



HAL
open science

Visual tracking with omnidirectional cameras: an efficient approach

François Rameau, Désiré Sidibé, Cedric Demonceaux, David Fofi

► **To cite this version:**

François Rameau, Désiré Sidibé, Cedric Demonceaux, David Fofi. Visual tracking with omnidirectional cameras: an efficient approach. *Electronics Letters*, 2011, 47 (21), pp.1183 - 1184. 10.1049/el.2011.2838 . hal-00637343

HAL Id: hal-00637343

<https://hal.science/hal-00637343>

Submitted on 2 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Visual tracking with omnidirectional cameras: an efficient approach

F. Rameau, D.D. Sidibé, C. Demonceaux and D. Fofi

An effective technique for applying visual tracking algorithms to omnidirectional image sequences is presented. The method is based on a spherical image representation which allows taking into account the distortions and nonlinear resolution of omnidirectional images. Experimental results show that both deterministic and probabilistic tracking methods can effectively be adapted in order to robustly track an object with an omnidirectional camera.

Introduction: Visual tracking is an important task in many computer vision applications such as video surveillance, smart rooms, mobile robotics, augmented reality and video compression. Different tracking methods have been developed in the literature and can be roughly divided into two groups: deterministic methods and stochastic methods. Methods of the former group, such as the Kanade-Lucas-Tomasi (KLT) tracker [1] and the mean-shift tracker [2], iteratively search for the local maxima of a similarity measure between a template of the target and the current image. In contrast, methods of the latter group use a state space representation of the moving object to model its underlying dynamics. These include the Kalman filter and particle filters [3].

Though these methods have been successfully employed in various application domains, they cannot be directly applied to images acquired by an omnidirectional camera. An omnidirectional camera provides a panoramic view of a scene by the association of a conventional perspective camera with a convex mirror. Omnidirectional sensors are useful in various applications such as 3D reconstruction, structure from motion, robot navigation and video surveillance. However, conventional visual tracking methods are not able to perform efficiently when applied to a sequence taken with an omnidirectional camera. This is mainly due to the geometrical distortion induced by the mirror. The mirror also makes the sensor highly sensitive to illumination changes.

Few tracking methods in omnidirectional images have been proposed in the literature. The simplest and widely used technique is based on image unwrapping. The unwrapping process creates a perspective image from the distorted one, and conventional tracking methods can be applied [4]. Other methods use a geometrical *a priori* knowledge about the object to track, for instance planes or lines [5, 6].

In this Letter, we show how to effectively adapt mean-shift and particle filters to omnidirectional images. We report experimental results which show a significant improvement in the tracking accuracy.

Adapted search window: To efficiently track a target, a distinctive model of its appearance is needed. The target model is usually computed within a bounding box (or search window) given by the object current position. With a perspective image this search window is defined by an $M \times N$ rectangle centred on the target's central pixel. Obviously, this window has to be adapted to take into account the distortion and the nonlinear resolution of omnidirectional images. Indeed, without adaptation, the target representation will contain a lot of background pixels which cause tracking drift and failure. Following the work presented in [7], omnidirectional images are considered as spherical images. A point X_S lying on the unit sphere S^2 can be defined by its spherical co-ordinates: $X_S = [\cos(\phi) \sin(\theta), \sin(\phi) \cos(\theta), \cos(\phi)]^T$, where $\phi \in [0, \pi]$ is the latitude, and $\theta \in [0, 2\pi]$ is the longitude. Then, the window centred on X_S is defined in the unit sphere using the following equation:

$$W(X_S) = \{X = (\theta', \phi') \in S^2; |\theta' - \theta| \leq \delta\theta \text{ and } |\phi' - \phi| \leq \delta\phi\} \quad (1)$$

Finally, the window is back-projected on the image plane to get the spatial neighbourhood on the image as illustrated in Fig. 1.

Colour histogram representation: The colour histogram is a widely used representation in tracking algorithms. It is robust against noise, partial occlusion and is scale invariant. To model the appearance of the object of interest, a weighted 3D colour histogram is used. The weights are assigned by a kernel which accounts for the spatial information in order to improve tracking accuracy. In conventional colour based tracking methods, the Epanechnikov kernel is used [2]. However, an adaptation of this profile is necessary to deal with the

distortion in omnidirectional images. To that end, the Euclidean distance is replaced by the geodesic distance [7]. The distance between two points $(X_{sph}, Y_{sph}) \in S^2$ is defined by:

$$d(X_{sph}, Y_{sph}) = \arccos(X_{sph}, Y_{sph}) \quad (2)$$

Tracking algorithms: In this Letter we consider two popular tracking methods and adapt them to omnidirectional images. The first one is the mean-shift algorithm [2] which is a deterministic method, and the second one is the probabilistic particle filter algorithm [3]. In both cases, the appearance of the object to track is represented by a colour histogram using the adapted window and the similarity between two colour distributions is obtained by the Bhattacharyya coefficient [2]. For the particle filter, we define the object's state vector in the unit sphere as $S = [\theta, \phi, \theta, \phi, \delta\theta, \delta\phi]^T$, where θ and ϕ are the co-ordinates of the window's centre, and $\delta\theta$ and $\delta\phi$ are the size of the window.

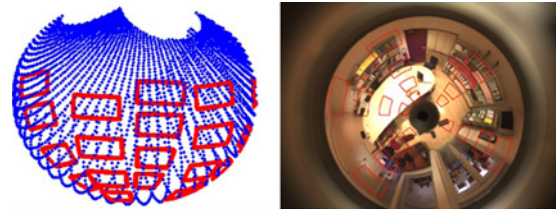


Fig. 1 Window adaptation on sphere
Left: window on unitary sphere; right: window projected on image plane

Results: We evaluate the performance of our adapted tracking methods using four sequences acquired in different conditions. The conditions include indoor and outdoor scenes, moving objects and persons, a moving or fixed camera, difficult illumination changes and occlusion. The image resolution is 640×480 and the size of sequences varies between 500 and 800 frames. The adapted tracking algorithms are compared with the conventional ones using three criteria: the spatial overlapping (percentage of common area between the ground-truth and the tracking bounding boxes), the distance between the real centre (from the ground-truth) and the centre estimated by the tracking algorithm, and the temporal overlapping (percentage of successfully tracked frames in a sequence).

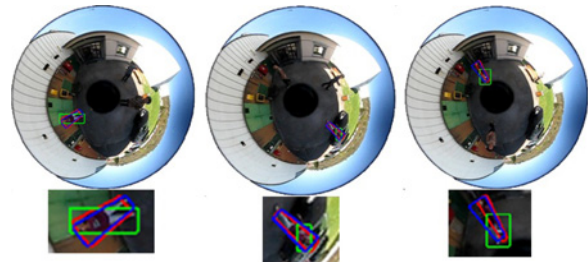


Fig. 2 Example of tracking results
Tracking results with conventional particle filter depicted in green. Results with proposed method shown in red, and ground truth is in blue

Table 1: Tracking performance evaluation

	Spatial overlapping (%)		Temporal overlapping (%)		Centres distance (in pixels)		
	Conventional method	Adapted method	Conventional method	Adapted method	Conventional method	Adapted method	
Seq. 1	45	71	100	100	5.7	2.6	Particle filter
Seq. 2	30	60	44	99.7	5.5	3.8	
Seq. 3	33	65	88	99.5	7.7	4.8	
Seq. 4	45	66	96.8	97.3	6	4.1	
Seq. 1	32.4	64.5	63.21	96.35	6.38	6.2	Mean-shift
Seq. 2	42.5	47.79	89	90	16.2	8.3	
Seq. 3	40.55	61.8	96	97.84	7.8	7.5	
Seq. 4	34.72	67.67	52.81	97.71	10.52	5.02	

A few tracking results are shown in Fig. 2. Note that in the images, the blue bounding box corresponds to the ground truth object location, while the red and green bounding boxes correspond to the estimated positions by the proposed tracking methods and the conventional

ones, respectively. As can be seen, the adapted particle filter provides a better localisation of the tracked object.

Quantitative performances are summarised in Table 1. As can be seen, our adapted methods outperform the conventional mean-shift and particle filter tracking methods for all four sequences. In particular, even with strong scale variation (all sequences), occlusion (sequence 3) and a strong change in illumination (sequence 2), the adapted algorithms accurately follow the targets. In particular, they give a better localisation of the object as shown by the mean distance between the true centre of the object and the centre given by the algorithms.

Conclusion: An efficient way to adapt conventional tracking methods to omnidirectional image sequences is proposed. The adaptation is based on using spherical image representation to account for the specific geometry of omnidirectional images. Experimental results with different sequences in different conditions show significant improvements in the tracking accuracy. In particular, the adapted methods are robust against occlusion, clutter background and large illumination variation.

© The Institution of Engineering and Technology 2011

6 September 2011

doi: 10.1049/el.2011.2838

One or more of the Figures in this Letter are available in colour online.

F. Rameau, D.D. Sidibé, C. Demonceaux and D. Fofi (*Université de Bourgogne, Laboratoire LE2I - UMR CNRS 5158, 12 rue de la Fonderie, 71200 Le Creusot, France*)

E-mail: dro-desire.sidibe@u-bourgogne.fr

References

- 1 Lucas, B.D., and Kanade, T.: 'An iterative image registration technique with an application to stereo vision'. Proc. Int. Joint Conf. on Artificial Intelligence, Vancouver, Canada, 1981, pp. 674–859
- 2 Comaniciu, D., Ramesh, V., and Meer, P.: 'Real-time tracking of non-rigid objects using mean shift'. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head Kland, SC, USA, 2000, pp. 142–149
- 3 Isard, M., and Blake, A.: 'Condensation – conditional density propagation for visual tracking', *Int. J. Comput. Vis.*, 1998, **28**, (1), pp. 5–28
- 4 Wallhoff, F., Zobl, M., Rigoll, G., and Potucek, I.: 'Face tracking in meeting room scenarios using omnidirectional views'. Proc. IEEE Conf. on Pattern Recognition, Washington, DC, USA, 2004, Vol. 4, pp. 933–936
- 5 Marchand, E., and Chaumette, F.: 'Fitting 3d models on central catadioptric images'. Proc. IEEE Int. Conf. on Robotics and Automation, Rome, Italy, 2007, pp. 52–58
- 6 Mei, C., Benhimane, S., Malis, E., and Rives, P.: 'Homography-based tracking for central catadioptric cameras'. Proc. Int. Conf. on Intelligent Robots and Systems, Beijing, China, 2006, pp. 669–674
- 7 Demonceaux, C., and Vasseur, P.: 'Omnidirectional image processing using geodesic metric'. Proc. Int. Conf. on Image Processing, 2009, pp. 221–224