Navigation in an autonomous flying robot by using a biologically inspired visual odometer

Fumiya Iida* and Dimitrios Lambrinos
AILab, Department of Information Technology, University of Zurich
Winterthurerstr. 190, 8057 Zurich, Switzerland

ABSTRACT
While mobile robots and walking insects can use proprioceptive information (specialized receptors in the insect’s leg, or wheel encoders in robots) to estimate distance traveled, flying agents have to rely mainly on visual cues. Experiments with bees provide evidence that flying insects might be using optical flow induced by egomotion to estimate distance traveled. Recently some details of this "odometer" have been unraveled. In this study, we propose a biologically inspired model of the bee's visual "odometer" based on Elementary Motion Detectors (EMDs), and present results from goal-directed navigation experiments with an autonomous flying robot platform that we developed specifically for this purpose. The robot is equipped with a panoramic vision system, which is used to provide input to the EMDs of the left and right visual fields. The outputs of the EMDs are in later stage spatially integrated by wide field motion detectors, and their accumulated response is directly used for the odometer. In a set of initial experiments, the robot moves through a corridor on a fixed route, and the outputs of EMDs, the odometer, are recorded. The results show that the proposed model can be used to provide an estimate of the distance traveled, but the performance depends on the route the robot follows, something which is biologically plausible since natural insects tend to adopt a fixed route during foraging. Given these results, we assumed that the optomotor response plays an important role in the context of goal-directed navigation, and we conducted experiments with an autonomous freely flying robot. The experiments demonstrate that this computationally cheap mechanism can be successfully employed in natural indoor environments.

Keywords: 3-D navigation, Visual odometer, Elementary Motion Detector, Flying robots, Biorobotics

1. INTRODUCTION
Navigation in complex 3-D environments is very challenging for both animals and robots. Flying insects solve this task in a robust and adaptive way despite their tiny brains. Being able to navigate in 3 dimensional space offers several advantages. Flying insects, for example, can reach target locations (e.g. food sources, or their nest) faster and easier, they have the birds eye perspective of their environment, and they have increased access to places, that are not reachable by land. However, this also introduces a number of additional issues that need to be addressed. For example, parsimony of control strategies is essential, since flying agents can not simply pause - they must make control decisions in a timely manner. Robustness of control strategies is another issue - while in wheel-based robots it is possible to make a sequence of bad decisions without endangering the integrity of the robot, this will be disastrous for a flying robot. There is also a much higher pressure to develop mechanisms that are able to cope with unanticipated situations. Abilities like flight stabilization, obstacle avoidance, safe landing, or target tracking, will often have to rely on vision. In some cases, required navigational accuracy and small size operating areas preclude the use of most commercially available equipment like GPS or radar. Especially dead reckoning in a small range can only be done supported by visual information, since proprioceptive errors are much larger compared to land-based vehicles.1

Bees make use of both landmark and dead reckoning information for navigating, and can communicate the distance and direction to a newly found food source to their nest mates. Use of dead reckoning requires that both directional and distance information be available. Whereas in direction estimation there is a lot of evidence (starting with the pioneering work of von

* Correspondence: E-mail: iida@ifi.unizh.ch, Telephone: +41-1-63-54343, Fax: +41-1-63-56809
Frisch\textsuperscript{2}) that celestial cues and especially the polarization pattern of the sky play crucial role\textsuperscript{3,4} the way that bees gain information about the distance traveled has been a point of dispute for many years.

Early studies suggested that the distance is gauged in terms of total energy consumption during a foraging journey, but recent studies of the bees’ behavior questioned this hypothesis, and suggested that visual cues, more specifically the amount of image motion, play an important role in estimating the distance traveled.\textsuperscript{5,6,7} In these studies, bees were initially trained to search for the food source in the vertically striped tunnel. Then the searching behavior of the trained bees was observed when the pattern of stripes on the walls was changed, by using different stripe periods, different width between right and left walls, and walls with horizontal stripes. With horizontal stripes, bees tend to fail searching for the food source at the right location, therefore it was concluded that vertical stripes are used for the visual odometer. With different period of stripes, bees succeed in finding the food source position, which implied that bees measure the distance independent of the spatial frequency of the stimulus in the environment. However, when the distance to both walls was increased or decreased compared to the training situation, bees tend to fail in searching at the right location. The main conclusion from these studies was that the bee’s visual odometer provides an estimate of distance that is independent of the spatial frequency of the visual stimulus, i.e., it only depends on the angular velocity of the image projected on the insect’s retina.

Although there is a lot of behavioral evidence that bees might be using optical flow for estimating distance, there is very little known of how this could be implemented neurally. A visual odometer for mobile robots was proposed, but this odometer is based on an image interpolation algorithm rather than optical flow.\textsuperscript{8} On the other hand, there is a lot of work done on insect motion perception and its use on stabilizing flight. The optomotor response of flying insects has been investigated for a long time. Reichardt reported on the optomotor response of the fly almost thirty years ago.\textsuperscript{9} he demonstrated that a tethered fly inside a striped drum tends to turn in the direction in which the drum is rotated. This response helps the insect maintain a straight course by compensating for undesired deviations. The landing response of the fly was investigated with case studies on the extending leg response for the preparation to land and deceleration of speed before landing.\textsuperscript{10} Those studies imply that flies might be using image motion of the texture on the ground for producing a landing response. Such a scheme would take advantage of the fact that texture in the image moves faster as the insect comes close to the ground. In addition, bees appear to use a similar strategy to regulate their flight speed,\textsuperscript{5} according to which bees decrease their flight speed as the vertically striped tunnel becomes narrower so as to keep the angular velocity of the image constant.

On the basis of above mentioned behavioral experiments as well as electrophysiological studies, a model of motion detection in the insect’s nervous system, the so-called Elementary Motion Detector (EMD), has been proposed.\textsuperscript{11} This biological motion detection mechanism is also very interesting from the engineering perspective, because traditional optical flow approaches require large computational resources.\textsuperscript{12} Some attempts to incorporate motion detector mechanisms to control artificial agents in both real-world robots and computer simulations have been reported in the last few years. An autonomous agent capable of avoiding obstacles, while moving very fast, by exclusively using optical flow was implemented by Franceschini et al.\textsuperscript{13} The visuomotor system of the robot was based on the study of the neural mechanisms of motion detection in the compound eye of the housefly. It consists of an array of EMDs capable of analog, continuous-time processing, which were used to steer the robot directly. In a similar way, EMDs were used in a simulated flying agent for altitude control, to avoid crushing on the ground.\textsuperscript{14,15} Due to their simplicity, analog VLSI implementation of EMDs has been also proposed.\textsuperscript{16}

Another advantage of EMDs as can be seen from the above mentioned applications, is that they can be used in different parts of the visual field for different purposes. EMDs at the periphery can be used to evaluate lateral flow as well as the motion of the ground for evaluating flight speed or distance covered, whereas EMDs tuned in the center of the visual field can be used for detecting objects and avoiding them. Being able to reuse the same basic circuit is very important for flying agents since there are normally very strict constraints on the amount of payload the agent can carry.

In this paper we introduce a biologically inspired visual odometer based on EMDs and present and analyze data of its performance on a flying robot that has to perform a navigation task in an indoor unstructured environment.
2. ELEMENTARY MOTION DETECTORS AND THE VISUAL ODOMETER

2.1 The basic EMD circuit

As mentioned earlier a large part of the flying insect’s behavior seems to rely on image motion. Their motion detection mechanism has been investigated in electrophysiological level for a long time,17 and a computational model of motion detection, the Elementary Motion Detector (EMD), has been proposed based on the results from these studies.11 There are several versions of the EMD model, one of the most commonly used is the so-called Reichardt or correlation model, which consists of two photoreceptors, high-pass, low-pass filters, multipliers, and summation functions, as shown in Figure 1. An important feature of this model is that it sensitive not only to the magnitude of image motion but also to the direction of the motion, i.e. EMD responses are positive when image moves preferred direction, and negative response with non-preferred.

Here we propose a method to measure distance traveled using the Reichardt model. Since the EMD response provides an angular velocity signal (more exactly, a spatio-temporal signal that depends on the speed of the stimulus as well as its spatial frequency, see next subsection), the distance traveled could be estimated by integrating the EMD outputs over time.

The information processing steps we employed in the experiments are as follows. At first, intensity values from two adjacent pixels on a CCD camera image are used to calculate activity of a photoreceptor pair (\( \text{Photo}_{1i}(t) \), \( \text{Photo}_{2i}(t) \) of the \( i \)th EMD). Subsequently the output of the High-Pass Filters \( \text{HPF}_{1i}(t) \) and \( \text{HPF}_{2i}(t) \) at time \( t \) is calculated as follows:

\[
\begin{align*}
\text{HPF}_{1i}(t) & = \text{Photo}_{1i}(t) - \text{Photo}_{1i}(t-1) \\
\text{HPF}_{2i}(t) & = \text{Photo}_{2i}(t) - \text{Photo}_{2i}(t-1)
\end{align*}
\]

Given output of HPFs, the output of the Low-Pass Filters \( \text{LPF}_{1i}(t) \), \( \text{LPF}_{2i}(t) \) is calculated as follows:

\[
\begin{align*}
\text{LPF}_{1i}(t) & = a \text{HPF}_{1i}(t) + (1-a)\text{HPF}_{1i}(t-1) \\
\text{LPF}_{2i}(t) & = a \text{HPF}_{2i}(t) + (1-a)\text{HPF}_{2i}(t-1)
\end{align*}
\]

![Diagram](image1.png)

Figure 1. Left: the Reichardt model of elementary motion detection. Photoreceptors, high-pass filters, low-pass filters, multipliers, and the subtraction module are wired in series. The output of the last step (subtraction) is an estimate of image speed (see text for details). Right: the visual odometer based on a wide field motion detector. The integrated response of an array of EMDs, each tuned to a different part of the visual field, is used to implement a wide field motion detector. Its outputs are then integrated over time to provide an estimate of the distance traveled (for explanations, see text).
where $a$ is a parameter for the time delay. Then each LPF output is multiplied with the output of its neighboring HPF:

$$\text{MLP}_1(t) = \text{LPF}_1(t) \times \text{HPF}_2(t)$$

(5)

$$\text{MLP}_2(t) = \text{LPF}_2(t) \times \text{HPF}_1(t)$$

(6)

The output of an EMD module is finally computed as:

$$\text{EMD}_i(t) = (1 - b) \times \text{EMD}_i(t - 1) + b \times (\text{MLP}_1(t) - \text{MLP}_2(t))$$

(7)

where $b$ is another delay parameter implementing a low-pass filter on the EMD's output to reduce noise.

By putting many EMD modules in parallel we can increase the visual field size, thus simulating the wide field motion sensitive H1 interneurons in the insect's brain:

$$H_1(t) = \sum_i \text{EMD}_i(t)$$

(8)

At the end the visual odometer response, $\text{VO}$, is calculated by integrating the $H_1$ output over time:

$$\text{VO} = \sum t H_1(t)$$

(9)

2.2 Initial experiments using a panoramic camera

To evaluate the performance of our model, we developed a miniature panoramic vision system, which consists of panoramic mirror, a CCD camera module, and the housing components. The system was designed to weight no more than 90 g in total so that we could later integrate it to the flying robot (Figure 2). The panoramic mirror was developed based on a panoramic optics study and it has a hyperbolic surface that provides a visual field of 360 degrees on horizontal plane and 260 degrees vertically. In the first experiment, this vision system was equipped with the caster wheel frame on a pair of rails, and a motor system which drives the vision system at constant speeds along a horizontal straight route (Figure 2).

Figure 2. Vision system and initial experiment setup. Left: The panoramic vision system we developed for this experiment consists of CCD camera, panoramic mirror, and housing components, which weighs 90 g in total so that it can be embedded in the flying robot. The visual field covers 360 degrees on the horizontal plane, and 260 degrees vertically. Middle and right: This vision system is equipped with the caster-frame on a pair of rails, and an external motor drives the vision system along a straight route at a constant speed. Initially walls with black and white stripes were installed along the rails (20cm width period of black and white, 60cm distance from vision system).
Initially walls with black and white stripes were installed along the railway (20cm width period of black and white, 60cm distance from vision system). Figure 3 shows that the typical responses of one EMD module, when the vision system moves 2 m at constant speed of 15 cm/sec. Black and white stripe patterns can be seen as big fluctuation of image intensities, i.e., image edges that activate the HPFs and LPFs. The EMD responses indicate image motion (as a result of egomotion), and eventually VO indicates the distance the vision system moves. Since only one EMD is used in this graph, the response of VO seems layered, because VO is incremented only when EMD goes through an edge in the image. This layered response in VO outputs could be improved by applying wide field of EMDs that cover wide range image angle such that changes in intensity (edges) continuously activate EMDs.

A characteristic feature of the EMD model is that its response depends not only on the speed of the image, but also on its spatial frequency. Thus, it acts more like a spatio-temporal filter rather than a pure speed detector. Its response attains a maximum at a certain speed that induces a specific temporal frequency. In order to test this characteristic, we have carried out another set of experiments, in which the vision system is driven 5 times at 3 different speeds (15cm/sec, 28cm/sec, and 40cm/sec) in the same setup as previous experiment. As shown in Figure 4, the visual odometer responses at different speeds indicate different distances even though it moves along the same absolute distances, i.e. the responses at middle speed indicate the longest distance. This result is not compatible with the results obtained from the bee experiments that show that distance estimation is independent of the spatial structure of the stimulus.

However, as shown in Figure 4, the odometer responses at each speed tend to indicate the same distance, which imply this odometer could be used for measuring distance to the target position as long as the agent follows the same route. Considering the fact that this assumption has to be made even in the case of a visual odometer which is independent of the spatial frequency (the bee still needs to encounter the same objects at the same distance), we can conclude that there is no obvious advantage of spatial-frequency independent visual odometer in this context.

With this assumption in mind, we tested this system in the natural office environment, in which the vision system moved 5 times at the same speed along the same routes without any artificial walls but by using the surroundings of the office. Photoreceptor inputs and odometer responses are shown in Figure 5, which demonstrates that the odometer has the ability to operate even in unstructured office environments when it follows the same route at the same speed.

2.3 3-D navigation with EMDs and the visual odometer

In the following, we focus on integration of the EMD-based visual odometer in an autonomous flying robot.

Flying agents also need to obtain the flight height information in order to maintain a certain flight altitude, to avoid crushing on the ground, and for reaching a certain target position in 3-D space. One possibility to realize such a mechanism is to use information from the ventral part of the visual field. EMDs receiving input from this part could measure the distance to the ground by using the optical flow induced by ground texture. A strategy of speed and altitude control of flying agents by using a similar mechanism has been evaluated in simulation. An alternative would be to use EMDs tuned to the vertical direction, in such a case the mechanism is the same as the one used in the visual odometer described in the previous section, but this time the EMDs are tuned to motion along the vertical plane (see Figure 6). According to this scheme, the vertical EMDs respond to the vertical motion of the agent (changes in altitude), thus this information could be used for altitude control. Note that, however, this approach does not measure the distance to the ground simultaneously but it depends on the accumulated response of the EMDs, and thus it is prone to cumulative errors. For the purpose for our experiments thought, it proved to be sufficient to keep the robot over a certain height and to avoiding crushing on the ground. This is not general purpose altitude control mechanism neither a model of how insects achieve this.

One prominent feature is that EMDs cannot accurately measure 2-D image motion, i.e. they cannot always distinguish horizontal from vertical motion depending on the image structure. We tested the performance of both vertically and horizontally tuned EMDs in the same experimental setup described in the previous subsection. During the experiment the robot had to move forward at constant height and speed in the office environment. Two arrays, consisting of 80 EMDs each, arranged along the horizontal, vertical directions, were used to provide both horizontal (forward motion) and vertical (changes in altitude) motion cues estimates, respectively. Figure 6 shows the output characteristics of the horizontally and vertically tuned EMDs in terms of the average and the standard deviation of the EMD responses. This result implies that the vertical EMDs could measure vertical motion with a 4% error.
Figure 3. Typical responses of the EMD and the visual odometer. Upper: Photoreceptor activity is calculated directly from the image captured by CCD camera. This graph shows the time-series of photoreceptor activity. Fluctuations on the signal correspond to the stripe pattern of the walls when the vision system moves horizontally along the corridor. Middle: The outputs of High-pass and low-pass filters. Fluctuations on these signals correspond to the edges in the image. Bottom: Responses of the wide field motion detectors and the accumulated visual odometer response.
Figure 4. Visual odometer responses at 3 different speeds. Left: Time series of photoreceptor activity for 3 different speeds (15cm/sec, 28cm/sec, 40cm/sec from upper to bottom). Right: Visual odometer response. Because the outputs of the EMDs depend on temporal frequency, the odometer produces different distance estimates at different speeds.

Figure 5. The response of the visual odometer in an experiment performed in a natural office environment. Left: Typical output values of four photoreceptors. Their output profile depends on the structure of environment and the light conditions. Right: Visual odometer response. The vision system moved along the same route 5 times at the same constant speed. This graph shows that the proposed odometer mechanism provides a robust distance estimate as long as it moves along the same route.
3. NAVIGATION EXPERIMENTS

3.1 Autonomous flying robot and 3-D goal-navigation

The flying robot Melissa is a blimp-like flying robot (Figure 7), which consists of a helium balloon, a gondola hosting the onboard electronics, and a host computer. The balloon is 2.3m long and has a lift capacity of approximately 400g. Inside the gondola, there are 3 motors for rotation, elevation and thrust control, a four-channel radio transmitter, a miniature panoramic vision system, and the batteries. The control process of Melissa can be decomposed in three basic steps. First, the video signal from the CCD camera attached to the gondola is transmitted to the host computer via a wireless video link. Second, the images are then digitized on the host computer that also performs the image processing in order to determine the target motor command. And third, the motor command is sent to the gondola also via radio transmission.

The control architecture of the robot is shown in Figure 8. Both left and right visual fields consist a two dimensional array of EMDs (resolution: 100 x 80 on each side). The EMD arrays provide input to the visual odometer as well as to a sensory-motor circuit that implements two basic reflexes. Responses from both horizontal and vertical wide field EMDs are extracted, and then used to drive two motor neurons that control to rotation and elevation motors. The right and left horizontal EMDs are connected to the rotation motor neuron in such a way that when the right EMDs have higher activation, i.e. high speed image motion, this rotation motor neuron will react to turn left, and vise versa. This scheme corresponds to a basic obstacle avoidance mechanism that will make the robot follow a route away from walls and obstacles, because image motion would be faster when the robot comes closer to objects. Right and left vertical EMD neurons are connected to elevation motor neuron, whereas the connection weights are chosen in such a way to suppress vertical motion, i.e. retaining height. The thrust motor neuron is connected to a bias neuron that drives the robot forward at a constant speed. The connection weights are set by hand, and not changed during the experiments.

3.2 Experiments and results

To evaluate the performance we conducted a set of experiments in an office environment. During each experiment the robot had to maintain a certain course direction and fly for 25 sec. All experiments were performed with the same initial condition, i.e. initial position, initial orientation, connection weights, and repeated 5 times. Figure 9 shows experimental setup, in which we used two video cameras to track and record the absolute trajectory of the robot for later analysis.
In Figure 10, the plots show 3-D coordinates of the robot in one second steps. Since the robot has the same neural connections through all of the 5 trials, the trajectories of the robot are similar. Figure 10 also shows the visual odometer responses that the robot obtains during each trial. Since the robot follows similar routes, the visual odometer measures almost the same distances even with natural stimuli in the office environment. In summary, considering that the robot follows the same route, and it measures the same "distance" robustly, this mechanism could be used for navigating between different places in the environment, i.e. for goal-directed navigation.

Figure 7. The autonomous flying robot, Melissa. Left: Melissa consists of a helium balloon, a gondola hosting the onboard electronics, and a host computer. The balloon is 2.3m long and has a lift capacity of approximately 400g. Right: Inside the gondola, there are 3 motors, a four-channel radio transmitter, a miniature panoramic vision system and image transmitter, and the battery.

Figure 8. Sensory-motor control circuit. Outputs of left and right EMD networks provide the vertical (EMD-V-L, EMD-V-R) and horizontal (EMD-H-L, EMD-H-R) image motion values of each side. The horizontally, vertically tuned EMDs are connected to a motor neuron that controls the robot’s rotation, elevation, respectively. The outputs of the horizontally tuned EMDs are also used by visual odometer. The thrust motor neuron has a connection to bias neuron that implements the default move-forward behavior.

In Figure 10, the plots show 3-D coordinates of the robot in one second steps. Since the robot has the same neural connections through all of the 5 trials, the trajectories of the robot are similar. Figure 10 also shows the visual odometer responses that the robot obtains during each trial. Since the robot follows similar routes, the visual odometer measures almost the same distances even with natural stimuli in the office environment. In summary, considering that the robot follows the same route, and it measures the same "distance" robustly, this mechanism could be used for navigating between different places in the environment, i.e. for goal-directed navigation.
4. DISCUSSION AND FURTHER WORK

The navigation experiments with the autonomous flying robot demonstrate that visual odometry is possible with a relatively cheap mechanism even in complex natural environments. The results imply that the optomotor response of flying agents...
would play an important role in a visual odometry scheme using EMDs, because their response tends to be dependent on the structure of environment. In this respect, the visual odometer that we describe here is not compatible with what we currently know about its biological counterpart, because the bee's response seems to be independent of the spatial structure of the environment. However, as we discussed earlier, this does not seem to have an advantage since the bee still has to follow the same route in order to experience the same distances to the objects, which is a prerequisite even for a pure image velocity visual odometer to work.

An interesting property of the navigation scheme described in this paper, is that the system not only achieves visual odometry but also avoids obstacles, controls its speed, and presumably this mechanism could also be used for safe landing, and these by using the same elementary mechanism, the EMD. For all of these behaviors, the sensory-motor response is, again, a very important issue to consider, because none of these behaviors could have been realized without it. The safe landing mechanism could be implemented by changing the connections between the EMDs and motor neurons. One interesting issue to be studied could be whether an action selection mechanism for switching between conflicting behaviors such as obstacle avoidance and landing would be necessary.

Another issue has to do with EMD design considerations about morphology and heterogeneousness. Considering fly's tendency of landing control to texture on the ground, it is not only intuitive but also biologically plausible that EMDs focusing on the ground could be used. This is a morphology issue of visual system design. Moreover it is natural to consider a heterogeneous design that would make use of different parameters among EMDs, such as the delay parameters or the connections to the motor neurons, because there would be different optimal parameters for lateral visual fields from ventral ones. Instead of manually tuning these parameters, one possibility would be to use artificial evolution to automatically generate them.

From the information theoretic perspective, this navigation algorithm has many advantages over traditional methods, because of its cheap and flexible design. In the extreme case, only 8 pixel values are required for navigation, because it can navigate as far as self-motion can be detected, which can be implemented, in principle, with only two EMD modules, for horizontal and vertical motion detection, on each side. The number of EMDs would be determined depending on the task-environment, because, for example, the agent does not need many EMDs when it navigates in a simple environment, with a limited number of visual patterns, like artificial gratings. And it could be also intuitive to imagine that design flexibility would be another advantage, since we, as designers (or by using evolutionary techniques), can easily increment or decrement the number of EMDs from one to thousands, with the same basic circuit. The relation between EMD design and task-environment could be another interesting issues to be investigated in future work.

To summarize, this paper presents a navigation algorithm using a biologically inspired visual odometer that is based on Elementary Motion Detectors, and evaluates its performance by using an autonomous flying robot equipped with a panoramic visual system. Since flying agents have to achieve a number of complex tasks in parallel with severe weight and real-time constraints, cheap design could be one of the most important design strategies. The result of the experiments with the autonomous flying robot show that the proposed algorithm can successfully be employed of navigation tasks in a complex indoor environment.

**ACKNOWLEDGEMENTS**

This work is supported by the Swiss National Science Foundation, grant no 20-53915.98, the Swiss Federal Office for Education and Science (VIRGO TMR network, BBW-No. 96.0148).

**REFERENCES**


