Tracking dragonflies in image sequences

Yara Bachalany[1] François Cabestaing[1] Christophe Vieren[1] Sébastien Ambellouis[2] Robert Olberg[3]

[1] LAGIS - UMR CNRS 8146
Cite Scientifique
59655 Villeneuve d'Ascq Cedex, France

[2] INRETS-LEOST
20, rue Elisee Reclus
59655 Villeneuve d'Ascq Cedex, France

[3] Department of Biological Sciences
Union College
Schenectady, NY 12308, USA

Abstract

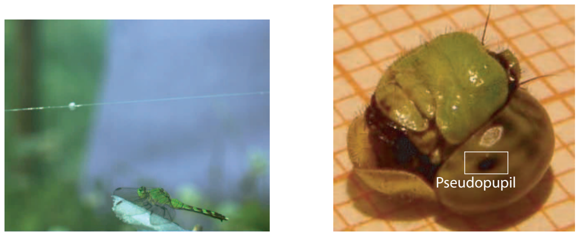
This paper investigates the problem of estimating the position and the 3D motion of a dragonfly from an image sequence. We propose to tackle this problem using a hybrid 2D/3D approach that deals with reflections constraints while requiring a limited computational cost. The algorithm is divided into two steps. First, a 2D alignment between two successive images computes a first estimation of the 3D parameters. Secondly, a refinement step is applied by matching the rendered image created using a complete 3D model of the dragonfly with each image of the sequence. Because image alignment fails to deal with large displacements, a prediction step is added. This step is based on a Taylor transform and takes into account a specific dynamic and kinematic model and a statistical behavior model to yield a first coarse approximation of the 3D model.

1 Introduction

1.1 Context

A complex foraging strategy : 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% [9]. 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.

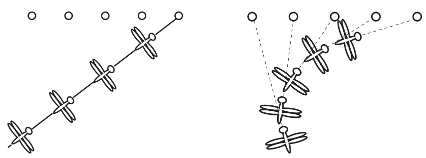
Experiments and initial analysis : Since dragonflies do not normally forage in captivity, in order to capture their chase maneuver on video (figure 1(a)), biologists had to reconstruct their natural environment inside a cage mounted outdoors. They attached a 2 mm white glass bead that resembles a small 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.



(a) A perching dragonfly (b) Pseudopupil phenomenon

Figure 1. Acquired images

Pursuit tracks: a first manual analysis [9] easily proved that the dragonfly steers to intercept its prey (figure 2(a)) instead of heading directly toward its prey (figure 2(b)). In fact, the straight line flight path indicates that the dragonfly predicts its prey's position.



(a) Interception (b) Tracking

Figure 2. Two possible pursuit tracks

Response latencies: inaccurate early studies [9] showed that the dragonfly's wing correction response to changes in its prey's trajectory comes after 33 ms. The dragonfly's course correction latency is of 50 ms. However, novel studies [7], applied on sequences with better time and spatial resolution, showed different latencies : 28 to 36 ms.

Head orientation: the biologist's purpose in [7] was to prove that the dragonfly moves its head so as to keep the image of its prey at a fixed position on the eye, in opposition to what had been proved in [9]. 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, pseudopupil), 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.

Estimating distances: the decision to take off is probably the result of a distance estimation performed by the dragonfly. An analysis, undertook in [8], aims at revealing it by trying two different hypotheses. The first suggests that the dragonfly moves its head and legs while perched on the vegetation to acquire stereoscopic images of its prey used to estimate its distance and its size. The second suggests that the sharp increase or decrease in angular velocity as the prey passes overhead is an indication of its vicinity since far away objects maintain a relatively constant angular velocity on the retina. This study requires an accurate estimation of the head position and orientation before take off that manual analysis can hardly reveal.

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.

1.2 Motion Estimation Related Work

Given a video sequence of a foraging dragonfly, our purpose is to determine accurately its position in each image. This looks like a classical motion estimation and object tracking problem, which has been highly explored in computer vision. However, new problems arise in the dragonfly tracking application.

One classical method is to define a perfect geometric model of the tracked object and use it to render synthetic images of the scene. Unknown model parameters are updated using an optimization algorithm that reduces the discrepancy between synthetic images and actual ones. The 3D model must be very accurate, and the rendering of synthetic images is computationally expensive. For a dragonfly, an accurate texture model is even harder to determine since, in addition to diffuse and specular components, one must take into account the pseudopupil phenomenon appearing in insects' eyes: some regions of the eye appear dark due to the absorption of light rays by the omatidiae that are oriented in the same direction as the camera (figure 1(b)).

Analyzing motion of standard image features, such as contours [10], is useless since this information is of no use for determining the 3D motion. For example, the dragonfly's head that can be modeled by a sphere will show a static external contour when rotating about one of its axes. Image alignment techniques, using texture information such as the Lucas-Kanade algorithm [5], seem more appropriate. Image alignment's goal is to warp an image or a patch into another that matches the most the analyzed image.

However, these methods used alone can yield bad results when tracking objects with reflective properties, because of the inconsistency that exists between the predicted patch and the actual one. In our application, in addition to specular reflection, pseudopupil phenomenon on the dragonfly's eyes is expected to drive the image alignment algorithm to fail. The approach that we suggest is a hybrid 2D/3D technique taking advantage of the high computational speed of the 2D alignment technique and the robustness relative to specular reflections and pseudopupil phenomenon of the 3D model based method (section 2.2).

Besides the challenges posed by the geometric model and the texture model, the rapid motion of the dragonfly results in a frame to frame image motion sometimes greater than 10 pixels, which is a problem for classical motion estimation approaches. We suggest to tackle this problem by introducing efficient predictors in the processing chain, based either on a statistical analysis of the dragonfly's behavior, on a Taylor transform or on a dynamic and kinematic model (section 2.3). In section 3, we conclude and we present some perspectives.

2 Description of our approach

2.1 Inputs

Image sequences showing the dragonfly capturing its prey are available, acquired by a static camera with known pose and optical parameters. The goal is to estimate the 3D motion of the dragonfly's head, thorax and abdomen, hence their position and orientation in each image. We assume that those parameters are known for the first frame. Its geometric model as well as its texture model are also known. Our method is recursive : assuming that the sought parameters have 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}). We first extrapolate the model thanks to the predictors and then apply the hybrid 2D/3D technique described in what follows. This hybrid approach allows us to update the 3D model and keeps it as realistic as possible avoiding therefore the accumulation of the matching error.

2.2 A hybrid 2D/3D approach

In their paper [5], B. Lucas and T. Kanade investigate the problem of aligning two successive images I_t and I_{t+1} by finding the warping matrix that transforms an image or a patch into another that matches the most the analyzed image by minimizing the SSD (sum of squared differences) error function over respective pixels of both patches. The minimization algorithm used is the steepest descent. This algorithm gives good results with low computational cost. However, all pixels undergo the same warping regardless of the kind of image component they represent. In our application, three different types of image components are considered: specular reflections, diffuse reflections as well as reflections due to the pseudopupil phenomenon. In reality, apparent motions of these components are different for the same movement of the object. For example, on a rotating sphere the specular component remains static in the image while the diffuse component moves.

Our approach is a combination of the image alignment estimation method and the model based matching technique. The goal is to reach the same accuracy as the model based algorithm without involving a high computational cost. A first coarse approximation of the 3D motion parameters is computed thanks to the alignment of I_t and I_{t+1}. This approximation is used to update the 3D model whose rendering yields image R^1_{t+1}. R^1_{t+1} and I_{t+1} are compared using the SSD error function. Until the error is small, the 3D motion refinement is repeated thanks to successive alignment and model match steps.

Until now we have validated our approach on synthetic images showing a highly reflective textured sphere. Figure 3 shows an example of the error convergence when matching images in figure 1 and 2. Each constant level represents an approximation level (i) of the 3D parameters and each peak represents the error between patches of the rendered image R^i_{t+1} and the real image. However, applying this approach to the dragonfly's images is still cumbersome since its success is closely dependent on the accuracy of the model of the dragonfly (geometry and texture modeling) and of the scene (lighting and camera parameters) which is still an unresolved problem.

2.3 Predictors

Taylor's Transform : As mentioned earlier, the rapid motion of the dragonfly implies a frame to frame motion projection sometimes greater than 10 pixels. This represents a challenge to classical motion estimation algorithms. Our proposed solution is to help our

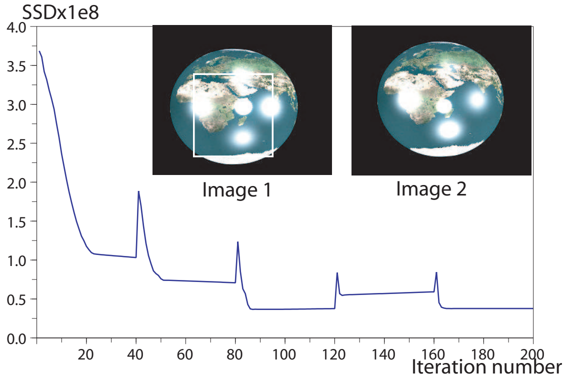


Figure 3. Example of convergence.
The error function convergence with respect to the iteration number when matching images 1 and 2. The rectangular patch is drawn in white.

motion estimation algorithm with predictors that offer a first approximation of the sought 3D parameters.

One of the predictors we endeavor to use is based on a Taylor transform described in what follows. As a first step validating this predictor, we start by manually extracting the tip of the dragonfly's tail coordinates over the different frames of a sequence.

$$\mathcal{D}(t)^T = (x(t), y(t))^T. \tag{1}$$

($\mathcal{D}$ as $\mathcal{D}$ragonfly's tip of tail). One should note that this manual extraction leads to a noisy signal (figure 4). We suppose that the successive positions of the tip of the tail is a sufficiently smooth series, which means that at any point of the equivalent time signal, the derivatives up to a certain order are continuous. So at a given time t we use past information of the signal on some interval $[t-T, t]$ to compute successive derivatives $\mathcal{D}^n(t)$, where n denotes the order of differentiation. Using the Taylor expansion, the prediction of the tip of the tail position on a finite interval of time $[t, t+\delta]$ is:

$$\tilde{\mathcal{D}}(t+\delta) = \mathcal{D}(t) + \dot{\mathcal{D}}(t)\delta + \ddot{\mathcal{D}}(t)\frac{\delta^2}{2} + \cdots + \mathcal{D}^{(n)}(t)\frac{\delta^n}{n!}. \tag{2}$$

Note that as δ gets bigger, the prediction becomes less accurate. The novelty in the approach is the derivative estimation. Recall that successive derivative estimation of a noisy signal is a longstanding ill-posed problem in numerical analysis and signal processing. Robust and fast time differentiation of noisy signals is now possible, thanks to an algebraic approach initiated in [2], and adapted to signal derivation in [1, 6]. Advanced results on algebra-based derivative estimation can be found in [6]. We use those results to estimate the derivatives of the abscissa which are used to estimate equation 2. The prediction of the abscissa on a finite future time $\delta = 2,\ 4,\ 5$ are shown in figure 4. By looking at this

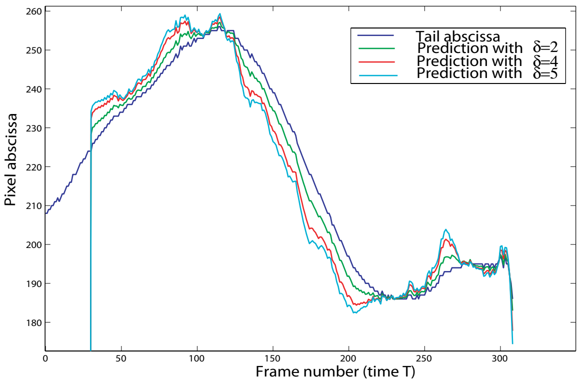


Figure 4. Tip of the tail prediction

figure, one can notice that the quality of prediction degrades with increasing δ. In addition, predictions start at frame 40 due to the fact that derivative estimation is accomplished over a sample of 40 observations.

Finally, note that our predictor ensures robustness to noise measurements, ease of implementation and low cost computation. This same prediction can be applied on the six 3D motion parameters instead of the tip of the tail's 2D coordinates. However, it is important to note that in the above computations the signal is supposed to be analytic which is a strong assumption.

Other predictors : We have studied other predictors, such as kinematic and dynamic model based predictor. A study of the dragonfly's anatomy was essential, where physical properties of the dragonfly such as its body parts masses as well as their dimensions were gathered. However, those parameters can significantly fluctuate between different dragonflies depending on their environment and their diets. As for the muscle forces applied over the articulations they can be modeled as a combination of spring and damper [3]. We lacked many parameters such as the stiffness gain, the tonic stiffness, and the damping coefficient of the muscles before being able to use this model. Finally, as for modeling the aerodynamic forces we have two possibilities: the first considering a steady state aerodynamics assumption as in [4] or unsteady aerodynamics. The first provides ease of computations but lacks in accuracy. The second is still not well developed.

Another predictor is under investigation: a statistical modeling of the dragonfly's behavior. In fact, given the number of available image sequences, this study is interesting to perform. can tackle this problem differently, i.e. taking the behavior model the biologists support and trying to validate this model by using it as a predictor.

3 Conclusion

We have presented in this article an approach to tracking foraging dragonflies on video sequences. It is a combination of a 2D image alignment (for its low computational cost) and a 3D model based approach (for its accuracy). However, this approach faces some challenges, such as the presence of an additional texture component due to the pseudopupil phenomenon. A condition for our technique to converge is to find an accurate model taking into consideration the 3D geometry as well as the diffuse, specular and pseudopupil components. Meanwhile, it has been validated over synthetic sequences showing a moving reflective sphere.

Dealing with the rapid motion of the dragonfly, which yields frame to frame displacements of more than 10 pixels, is also a challenge. We endeavor to tackle this problem with the aid of predictors of which Taylor's transform has been investigated and validated over the signal representing the 2D position of the tip of the dragonfly's tail manually extracted from one of the available sequences. Further studies are to be applied so as to acquire a kinematic and dynamic model as well as a statistical behavior model. An essential step is yet to be accomplished in order to validate the whole approach over real images which is modeling the scene i.e. geometric and texture model of the dragonfly as well as the camera and the lighting parameters.

References

[1] M. Fliess, C. Join, M. Mboup, and H. Sira-Ramirez. Compression differentielle de transitoire bruités. *C.R.Acad. Sci.*, 339(1):821–826, 2004.

[2] M. Fliess and H. Sira-Ramirez. An algebraic framework for linear identification. *ESAIM:COCV 9*, pages 151–168, 2003.

[3] A. J. Ijspeert. A 3-d biomechanical model of the salamander. In *Lecture notes in computer science*, volume 1834, pages 225–234.

[4] W. Lai, J. Yan, M. Motamed, and S. G. and. Force measurements on a scaled mechanical model of dragonfly in forward flight. In *12th ICAR '05*, pages 595– 600.

[5] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI81*, pages 674–679, 1981.

[6] M. Mboup, C. Join, and M. Fliess. A revised look at numerical differentiation with an application to nonlinear feedback control. In 15^{th} *MED 2007*, june 2007.

[7] R. M. Olberg, R. C. Seaman, M. I. Coats, and A. F. Henry. Eye movements and target fixation during dragonfly prey-interception flights. *J Comp Physiol A*, Volume 193(7):685–693, July 2007.

[8] R. M. Olberg, A. H. Worthington, J. L. Fox, 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, Sept. 2005.

[9] R. M. Olberg, A. H. Worthington, and K. R. Venator. Prey pursuit and interception in dragonflies. *J Comp Physiol A*, 186(2):155–162, Feb. 2000.

[10] S. M. Smith. Asset-2: real-time motion segmentation and shape tracking. In *ICCV '95*, page 237, Washington, DC, USA, 1995. IEEE Computer Society.