

Optical Flow Estimation of Large Displacements from Real Sequential Images

Jin-Woo Kim, *Member, KIMICS*

Abstract— In computing the optical flow, Horn and Schunck's method which is a representative algorithm is based on differentiation. But it is difficult to estimate the velocity for a large displacement by this algorithm. To cope with this problem multigrid method has been proposed. In this paper, we have proposed a scaled multigrid algorithm which the initial flow for a level is calculated by the summation of the optimally scaled flow and error flow. The optimally scaled flow is the scaled expanded flow of the previous level, which can generate an estimated second image having the least RMS error with respect to the original second image, and the error flow is the flow between the estimated second image (generated by the optimally scaled flow) and the original second image. The flow for this level is then estimated using the original first and second images and the initial flow for that level. From among the various coarsest starting levels of the multigrid algorithm, we select the one that finally gives the best estimated flow. Better results were achieved using our proposed method compared with Horn and Schunck's method and a conventional multigrid algorithm.

Index Terms— Optical flow, Multigrid method, RMSE, Horn and Schunck algorithm.

I. INTRODUCTION

THERE are many different methods for calculating an optical flow field. [1] discusses several of these methods, and suggests that most methods of calculating optical flow can be broken into three general stages of processing. First, the image is prefiltered to extract signals of interest and to reduce or smooth image noise. Second, basic image structures such as image derivatives are extracted from the image data.

The measurement of optical flow, which is a 2D motion or velocity field, is a fundamental problem in computer vision with a variety of applications. For instance, 3D motion and structure can be inferred from 2D velocity fields (or 2D displacement fields). Motion detection and object segmentation, computing the focus of expansion and time to collision, and computing stereo disparity are among other important uses of optical flow[2-6].

Two main approaches have been proposed for determining optical flow: intensity-based methods, and methods based on the matching of tokens, such as zero-crossings or other high-level features [7,8]. One of the subclasses of intensity-based algorithms is comprised of differential methods which exploit the relationship between velocity and spatio-temporal gradients in the image brightness [9-12]. Horn and Schunck's method [10] is a well-known algorithm of this type, and is the method we employ here to compute optical flow. In this method, the optical flow field is assumed to vary smoothly as a function of the image coordinate vector, and as a result, flow of images with large displacements may converge to give an incorrect solution. To address this problem, multiscale determination of optical flow has been considered by many researchers. In this approach, the essentially unsmoothed intensity variations at a fine level become smoothed at a coarser level, thus satisfying the smoothness constraint. Multigrid relaxation methods utilize a continuity of velocity through a hierarchy of resolution levels (P_0, P_1, \dots, P_M) [14]. The corresponding number of pixels is $h_0 < h_1 < \dots < h_M$. We assume that the grid points of each level are uniformly distributed and the ratio of the number of pixels of adjacent levels is a multiple of four, i. e., $h_{k+1}:h_k = 4:1$. We can classify these multiscale determination procedures into homogeneous methods and a non-homogeneous one. Homogeneous methods estimate the optical flow using a hierarchy of resolution grids and a solution process that is the same for all points in the image [13-15], while a non-homogeneous method that employs different weighting factors to expand the flow for a finer level has also been proposed [16]. Other relevant work can be found in [17-19].

II. PROPOSED ALGORITHM

Our algorithm consists of the following step: the estimated image is obtained by shifting the first image pixels according to the estimated flow;

STEP 1:

Build the image pyramid $(P_m, P_{m+1}, \dots, P_M)$.

STEP 2:

Start with the coarsest level $i = m$. The initial flow

Manuscript received May 17, 2011; revised May 23, 2011; accepted June 1, 2011.

Jinwoo Kim is with the Department of Information Communication Engineering, Kyungsung University, Busan, 607-736, Korea (Email: jinwoo@ks.ac.kr)

value for this level is considered to be equal to zero ($S_i^0 = 0$). Go to step 5.

STEP 3:

Find the best estimated second image which has the smallest RMS error with respect to the real second image by changing the scale ($1 \leq \alpha \leq 3.3$) of the flow s_i , and thus obtain the optimally scaled flow $\alpha^* s_i$ ($*$ indicates the optimal factor).

STEP 4:

Using the estimated second image mapped by $\alpha^* s_i$, compute the flow (s_e) between the estimated and read second images. Obtain the initial flow S_i^0 by adding s_e to $\alpha^* s_i$.

STEP 5:

Compute the flow (S_i) for level L_i by using the original first and second images and with initial flow S_i^0 . If $i=M$ then stop.

STEP 6:

Expand the flow (S_i) to (S_{i+1}) for the next fine level (P_{i+1}). In this expansion, flow S_i is copied only to the horizontally and vertically odd coordinates of s_{i+1} and other coordinates are interpolated. Increase i by one and go to step 3.

When we make a finer image we apply the best scaling factor giving an estimated second image with the smallest RMS error between that image and the real second image:

$$\text{RMS error} = \sqrt{\frac{\sum_x \sum_y (\rho_2(x,y) - \hat{\rho}_2(x,y))^2}{\sum_x \sum_y \rho_2^2(x,y)}} \quad (1)$$

where $\rho_2(x,y)$ and $\hat{\rho}_2(x,y)$ are the actual and estimated images.

In the above procedure, the preliminary processing consists of building an image pyramid with a 2:1 resolution for the first and the second images. To estimate the flow we used only two images without filtering them. The coarser versions of these images were obtained by local four-point averaging. Expansion and the generation of estimated images was done using bilinear interpolation. The methods procedure is described in following steps as illustrated in Fig. 1.

III. EXPERIMENTAL RESULTS

For comparative evaluation, we applied Horn and Schunck's method, a standard multigrid method, and our proposed algorithm to of real sequential images (see Fig. 2).

A. Hamburg Taxi sequence

Hamburg Taxi are shown in Fig. 2(a) whose sizes are both 128×128 . We make the No. 3 the base first image. Hamburg Taxi has motions of about four pixels left and

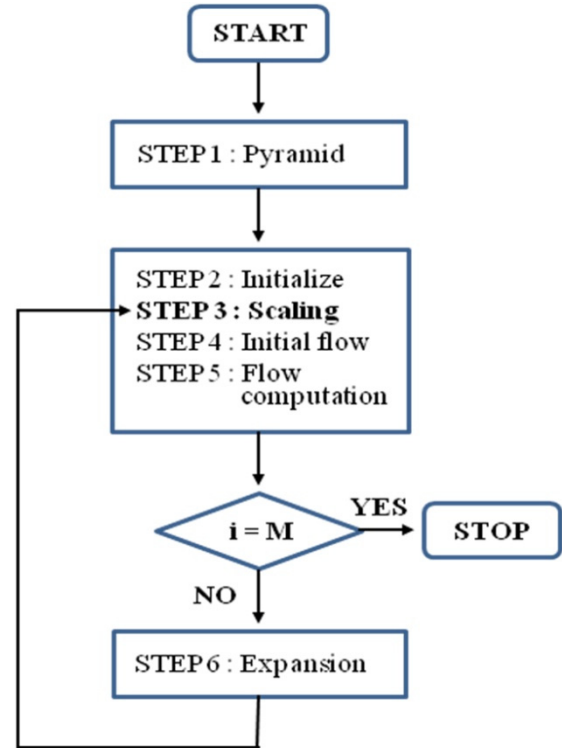


Fig. 1. Method's Procedure.



(a) Hamburg Taxi sequence



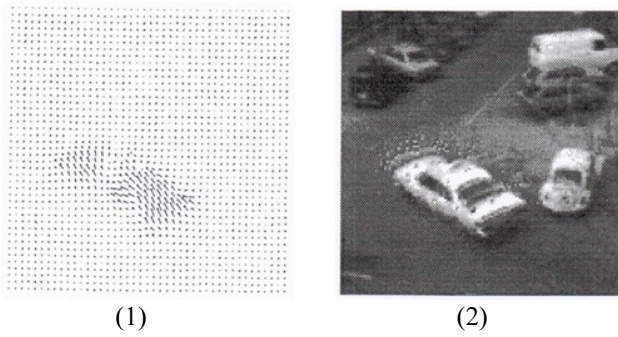
(b) Rubic cube sequence



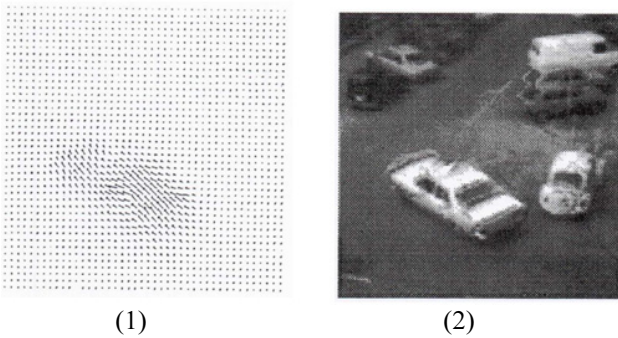
(c) SRI tree sequence

Fig. 2. Three frames of the experimented real image sequence.

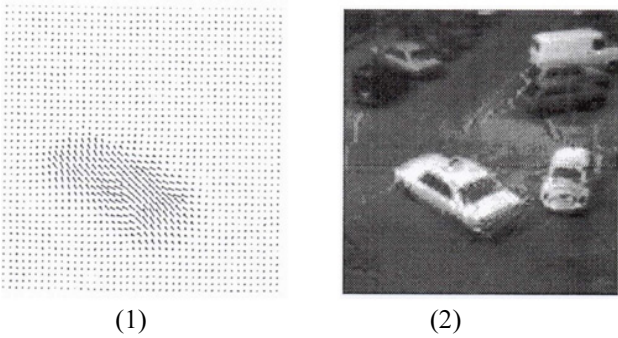
three pixels up at the center, and about 2.5 degrees of right-handed rotation. Fig. 3 shows the Estimated flow between frame numbers 3 and 6 ((1)'s), and estimated image No. 3 ((2)'s), using three optical flow.



(a) Horn & Schunck's method



(b) Multigrid method



(c) Proposed method

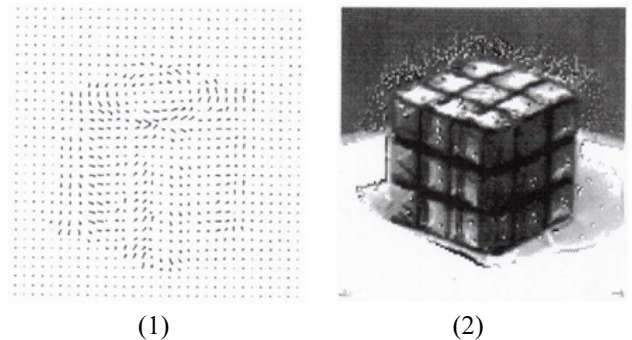
Fig. 3. Estimated flow between frame numbers 3 and 6 ((1)'s, and estimated image No. 3 ((2)'s), using three optical flow.

TABLE I
COMPARISON OF RMS ERROR'S AND CORRELATIONS BETWEEN ORIGINAL AND ESTIMATED SECOND IMAGES USING ESTIMATED FLOWS.

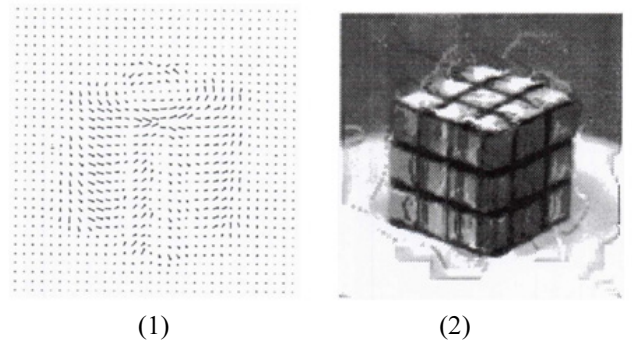
Comparison of the methods		Sample image Hamburg Taxi
Horn & Schunck method	RMS error	12.328
	Correlation	0.987
Multigrid method	RMS error	11.224
	Correlation	0.992
Proposed method	RMS error	9.591
	Correlation	0.995

B. Rubic cube sequence

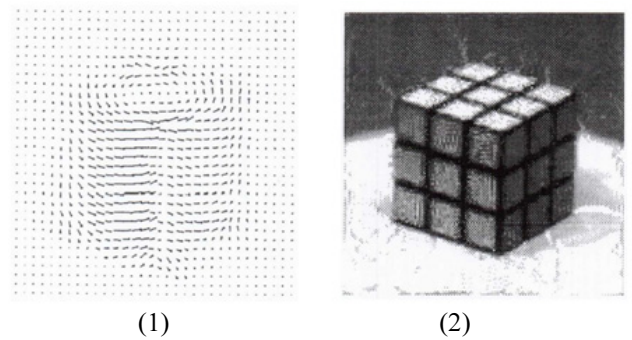
In this sequence (21 frames, 256×240) a Rubic's cube is rotating counter-clockwise on a turntable. The velocities induced by the rotating of the cube are less than 2 pixels / frame. We work with the center parts (128×128) of frames No. 1-18 with considering frame No. 1 as the first image and others as the second image. Fig. 4 shows the Estimated flow between frame numbers 1 and 18 ((1)'s, and estimated image No. 18 ((2)'s), using three optical flow.



(a) Horn & Schunck's method



(b) Multigrid method

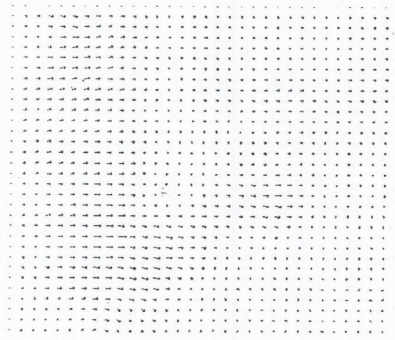


(c) Proposed method

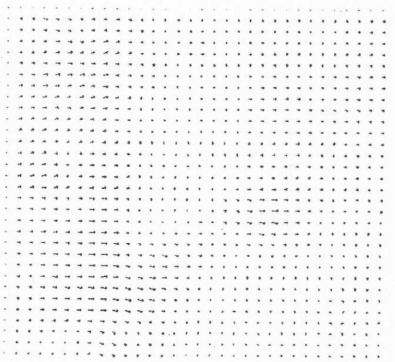
Fig. 4. Estimated flow between frame numbers 1 and 18 ((1)'s, and estimated image No. 18 ((2)'s), using three optical flow.

C. SRI Trees sequence

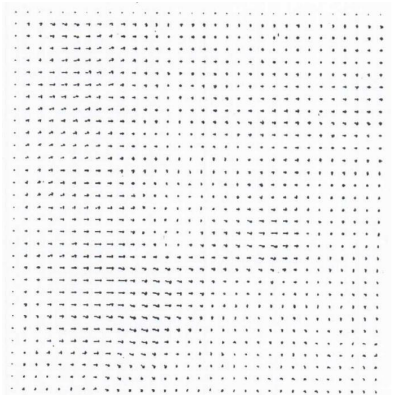
In this sequence (21 frames, 256×233) the camera moves from left to right, parallel to the ground plane and perpendicular to its line of sight. Velocities are as large as 2 pixels / frame. The frames (No. 1 to 10, 128×128) of this sequence are examined as the second image against the frame No. 0 as the first image. Fig. 5 shows the estimated flows between image frames No. 0 and No. 3 of SRI tree sequence using three optical flow.



(a) Horn & Schunck method



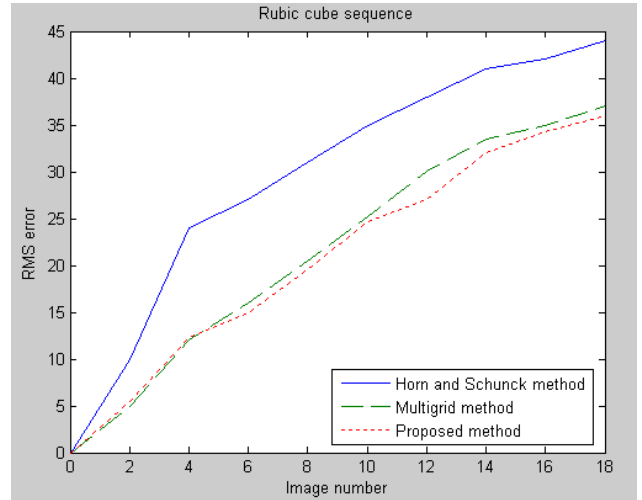
(b) Multigrid method



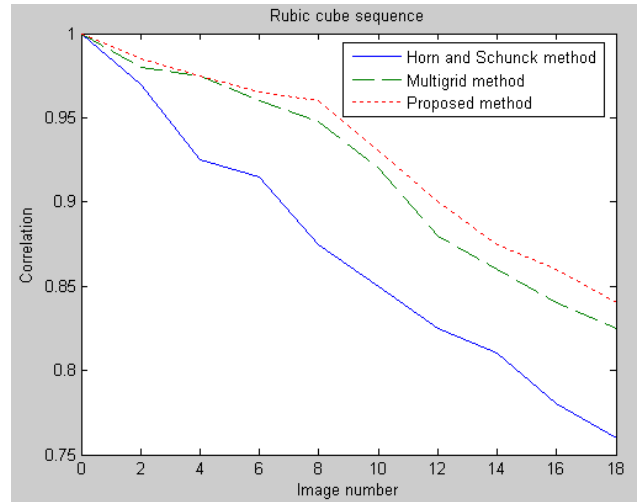
(c) Proposed method

Fig. 5. The estimated flows between image frames No. 0 and No. 3 of SRI tree sequence using three optical flow algorithms.

As mentioned before, the RMS error can be a good measure of optical flow error in real-image sequences with unknown correct flow. The similarity of two images



(a) RMS error



(b) Correlation

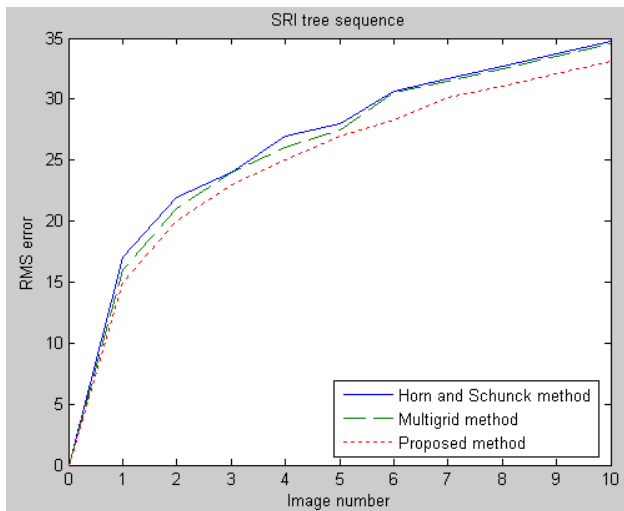
Fig. 6. Comparisons of the original and estimated second images for Rubic cube sequence using RMS error, and correlation.

can be expressed by their correlation, also. So we use these factors to evaluate the methods.

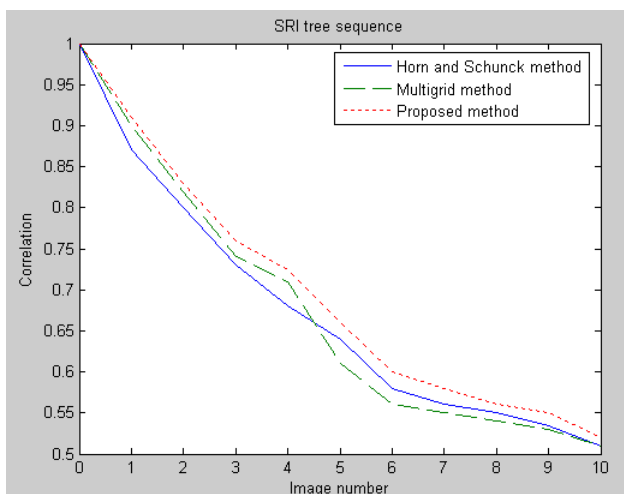
CORRELATION=

$$\frac{\sum_x \sum_y [\rho_2(x, y) - \bar{\rho}] [\hat{\rho}_2(x, y) - \bar{\hat{\rho}}_2]}{\sqrt{\sum_x \sum_y [\rho_2(x, y) - \bar{\rho}]^2} \sqrt{\sum_x \sum_y [\hat{\rho}_2(x, y) - \bar{\hat{\rho}}_2]^2}} \tag{2}$$

Where $\bar{\rho} = \sum_x \sum_y [\rho_2(x, y)] / M \times N$, and $\bar{\hat{\rho}}_2 = \sum_x \sum_y [\hat{\rho}_2(x, y)] / M \times N$ are the actual and estimated images of size $M \times N$. The comparison of three methods using RMS error and correlation between the original and the estimated second images for Hamburg Taxi, Rubic cube, and SRI tree sequences are shown in TABLE I, Figs. 6 and 7, respectively. According to these comparisons here also, generally, the better results belong to the proposed algorithm, especially for image pairs which contains large displacements. Figs. 3, 4, and 5 show the estimated flows for selected image pairs of mentioned sequences, using three optical flow methods.



(a) RMS error



(b) Correlation

Fig. 7. Comparisons of the original and estimated second images for SRI tree sequence using RMS error, and correlation.

IV. CONCLUSIONS

We have introduced as a criterion for evaluating optical flow the RMS error between the estimated second image (obtained from the first image and the estimated flow) and the real (original) second image. Based on this criterion, we have developed a new algorithm which is an improvement on the multigrid method. The algorithm calculates the error flow between the estimated and real second images, adds it to the scaled expanded flow of the previous level, and sets this summation as the initial flow for the present level flow calculation. Using this initial value, the flow for the present level is calculated using Horn and Schunck's method. The scaling range was set between one and the root of ten, although in our experiments the scales that gave the best (the least RMS error) results never exceeded 2.3. Working with two kinds of synthetic and real image sequences with known and unknown flows, we could achieve better results using our proposed algorithm in comparison with those obtained using the Horn and Schunck or standard multigrid methods. Based on the proposed evaluation criterion, we may be able to further improve the algorithm by changing adaptively the coarsest starting levels (m) of the multigrid method. In this paper, we evaluated purely the improvement by the optimal scaling in the standard multigrid method.

ACKNOWLEDGMENT

This research was supported by Kyungseong University Research Grants in 2011.

REFERENCES

- [1] S. S. Beauchemin, and J. L. Barron, "The computation of optical flow," *ACM Comput. Surv.* vol. 27, no. 3, pp. 433-466, 1995.
- [2] D. W. Murray, and B. F. Buxton, "Experiments in the Machine Interpretation of Visual Motion," *MIT Press, Cambridge, Massachusetts*, Sep. 1990.
- [3] H. C. Longuet-Higgins, and K. Prazdny, "The interpretation of a moving retinal image," *Proc. Roy. Soc. London B-208*, pp. 385-397, 1980.
- [4] A. M. Waxman, and S. Ullman, "Surface structure and three dimensional motion from image-flow kinematics," *Int. J. Robotics*, no. 4, pp. 72-94, Sep. 1985.
- [5] A. R. Bruss, and B. K. P. Horn, "Passive navigation," *Computer Vision Graphics Image Processing*, vol. 21, pp. 3-20, 1983.
- [6] D. Heeger, and A. Jepson, "Simple method for computing 3D motion and depth," *Proc. Third Int' Conf. Computer Vision, Osaka, Japan*, pp. 96-100, April. 1990.
- [7] S. Ullman, "Analysis of vision motion by biological and computer systems," *IEEE Trans. Comput.*, vol. 14, no. , pp. 57-69, 1981.
- [8] E. Hildreth, and C. Koch, "The analysis of visual motion: from computational theory to neuronal mechanisms," *Ann. Rev. neurosci.*, pp. 477-533, 1987.
- [9] E. Hildreth, "Computations underlying the measurement of visual motion," *Artificial Intell.*, vol. 23, pp. 309-354, 1984.
- [10] B. K. P. Horn, and G. Schunck, "Determining optical flow," *Artificial Intell.*, vol.17, pp. 185-203, 1981.

- [11] H. Nagel, "Analysis techniques for image sequences," *Proc. 4th Int. Joint Conf. Patt. Recog. Kyoto, Japan*, 1987.
- [12] S. Uras, F. Girosi, A. Verri, and V. Torre, "A computational approach to motion perception," *Biological Cybernet*, vol. 60, pp. 79-87, 1988.
- [13] A. Brandt, "Multi-level adaptive solutions to boundary value problems," *Math. Comput.*, vol. 31, pp. 333-390, 1877.
- [14] W. Enkelmann, "Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences," *Computer Vision Graph Image Process*, vol. 43, pp. 150-177, 1988.
- [15] D. Terzopoulos, "Image analysis using multigrid relaxation methods," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 8, pp. 129-139, 1986.
- [16] R. Battiti, E. Amaldi, and C. Koch, "Computing optical flow across multiple scales: An adaptive coarse-to-fine strategy," *Int. J. Comput. Vision*, vol. 6, no. 2, pp. 133-145, 1991.
- [17] D. Marr, and S. Ullman, "Directional selectivity and its use in early visual processing," *Proc. Roy. Soc. London*, vol. B-211, pp. 151-180, 1981.
- [18] T. Lin, and J. L. Barron, "Image reconstruction error for optical flow," *CVRP*, 1992.
- [19] D. J. Fleet, and A. D. Jepson, "Computation of component image velocity from local phase information," *IJCV*, vol. 5, no. 1, pp. 77-104, 1990.



Jinwoo Kim received the B.S degree in Electrical Engineering from Myongji University in 1992 and the M.S and Ph.D. degrees in Electronic Engineering and System design Engineering from Fukui University, Fukui, Japan, in 1996 and 1999, respectively. From 2000 to 2003, he was a contract Professor in the Department of Information Communication and Computer Engineering at Hanbat National University, Daejeon, Korea.

Since 2003 he has been with the Department of Information Communication Engineering at Kyungsoong University, Busan, Korea, where he is currently an associate professor. From Dec., 2007 to Mar., 2009, he was a visiting researcher in the Department of Bioengineering at Tokyo University, Japan. His current research interests include image processing, pattern recognition, and medical imaging technology.