

Motion Estimation for Omnidirectional Images using the Adapted Block-Matching

ALOUACHE Djamel and AMEUR Zohra

Laboratory Analysis and Modeling Random Phenomena (LAMPA) University of Mouloud Mammeri UMMTO
Tizi-Ouzou, Alegria
djamel.alouache@yahoo.fr; amzohra@yahoo.fr

KACHI Djemaa

Laboratory Modeling, Information & system (MIS) University of Picardie Jules Verne, France
33 rue Saint Leu 80039 Amiens Cedex 1.
djemaa.kachi@u-picardie.fr

Abstract—The Block-Matching (BM) method for motion estimation in most video coding is largely discussed in the case of perspective images. The omnidirectional cameras provide images with large field of view. These images contain global information about motion and permit to remove the ambiguity present with little camera motion in perspective case. Nevertheless, these images contain significant radial distortions. The Block-Matching in these catadioptric images is not a resolved problem, and still a challenging research field. A rectangular block representing the neighborhood in BM of a point and used in the perspective images is not appropriate for catadioptric cameras. The work presented in this article concerns the local motion estimation in catadioptric videos with the Adapted Block-Matching (ABM). The ABM based on an adapted neighborhood, the local motion estimation allows successful compensation prediction in catadioptric images. The Adapted Block-Matching is obtained from the equivalence between the omnidirectional image and the projection of scene points on a unit sphere.

Index Terms—Catadioptric images, adapted neighborhood, motion estimation, Adapted Block-Matching.

I. INTRODUCTION

Block-Matching methods are the most used in practice for motion estimation [1], [2]. They are found in almost all Modern video coding standards (H.261, MPEG-1, MPEG-2-4...). In the methods of the Block-Matching, two rectangular blocks are mapped according to the information distribution of their pixels.

The omnidirectional image has a non-homogeneous resolution. A rectangular block representing the neighborhood in BM of a point used in perspective images is not appropriate for catadioptric cameras. The Block-Matching method cannot be applied directly on catadioptric images because it necessarily leads to errors. Some authors seek to adapt the classical methods (applied

to perspective images) in the omnidirectional images, by seeking the adequate neighborhoods [9], [10]. Tomic et al [3] and Bogdanova et al [4] propose to use Block-Matching method in terms of spherical coordinates. The omnidirectional image is projected on the sphere of equivalence and the Spherical Block Matching (SBM) is developed to estimate motion in catadioptric sequences. To estimate the motion, the algorithm simply puts in correspondence two solid angles between two spherical images. Other researchers have also used Spherical Block Matching in catadioptric videos compression application [5], [6]. For example, in article [6] the authors used the Spherical Block Matching in loop motion compensation in the Slepian-Wolf decoder. The correlation between images is then assessed by estimating the motion between the two spherical images.

Because of interpolation noise introduced by the projection on the sphere some authors have underlined the incompatibility of treatments on the sphere (see for example [7], [8] and [9]). The objective of this paper is therefore to propose a direct adaptation of the classical Block-Matching in the omnidirectional images from the definition of a neighborhood system taking distortions into account. The remainder of this paper is organized as follows: in section 2, we present classical Block-Matching. In section 3 the catadioptric vision and the equivalence projection are presented. Section 4 describes, our approach to estimate motion using the Block-Matching with appropriate neighborhood. In section 5, the experimental evaluation is given and comparative measurements are discussed.

II. CLASSICAL BLOCK-MATCHING

Classical Block-Matching computes the optical flow on pixel $p(x,y)$ considering that the motion is constant in a fixed rectangular neighborhood. Two blocks $B(x,y,t)$ and $B'(x,y,t+I)$ are correlated according to the information distribution of their pixels in two successive images $I(x,y,t)$ and $I(x,y,t+I)$. Let us note $V = (v_x, v_y)$ the vector parameters of motion, B and B' the rectangular

neighborhood of the point in position $p(x,y)$ in successive images and $W(x,y)$ the search window as depicted in Fig. 1.

The Block-Matching method consists to estimate velocity vector $V(x,y)$ at the neighborhood $B(x,y)$ in search window $W(x,y)$, generally the best candidate block is selected with the minimum SAD error (1):

$$\vec{V}(x,y) = \arg \min_{v_x, v_y} \sum_{W(x,y)} |B(x,y,t) - B'(x+v_x, y+v_y, t+1)| \quad (1)$$

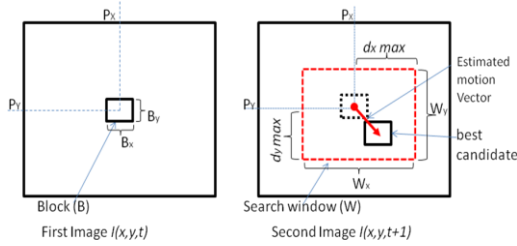


Fig. 1. Classical Block Matching

III. CATADIOPTRIC VISION

Catadioptric vision consists in associating a convex mirror with a camera whose optical axis is coincident with the axis of revolution of the mirror. Many authors have studied in their works equivalence between a catadioptric projection and stereographic projection [10, 11].

Geyer and Daniilidis have introduced a unifying theory for all central catadioptric sensors [11]. They proved that central catadioptric projection is equivalent to a central projection on a virtual sphere followed by projection from the sphere to the retina. This second projection depends on the shape of the mirror.

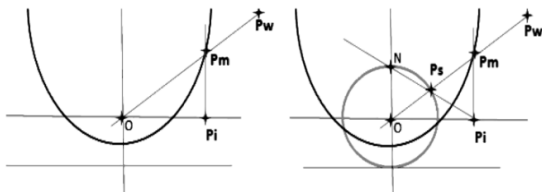


Fig. 2. Equivalence between the catadioptric projection and the two-step mapping via the unit sphere

Firstly the 3D point is projected in sphere from the center of sphere "O"; the second step consists in projecting point P_s on the sphere to the image plane from point "N" placed on the optical axis (Fig. 2). These projection points are obtained by calibrating the sensor.

Let (X_s, Y_s, Z_s) , Cartesian coordinate's point P_s and (θ, φ) the equivalent spherical coordinates.

$$P_s(X_s, Y_s, Z_s) = \begin{cases} \sin \theta \cos \theta \\ \sin \theta \sin \varphi \\ \cos \theta \end{cases} \quad (2)$$

With $\varphi \in [0, 2\pi[$, longitude angle and $\theta \in [0, \pi/2]$, latitude angle, the stereographic projection of P_s on the image plane yields point $P_i(x, y)$ given by:

$$\begin{cases} x = X_s / (1 - Z_s) \\ y = Y_s / (1 - Z_s) \end{cases} \quad (3)$$

By combining Equations (4) and (5) we obtain the spherical coordinates of point $P_i(x(\theta, \varphi), y(\theta, \varphi))$

$$\begin{cases} x = \cot \frac{\theta}{2} \cos \varphi \\ y = \cot \frac{\theta}{2} \sin \varphi \end{cases} \quad (4)$$

IV. ADAPTED BLOCK-MATCHING

A. Proposed Neighborhood for catadioptric images

In Démonceaux et al [7] the adapted neighborhood has been used for segmentation of omnidirectional images using Markov fields. The adapted neighbourhood is obtained from the equivalence between the omnidirectional image and the projection of scene points on a sphere as defined in Fig. 3. Radgui et al [8], [12] propose to use the spherical neighborhood in order to adapt the Lucas-Kanade method to estimate the optical flow in paracatadioptric images. The method proposed by Radgui et al [8], [12] allows a good estimate of the optical flow in the case of small displacements of the catadioptric camera. We propose, in this paper, to use the neighbourhood initially proposed in [7]. The adapted neighborhood noted $\vartheta(x, y)$ at point (x, y) , is defined as follows :

$$(x_i, y_i) \in \vartheta_{x,y} \Leftrightarrow |\theta - \theta_i| < d\theta \text{ et } |\varphi - \varphi_i| < d\varphi \quad (5)$$

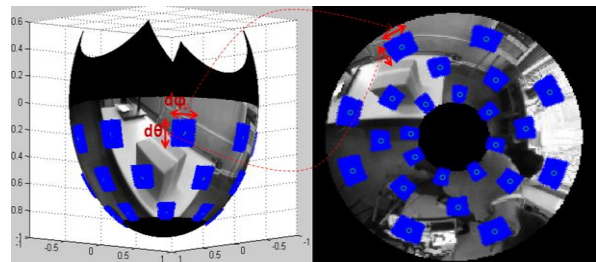


Fig. 3. Adapted neighborhood in omnidirectional image

Where θ_i and φ_i are the spherical coordinates of pixel (x_i, y_i) . Constants $d\theta$ and $d\varphi$ define the size of adapted neighborhood. In the proposed approach, $(d\theta_B, d\varphi_B)$ and $(d\theta_W, d\varphi_W)$ define respectively the size of block $B_{x,y}$ and search window $W_{x,y}$. With the neighborhood in (5) block $B_{x,y}$ and search window $W_{x,y}$ depend on the position of pixels in the omnidirectional image as depicted in Fig. 3. In the following section, we describe the Block-Matching method which uses this kind of neighborhood.

B. Adapted neighborhood used in Block-Matching

The Adapted Block-Matching puts in correspondence two adapted blocks between two successive omnidirectional images. Adapted search window $W_{x,y}$ and adapted block $B_{x,y}$ are obtained from equivalence between points of the omnidirectional image and points of equivalent sphere as depicted in Fig. 3.

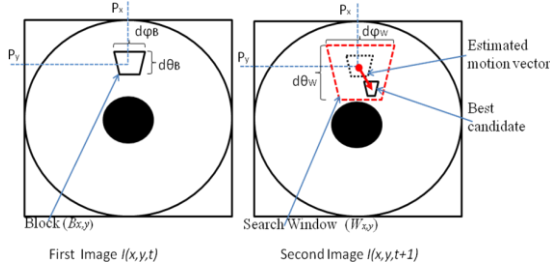


Fig. 4. Adapted Block Matching

First, the Algorithm consists in projecting two successive omnidirectional images on the unit sphere by using inverse stereographic projection and in addition divide first spherical Image $I(\theta, \varphi, t)$ into uniform solid angles of size $d\theta_B \times d\varphi_B$ that form blocks with a constant motion (See Fig. 5). The second step is to define the neighborhood $B_{x,y}$ in omnidirectional image $I(x,y,t)$ using stereographic projection as described in (6). The last step is to find best block $B'_{x,y}$ belonging in target adapted search window $d\theta_w \times d\varphi_w$ on second omnidirectional image $I(x,y,t+1)$ as depicted in Fig. 4.

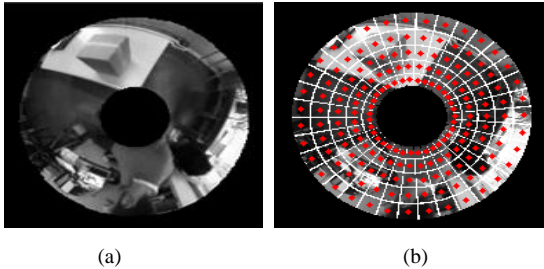


Fig.5. Adapted Block in Omnidirectional Image
(a) Reference Image (t) (b) Dividing in no overlap blocks

Adapted block of point $p(x_i, y_i)$ noted $B_{x,y}$ is defined as follows:

$$(x_i, y_i) \in B_{x,y} \Leftrightarrow |\theta - \theta_i| < d\theta_B \text{ et } |\varphi - \varphi_i| < d\varphi_B \quad (6)$$

Blocks $B_{x,y}$ and $B'_{x,y}$, minimizing the well known Sum of Absolute Differences (SAD) are used to compute the motion vector (dx, dy) :

$$SAD(x, y) = \sum_{W_{x,y}} |B_{x,y,t} - B'_{x+dx, y+dy, t+1}| \quad (7)$$

$$(d\hat{x}, d\hat{y}) = \arg \min_{dx, dy} (SAD(x, y)) \quad (8)$$

The adapted search window of point $P(x_i, y_i)$ noted $W_{x,y}$ is defined as follows:

$$(x_i, y_i) \in W_{x,y} \Leftrightarrow |\theta - \theta_i| < d\theta_w \text{ et } |\varphi - \varphi_i| < d\varphi_w \quad (9)$$

Omnidirectional images $I(x,y)$ are defined on $L \times L$, ($L \in \mathbb{N}$) squared grid and plotted with the spherical coordinate θ and φ by the equation of an equiangular grid defined as follows:

$$\zeta_j = \left\{ (\theta_p, \varphi_q) \in S^2 : \theta_p = \frac{(2p+1)\pi}{2L}, \varphi_q = \frac{2q\pi}{L}, p, q \in \mathbb{Z}[L] \right\} \quad (10)$$

Algorithm: Adapted Block Matching (ABM)

$$I(x, y) \in \mathbb{R}^2 \xrightarrow{\Pi^{-1}} I(\theta, \varphi) \in S^2$$

$$V_i = [0, 0], \forall i,$$

(divide $I(\theta, \varphi)$ into Nb uniform blocks of size $d\theta_B \times d\varphi_B$)

$$i = 0;$$

repeat

$$(\theta_i, \varphi_i) \leftarrow \text{position of } I(\theta, \varphi);$$

$$B_i = \{ p(\theta, \varphi) / |\theta - \theta_i| < d\theta_B, |\varphi - \varphi_i| < d\varphi_B \};$$

$$W_i = \{ p(\theta, \varphi) / |\theta - \theta_i| < d\theta_w, |\varphi - \varphi_i| < d\varphi_w \};$$

$$B_i(\theta, \varphi) \xrightarrow{\Pi} B_i(x, y);$$

$$W_i(\theta, \varphi) \xrightarrow{\Pi} W_i(x, y);$$

$$S^2 \rightarrow \mathfrak{R}^2;$$

$$B_i : \{ p(x, y) / |x - x_i| < M_x, |y - y_i| < M_y \}$$

$$W_i : \{ p(x, y) / |x - x_i| < W_x, |y - y_i| < W_y \}$$

$$(x_i, y_i) \leftarrow \text{position of } B_i(x, y);$$

$$W_i(x, y) \leftarrow \{ (x, y) \} \text{ such that}$$

$$x \in \left[x_i + V_i(1) - \frac{W_x}{2} + 1, x_i + V_i(1) + \frac{W_x}{2} \right] \text{ and}$$

$$y \in \left[y_i + V_i(2) - \frac{W_y}{2} + 1, y_i + V_i(2) + \frac{W_y}{2} \right];$$

$$B'_i(x, y) \leftarrow \text{position of } (x, y)$$

$$B'_i = \arg \min_{W_i} SAD(B_i, B'_i);$$

$$(dx_i, dy_i) \leftarrow \text{position of } B'_i;$$

$$V_i \leftarrow [x_i + dx_i, y_i + dy_i];$$

$$i \leftarrow i + 1;$$

until $i > Nb$

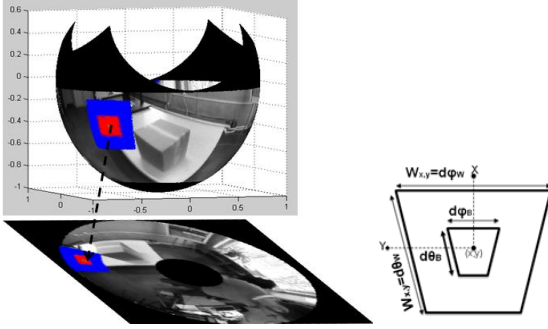


Fig. 6. Adapted block with fixed value $d\phi_B=14^\circ$, $d\theta_B=11^\circ$ and adapted search window with fixed value $d\phi_W=28^\circ$, $d\theta_W=25^\circ$ on the sphere and their projection on the image plane.

V. EXPERIMENT SIMULATION AND RESULT ANALYSIS

The sequences are obtained using a camera embedded on a mobile robot which is moving on a plan perpendicular to its optical axis. We consider that the camera is calibrated. In order to assess the proposed method, different types of motion are applied to the camera. The translation motion in axis-Y, rotation motion in axis-Z and the moving objects in the scene have been tested. We have chosen three types of sequences as follows: indoor sequence.1 (Fig. 7.a), with pure rotation of the camera, indoor sequence.2 (Fig. 7.b), with a fixed camera and moving objects in the scene and outdoor sequence.3 (Fig. 7.c), a camera placed on a moving car. The block sizes and search window used in the classical Block-Matching are $(B=15 \times 15$ and $W=41 \times 41$ for sequence.1), $(B=13 \times 13$ and $W=41 \times 41$ for sequence.2) and $(B=11 \times 11$ and $W=41 \times 41$ for sequence.3). The two delimiter angles in our Adapted Block-Matching are $(d\theta_B=6^\circ$, $d\phi_B=9^\circ$, and $d\theta_W=15^\circ$, $d\phi_W=27^\circ$ for sequence.1), $(d\theta_B=5^\circ$, $d\phi_B=8^\circ$, and $d\theta_W=15^\circ$, $d\phi_W=27^\circ$ for sequence.2) and $(d\theta_B=4^\circ$, $d\phi_B=7^\circ$, and $d\theta_W=15^\circ$, $d\phi_W=27^\circ$ for sequence.3).

To evaluate the performance of our method on three sequences, we use the PSNR (11) and MSSIM (13). The PSNR in Fig. 8, MSSIM in Table 1 and Error Image are computed from the predicted and the Reference images. The predicted image is estimated by using estimated motion for each pixel in the Reference image. The results of both indoor and outdoor sequences are illustrated respectively in Figures: Fig. 9, Fig. 10 and Fig. 11. In all these sequences, the motion estimated by our proposed method allows an estimation of accurate motion field in various scenes and different kinds of camera motions.

- The PSNR is expressed in the case of grayscale images by:

$$PSNR = 10 \log_{10} \cdot \frac{(255)^2}{\frac{1}{N \times N} \sum_{j=0}^{N-1} \sum_{i=0}^{N-1} [I(i, j) - \tilde{I}(i, j)]^2} \quad (11)$$

Where: $N \times N$ is the size of the image, $I(x, y)$ represents the pixels of the reference image, $\tilde{I}(x, y)$ is the predicted image obtained using the estimated vectors and the target image.

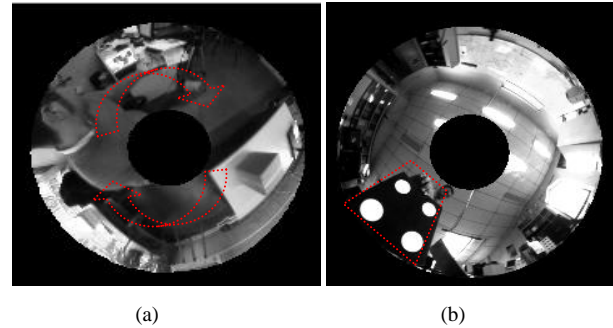
- The Structural Similarity (SSIM) measure is calculated on various windows of an image. The measure between two windows x and y of common size $N \times N$ is calculated as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (12)$$

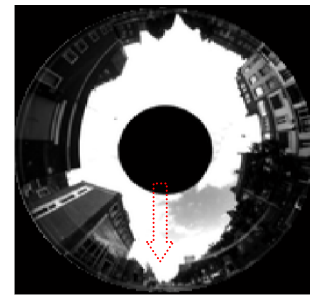
Where, μ_x and μ_y represent the mean of windows x , and y . σ_x^2 and σ_y^2 are the standard deviation; and C_1 and C_2 are included to avoid instability when $\mu_x^2 + \mu_y^2$ and $\sigma_x^2 + \sigma_y^2$ are very close to zero. The mean SSIM (MSSIM) is used to evaluate the overall quality:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{i=0}^M SSIM(x_i, y_i) \quad (13)$$

Where, X and Y are the reference and the predicted images. x_i and y_i are the intensity at the local window; and M is the number of local windows of the image [13].



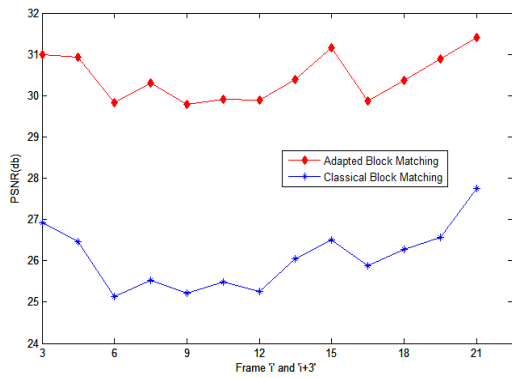
(a) (b)



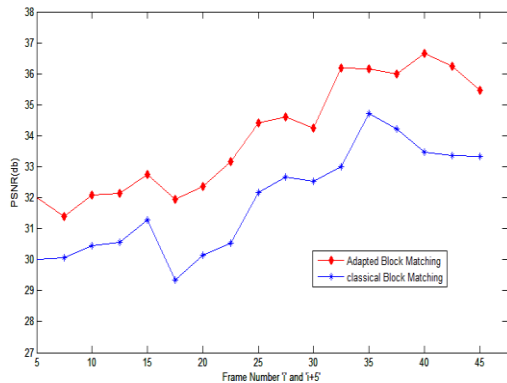
(c)

- (a) Sequence.1: pure rotation camera (in Z-axis)
- (b) Sequence.2: fixed camera and moving object
- (c) Sequence.3: camera placed on a moving car (translation Y-axis)

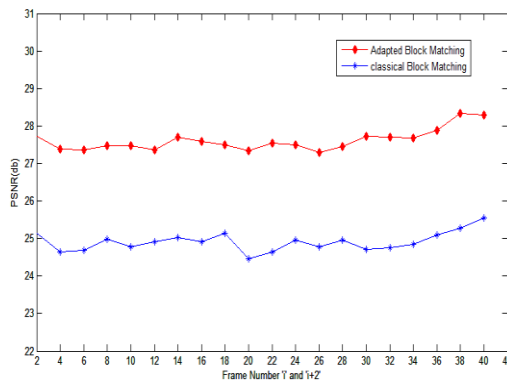
Fig. 7. The Omnidirectional Sequences



(a)



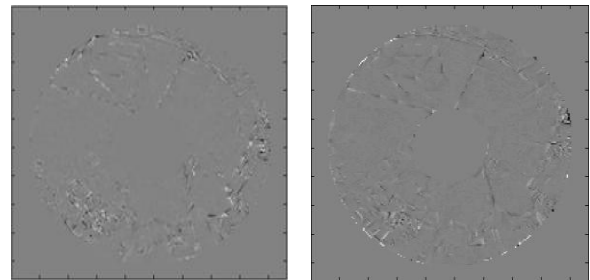
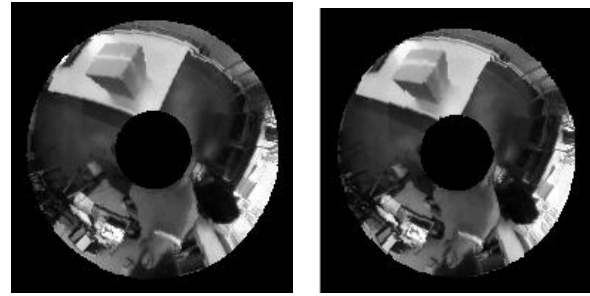
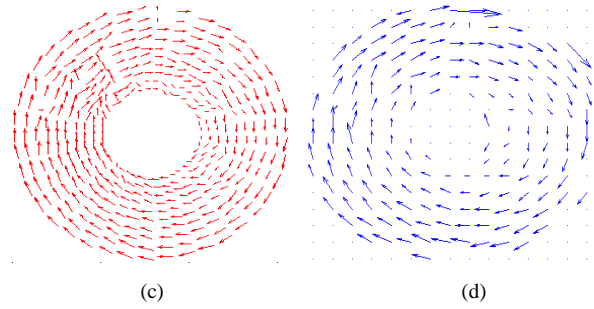
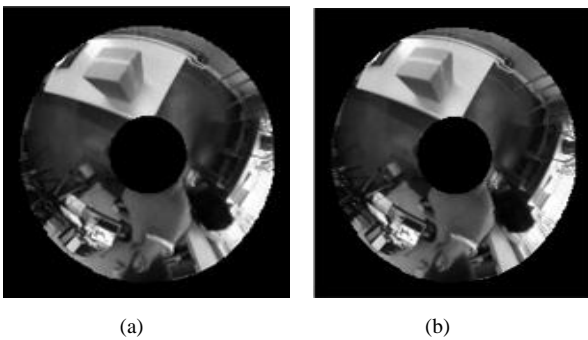
(b)



(c)

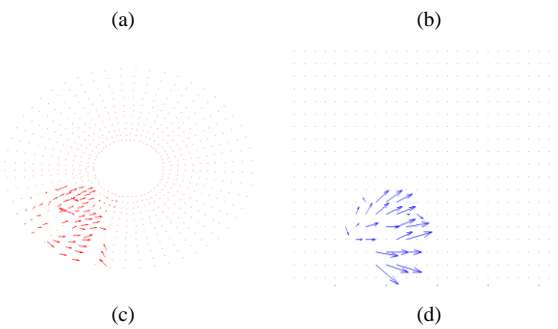
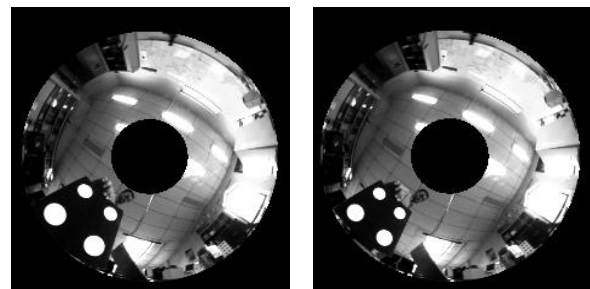
(a) Sequence.1 was used with a frame distance of 3
 (b) Sequence.2 was used with a frame distance of 5
 (c) Sequence.3 was used with a frame distance of 2

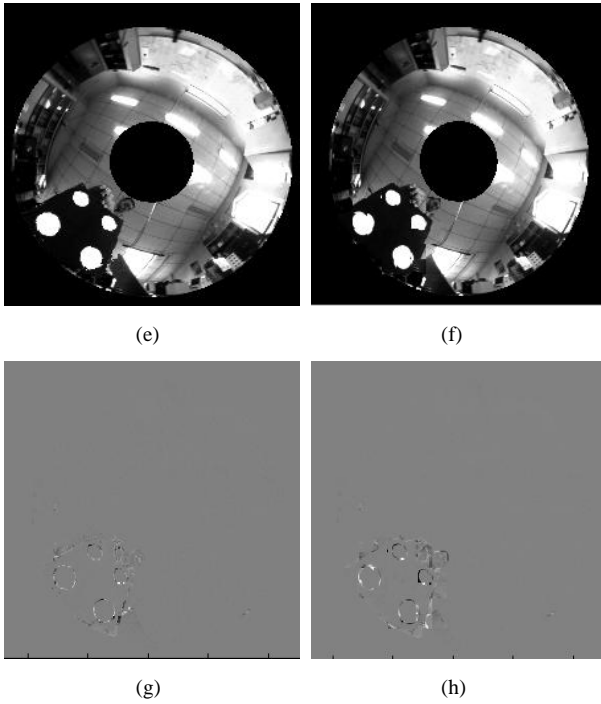
Fig. 8.PSNR performances of Block Matching Algorithms



(a) Reference image (t) (b) Target image (t+3)
 (c) Motion estimated (ABM) (d) Motion estimated (BM)
 (e) Predicted image (ABM) (f) Predicted image (BM)
 (g) Error image (ABM) (h) Error image (BM)

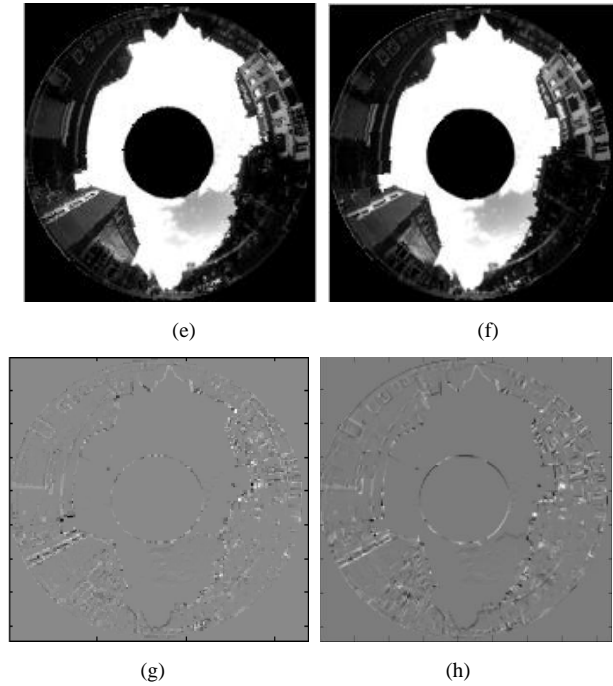
Fig. 9.Moving camera in pure rotation





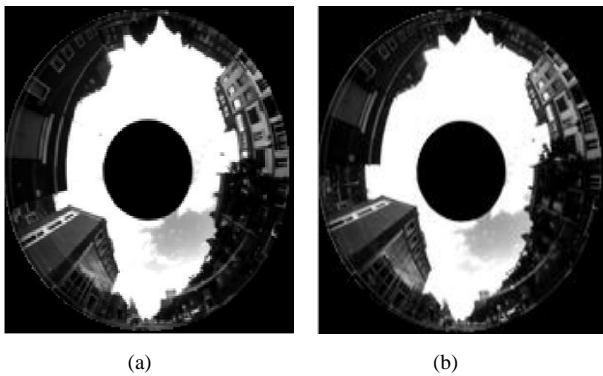
(a) Reference image (t) (b) Target image (t+5)
 (c) Motion estimated (ABM) (d) Motion estimated (BM)
 (e) Predicted image (ABM) (f) Predicted image (BM)
 (g) Error image (ABM) (h) Error image (BM)

Fig. 10. Fixed camera and moving object



(a) Reference image (t) (b) Target image (t+2)
 (c) Motion estimated (ABM) (d) Motion estimated (BM)
 (e) Predicted image (ABM) (f) Predicted image (BM)
 (g) Error image (ABM) (h) Error image (BM)

Fig. 11. Moving camera in pure Translation (Y-axis)



(a) Reference image (b) Target image (t+5)
 (c) Motion estimated (ABM) (d) Motion estimated (BM)

Table 1. Average of PSNR (db) and MSSIM of sequences used for verifying our motion model (ABM)

Sequences	Mesures (Average)	BM	ABM
Sequence.1	PSNR (db)	26.28	31.90
	MSSIM	0.902	0.973
Sequence.2	PSNR (db)	32.28	34.78
	MSSIM	0.979	0.990
Sequence.3	PSNR(db)	25.94	28.12
	MSSIM	0.896	0.949

The mean PSNR and SSIM obtained from different sequences using the two implemented methods (BM) and (ABM) are indicated in Table 1. These results show that (BM) method is not appropriate in the case of omnidirectional images. The results obtained using our proposed method is superior to that obtained by classical methods. These results give the argument that methods dedicated to omnidirectional images are more robust and give more precise results. Finally, we can see that the measures are superior using approach with different camera motions. This means that the fixed support of neighborhood is not appropriate to omnidirectional images.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an adaptation of the Block Matching method in omnidirectional sequences using adapted blocks. The use of an adapted block instead of a rectangular block of fixed size improves the results of

estimating motion in omnidirectional sequences. The experimental results have shown the effectiveness of blocks that are appropriate for motion estimation in omnidirectional images. Our Block-Matching method (ABM) allows interesting results in the different movements of the catadioptric camera contrary to the results obtained with the classical Block-Matching method proposed on the perspective images.

Indeed, in our various experiments, this method has enabled us to rebuild the various areas and contours in the predicted images and to reach PSNR and MSSIM of quality exceeding the results obtained with classical Block-Matching.

To make our Adapted Block Matching method more robust, it is desirable to integrate and to integrate other Block Matching research strategies [14] [2]. Nonetheless improvements can be introduced to increase the rate calculated and a good estimate of the movement for example by applying the method in a multi-resolution arrangement which consists in performing as many motion estimates as levels of decomposition [15].

REFERENCES

- [1] A. Barjatya, "Block Matching Algorithms for Motion Estimation" *IEEE Trans on Digital Image Proc* DIP 6620 Spring 2004 Final Project Paper
- [2] A. Nayak and B. Biswal "Evaluation and Comparison of Motion Estimation Algorithms for Video Compression" *International Journal. Image, Graphics and Signal processing*, Vol.10, pp.9-18. 2013
- [3] I. Tosic, I. Bogdanova, P. Frossard, and P. Vanderghyest. "Multiresolution motion estimation for omnidirectional images". *Proceedings of EUSIPCO 2005*, Antalya, September, 2005.
- [4] I. Bogdanova, A. Bur, P. Farine, and H. Hügli. "Dynamic visual attention on the sphere". *Journal of Computer Vision and Image Understanding*, Vol.114, pp.100-110, 2010.
- [5] V. Thirumalai, I. Tosic, and P. Frossard. "Distributed coding of multiresolution omnidirectional images". *Image Processing, ICIP 2007*, Vol.2, pp.II – 345 – II – 348, 2007.
- [6] I. Tosic and P. Frossard. "Low bit rat compression of omnidirectional images". *Proceedings of PCS*, pp.1-4, 2009.
- [7] C. Démonceaux and P. Vasseur. "Markov random fields for catadioptric image processing". *Pattern Recognition*, Vol.27 pp.1957-1967, 2006.
- [8] A. Radgui, C. Démonceaux, E. Mouaddib, D. Aboutajdine and M. Rziza. "An adapted lucas-kanade's method for optical flow estimation in catadioptric images". *Proceedings of the eighth workshop on Omnidirectional Vision and Camera Networks (OMNIVIS 08)*, 2008.
- [9] Rameau. F., Sidibe. D., Démonceaux., C., Fofi, D. "Tracking Moving Objects With a Catadioptric Sensor Using Particle Filter" *Computer Vision Workshops (ICCV Workshops)*, *IEEE International Conference*. pp. 328 – 334, Barcelona 2011.
- [10] C. Geyer and K. Daniilidis, "Paracatadioptric camera calibration," *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol. 24, (5), 2002.
- [11] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," *International Journal of Computer Vision*, vol. 45, (3), pp. 223 – 243, December 2001.
- [12] A. Radgui, C. Démonceaux, E. Mouaddib, M. Rziza, and D. Aboutajdine. "Optical flow estimation from multichannel spherical image decomposition". *Computer Vision and Image Understanding*, Vol.115, pp.1263-1272, 2011.
- [13] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans on Image Processing*, vol. 13, (4), pp. 600-612, avril 2004.
- [14] S. Acharjee and S.S. Chaudhuri "Fuzzy Logic Based Three Step Search Algorithm for Motion Vector Estimation." *International Journal. Image, Graphics and Signal Processing*, Vol. 2 pp.37-43, 2012.
- [15] M. J. Kim, Y. G. Lee, and J. B. Ra, "A fast multi-resolution block matching algorithm for multiple-frame motion estimation". *IEICE Trans on Information and Systems*, Vol. E88-D, (12), pp. 2819-2827, 2005.

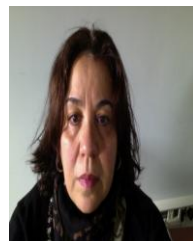
AUTHOR'S PROFILES



ALOUACHE Djamel received his Engineering Degree in Electronic in 2007, and then received his Magister diploma in Remote Sensing in 2009 at University Mouloud Mammeri in 2010. Alouache is now PhD student in University of Mouloud Mammeri Tizi-Ouzou and member in two laboratories, laboratory MIS University of Jules Verne France, team Perception Robotics, and Laboratory LAMPA University of Mouloud Mammeri , Algeria, team image processing. His research interests are analysis and processing of spherical images.



Ms. **AMEUR Zohra** teacher is a researcher at the University of MOULOUD MAMMERI Tizi-Ouzou, Algeria since September 1986. She obtained her PhD at the Faculty of Electrical and Computer Engineering in 2005 in specialty image processing. Currently a professor and head of a laboratory: Analysis and Modeling Random Phenomena "LAMPA" since June 2010.



KACHI Djemaa received her Ph.D. in System Control from the Technology University of Compiègne, France in 1996. She is an associate professor at the University of Picardy, Amiens and member of M.I.S laboratory since 1999. Her research interests are image processing, omnidirectional vision, and their applications to mobile robots.