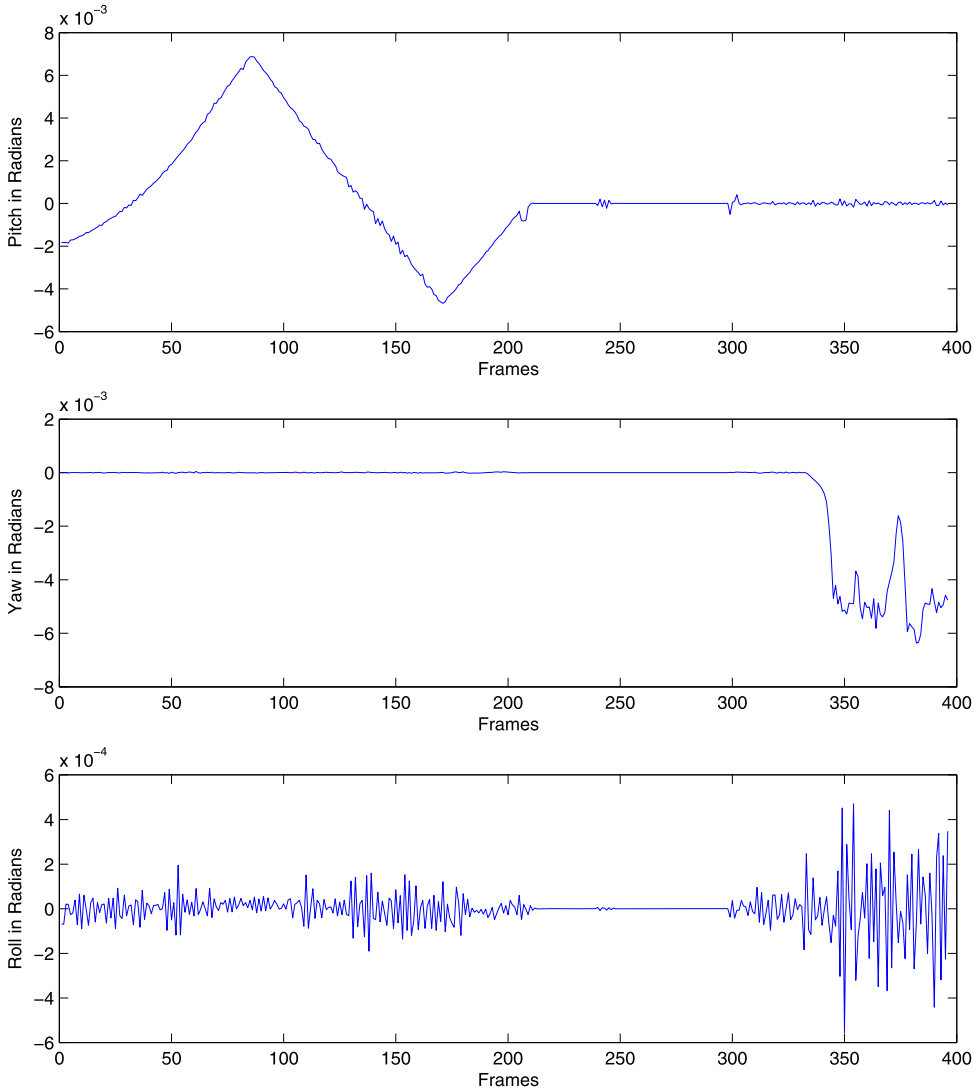


**Fig. 9** Plots of ground-truth *pitch*, *yaw* and *roll* angles through the whole sequence

of errors in *pitch*, *yaw* and *roll* angles by ground-truth optical flow are smaller than those computed by using optical flow estimated in Subsampled-Cartesian, LPR and RLPR. Both space-variant representations (LP and RLP) use only a quarter of the information used in the original Cartesian representation. Hence, in order to have a fair comparison, we subsampled the original Cartesian images to the resolution of the space-variant representations (LPR and RLPR), i.e., to $240 \times 360$; this is referred hereinafter to as Subsampled-Cartesian representation. The orientation parameters are estimated using the optical flow computed on these representations. The results in Table 3 shows that LPR and RLPR are more accurate than Subsampled-Cartesian results. The mean and standard-deviation of *pitch* and *yaw* angles' errors are lower in RLPR than those in LPR. Note that *roll* angle do not change so much through the whole sequence (Fig. 9 depicts ground-truth of *pitch*, *yaw* and *roll* angles in radians). Hence, the mean error of *roll* estimation, though

**Table 4** Mean and standard-deviation of errors in *rotation matrices* [25] in radians of sequence-2 in set-2 [31]

| | Rotation matrix | |
|---|---|---|
| | Mean | Std |
| GTOF | 2.46E-04 | 1.41E-04 |
| Sub. cart. | 4.77E-03 | 7.60E-03 |
| LPR | 2.62E-03 | 4.04E-03 |
| RLPR | 2.15E-03 | 1.93E-03 |

lower in LPR than those in RLPR, is not as much significant as in the case of *pitch* and *yaw* angles. In Table 3, it can be observed that roll angle from GTOF is in the same order (E-04) as those from LPR and RLPR, whereas *pitch* and *yaw* angles from GTOF are far more accurate as compared to those from LPR and RLPR. Regarding *roll* angle variations, up to our understanding, visually there is no variation all over the original sequence. Hence, it seems to be an error introduced in the process of generating ground-truth values provided by [31]. Estimation of motion parameters can be ambiguous (e.g., translation on X axis can be incorrectly estimated as rotation on Y axis and vice-versa) under a reduced field of view or with insufficient depth variation (e.g., [36, 37]) because different motions can induce similar flow vectors. Therefore, another way to evaluate these three orientations is just by comparing the rotation matrices directly instead of comparing individual orientations [25]. Table 4 provides the error values with rotation matrices. These values also reproduce the same conclusion that RLPR is better than LPR and Subsampled-Caretesian. Overall it can be concluded that orientations estimated using optical flow from RLPR is more accurate than those estimated from LPR. Additionally, results are better than those obtained by the Subsampled-Cartesian representation that contains the same number of data points.

## 6 Conclusion

The current work proposes to change image sampling, resulting in a novel space variant representation. This validates the initial intuition that more accurate results can be obtained if we change the philosophy of oversapling fovea region to the higher sampling in peripheral region. This improvement is useful when forward faced on-board vision systems are considered, where translation in the optical axis is the predominant motion (e.g., mobile robotics, driver assistance). The proposed RLP representation is evaluated and compared with classical LP by computing optical flow on them. The rectangularized representations of space-variant sampling may introduce systematic errors in flow fields when traditional optical flow approaches are used on them. Hence we can say that there is a need of research in variational optical flow methods to be applied directly on the space-variant sampled images. The experimental analysis is performed on both synthetic and real video sequences. Additionally, results from the estimation of egomotion are provided as an illustrative application. Experimental results shows that RLPR is better than LPR in navigational applications and is also better than the corresponding Subsampled-Cartesian representation.

**Acknowledgements** 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. Authors gratefully acknowledge discussions with Henning Zimmer and Prof. Joachim Weickert. The authors would like to thank Oisin Mac Aodha for providing the Python code for raytracing with Maya, and Diego Cheda for the help in egomotion estimation.

## References

1. Adato, Y., Zickler, T., Ben-Shahar, O.: A polar representation of motion and implications for optical flow. In: IEEE Int. Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, pp. 1145–1152 (2011)
2. Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M., Szeliski, R.: A database and evaluation methodology for optical flow. In: IEEE Int. Conference on Computer Vision, Rio de Janeiro, Brazil, pp. 1–8 (2007)
3. Barnes, N., Sandini, G.: Direction control for an active docking behaviour based on the rotational component of log-polar optic flow. In: European Conference on Computer Vision, Dublin, Ireland, pp. 167–181 (2000)
4. Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. Int. J. Comput. Vis. **12**(1), 43–77 (1994)
5. Bolduc, M., Levine, M.D.: A review of biologically motivated space-variant data reduction models for robotic vision. Comput. Vis. Image Underst. **69**(2), 170–184 (1998)
6. Boluda, J.A., Domingo, J., Pardo, F., Pelechano, J.: Detecting motion independent of the camera movement through a log-polar differential approach. In: Computer Analysis of Images and Patterns, Kiel, Germany, pp. 702–709 (1997)
7. Bruhn, A.: Variational optic flow computation: accurate modelling and efficient numerics. Ph.D. thesis, Department of Mathematics and Computer Science, Saarland University, Saarbrcken (2006)
8. Cheda, D., Ponsa, D., Lopez, A.: Camera egomotion estimation in the ADAS context. In: IEEE Int. Conference on Intelligent Transportation Systems, Madeira Island, Portugal, pp. 1415–1420 (2010)
9. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-Interscience, New York (1991)
10. Daniilidis, K.: Computation of 3-d-motion parameters using the log-polar transform. In: Computer Analysis of Images and Patterns, Prague, Czech Republic, pp. 82–89 (1995)
11. Daniilidis, K., Krüger, V.: Optical flow computation in the log-polar-plane. In: Computer Analysis of Images and Patterns, Prague, Czech Republic, pp. 65–72 (1995)
12. Dias, J., Araújo, H., Paredes, C., Batista, J.: Optical normal flow estimation on logpolar images. A solution for real-time binocular vision. Real-Time Imaging **3**(3), 213–228 (1997)
13. Fischl, B., Cohen, M.A., Schwartz, E.L.: Rapid anisotropic diffusion using space-variant vision. Int. J. Comput. Vis. **28**, 199–212 (1998)

14. Fischler, M., Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Graph. Image Process. **24**(6), 381–395 (1981)
15. Guiasu, S.: Information Theory with Applications. McGraw-Hill Inc., New York (1977)
16. Horn, B.K.P., Schunk, B.G.: Determining optical flow. Artif. Intell. **17**, 185–203 (1981)
17. Kukelova, Z., Bujnak, M., Pajdla, T.: Polynomial eigenvalue solutions to the 5-pt and 6-pt relative pose problems. In: British Machine Vision Conference, Leeds, UK (2008)
18. Liu, C., Freeman, W.T., Adelson, E.H., Weiss, Y.: Human-assisted motion annotation. In: IEEE International Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska, USA, pp. 1–8 (2008)
19. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision (darpa). In: DARPA Image Understanding Workshop, pp. 121–130 (1981)
20. Onkarappa, N., Sappa, A.D.: On-board monocular vision system pose estimation through a dense optical flow. In: Int. Conference on Image Analysis and Recognition, Póvoa de Varzim, Portugal, pp. 230–239 (2010)
21. Onkarappa, N., Sappa, A.D.: An empirical study on optical flow accuracy depending on vehicle speed. In: IEEE Intelligent Vehicles Symposium, Alcalá de Henares, Spain, pp. 1138–1143 (2012)
22. Roth, S., Black, M.J.: On the spatial statistics of optical flow. In: IEEE Int. Conference on Computer Vision, Beijing, China, pp. 42–49 (2005)
23. Schwartz, E.L., Greve, D.N., Bonmassar, G.: Space-variant active vision: definition, overview and examples. Neural Netw. **8**(7–8), 1297–1308 (1995)
24. Sun, D., Roth, S., Black, M.J.: Secrets of optical flow estimation and their principles. In: IEEE Int. Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, pp. 2432–2439 (2010)
25. Tian, T.Y., Tomasi, C., Heeger, D.J.: Comparison of approaches to egomotion computation. In: IEEE Int. Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, pp. 315–320 (1996)
26. Tistarelli, M., Sandini, G.: On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. IEEE Trans. Pattern Anal. Mach. Intell. **15**(4), 401–410 (1993)
27. Traver, V.J., Bernardino, A.: A review of log-polar imaging for visual perception in robotics. Robot. Auton. Syst. **58**(4), 378–398 (2010)
28. Traver, V.J., Pla, F.: The log-polar image representation in pattern recognition tasks. In: Iberian Conference on Pattern Recognition and Image Analysis, pp. 1032–1040 (2003)
29. Traver, V.J., Pla, F.: Motion analysis with the radon transform on log-polar images. J. Math. Imaging Vis. **30**(2), 147–165 (2008)
30. Tunley, H., Young, D.: First order optic flow from log-polar sampled images. In: European Conference on Computer Vision, Stockholm, Sweden, vol. 1, pp. 132–137 (1994)
31. Vaudrey, T., Rabe, C., Klette, R., Milburn, J.: Differences between stereo and motion behaviour on synthetic and real-world stereo sequences. In: Proc. Image and Vision Computing New Zealand, Christchurch, New Zealand, pp. 1–6 (2008)
32. Wallace, R.S., Ong, P.W., Bederson, B.B., Schwartz, E.L.: Space variant image processing. Int. J. Comput. Vis. **13**, 71–90 (1994)
33. Yeasin, M.: Optical flow in log-mapped image plane-a new approach. IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), 125–131 (2002)
34. Yeung, A., Barnes, N.: Efficient active monocular fixation using the log-polar sensor. Int. J. Intell. Syst. Technol. Appl. **1**, 157–173 (2005)
35. Zach, C., Pock, T., Bischof, H.: A duality based approach for realtime TV-$L^1$ optical flow. In: Annual Symposium of the German Association for Pattern Recognition, Heidelberg, Germany, pp. 214–223 (2007)
36. Adiv, G.: Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field. IEEE Trans. Pattern Anal. Mach. Intell. **11**(5), 477–489 (1989)
37. Fermuller, C., Aloimonos, Y.: Ambiguity in structure from motion: sphere versus plane. Int. J. Comput. Vis. **28**(2), 137–154 (1998)


**Naveen Onkarappa** received B.Sc. degree in Computer Science from Kuvempu University, India, and M.Sc. degree in computer science from University of Mysore, India, in 1999 and 2001 respectively. From 2001 to 2005 he was a guest lecturer at Kuvempu University and University of Mysore, India. In 2007 he received Master of Science Technology (by research) in computer science and technology from University of Mysore, India. From 2007 to 2009 he was with HCL Technologies, Bangalore, India. Since 2009 he is a Ph.D. student in the Computer Vision Centre, Autonomous University of Barcelona, Barcelona, Spain, under the supervision of Dr. Angel D. Sappa. He is a member of the Advanced Driver Assistance Systems Group. His research work is focused on optical flow estimation and its applications to driver assistance systems.

**Angel D. Sappa** received the electromechanical engineering degree from 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 UK, and Greece, he joined the Computer Vision Center, where he is currently a senior researcher. He is a member of the Advanced Driver Assistance Systems Group. His research interests span a broad spectrum within the 2D and 3D image processing. His current research focuses on stereo image processing and analysis, 3D modeling, and dense optical flow estimation.