GRADIENT-BASED OPTICAL FLOW FOR LARGE MOTION USING MULTI-RESOLUTION SMOOTHING OPERATION PRE-PROCESSING TECHNIQUE

***Md. Mosharrof Hossain Sarker*[1], *Kamal Bechkoum*[2]**
[1] Department of Electrical & Electronic Engineering,
Islamic University of Technology, Gazipur-1704, Bangladesh. Email: MHSarker@iut-dhaka.edu
[2] School of Computing and Information Technology,
University of Wolverhampton,35/49 Lichfield Street, WV1 1EQ ,UK. Email: K.Bechkoum@wlv.ac.uk

ABSTRACT

Motion estimation is a key problem in the analysis of image sequences. From image sequences, we can only estimate an approximation of the image motion called optical flow. In this paper, we present the gradient-based optical flow method that estimates the two-dimensional velocity of object motion. A multi-resolution smoothing operation proposes in this paper as a preprocessing step for overcoming the difficulty of large motion estimation by gradient-based optical flow techniques. The effectiveness of the proposed method has confirmed by applying image sequence of large motion. Experimental results with an image sequence show a qualitative improvement.

Keywords: **Motion estimation, large motion, optic flow**

1.0 INTRODUCTION

The analysis of image sequence is one of the most active study areas of computer vision and image processing. Motion estimation is a key problem in the analysis of image sequence. Usually, this is accomplished by determining the motion by which an object moves from one frame to the next. The motion field is the 2D vector field, which is perspective projection on the image plane of the 3D velocity field in the scene. From the information available from an image sequence (e.g.; the spatial and temporal variation of the brightness pattern), it is only possible to derive an estimate of the motion field, which is called optical flow or image flow [1],[2],[3],[4],[6]. In recent research much attention has been paid to more precise estimation techniques of optical flow, which are computationally expensive. Most of these techniques use more than two frames to estimate the flow and cannot achieve real-time performance. However, the computational efficiency and the property to estimate large displacement accurately are important preconditions for optical flow techniques in human motion, navigation and so forth.

Optical flow methods [1],[9] have the advantages that they can estimate velocity field with details local variations using two or three frames. However, they behave poorly in the estimation of large motion due to Taylor's approximation of the gradient constraint equation that is validity only in the small vicinity of motion. For a better understanding of the dependencies between image, flow vectors, flow regions are shown in Fig. 1.

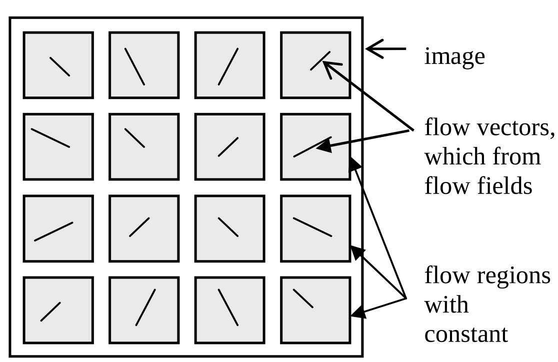

Fig.1: Relationship between image, flow regions, flow vectors and field

Definition 1:
"Large Motion in image sequences involves the subtraction of successive frames, and/or it affects to enlarge the flow region size where the presentation of pixels can be changed".

Logical aspects on behalf of large motion can be presented as: estimating large camera motions for omni-directional vision including combinations of large displacements, estimating the large displacement on the prediction of global flow field parameters and so forth [4],[5],[8]. In this work, monocular image sequence is considered.

The rest of the paper is organized as follows: Firstly, we give a brief introduction of motion estimation of objects. Secondly, we discuss the treatment of large motion estimation, multi-resolution smoothing operation technique and the error estimation due to motion of the object. Thirdly, we analyze the proposed method by using some image sequences and we show some experimental results. Finally, we conclude the paper.

2.0 RELATED WORKS

A number of researchers [4],[5],[6],[7],[8],[9] had referred to large motion. There are several research projects in the field of motion of object in image sequences. Lee *et al.* [4] presented a method for estimating large camera motions for omni directional vision, including combinations of large rotations and displacements. Radtke and Salzwedel [5] discussed the large displacement on the prediction of global flow field parameters. George and Tjahjadi [6] addressed the problems of high spatial resolution in the estimated optical flow velocity vector map and robustness to large object displacement and presented a novel Multi-Resolution Adaptive Shifting method (MRAS).

Kim *et al.* [7] proposed a fast scheme for partitioning an image based on the feature distribution and multi-resolution analysis. The grayscale intensity is considered as a feature for image segmentation. At the end, the optimum feature values are expanded into the original feature space. Wink and Roerdink [8] presented a general wavelet-based de-noising scheme for functional Magnetic Resonance Imaging (fMRI) data and compared it to Gaussian smoothing. They discussed elaborately about de-noising methods. Wang *et al.* [9] presented a comparison between ID grayscale histogram analysis and 2D entropic threshold, and proposed an algorithm using wavelet decomposition. The proposed technique saved computation costs, processing time and so forth.

The previous works paid no attention to a gradient-based motion estimation of an object (e.g.; usually human motion) with large motion in the image sequence using multi-resolution smoothing operation pre-processing technique. Another disadvantage of previous works is they work well for small optical flow.

3.0 LARGE MOTION ESTIMATION

In this section, we briefly discuss the spatio-temporal derivative technique for motion estimation and drawback of gradient-based technique.

Consider $p(\vec{x},t)$ is the usual, continuous, space time brightness function of a point $\vec{x}(x,y)$ in the image plane at time t. If the brightness remains practically constant along a motion trajectory, we have: $\frac{dp(\vec{x},t)}{dt}=0$, where $\vec{x}$ varies with respect to t according to motion trajectory. This is the total derivative expression and denotes the rate of change of intensity along the motion trajectory. Based on constant brightness assumption, the brightness of the object point in the images sampled at times t and $t+\delta t$ is the same, which means $p(\vec{x},t)=p(\vec{x}+\delta\vec{x},t+\delta t)$. Expanding by Taylor's series and omitting higher order terms, we have the following gradient constraint equation or optical flow constraint (OFC) equation [2],[3]:

$$\frac{\partial p(\vec{x},t)}{\partial x}v_x(\vec{x},t)+\frac{\partial p(\vec{x},t)}{\partial y}v_y(\vec{x},t)+\frac{\partial p(\vec{x},t)}{\partial t}=0 \qquad (1)$$

Where $v_x(\vec{x},t) = \dfrac{dx}{dt}$ and $v_y(\vec{x},t) = \dfrac{dy}{dt}$ denotes the component of image velocity vector in term of the continuous image coordinates, and the partial spatial derivatives of the image brightness are the components of the spatial gradient ∇p. Equation (1) can be written as the image brightness constancy equation:

$$(\nabla p)^{\mathrm{T}} \vec{v} + p_t = 0 \qquad (2)$$

Where p_t denotes partial differentiation with respect to time. Equation (2) provides one linear equation for two unknown components of velocity vector; hence, further constraints are necessary to solve the value $\vec{v}$. Global or local optimization techniques can be used to overcome this problem. In global optimization [2], the velocity field is determined by using an error function based on the gradient constraint and a global smoothing term. On the other hand, least square solution of the gradient constraint equation uses the local optimization [3],[5],[7] without assigning any smoothness.

The simplest constraint is to assume that the motion is the same on a small spatial neighborhood: the optical flow can estimate within a local region as the vector, $\vec{v}$ that minimizes the parameter Ψ defines as follows:

$$\Psi = \int\int_{(\vec{x},t)\in\Gamma} \left((\nabla p)^{\mathrm{T}} \vec{v} + p_t\right)^2 d\vec{x}dt \qquad (3)$$

Where Γ is the local region or neighborhood of the observation point. The velocity is estimated by minimizing the parameter Ψ. Differentiating with respect to v_x and v_y, and then setting it to zero, we have the following set of equations:

$$v_x p_{xx} + v_y p_{xy} + p_{tx} = 0 \qquad (4)$$

$$v_x p_{xy} + v_y p_{yy} + p_{ty} = 0 \qquad (5)$$

Where p_{ij} is given by

$$p_{ij} = \int\int_{(i,j:\vec{x},t)} p_i(\vec{x},t)\, p_j(\vec{x},t) d\vec{x}dt$$

From the equations (4) and (5), the component of image velocity $v_x(\vec{x},t)$ and $v_y(\vec{x},t)$ at any position $(\vec{x},y)$ in Γ is determined. The velocity obtained using above equations is:

$$v_x(\vec{x},t) = \frac{p_{yt}p_{xy} - p_{xt}p_{yy}}{p_{xx}p_{yy} - p_{xy}^2} \qquad (6)$$

$$v_y(\vec{x},t) = \frac{p_{xt}p_{xy} - p_{yt}p_{xx}}{p_{xx}p_{yy} - p_{xy}^2} \qquad (7)$$

The magnitude and phase angle correspond in Γ is $V(\vec{x},t) = \sqrt{v_x^2 + v_y^2}$ and $\phi = \tan^{-1}\dfrac{v_y}{v_x}$.

In the above motion estimation, the OFC equation derives by taking the Taylor expansion of the image constancy equation $dp(\vec{x},t)/dt = 0$ up to first order terms. The higher-order terms neglect under the assumption that the time step between consecutive frames is arbitrary small, which is most often violated in practice. In reality, time step is usually fixed. Therefore, the lack of higher order terms becomes main sources of error in the data constraint. The reliability of the image from constraint equation depends on the magnitude of the higher order derivatives of the image brightness function, $p(\vec{x},t)$. If the image brightness function in the neighbourhood of a point is well approximated by a linear function, i.e. its higher order derivatives are small, then the flow constraint is said to be very reliable at this point. On the contrary, the image flow constraint is very unreliable at the locations with significant higher-order derivatives, which is mainly in the neighbourhood of brightness discontinuities.

4.0 LEARNING STAGES

There are several possibilities to estimate large motions of single flow region with sufficient accuracy. The possible ways are as follows:

- Increasing the size of flow regions
- Predicting the global flow field parameters
- Computing the flow at multiple resolutions

4. 1 Increasing Flow Region Size

By increasing the size of flow regions, large motion can estimate. Unfortunately, the region cannot be enlarged too much, because in gradient-based motion estimation, we assume that optical flow is constant for a small region or very small areas.

4. 2 Prediction of Global Flow Field Parameters

For a given flow field it is possible to estimate parameters describing the movement of the camera (ego-motion). It allows the motion estimation by predicting of the next quadruple of parameters [5].

4. 3 Multi-Resolution Smoothing

The optical flow can estimate at different levels of an image hierarchy, starting at top level. For each level of hierarchy, image rows and columns are down sampled by a factor of two and the measurable motions double.
The multi-resolution smoothing operation with a window size $w = m \times m$ is shown in Fig. 2 in which the image area and the velocity are reduced by the factors of m^2 and m respectively. Due to reduced image area, the motion reduces in smoothing image. Therefore, it is possible to get smoothing estimated motion.

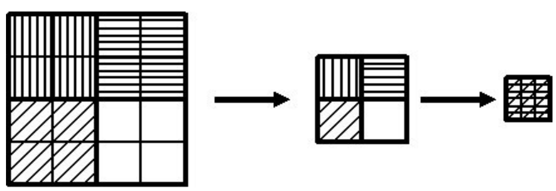

Fig. 2 : Multi-resolution smoothing operation with a window size of w = 2×2

Therefore, the multi resolution smoothing operation has some advantages stated below:

- Noise reduces due to averaging filter effect
- Image area reduces as well as motion doubles, so, computation complexity decreases
- Large motion reduces due to multi-resolution smoothing, therefore large motion estimation is possible by smoothing

4. 4 Error Estimation

It is mentioned that in the neighbourhood, the number of gradient constraint equations is larger than the unknown variables. Therefore, an over determined system of equations is formed in the neighbourhood. If an over determined system of equations is used to estimate motion, the measurement error in the gradients and in compatibilities among the constraint equations due to differential motion will be reflected in the residual of the solution [3],[8]. The residual vector r can be estimated by:

$$r \equiv (\nabla p)^{\mathrm{T}} \vec{v} + p_t \qquad (8)$$

Now, in order to estimate the error (ε) at each pixel, we have normalized the residual by the spatial gradient of brightness at each pixel. The error specifies the degree at which each pixel satisfies the image brightness constancy equation and it is given by the following equation:

$$\varepsilon = \frac{\left|(\nabla p)^{T} \vec{v} + p_t\right|}{\sqrt{\left|(\nabla p)^{T}\right|}} \tag{9}$$

When the constraint equation is completely satisfied over a local region, then ε becomes zero. The error will be increases if ε increase.

4. 5 Computer Environment

We use WINDOWS environment in the computer system for experimental implementation of the human motion detection where AVS / EXPRESS and MATLAB are used for this work.

5.0 EXPERIMENT AND DISCUSSION

To estimate the large motion of object, first multi-resolution smoothing operation over the image sequences is used. The gradient-based local optimization method is applied as discussed in the previous section with Γ = 4 pixels × 4 pixels × 2 frames. The gradients, p_x, p_y ,and p_t are determined considering a 2 × 2 × 2 frames spatio-temporal neighbourhood as depicts by [2],[9].

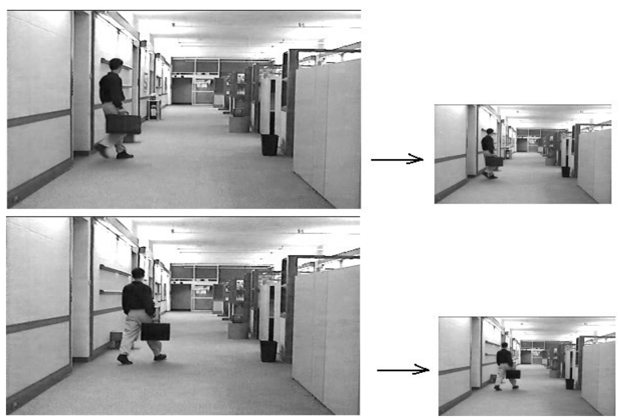

Fig. 3: Original image frames and smoothed image frame in the hall monitor sequence.

We have applied our algorithm to "*hall monitor*" image sequences and "*mob*" image sequences with fast walking. These image sequences contain the moving objects as humans with large motion and complex background shown in Fig. 3 and Fig. 4. The *hall monitor* image sequence size is 352 pixels × 240 pixels with 300 frames, and the *mob* image sequence size is 480 pixels × 640 pixels with 60 frames. In the pre-processing steps of multi-resolution smoothing, the pixels within the window averages. The reduced image frame of the first sequence is shown along with the Fig.3 on the right side.

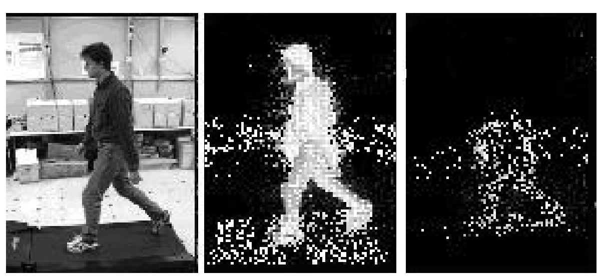

Fig. 4: Original image frame and smoothed image frame in the "mob sequence" of fast walking [top] original image, motion image, and error image, [bottom] smoothed image, motion image and error image.

Fig. 4 shows the second frame of *mob* sequence image with fast walking. The top three images show the original image, motion image and error image respectively. Similarly, the bottom three images are smoothed image, corresponding motion image and error image after multi-resolution and applying optical flow. The motion and error is calculated by using 2 pixels × 2 pixels in consecutive 2 frames without any overlapping in spatial dimension of the image.

In Fig. 4, it is clear that the error decreases due to multi resolution smoothing. Fig. 5 shows the error image caused by motion in 34 and 45 frames of "hall monitor" sequence. The magnitude of errors normalizes to compare easily. We estimated the optical flow on the intensity and calculated the error in each pixel. We have shown the motion as well as the error in a typical position of the image sequence over several frames in Fig. 6 and Fig. 7. In the figures "original" and "smooth" represent the data for original image sequence and sequence after multi-resolution smoothing treatment respectively. From the figures, it is found that estimation of velocity improves after multi resolution smoothing operation. It is shown that the error is reduced and the problem of gradient-based method is hence improved.

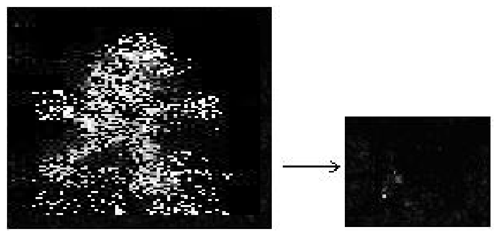

Fig. 5: Error images of the original and smoothed image frames 34 and 45 of the image sequence.

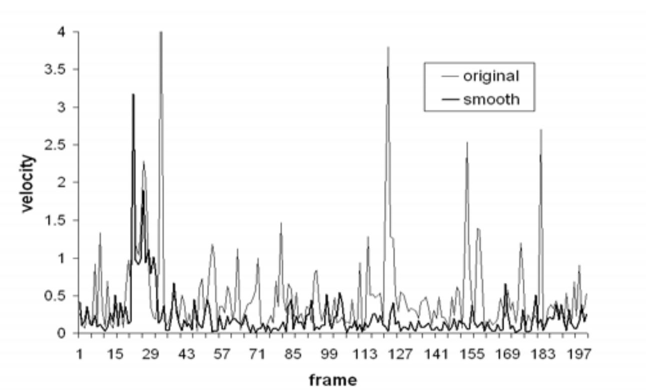

Fig. 6: Motion in a typical position of the original and smoothed image sequence of hall monitor sequence.

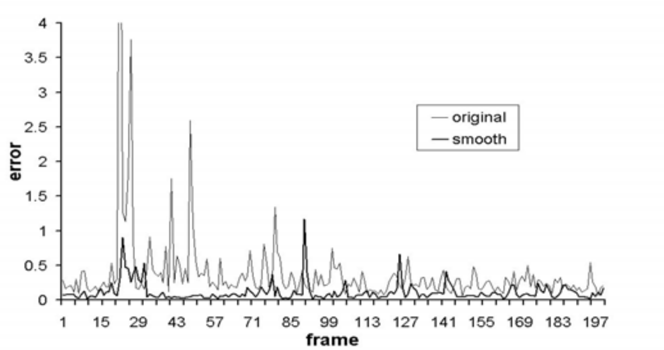

Fig. 7: Estimated error in the image sequence before and after smoothing operation.

Fig. 8 and Fig. 9 show the normalized motion and normalized error over several image frames in the *hall monitor* sequence. We have shown a limited number of frames in the figures. The error over the whole image sequence is reduced in the smoothed sequence of the image. In the same way, the velocity as well as the error in the fast walking of the mob sequence is shown in Fig. 10 and Fig. 11. The velocity reduces in Fig.10 is due to reduced size of image area for the smooth image which overcomes the difficulty of large motion estimation in gradient-based technique. The error reduces significantly with the smoothing operation.

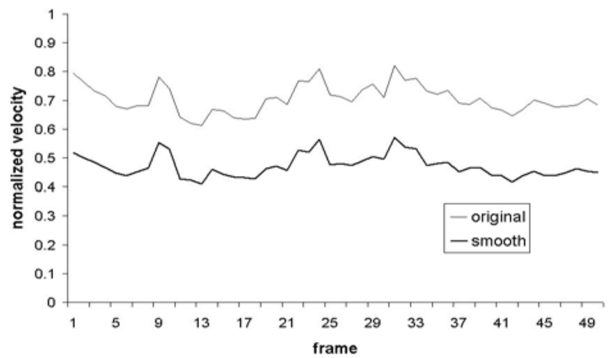

Fig. 8: Normalized motion of each frame in the hall monitor image sequence [50 frames are shown].

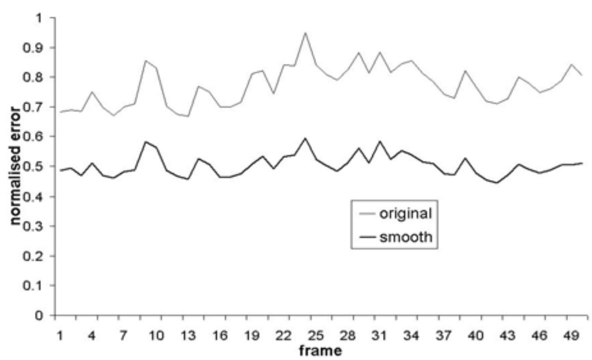

Fig. 9: Normalized error of each frame in the hall monitor image sequence.

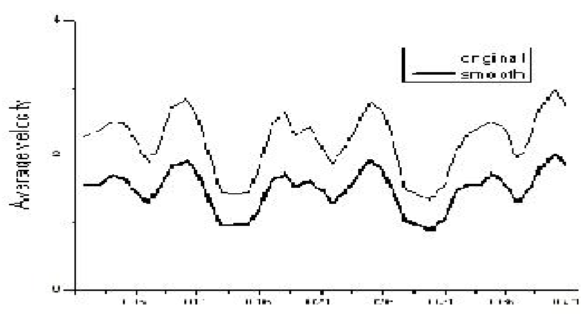

Fig. 10: Average motion of each frame in the “mob” image sequence [40 frames are shown].

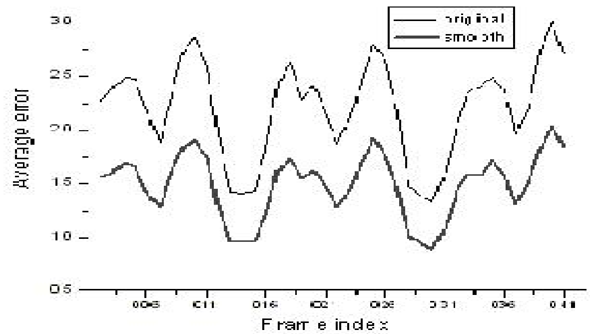

Fig. 11: Average error of each frame in the mob image sequence of fast walking.

In this work, $p(\vec{x},t)$ has been expanded using Taylor's series later on it is presented to linear equation for two unknown components of velocity vector, further it is presented to spatial neighbourhood (e.g.; but Laplacian concept is considered for smoothing in [2], and Gaussian smoothing in [8]). Based on this mathematical and AI logic, multi-resolution smoothing concept operates (e.g.; in section 4.3) and the computer program is structured. Therefore, "pre-processing" term crept into existence tacitly. We introduce here a pre-processing multi-resolution smoothing operation in which image velocity is reduced due to a decrease in image area and improve the overall flow estimation. We estimate the error for evaluating the results of original image sequence and smoothed sequence.

6.0 CONCLUSION

In this paper, we have shown that gradient-based optical flow estimates can be improved by including the multi-resolution smoothing operation pre-processing technique. This pre-processing overcomes the difficulty of estimation large motion by gradient technique. The error estimation evaluates the performance of the technique and show the improvement. The algorithm is easy, noise is reduced, and the overall computational cost is inexpensive.

ACKNOWLEDGEMENTS

We would like to thank Prof. Dr. Shahid Ullah, AVS/EXPRESS and MATLAB expert Mr. Shahriar Rahman for their helpful discussions.

REFERENCES

[1] J. L. Barron et al., "Performance of optical flow technique". *International Journal of Computer Vision*, Vol. 12 No. 1, 1994, pp. 43-77.

[2] B.K.P.Horn and B.G. Schunck, "Determining Optical Flow''. *Artificial Intelligence*, Vol.17,1981, pp. 185-203.

[3] J. K. Kearney et al., "Optical flow estimation: An Error Analysis of Gradient-Based Methods with Local Optimization" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, 1987, pp.229-244.

[4] J.W. Lee et al., "Large motion estimation for omnidirectional vision" in *Proceedings of the IEEE workshop on omni-directional vision*, 2000, pp. 161-168.

[5] T. Radtke and H. Salzwedel, "Estimation of optical flow for large displacements" in *Proceedings of the International Conference on Signal and Image Processing*, Kawaii, USA, 13 August 2001, pp. 55-58.

[6] M. George and T. Tjahjadi, “Multi-resolution optical flow estimation using adaptive shifting”, in *Proceedings of 1999 International Conference on Image Processing*, Volume 3, 24-28 Oct. 1999, pp. 717 – 721.

[7] B.G. Kim et al. "Fast image segmentation based on multi-resolution analysis and wavelets", *Pattern Recognition Letters 24*, Elsevier, 2003, pp.2995–3006.

[8] A.M. Wink and J.B.T.M. Roerdink, "Denoising functional MR images: a comparison of wavelet denoising and Gaussian smoothing" *IEEE Transactions on Medical Imaging* ,Vol. 23, No.3, March 2004, pp. 374 – 387.

[9] Q. Wang et al. "A fast 2D entropic thresholding method by wavelet decomposition"*International Conference on Image Processing. 2002*, Vol. 3, 24-28 June 2002, pp. III-265 - III-268.

BIOGRAPHY

Md. Mosharrof Hossain Sarker received a B. Sc. Engg. in EEE from Rajshahi University of Engineering & Technology (RUET) in 2001, and he received his M. Sc. in Computer Science from University of Wolverhampton (UK) in 2005. He is currently a lecturer in the Department of EEE, Islamic University of Technology (IUT), Bangladesh. He has published several papers in international and national journals as well as international conference. He is a member of IEB and BCS. His research interests are intelligent agent, machine graphics and vision.

Kamal Bechkoum holds a PhD from Cranfield University, UK, in Software Techniques for Computer Aided Engineering and has over twenty years of experience working in the field of artificial intelligence and its applications to business and industry. He published several papers in refereed international conferences and journals. He teaches Advanced Topics in Computer Science to MSc classes and supervises a number of PhD students in the field of Intelligent Software Agents and Distributed Artificial Intelligence. He is a Fellow of the British Computer Society (BCS), Chartered IT Professional and a member of the IEEE Computer Society.