Hierarchical Properties of Multi-resolution Optical Flow Computation

Yusuke Kameda, Atsushi Imiya, and Tomoya Sakai

Graduate School of Advanced Integration Science, Chiba University
Institute of Media and Information Technology, Chiba University
1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan
Graduate School of Engineering, Nagasaki University
1-14, Bunkyo-cho, Nagasaki, 852-8521 Japan
yu-kameda@graduate.chiba-u.jp, imiya@faculty.chiba-u.jp,
tsakai@cis.nagasaki-u.ac.jp

Abstract. Most of the methods to compute optical flows are variational-technique-based methods, which assume that image functions have spatiotemporal continuities and appearance motions are small. In the viewpoint of the discrete errors of spatial- and time-differentials, the appropriate resolution for optical flow depends on both the resolution and the frame rate of images since there is a problem with the accuracy of the discrete approximations of derivatives. Therefore, for low frame-rate images, the appropriate resolution for optical flow should be lower than the resolution of the images. However, many traditional methods estimate optical flow with the same resolution as the images. Therefore, if the resolution of images is too high, down-sampling the images is effective for the variational-technique-based methods. In this paper, we analyze the appropriate resolutions for optical flows estimated by variational optical-flow computations from the viewpoint of the error analysis of optical flows. To analyze the appropriate resolutions, we use hierarchical structures constructed from the multi-resolutions of images. Numerical results show that decreasing image resolutions is effective for computing optical flows by variational optical-flow computations in low frame-rate sequences.

1 Introduction

In this paper, we analyze the appropriate resolutions for optical flows estimated by variational optical-flow computations from the viewpoint of the error analysis of optical flows. To analyze the appropriate resolutions, we use hierarchical structures constructed from the multi-resolutions of images. In the error analysis, we measure the average spatiotemporal angle errors (ASAE) between the ground truths of optical flows in each hierarchy of resolutions and estimated optical flows from the corresponding hierarchy of resolutions. Since the spatiotemporal angle error normalizes the norms of flow vectors, the effects of the decreases of the norms by making lower resolutions. Candidates for computation methods

A. Fusiello et al. (Eds.): ECCV 2012 Ws/Demos, Part II, LNCS 7584, pp. 576–585, 2012.

are [1, 2] that both assume that optical flows are small and approximate space derivatives from the previous and next frames.

Optical flows are appearance motions computed from image sequences and contain some feature-quantities such as self-motions of cameras and motions of objects. By computing optical flows from the image sequences taken by car-mounted cameras, we can obtain information about self-motions and obstacles [3, 4]. Furthermore, there are some researches about analyses and classifications of the motion properties of objects by using optical flows [5]. Thus, optical flows are widely applied to motion analyses such as motion detections and recognitions.

Optical flows are estimated under an assumption that the brightness of corresponding points on image sequences does not change temporally. However, optical flows are underspecified only by this assumption since it is an ill-posed problem. To avoid this problem, many researchers have been proposing block-matching based methods and variational-technique based methods [6] since 1980s. These methods adds another assumption that optical flows are locally smoothed.

One of the advantages of variational-technique based methods is that the partial differential equations (PDE) to be solved are analytically derived by defining an energy-minimization problem from some assumptions about optical flows. Thus many researchers have been proposing various assumptions since 1980s [6–9]. Most of these assumptions are expressed by PDE for images and optical flows. Therefore, images and optical flows should have spatiotemporal continuities naturally. Moreover, optical flows in itself are expressed by spatiotemporal derivatives and therefore it is assumed that the flow vectors are very small.

Because the images on computers are discrete, optical flows are also discrete. Then its physical unit is denoted as pixel/frame from the pixel of images and frame of sequences. In the discretization of the PDE above, space- and time-derivatives are approximated by using step sizes of pixels and frames respectively. The frames of optical flows should correspond to the intermediate frames of two successive frames in an image sequence. To synchronize the frames of images and optical flows, the space derivatives of images should be approximated from the previous and next frames of each intermediate frame.

Although image resolutions become higher and higher by the performance upgrades of cameras, frame rates of image sequences do not upgrade so much. This induces the image sequences with high resolutions and low frame-rates in the viewpoint of approximations of spatiotemporal derivatives. The optical flows computed from such image sequences have larger appearance velocities in pixel unit than those computed from images with lower resolutions in the same scenes. Therefore, in this case the assumption that optical flows are small is not satisfied. Furthermore, there is a problem about the accuracy of the approximations of space derivatives since the approximations from the previous and next frames assumes implicitly that images does not change so much around the intermediate frames. Thus in the case of the images that have higher resolutions and lower frame-rates, the optical-flow computations by using PDE have less accuracy. Therefore, from the viewpoint of spatiotemporal derivatives, we should compute optical flows with appropriate resolutions against its frame rates.

2 Mathematical Preliminary

2.1 Optical Flow

For a time dependent image $f(\boldsymbol{x}, t)$ defined in $\mathbb{R}^n \times \mathbb{R}_+$, where n is the dimension of images, the total differentiation for time t is give as

$$\frac{d}{dt} f = \nabla f^\top \boldsymbol{v}(\boldsymbol{x}) + \partial_t f \tag{1}$$

where $\nabla = (\partial_x, \partial_y)^\top$ and $\boldsymbol{v}(\boldsymbol{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^3$ is an image velocity or optical flow function on the image, and in $n = 2$, $\boldsymbol{v}(\boldsymbol{x}) = (u(\boldsymbol{x}), v(\boldsymbol{x}))^\top = (\dot{x}, \dot{y}) = \frac{d\boldsymbol{x}}{dt}$. Optical flow consistency $\frac{d}{dt} f = 0$ implies that the flow vector $\boldsymbol{v}$ of the point $\boldsymbol{x}$ is the solution of the singular equation,

$$F(f, \boldsymbol{v}) = \nabla f^\top \boldsymbol{v} + \partial_t f = 0. \tag{2}$$

To solve the eq. (2), the additional constraints are required. We generalize the data term such as the square of eq. (2) to $E_1(\boldsymbol{x}, t, \boldsymbol{v}, f)$.

Equation (2) is an ill-posed problem and therefore $\boldsymbol{v}$ is underspecified. Thus $\boldsymbol{v}$ is solved by using the energy minimization problem

$$\min \iint_\Omega E_1(\boldsymbol{x}, t, \boldsymbol{v}, f) + \alpha E_2(\boldsymbol{x}, t, \boldsymbol{v}) d\boldsymbol{x} \tag{3}$$

in variational method [6, 10]. Here, E_1 is a data term and the positive constant α is the weight coefficient of a prior term E_2, which is a convex function [1] for $\boldsymbol{v}$ and regularizes $\boldsymbol{v}$ under some assumption. Some examples of E_2 are the smoothness regularizer [6]

$$(\partial_x u)^2 + (\partial_y u)^2 + (\partial_x v)^2 + (\partial_y v)^2, \tag{5}$$

total variation regularizer $|\nabla u| + |\nabla v|$, and deformable-model regularizer $u_{xx}^2 + 2u_{xy}^2 + u_{yy}^2 + v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2$, respectively. In the method [1], $E_1 = (F(f, \boldsymbol{v}))^2$ and E_2 is eq. (5) and then its energy functional is

$$\min \iint_\Omega (F(f, \boldsymbol{v}))^2 + \alpha \left((\partial_x u)^2 + (\partial_y u)^2 + (\partial_x v)^2 + (\partial_y v)^2 \right) d\boldsymbol{x}. \tag{6}$$

In the conventional methods expressed as eq. (3), α is a given parameter and its appropriate value depends on images. Although, in general, the appropriate value depends on $\boldsymbol{x}$ and t, α is a constant value in eq. (3).

The boundary condition of the optical-flow computation is the free boundary condition

$$\forall \boldsymbol{x} \in \partial\Omega, \forall t, \forall \boldsymbol{n} \quad \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{n}}(\boldsymbol{x}, t) = \boldsymbol{0} \tag{7}$$

[1] A function g that satisfies

$$\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2, \forall k \in [0, 1] \quad g(k\boldsymbol{a} + (1-k)\boldsymbol{b}) \leq k g(\boldsymbol{a}) + (1-k) g(\boldsymbol{b}). \tag{4}$$

that is derived from variational method. Here, $\boldsymbol{n}$ is a unit normal vector for the image boundary $\partial\Omega$. The meaning of this condition is that optical flow continues smoothly beyond the image boundary.

2.2 Pyramid-Based Multiresolution

We define $g(x, y) = R^k f(x, y)$ of image $f(x, y, t)$ for an integer $k \geq 1$, such that

$$R^k f(x,y) = \iint_{\mathbb{R}^2} w_k(i) w_k(j) f(2^k x - i, 2^k y - j) didj, \tag{8}$$

for

$$w_k(i) = \begin{cases} \frac{1}{2}^k (1 - \frac{|i|}{2}^k), & |i| \leq 2^k \\ 0, & |i| > 2^k \end{cases}, \tag{9}$$

and then there is the relation

$$R^{k+1} f = R(R^k f). \tag{10}$$

The dual operation of R^k is defined as

$$E^k g(x,y) = 4^k \iint_{\mathbb{R}^2} w_k(i) w_k(j) g(\frac{x-i}{2^k}, \frac{y-j}{2^k}) didj. \tag{11}$$

3 Variational Problem for Adaptive Optimization

We describe about the summary of variational optical-flow computation using the Lagrange Multiplier Method [2]. This method assumes the higher frame-rate of image sequences. However, ordinary image sequences are not so. Therefore we use this method to analyze the effects of resolutions of images in variational optical-flow computation.

To generalize weight coefficient α to the function of $\boldsymbol{x}$ and t and determine it mathematically, they define $(F(f, \boldsymbol{v}))^2$ as a constraint equation and use the energy-minimization problem with the constraint, which is

$$\min_{\boldsymbol{v}(\boldsymbol{x},t)} \int_\Omega E_2(\boldsymbol{x}, t, \boldsymbol{v}) d\boldsymbol{x} \quad \text{subject to} \quad (F(f, \boldsymbol{v}))^2 = 0. \tag{12}$$

Converting this problem using the Lagrange Multiplier Method, they get the energy-minimization problem without constraint

$$\min_{\boldsymbol{v}(\boldsymbol{x})} \int_\Omega J(\boldsymbol{u}, f, \lambda) d\boldsymbol{x}, \quad \text{w.r.t.} \quad J(\boldsymbol{u}, f, \lambda) = \lambda(\boldsymbol{x}, t)(F(f, \boldsymbol{v}))^2 + E_2(\boldsymbol{x}, t, \boldsymbol{v}) \tag{13}$$

where $\lambda(\boldsymbol{x}, t) : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ is the Lagrange multiplier. In the method, E_2 is eq. (5) and then its Euler-Lagrange equations are

$$(F(f, \boldsymbol{v}))^2 = 0 \tag{14}$$

$$\lambda F(f, \boldsymbol{v}) \nabla f - \nabla^2 \boldsymbol{v} = \boldsymbol{0}, \tag{15}$$

where, $\nabla^2 = \partial_x^2 + \partial_y^2$ is Laplacian and $\nabla^2 \boldsymbol{v} = (\nabla^2 u, \nabla^2 v)^\top$.

Tab. 1. Experimental Methodology. We use two methods that assume both spatiotemporal continuities to compute optical flows by variational method. We analyze the relationship between the number of down-sampling operation R of images and the evaluation of estimated optical flows.

	weight	number of R	initial
Lagrange method [2]	$\lambda(\boldsymbol{x}, t)$	0,1,2,3,4,5	$\{\boldsymbol{0}\}$
competitive [1]	1 or 0.1, $^\forall\boldsymbol{x}$,$^\forall t$	0,1,2,3,4,5	results of previous frame

The numerical computation of the PDE above is defined as

$$\lambda^{(m+1)} = \lambda^{(m)} + \tau_1 \left(F(f, \boldsymbol{v}^{(m)}) \right)^2 \tag{16}$$

$$\boldsymbol{v}^{(m+1)} = \boldsymbol{v}^{(m)} - \tau_2 \left(\lambda^{(m+1)} F(f, \boldsymbol{v}^{(m+1)}) \nabla f - \nabla^2 \boldsymbol{v}^{(m)} \right), \tag{17}$$

where, $\lambda^{(m)}, \boldsymbol{v}^{(m)}$ is the variables at the iterative number $m \geq 0$ and $\tau_1 > 0, \tau_2 > 0$ are small constants.

4 Experiments

4.1 Evaluations

We analyze the relation of the errors and the resolutions of images, where we use all frames of EISATS set 2 sequence 1 and 2 [11]. Table 1 shows the experimental methodology of it.

Defining i, j as the coordinates of square grid points on discrete images and h as the step size between grid points, we denote $\boldsymbol{v}_{i,j}$ as the value $\boldsymbol{v}(h(i,j)^\top, t)$ at i, j and define

$$(i^*, j^*)^\top = \arg\max_{i,j} |\boldsymbol{v}_{i,j}^{(m+1)} - \boldsymbol{v}_{i,j}^{(m)}|. \tag{18}$$

The iteration ends when the convergence condition

$$\frac{|\boldsymbol{v}_{i^*,j^*}^{(m+1)} - \boldsymbol{v}_{i^*,j^*}^{(m)}|}{\boldsymbol{v}_{i^*,j^*}^{(m+1)}} < 10^{-5} \tag{19}$$

is satisfied or m reaches 10^3.

We choose the large step sizes $\tau_1 = 10^3, \tau_2 = \frac{1}{4}$ of the iteration for the Lagrange method [2]. On the other hand, we choose 10^3 and 10^4 for α, and $\frac{1}{4}$ for the step size corresponding to τ_2 for the method [1]. The initial optical flows of both methods are $\{\mathbf{0}\}$. After second frame for the method [1], we use the estimated optical flows of previous frames as initial flows of next frames.

The evaluation is based on spatiotemporal angle error [12, 13]

$$\arccos\left(\frac{\boldsymbol{w}_{i,j}^{\top}\boldsymbol{v}_{i,j}^{(m)}+1}{\sqrt{(|\boldsymbol{w}_{i,j}|^2+1)(|\boldsymbol{v}_{i,j}^{(m)}|^2+1)}}\right), \tag{20}$$

where $\boldsymbol{w}_{i,j}$ is the flow vector at i, j of the ground truth. Spatiotemporal angle error is the angle between $\frac{(\boldsymbol{w}^{\top},1)^{\top}}{|(\boldsymbol{w}^{\top},1)|}$ and $\frac{(\boldsymbol{v}^{(m)\top},1)^{\top}}{|(\boldsymbol{v}^{(m)\top},1)|}$, and the errors in a region of smooth non-zero motion are penalized less than errors in regions of zero motion. [12]. We compute the average values at each frame. However, we do not include the region in which the ground truth is undefined.

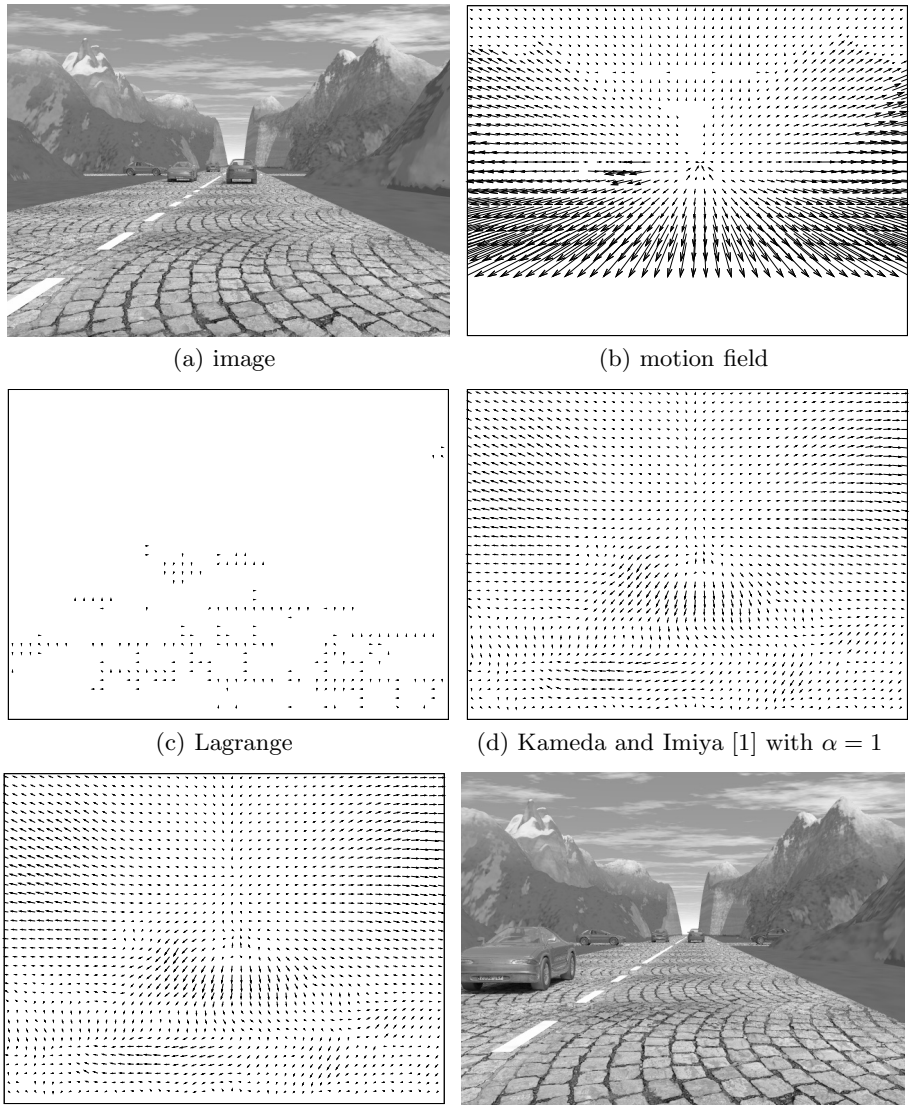
4.2 Results and Discussions

Figure 1 and 2 shows the computed optical flows and the evaluations of it, respectively. From the figure 1, the result of the Lagrange method is better than the competitive method to compute the spatial differences of appearance motions which are induced by the leading and oncoming vehicles. From the result of down-sampling, the errors decreased when the resolutions reduced. Similarly, in Fig. 3, the error analysis of EISATS set 2 sequence 2 shows the same effects.

The errors from the 40th frame to the 50th frame are larger than those of other frames. In these frames shown in Fig. 1 (f), there are much larger appearance motions since the oncoming vehicle falls out to the image boundary. It is hard to express such larger appearance motions even if the image resolution is decreased. Thus there are large errors comparatively.

Figure 4 shows the comparison of the Lagrange and competitive methods. The Lagrange method indicates the performance as much as that of the competitive method in low-resolutions. From the results, if the appearance motions are made small by decreasing the resolution, the Lagrange method satisfies the assumption.

From these resluts, $R4$ and $R5$ have lower errors than higher-resolutions, however, the stability for time-direction is also lower. Therefore, to find an appropriate resolution we have to analyze the stability of time-direction of the errors.

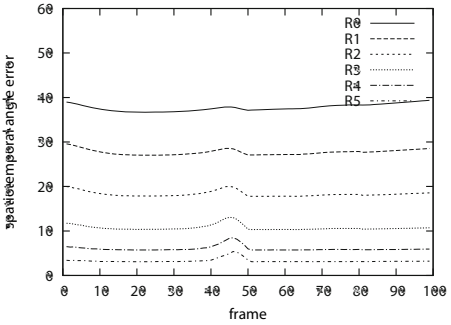


(a) image (b) motion field

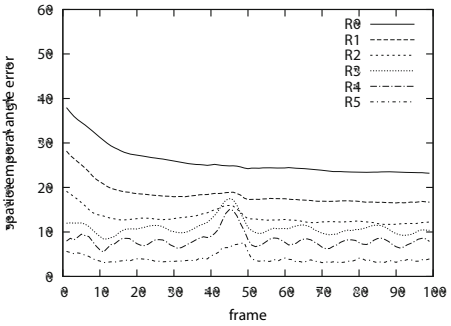
(c) Lagrange (d) Kameda and Imiya [1] with $\alpha = 1$

(e) Kameda and Imiya [1] with $\alpha = 0.1$ (f) The 45th frame

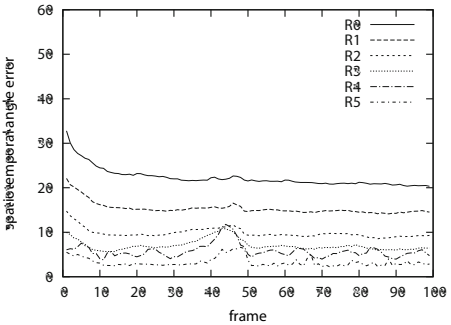
Fig. 1. The optical flows between the 24th and 25th frames of EISTAS set 2 sequence 1. The Lagrange method is better than the competitive method to compute the spatial differences of appearance motions which are induced by the leading and oncoming vehicles. In (f), since the oncoming vehicle falls out to the image boundary, the appearance motions are larger than the other frames.



(a) Lagrange

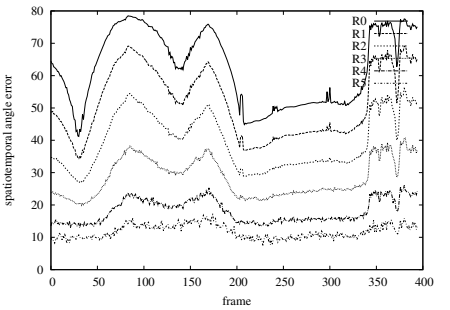


(b) Kameda and Imiya [1] with $\alpha = 1$

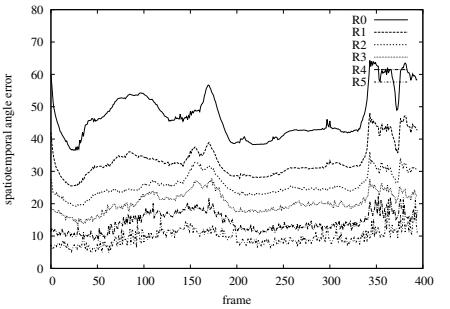


(c) Kameda and Imiya [1] with $\alpha = 0.1$

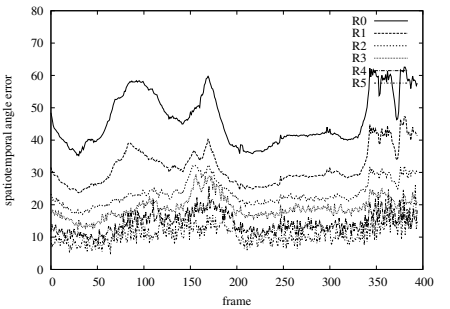
Fig. 2. The spatiotemporal-angle error of all frames of EISATS set 2 sequences 1. R* means the number of down-sampling operator R and the error functions of those are shown in the order corresponding to the numbers of R. Thus the errors decreased when the resolutions reduced.



(a) Lagrange

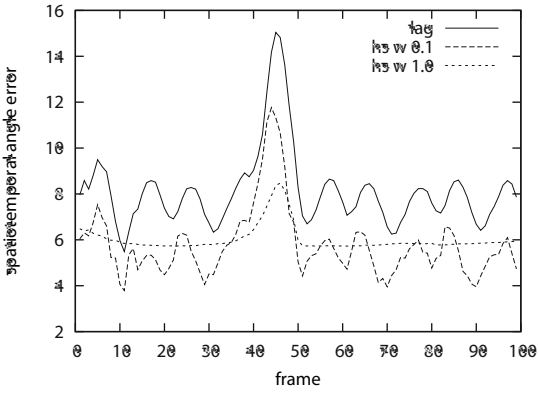


(b) Kameda and Imiya [1] with $\alpha = 1$

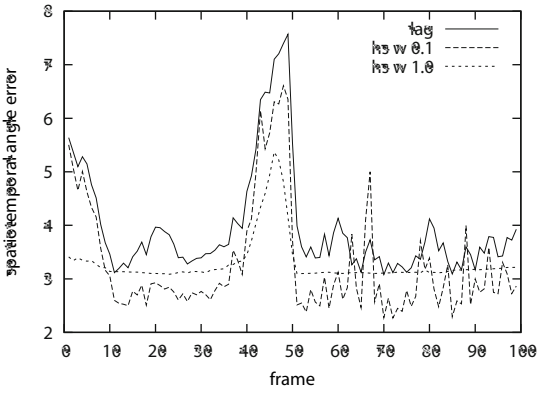


(c) Kameda and Imiya [1] with $\alpha = 0.1$

Fig. 3. The spatiotemporal-angle error of all frames of EISATS set 2 sequence 2. R* means the number of down-sampling operator R and the error functions of those are shown in the order corresponding to the numbers of R. Thus the errors decreased when the resolutions reduced.



(a) for R4, number of operator R is 4



(b) for R5, number of operator R is 5

Fig. 4. The comparison of the Lagrange and competitive methods in the cases of low resolutions of EISATS set 2 sequence 1.

5 Conclusion

In this paper, we analyzed the appropriate resolutions of optical flows estimated by variational optical-flow computations from the viewpoint of the error analysis of optical flows. To analyze appropriate resolutions, we used hierarchical structures constructed from the multi-resolutions of images. Numerical results show that decreasing image resolutions is effective to compute optical flows by variational optical-flow computations in low frame-rate sequences. From the results, if the appearance motions are made small by decreasing the resolution, the variational optical-flow computations satisfies the assumption of spatiotemporal continuities of images and optical flows.

References

1. Kameda, Y., Imiya, A.: Classification of Optical Flow by Constraints. In: Kropatsch, W.G., Kampel, M., Hanbury, A. (eds.) CAIP 2007. LNCS, vol. 4673, pp. 61–68. Springer, Heidelberg (2007)
2. Kameda, Y., Imiya, A., Sakai, T.: Adaptive estimation of Lagrange Multiplier functions of constraint terms for variational computation of optical flow. The Transactions of the Institute of Electronics, Information and Communication Engineers D J95-D, 1644–1653 (2012) (Japanese)
3. Ohnishi, N., Imiya, A.: Featureless robot navigation using optical flow. Connection Science 17, 23–46 (2005)
4. Klette, R., Kruger, N., Vaudrey, T., Pauwels, K., van Hulle, M., Morales, S., Kandil, F.I., Haeusler, R., Pugeault, N., Rabe, C., Lappe, M.: Performance of Correspondence Algorithms in Vision-Based Driver Assistance Using an Online Image Sequence Database. IEEE Transactions on Vehicular Technology 60, 2012–2026 (2011)
5. Jagadeesh, V., Karthikeyan, S., Manjunath, B.S.: Spatio-temporal optical flow statistics (STOFS) for activity classification. In: Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP 2010, pp. 178–182. ACM, New York (2010)
6. Horn, B.K., Schunck, B.G.: Determining optical flow. Artificial Intelligence 17, 185–203 (1981)
7. Nagel, H.H., Enkelmann, W.: An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. IEEE Transaction on Pattern Analysis and Machine Intelligence 8, 565–593 (1986)
8. Weickert, J., Schnörr, C.: A theoretical framework for convex regularizers in PDE-based computation of image motion. International Journal of Computer Vision 45, 245–264 (2001)
9. Zimmer, H., Bruhn, A., Weickert, J.: Optic Flow in Harmony. International Journal of Computer Vision 93, 368–388 (2011)
10. Pock, T., Unger, M., Cremers, D., Bischof, H.: Fast and exact solution of Total Variation models on the GPU. In: CVPR Workshop on Visual Computer Vision on GPU's, pp. 1–8 (2008)
11. Vaudrey, T., Rabe, C., Klette, R., Milburn, J.: Differences between stereo and motion behaviour on synthetic and real-world stereo sequences. In: 2008 23rd International Conference Image and Vision Computing, New Zealand, pp. 1–6. IEEE (2008)
12. Baker, S., Scharstien, D., Lewis, J.P., Roth, S., Black, M.J., Szeliski, R., Scharstein, D.: A Database and Evaluation Methodology for Optical Flow. International Journal of Computer Vision 92, 1–31 (2010)
13. Barron, J.L., Fleet, D.J., Beauchemin, S.S., Burkitt, T.A.: Performance of optical flow techniques. International Journal of Computer Vision 12, 43–77 (1994)