



HAL
open science

An Altered Image Alignment Technique for 3D Motion Estimation of a Reflective Sphere

Yara Bachalany, François Cabestaing, Christophe Vieren, Sébastien Ambellouis

► **To cite this version:**

Yara Bachalany, François Cabestaing, Christophe Vieren, Sébastien Ambellouis. An Altered Image Alignment Technique for 3D Motion Estimation of a Reflective Sphere. First Workshop on Image Processing Theory, Tools and Applications, IPTA'2008, Nov 2008, Sousse, Tunisia. pp.1-8, 10.1109/IPTA.2008.4743735 . hal-00521100

HAL Id: hal-00521100

<https://hal.science/hal-00521100>

Submitted on 25 Sep 2010

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.

An Altered Image Alignment Technique for 3D Motion Estimation of a Reflective Sphere

Yara Bachalany*, François Cabestaing*, Christophe Vieren* and Sébastien Ambellouis†

* LAGIS - UMR CNRS 8146, Cite Scientifique, 59655 Villeneuve d'Ascq Cedex, France.

e-mail: yara.bachalany@ed.univ-lille1.fr, fcab@ieee.org, Christophe.Vieren@univ-lille1.fr

† INRETS-LEOST, 20, rue Elisee Reclus, 59655 Villeneuve d'Ascq Cedex, France.

e-mail: Ambellouis@inrets.fr

Abstract—Recovering 3D motion of reflective objects in image sequences is still a cumbersome problem for computer vision. One common approach is to track geometric features of the object such as contours and edges since they are rather insensitive to light reflections. However, such basic features fail to recover the actual 3D motion in some cases. For example, the external contour of a sphere rotating about one of its axes remains static in the image. In this paper, we propose a new approach to 3D motion recovery of a reflective sphere visible in an image sequence. Instead of tracking only geometric features, our technique makes use of texture information in a slightly modified image alignment method. Unlike in classical image alignment methods, texture information is processed differently whether it comes from a diffuse or a specular light component. Using this technique, we show that motion estimation is possible when dealing with reflective objects, but also that specular components can offer information about the 3D motion. Finally, we present some results obtained from the analysis of image sequences.

I. INTRODUCTION

Motion estimation and object tracking is a highly explored task in computer vision. New techniques evolve constantly and depend on the kind of motion (2D or 3D) to be analyzed, the number and state (steady or in motion) of used cameras, the constrained level of the environment (lighting sensitivity), the required accuracy of results, the number of tracked objects. . .

One common technique is to predict what will be observed and then match this prediction with the actual observation. The prediction can be applied on different types of features of the tracked object, such as geometric features or appearance information (intensity, texture, color).

Image alignment is a commonly used texture-based technique whose goal is to warp an image or a patch into another one that matches the most the analyzed image. Image alignment techniques emerged with the Lucas-Kanade algorithm [1]. Since, many extensions have been proposed to decrease its computational cost [2]. However, these methods yield bad results when we focus on tracking objects with reflective properties. In this case, the warping must be different for pixels showing diffuse or specular components since their apparent motions are not the same. For example, a sphere in rotation about one of its axes will show a static specular component while its diffuse component will apparently move.

Geometric model-based matching methods use geometric features of the tracked object such as corner-like features or

feature points [3], [4]. These methods avoid the matching inconsistency problem induced by the specular components. However, in some cases, geometrical information is not enough to reconstruct the 3D motion parameters. In addition, feature extraction and correspondence can be computationally expensive specially when the object is complex.

If available, it is possible to use the object texture and geometric model, the camera parameters as well as the lighting information. The scene is fully rendered and the set of unknown parameters is updated based on an optimization algorithm, such as the steepest descent, whose goal is to reduce the discrepancy between the rendered image and the analyzed one. The 3D model must be accurate in order to obtain satisfying results but, in this case, the rendering is slow. Moreover, this technique is even more time consuming when reflective properties are taken into account in the object model.

In this paper, our purpose is to estimate the 3D motion of a reflective sphere in every image of a sequence based on an image alignment technique chosen for its low computational cost. To avoid inconsistency problems between the transformed patch and the actual patch, we warp separately specular and diffuse components and then combine them to get the synthetic image. Those components are obtained from a priori information about the scene. Our system outputs the six parameters representing 3D motion of the sphere for every image of the sequence.

In the next section we discuss related work. The image alignment technique is described in section 3.2. Our algorithm is described in section 3.3. We present the results in section 4 and conclude in section 5.

II. PREVIOUS WORK

During a study undertook on image sequences showing foraging dragonflies, we were asked to locate precisely the dragonfly's head (*cf.* appendix). We noticed that given the reflective surface of the head and its geometrical properties, the problem was close to that of estimating motion of a textured reflective sphere. However, no reference was found that dealt explicitly with motion estimation of reflective objects. In this section, we present briefly some classical motion estimation approaches and some other innovative techniques.

As mentioned earlier, classical motion estimation methods are based either on shape information or on appearance information.

Appearance-based techniques, such as optical flow techniques, assume that the brightness of a given scene point is constant during the sequence. Their goal is to reconstruct the image velocity field representing the apparent motion of the object in the analyzed image, i.e. the optical flow. This field can then be used to reconstruct partially the 3D motion. Many techniques are used to estimate the optical flow. Barron and Fleet [5] divide them into four categories : differential methods which are based on the assumption of intensity conservation locally [2], globally [6] or a combination of both [7]; region based matching where several blocks of two successive images are compared yielding displacement vectors representing a more or less accurate result [8]; energy based methods and phase based techniques where motion is estimated by applying spatiotemporal filters such as Gabor filters [9].

All these methods provide good results when tracking objects with lambertian surfaces. However, given that the information gathered from the specular component is processed the same way as the information gathered from the diffuse component, the accuracy of these methods is expected to decrease as the specular component increases. In addition, since appearance-based techniques yield only the optical flow, further computations are required to estimate the 3D motion.

Shape-based techniques suffer from a high computational cost specially when a complex geometric model is involved. Hence, they are often preceded by an information extraction step in order to reduce the number of tracked features and therefore decrease the computational cost. A multitude of methods have been proposed based on tracking the image contours of a known model. A Hough transform has been used in [10] to extract lines which were then matched with the model line segments. However, these techniques fail to estimate rotations in the case of a spherical object.

One interesting approach is the use of a 2D (edge map) 3D geometric model-based approach in order to track complex objects. Polat *et al.* [11] use a multiple hypothesis tracking (MHT) along with a Hausdorff distance for analyzing the motion of multiple objects. This method is mainly used to track multiple objects while dealing with partial occlusion. However, as long as the texture is not modeled, this approach can't be used in our case.

Novel approaches endeavor to recover the 3D motion under variable illumination ([15], [14], [12], [13]). Most of those methods [15], [12] consider the case of a lambertian surface. Yang *et al.* [15], and Basri and Jacobs [12], consider the case of a lambertian object under directional illumination aiming respectively to compute the object motion, the illumination and the model texture. They have proved that even under arbitrary complex illumination a nine dimensional linear subspace is sufficient to recover the illumination function using spherical harmonics. Freedman and Turek [13] seek to estimate an

illumination invariant optical flow using graph cuts. Hager and Belhumeur [14] study the case of tracking (but not motion estimation) when complications such as changes in pose relative to the camera, changes in illumination relative to the light source and occlusions arise.

III. DESCRIPTION OF OUR APPROACH

A. Inputs

We assume that an image sequence showing a highly reflective sphere with center P_c in the camera coordinate and radius R is available. The camera is static with known pose and parameters. The goal is to estimate the 3D motion of the sphere, hence its position and orientation in every image. We assume that the sphere pose is known for the first frame. The sphere diameter (i.e. geometric model) and its surface properties (i.e. texture model) are also supposed to be known. Phong lighting is provided by a light spot of known intensity and position.

Our method is recursive : assuming that the sphere pose has been disclosed till image at time t (I_t) our algorithm seeks the 3D motion between times t and $t + 1$ by analyzing image at time $t + 1$ (I_{t+1}). In the following, we first explain the classical image alignment technique applied on two successive images and then describe our technique that derives from image alignment. The image reference, i.e. the first image of the analyzed sequence, is shown in figure 1 where the used texture is the earth texture. The specular reflection appears clearly near the middle of the sphere image as a large white spot.

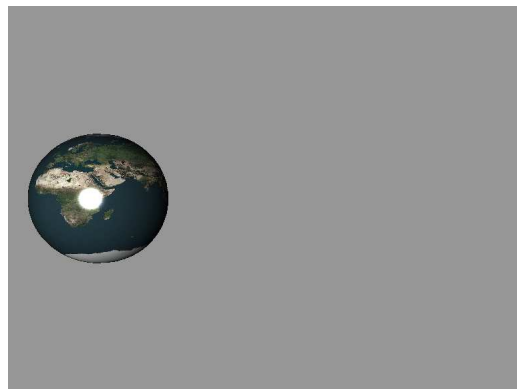


Fig. 1. Reference image

B. Image alignment technique

In their paper [1], B. Lucas and T. Kanade investigate the problem of aligning two functions $F(x)$ and $G(x)$ over some interval R (figure 2). This is done by minimizing the following expression with respect to h :

$$L_2norm = \left(\sum_{x \in R} [F(x+h) - G(x)]^2 \right)^{\frac{1}{2}}, \quad (1)$$

where $x' = x + h$ is the warping vector x undergoes.

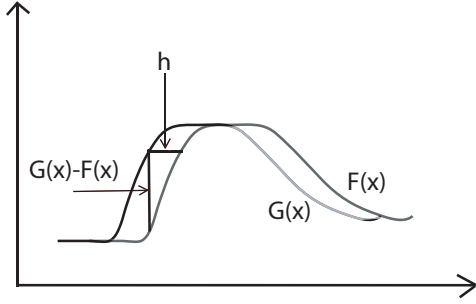


Fig. 2. Two curves to be aligned

When the functions to be aligned are bidimensional functions, $I_t(x, y)$ and $I_{t+1}(x, y)$ that represent the pixel intensities at position (x, y) of respectively image at time t and image at time $t + 1$ of a given sequence, the purpose is to find the warping matrix that minimizes the following error within a domain $R = [x_{min}, x_{max}] \times [y_{min}, y_{max}]$ of I_t :

$$D = \frac{1}{N_{pixels}^2} \sum_{(x,y) \in R} [I_t(x', y') - I_{t+1}(x, y)]^2, \quad (2)$$

where $(x', y')^T = M_{2D} \cdot (x, y)^T$, M_{2D} is the warping matrix and $N_{pixels}^2 = (x_{max} - x_{min} + 1) \cdot (y_{max} - y_{min} + 1)$ is the size of the domain or the total number of pixels.

In [1] the optimization method used is the Gauss-Newton algorithm. In our application, we use a steepest descent optimization which yields the same result more efficiently as soon as the warping matrix initially estimated is close to the actual one.

To reach convergence, the warping matrix M_{2D} must be the projection onto the image plane of a 3D motion defined by $P = (T_x, T_y, T_z, \theta_x, \theta_y, \theta_z)^T$, where T_x, T_y, T_z are the translational displacement of the object relative to the sphere reference and $\theta_x, \theta_y, \theta_z$ represent its angular displacement. Therefore, M_{2D} is a function of the parameters P of the 3D motion. The 2D warping matrix is computed as follows.

A pixel (x, y) (cf. equation 3) in the image reference is the projection of a point (X, Y, Z) in the camera reference. This can be expressed in a homogeneous representation :

$$\begin{pmatrix} sx \\ sy \\ s \\ 1 \end{pmatrix} = T_{2D} \cdot H_{2D} \cdot P_{3D} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad (3)$$

where P_{3D} is the 3D projection matrix onto the image plane of an ideal pinhole camera, H_{2D}, T_{2D} are respectively the 2D translation-rotation matrix and the scaling matrix. Z is computed using a sphere equation with known center position P_c and radius R . This is possible since for the first image the sphere position and geometrical model are perfectly known, and then they are updated for each image.

We consider that the 3D motion, with parameters P , transforms the point (X', Y', Z') into the point (X, Y, Z) following:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = T_c \cdot M_{3D}(P) \cdot T_c^{-1} \cdot \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}, \quad (4)$$

where T_c is the matrix translating the origin to the sphere position at time t and $M_{3D}(P) = T(P) \cdot R_y(P) \cdot R_x(P) \cdot R_z(P)$. The homogeneous rotation matrices $R_x(P), R_y(P)$ and $R_z(P)$ and the homogeneous translation matrix $T(P)$ are computed using the parameter vector P .

Combining (3) and (4) yields:

$$\begin{pmatrix} sx' \\ sy' \\ s' \\ 1 \end{pmatrix} = M_{2D}(P) \cdot \begin{pmatrix} sx \\ sy \\ s \\ 1 \end{pmatrix}, \quad (5)$$

where $M_{2D}(P)$ denotes the following warping matrix:

$$M_{2D}(P) = T_{2D} \cdot H_{2D} \cdot P_{3D} \cdot T_c \cdot M_{3D}^{-1}(P) \cdot (T_{2D} \cdot H_{2D} \cdot P_{3D} \cdot T_c)^{-1}. \quad (6)$$

As mentioned earlier the purpose is to minimize equation (2) with respect to P by applying the iterative steepest descent algorithm. The parameter at iteration $i + 1$ is computed from the parameter at iteration i using $P_{i+1} = P_i - \mu \nabla I_t$, where μ is the step size and ∇I_t denotes the gradient of $I_t(M_{2D}(P)(x, y))$, i.e. the partial derivatives of $I_t(M_{2D}(P)(x, y))$ along each of the 6 coordinates of P .

Generally, this algorithm yields good results with a low computational cost. However, all pixels undergo the same warping process regardless of whether they represent specular or diffuse reflection. In reality, the two light components move with two different apparent motions under the influence of object motion with parameter P . For a rotating sphere, the specular component remains static while the diffuse component apparently rotates.

Let us consider the case of a pure rotation $\theta_y = 0.01$ rad about the vertical axis y of the sphere. To show the error induced by the specular component, we consider the case of a 40×40 (figure 3(a)) image patch centered on the white spot in the reference image. In this patch, we notice that the "reliable" information provided by the diffuse component is poor.

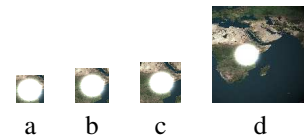


Fig. 3. Different patches used to induce different specular over diffuse ratios : (a) Patch 40×40 , (b) Patch 50×50 , (c) Patch 60×60 , (d) Patch 144×144

Indeed, when we compute the error vector relative to θ_y while the 5 other parameters are set to their real value (zero)

(dash dotted blue line in figure 4), we notice that the position of the minimum value of the error function does not coincide with the real value $\theta_y = 0.01$ rd. We can also compute the same error while considering different patch sizes so as to obtain different specular over diffuse ratios. In figure 4, the big dashed red line corresponds to the patch with size 50×50 of figure 3(b), the small dashed green line corresponds to the patch with size 60×60 of figure 3(c) and the solid black line corresponds to the patch with size 144×144 of figure 3(d). We can see that the minimum of the error function approaches the real value (0.01) of θ_y when the level of information contained in the diffuse component increases, i.e. when the patch becomes larger. Therefore, in order to guaranty a convergence toward the real value, we must consider a patch containing enough diffuse information. This can hardly be achieved when the patch position is automatically set or when the surface is highly specular.

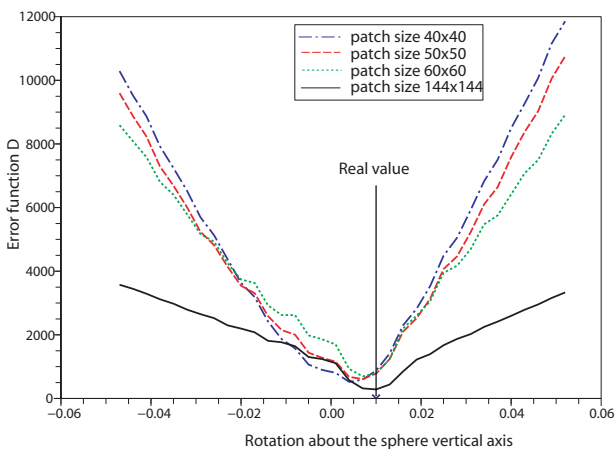


Fig. 4. Error D with respect to θ_y . Real value of θ_y which is equal to 0.01 rd, is not compatible with the minimum of error when considering small patches.

In fact, with this method, since both specular and diffuse components are processed with the same warp they are supposed to represent the same image motion. This is shown in figure 7(b) where "realistic" images of a rotating sphere are shown. Figure 7(a) shows the warped images relative to real values of θ_y : we can see how the warping matrix induces a rotation of the specular component. Therefore, this will cause inconsistency problems between the transformed image and the real one, which explains why the minimum value of the error function is not a good indication of the actual value.

C. Our approach

In the previously described technique we endeavor to align I_t and I_{t+1} by searching for the 3D motion parameters. Hence, inconsistencies between the warped patch and the real one are expected to drive the algorithm to diverge when diffuse information is not sufficient. In the following section we present an approach that enables us to overcome this difficulty in the case of a reflective sphere.

Given the geometry of the tracked object, in addition to

its texture model and a priori information about the lighting conditions and the camera parameters, it is possible to render a realistic image of this model. Here, we render it twice: the first time considering a pure lambertian surface to obtain the diffuse image (R_d , see figure 6(b)) and the second time considering a black highly reflective surface to obtain the specular image (R_s , see figure 6(a)). These images are then warped using two different warping matrices since their corresponding apparent motions are different. The diffuse image warping matrix $M_{2D}^d(P)$ is the same as $M_{2D}(P)$ of the regular method. The specular image warping matrix $M_{2D}^s(P)$ is computed using vector P as follows.

First, $M_{3D}(P)$ is computed using P and represents motion relative to the sphere reference. Then, the homogeneous matrix $M_{3D}(P)$ is modified by considering only the translational values, i.e. by setting the rotational parameters of the 3D motion to zero. This corresponds to computing the same matrix as $M_{2D}(P)$, but for a parameter vector P' equal to $(T_x, T_y, T_z, 0, 0, 0)^T$. The warping matrix $M_{2D}^s(P)$ for the specular component is therefore given by:

$$\begin{aligned} M_{2D}^s(P) &= M_{2D}(P') \\ &= T_{2D} \cdot H_{2D} \cdot P_{3D} \cdot T_c \cdot M_{3D}^{-1}(P') \cdot \\ &\quad (T_{2D} \cdot H_{2D} \cdot P_{3D} \cdot T_c)^{-1}. \end{aligned} \quad (7)$$

The final synthetic image used for parameter estimation by minimizing equation (2) combines information from both specular and diffuse sources and corresponds in our case, to a simple pixel to pixel addition of the two warped components:

$$I_t(x', y') = R_s(M_{2D}^d(P)(x, y)) + R_d(M_{2D}^s(P)(x, y)).$$

Using this approach we notice that the problem generated by the presence of specular components is now solved. The error function's minimum corresponds better to the real value of θ_y than in the previous case. Figure 5 shows this error function relative to the different patches shown in figure 3.

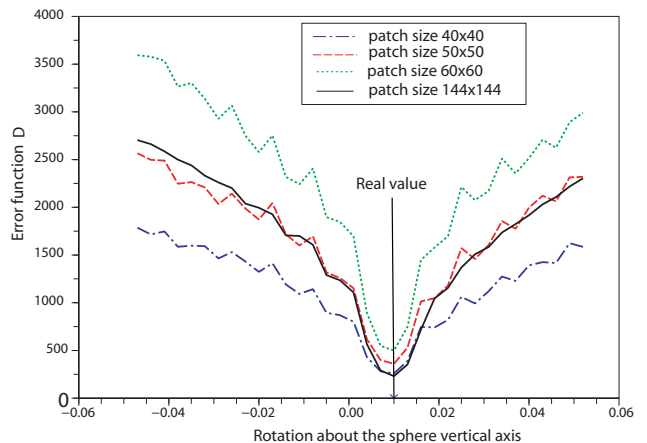


Fig. 5. Error D with respect to θ_y . Real value of θ_y which is equal to 0.01 rd, corresponds to the minimum of error even when considering small patches.

We see that even with a high level of specular reflection the error function still has its minimum marked at $\theta_y = 0.01$ rd. In fact, we can see in figure 7(c) that when we process the two components separately we can obtain a more realistic estimation of the transformed image.

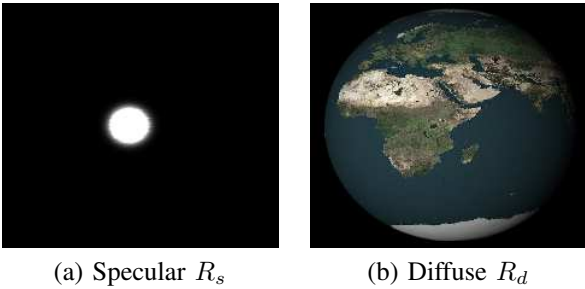


Fig. 6. Specular and diffuse images obtained by rendering 3D model set on the initial known location values.

IV. RESULTS

We first tested our algorithm on a series of six situations observed by two successive images. In those image pairs, only one basic motion was considered each time (arbitrary limited motion of 0.01 rd for rotations and between 1 and 2 pixels for translations). We tried recovering this motion by first comparing the two successive images then by comparing the sum of the transformed specular and diffuse images under fixed number of iterations (200 iterations). To do this we tried different patch sizes so as to have different specular over diffuse ratios.

We notice that when comparing successive images, if the chosen patch shows a high ratio of specular reflection, the tracking fails with the classical alignment function while our algorithm continues to give good results. Overall results are better with our algorithm than with the classical alignment approach even when we consider a purely translational motion. In this case, the warping matrix applied to R_d , $M_{2D}(P)$, is equal to the warping matrix applied to R_s , $M_{2D}(P')$. This means that we were able to use the specular component as an information about the real motion. It is clear that when the information provided by the diffuse component increases, hence when the patch gets larger, results are generally more accurate. Anyway, estimating the depth information is always troublesome using a single camera and errors are always rather important.

Please refer to <http://lagis-vi.univ-lille1.fr/~yb/recherche.php?langue=uk> to view videos showing these results.

Then, we tried our algorithm on a sequence of 150 images showing a small frame to frame motion. The patch size used is 40×40 . A bigger patch means more computational cost specially since Z is computed for each pixel (If we use the assumption of a plane at known Z the patch should also be small enough to limit the depth error). In this case, the patch position is chosen automatically as centered on the projection

of the previously computed center of the sphere. This means that the amount of specular component is not controlled and the patch can therefore either be with high specular presence or the opposite. The sphere texture should therefore be adequate (not homogeneous) for the matching to succeed. Our approach allows us to reconstruct the 3D motion rather accurately.

However, we notice from figure 9 that the error is cumulative. Fully rendering of the diffuse and specular components after a certain number of iterations should normally correct this error. In this case, the estimated parameters can be expected to be more accurate. This is the subject of a future work.

In order to validate this approach, we also tried different scenarios.

- The sphere diameter was approximated with an error of 10 and 15% yielding a loss of respectively 3 and 4.7% of accuracy.
- A Gaussian noise with a variance of 10 was added to the model's texture yielding 18% of loss of accuracy.
- A 10 degrees error in the illumination source position yielded less than 1% additional error.
- An error in the initial position of the sphere gave a 12% loss of accuracy.

With challenges imposed by the fact that for small motions, the x-translation (resp. y-translation) cannot be distinguished from the y-rotation (resp. x-rotation), results are better when specular reflections are present and used explicitly like in our approach. To verify this assumption, we have considered a single parameter motion, i.e. a translation along the y axis $T_y = 0.01$. We endeavor to estimate this motion on the same patch for the same gradient-descent parameters considering the case of a lambertian surface using the conventional image alignment technique and the case of a reflective surface using our approach. Results with a precision of 10^{-3} are shown in the following table :

	Conventional approach	our approach	sought values
T_x	0	0	0
T_y	-0.001	-0.008	-0.01
T_z	0.001	0	0
θ_x	-0.007	0	0
θ_y	0	0.001	0
θ_z	-0.004	-0.002	0

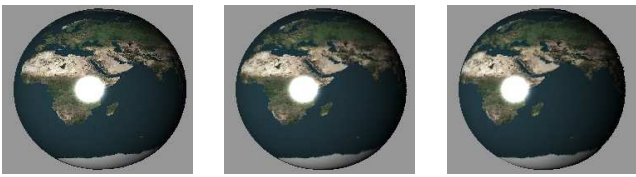
We can clearly see that the specularity reveals part of the ambiguity imposed by the similar 2D projection of the y-translation and the x-rotation 3D motions. In fact, T_y is estimated with an error of only 20%. When the surface is lambertian, this error reaches 90% and the y-translation is confused with an x-rotation which is approximated to -0.007 . In addition, in this case, overall errors (T_z , θ_z) are bigger when the surface is purely lambertian.

V. CONCLUSION AND FUTURE RESEARCH

The main contribution of this work is that it recovers the 3D motion of a highly reflective sphere. It consists in an image alignment technique processing differently specular and

diffuse components. We were able to obtain good results while considering the worst case scenario where the patch is positioned on a region of the image with a high level of specular reflections. We even noticed that the specular reflection is a source of information about the 3D motion not to be discarded. The extension to other objects will be taken into consideration in future works where we will consider a coarse to fine approach : given the geometric model as well as the texture model and the scene a priori information, we endeavor to match 2D rendered image with the analyzed image for more complex convex shapes.

Future works aim at using less prior information about the scene by automatically separating specular and diffuse components using the property of polarization [19] or color [20] instead of rendering the model twice. The challenge consists then in computing $M_{2D}^s(P)$.



(a) The 2D warping affects specular and diffuse components similarly here rotation of a sphere about θ_y on a patch of the size of the image (zoomed image)). The texture and the white spot move with the same apparent motion during the sequence.



(b) In reality, specular component remains static when the sphere rotates about one of its axes, i.e. the texture moves with an apparent motion to the left whereas the white spot remains at the same position.



(c) Our warping applied on a patch of the size of the image showing the same motion as in (b). Specular component remains static under pure rotation.

Fig. 7. Specular versus diffuse

VI. APPENDIX

Dragonflies are excellent predators. They wait, perched on a vegetation stick, and choose the appropriate moment to take

off after small insects. Prey pursuit strategy is remarkable, with success rates as high as 97% [17]. In addition, prey interception happens during the flight as the dragonfly swoops upwards from underneath its flying prey, grabbing it with its outstretched legs. This complex behavior is an example of visually guided interception, which is composed of at least three different processes: decision to take off after the prey, steering towards the prey and coordinating leg movements in time and space to grab the prey. The biologists are interested in understanding those separate but interdependent processes. The neural guidance system is also interesting for control scientists who look forward to developing effective biomimetic guidance mechanisms.

Since dragonflies do not normally forage in captivity, in order to capture their chase maneuver on video (figure 8), biologists had to reconstruct their natural environment inside a cage mounted outdoors. They attached a 2 mm white glass bead that resembles a small insect and moved it above the perching dragonfly. This setting attracted the dragonfly and drove it to start the pursuit. It was then possible to acquire high-speed videos (500 frames/s) with a camera fixed in a position and orientation allowing the dragonfly and its prey to remain in the camera viewfield during the whole pursuit. To restrict the behavior to a single plane, the bead was moved in the same plane as the dragonfly, on a path orthogonal to the camera optical axis. Biologists analyzed those sequences manually, while trying to validate a certain number of hypotheses in which the head location was essential [17], [16], [18].



Fig. 8. A perching dragonfly

This task is cumbersome due to several obstacles: it is not evident to define marks on the dragonfly's head due to the non consistency of its texture (specular reflections, etc. . .), the small number of sequences where the head is not occluded by other parts of the body or blurred, assumptions made on the size of the dragonfly's head and body used to estimate the depth.

Contributions of automatic analysis : automatic analysis of the image sequences would therefore be very useful. These analysis will have the advantage of avoiding the time cost required to extract different information from the sequence images. In addition, we expect it to be more accurate by avoiding human errors induced by manual analysis. Finally, automatic analysis have the advantage of reconstructing directly the 3D motion and do not require further computations based on point correspondence.

REFERENCES

- [1] B. D. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. In *Int. Joint Conf. on Artificial Intelligence*, 674–679, Vancouver, British Columbia, Canada, August 1981.
- [2] S. Baker and R. Gross and I. Matthews. Lucas-Kanade 20 Years On: A Unifying Framework: Part 4. Robotics Institute, Carnegie Mellon University, 2004.
- [3] S. M. Smith. ASSET-2: real-time motion segmentation and shape tracking. In *Proceedings of the Fifth International Conference on Computer Vision*, 237, 1995.
- [4] R. Polana and R. Nelson. Low level recognition of human motion. In *Proc. of IEEE CS Workshop on Motion of Non-Rigid and Articulated Objects*, 77–82, 1994.
- [5] J. L. Barron and D. J. Fleet and S. S. Beauchemin and T. A. Burkitt. Performance Of Optical Flow Techniques. In *Computer Vision and Pattern Recognition*, 236–242, 15-18 Jun 1992.
- [6] B. K. P. Horn and B. G. Schunck. Determining Optical Flow. Massachusetts Institute of Technology, Cambridge, MA, USA, 1980.
- [7] A. Bruhn and J. Weickert and C. Schnörr. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. *Int. J. Comput. Vision*, 61(3) :211–231, 2005.
- [8] P. Anandan. A Computational Framework and an Algorithm for the Measurement of Visual. University of Massachusetts, Amherst, MA, USA, 1987.
- [9] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1) :77–104, 1990.
- [10] P. Wunsch and G. Hirzinger. Real-time visual tracking of 3d objects with dynamic handling of occlusion. In *IEEE International Conference on Robotics and Automation*, Albuquerque, Mexico, apr 1997.
- [11] P. Ediz Y. Mohammed and S. Rajeev. A 2D/3D modelbased object tracking framework. *Pattern recognition*, 36(9) :2127–2141, 2003.
- [12] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2) :218–233, 2003.
- [13] D. Freedman and M.W. Turek. Illumination-invariant tracking via graph cuts. In *Comp. Vision and Pattern Recognition*, San Diego, CA, USA June 20-25.
- [14] G. D. Hager and P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(10) :1025–1039, 1998.
- [15] H. Yang, G. Welch, and M. Pollefeys. Illumination insensitive model-based 3d object tracking and texture refinement. In *Third International Symposium on 3D Data Processing, Visualization, and Transmission*, University of North Carolina, Chapel Hill, USA June 14-16, 2006.
- [16] R. M. Olberg and A. H. Worthington and J. L. Fox and C. E. Bessette and M. P. Loosemore. Prey size selection and distance estimation in foraging adult dragonflies. *J Comp Physiol A*, 191(9) :791–797, 2005.
- [17] R. M. Olberg and A. H. Worthington and K. R. Venator. Prey pursuit and interception in dragonflies. *J Comp Physiol A*, 186(2) :155–162, 2000.
- [18] R. M. Olberg and R. C. Seaman and M. I. Coats and A. F. Henry. Eye movements and target fixation during dragonfly prey-interception flights. *J Comp Physiol A*, 193(7) :685–693, 2007.
- [19] S. Umeyama and G. Godin. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5) :639–647, 2004.
- [20] G.J. Klinker and S.A. Shafer and T.Kanade. Using a Color Reflection Model to Separate Highlights from Object Color. In *International Conference on Computer Vision*, 145–150, London, June 8-11, 1987.

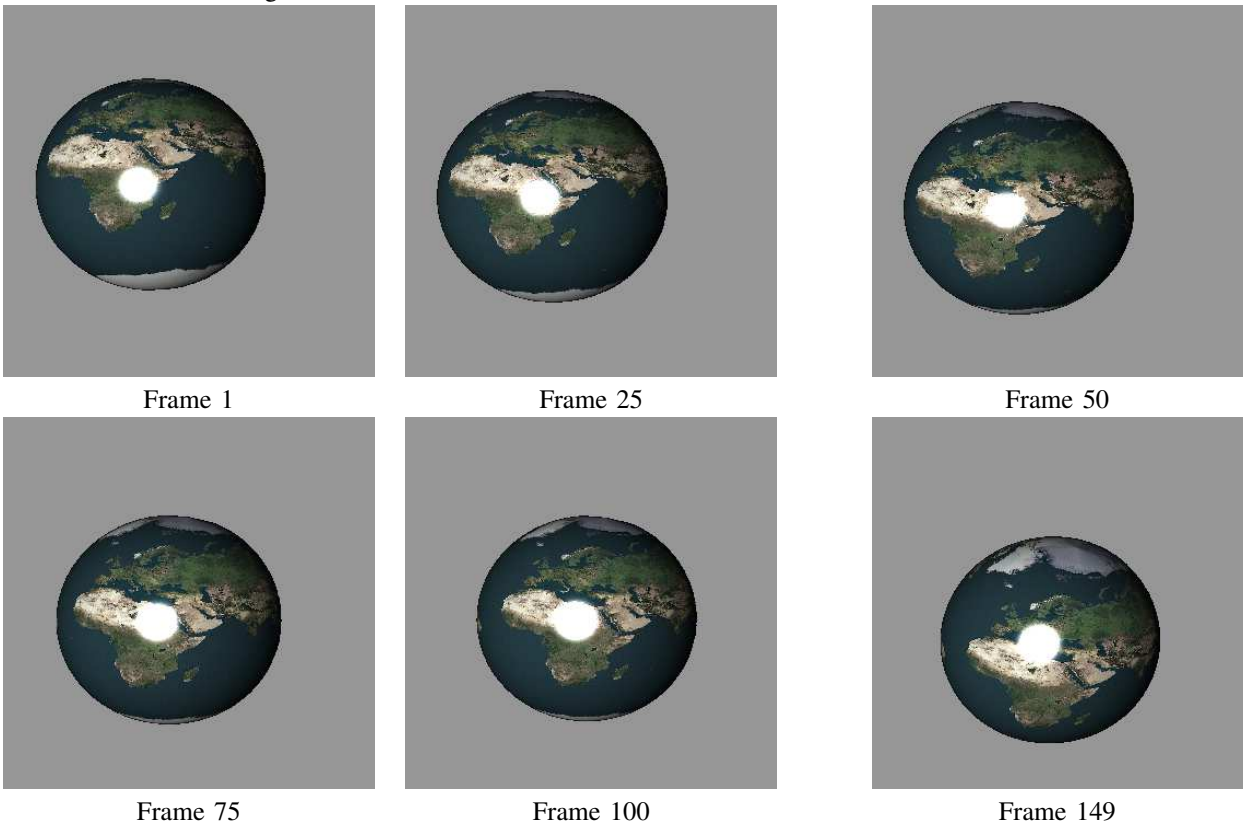
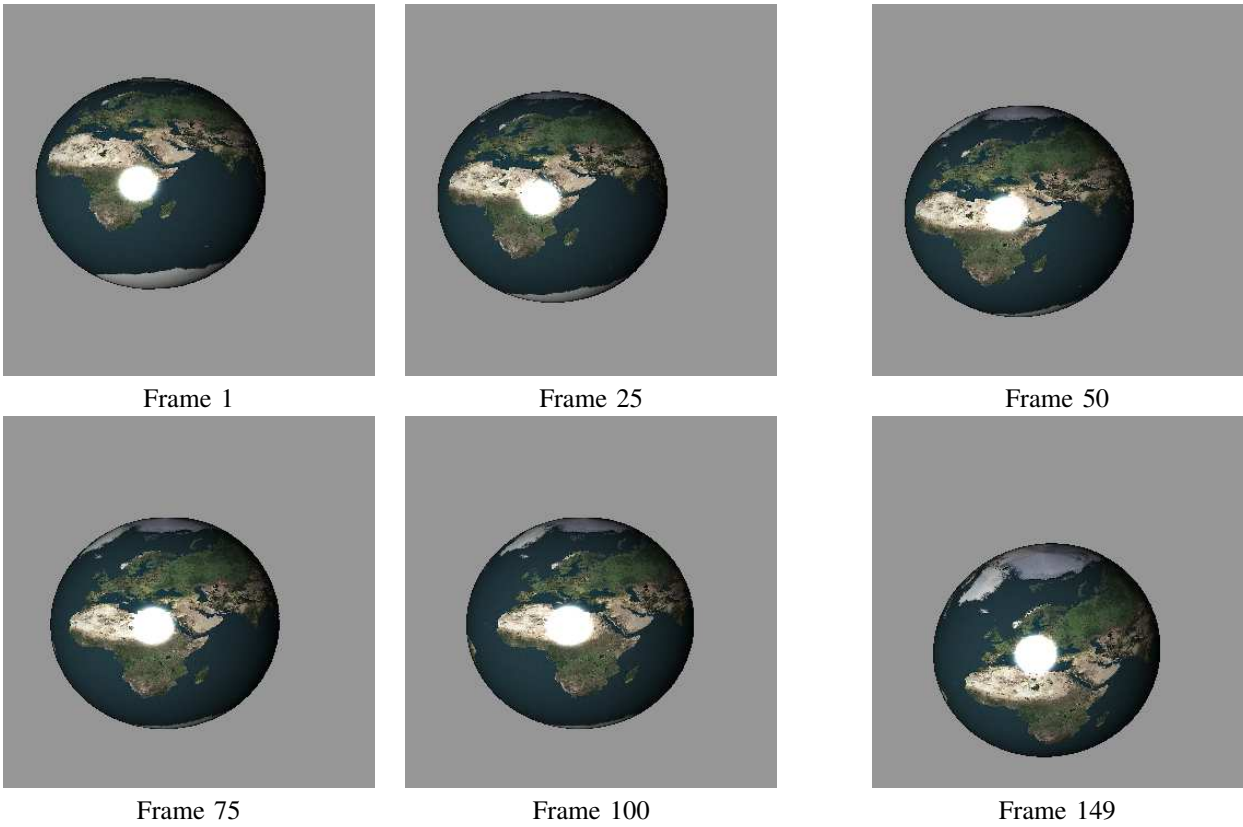


Fig. 9. Comparison between the actual sequence and the sequence reconstructed based on our approach for a patch size of 40×40 .