View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Infoscience - École polytechnique fédérale de Lausanne

Vision-based Navigation from Wheels to Wings

Jean-Christophe Zufferey Antoine Beyeler Dario Floreano

Autonomous Systems Lab (http://asl.epfl.ch)
Swiss Federal Institute of Technology in Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
email: name.surname@epfl.ch

Abstract - We describe an incremental approach towards the development of autonomous indoor flyers that use only vision to navigate in textured environments. In order to cope with the severe weight and energy constraints of such systems, we use spiking neural controllers that can be implemented in tiny micro-controllers and map visual information into motor commands. The network morphology is evolved by means of an evolutionary process on the physical robots. This methodology is tested in three robots of increasing complexity, from a wheeled robot to a dirigible to a winged robot. The paper describes the approach, the robots, their degrees of complexity, and summarizes results. In addition, three compatible electronic boards and a choice of vision sensors suitable for these robots are described in more details. These boards allow a comparative and gradual development of spiking neural controllers for flying robots.

I. INTRODUCTION

A. Overview

In this paper we describe an incremental methodology towards the development of an autonomous, vision-based robot capable of indoor flight (see video). Despite active research in mechatronics and bio-mechanics of micro flying robots [2][7], there are not yet flying devices capable of autonomously navigating in small cluttered environments, such as offices or house rooms. Strong constraints of size, weight, and energy consumption restrict the choice of technologies and control strategies that can be used for these flying robots. In order to understand the typical constraints of an indoor flying robot, we have developed a prototype [10] capable of flying in rooms at walking speed for 15 minutes. This robot has a maximum payload of 10 grams for electronics and sensors.

B. Methodology

In order to meet these severe weight and energy constraints, we decided to use vision sensors and networks of artificial spiking neurons [8] that can be implemented in low-power micro-controllers. We resort to an evolutionary process in order to generate small functional networks that map visual information into suitable control actions [11]. The project is articulated in a number of sub-goals of incremental complexity where we start with wheeled robots, then move to a blimp, and finally to a winged robot (Fig.1).

In the first stage of the project, we showed that artificial evolution could quickly discover functional networks of spiking neurons to control both a vision-based Khepera robot [3] and a vision-based blimp [12] that were asked to move straight forward as much as possible within a rectangular area (Fig.2).

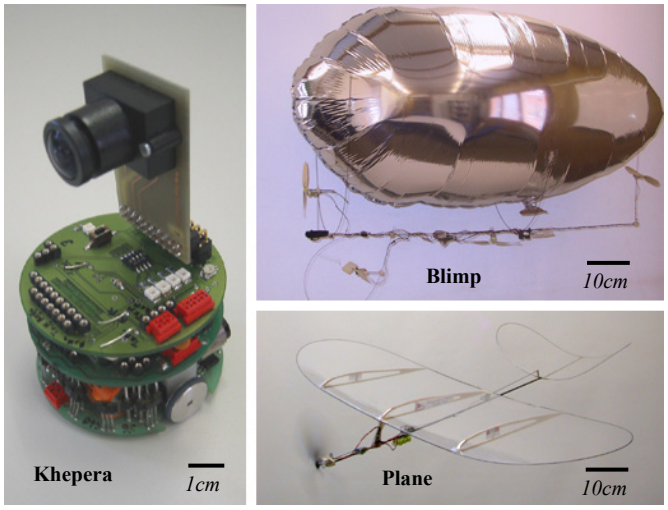

Fig.1 Khepera with a custom turret supporting a linear camera, Blimp, and Plane (indoor slow flyer).

Fig.2 Khepera and Blimp in their environments with randomly arranged black and white patterns providing contrasts to the embedded vision sensors.

In those preliminary investigations, the evolutionary algorithm and the neural networks were implemented on a desktop computer that could access the evolving robot every 100 ms (through cables and rotating contacts in the case of the Khepera and through a Bluetooth™ radio module in the case of the Blimp). This solution turned out to be very useful to evaluate the evolutionary and neural models as well as record all data of the evolving robots for analysis. However, the time delay caused by this remote strategy (read sensors

from robot, compute network activation on desktop computer, send motor commands back to robot) makes the detection of visual motion cues much more difficult by limiting the image acquisition rate. Motion cues, such as optical flow, are known to be exploited by visual insects for flight stabilization and obstacle avoidance [1], but could not be exploited by our evolved robots, which instead discovered a much simpler control strategy based on spatial frequency. Furthermore, delegating the neural algorithm to the processor of an external desktop computer is not a viable solution for robots that are expected to fly autonomously. An ideal solution would be to run the neural controller onboard the robot. Considering the small payload available in indoor flyers, the onboard processor must be extremely small and energy efficient. Therefore, in a second stage, we developed a methodology to implement an evolvable spiking circuit in microcontrollers with less than 50 bytes of memory. The entire system was tested in the 1-inch mobile robot Alice with infrared distance sensors [4].

C. Incremental approach

Porting this methodology to our vision-based robots in an incremental fashion from wheels to wings requires the development of a series of compatible embedded boards equipped with micro-controllers, a vision device, and wireless communication abilities (for supervision, debugging and occasional data logging). In this paper we describe the three robotic platforms, how they fit in our incremental development strategy, and the three respective embedded boards compatible with a set of vision sensors for real-time, vision-based navigation. Preliminary tests show that these boards can maintain visual structure at higher motion speeds and are therefore suitable for detection of motion cues.

II. THREE ROBOTIC PLATFORMS

Controlling a 6 DOF (degrees of freedom) flying robot in a 3D space can be divided into four mechanisms: attitude control (ATC) around pitch and roll axes, course stabilization (CS) around yaw axis, obstacle avoidance (OA), and altitude control (ALC). In our incremental approach we introduce the four mechanisms one after another by using different robotic platforms. The underlying idea is to capitalize and build upon each stage before moving to the next. To this idea, we developed a set of three robots (Fig.1) featuring increasing speed range, dynamic complexity, and degrees of freedom. For sake of comparison, Table I provides an overview of the characteristics and equipments of those three platforms.

A. The Khepera robot with a custom extension turret

The first one is the miniature differential-drive wheeled robot Khepera [9] with a custom-made microcontroller extension turret named KEvoPic (Khepera Evolutionary PIC). As explained in more details later on, this board features the same embedded microcontroller as the two other platforms. Different kinds of vision sensors (see § IV) can be plugged in directly and processed onboard without using the main processor available on the Khepera. This wheeled robot moves on a flat surface and has 3 DOF. Therefore, it is an ideal candidate for testing vision-based OA without requiring CS because it is in contact with the floor and has negligible inertial forces (wheel speeds fully determine the trajectory). ATC and ALC are not required since the robot moves on a flat surface.

B. The Blimp – a lighter-than-air robot

The second platform is the Blimp [12]. It represents an intermediate step between wheeled robot and the winged aircraft. It has a helium-filled envelope that produces a lift of about 100g. Although the Blimp can move in 3D, roll and pitch movements are passively stabilized around horizontal since the buoyancy center is positioned above the gravity center. Consequently, the Blimp has only 4 DOF. To further simplify the control problem, we implemented an automatic (non vision-based) altitude control using a simple vertical distance sensor in order to reduce the maneuvering space to 2D and the control to 3 DOF, when activated. Even with this simplification, the Blimp displays much more complex dynamics than the Khepera and there is no trivial relation between the voltages applied to the motors and the resulting trajectory because of inertia and aerodynamic forces. Therefore, in addition to OA, the Blimp requires CS (and ALC when the vertical distance sensor is disabled).

C. The Plane – a 50 g indoor slow flyer

The third platform, the 50g Plane [10], is the final demonstrator of the project. It is a full 3D, 6-DOF robot requiring all four control mechanisms (ATC, CS, OA and ALC) in order to stay airborne and autonomously maneuver in a room. It is also much more delicate than the Blimp since it is not able to fly at zero speed and recover from collisions without damage.

TABLE I

CHARACTERISTICS OF THE THREE ROBOTIC PLATFORMS

	Khepera	**Blimp**	**Plane**
Actuators	2 wheels	3 propellers	1 propeller + 2 servos
Consumption (W)	~ 4	~ 1	~ 4
Weight (g)	120	100	50
Typical speed (m/s)	0 to 0.2	0 to 1	1.5
Typical maneuver space (m)	0.6 x 0.6	5 x 5	15 x 15
Compatible vision sensors (see § IV)	K213, TSL3301, OV7645, SMDM1	TSL3301, OV7645, SMDM1	OV7645, SMDM1
Additional sensors (not available to the neural control systems)	wheel encoders, 8 proximity	anemometer, bumpers, altitude	-
Communication	RS232	Bluetooth	Bluetooth
Power supply	cables	Li-Ion batt.	Li-Poly batt.
Autonomy	-	2 - 3 hours	~ 15 min

Table II below summarizes the incremental dynamic complexity and control requirements as discussed above and highlights the smooth increase in control complexity.

TABLE II

INCREASING CONTROL COMPLEXITY

	Khepera	Blimp 2D limited	Blimp	Plane
DOF	3	3	4	6
ATC	-	-	-	required
OA	required	required	required	required
CS	-	required	required	required
ALC	-	-	required	required

III. EMBEDDED CONTROLLER BOARDS

Aiming to make the three robots similar from a software point of view and compatible with a number of different visual sensors, we developed three corresponding microcontroller boards (Fig.3) that can be programmed using the same tools and among which software modules can easily be reused. A common aspect of these boards is that they are all based on a Microchip™ PIC microcontroller for which we have developed a self-contained, evolutionary spiking network [4]. Different kinds of vision sensor interfaces (analog, serial, parallel) are possible thanks to a flexible extension port, which gathers a large number of digital and analog inputs/outputs of the microcontroller. Since the flying platforms can lift no more than one lightweight Li-Ion battery, all these boards are 3V-compatible. However, in order to provide enough voltage for the engines, a DC-DC step-up converter is provided on the aerial versions, that can deliver up to 1.5A at 6V. This also avoids decrease of thrust along with the battery level over time.

A. KEvoPic – a microcontroller board for the Khepera

The first board, KEvoPic (Khepera Evolutionary PIC), is slightly different from the two next ones. As a Khepera add-on, it is not directly connected to some of the robot's peripherals (motors, wheel encoders, and proximity sensors), but relies on the Khepera base module as a slave. It has only a serial communication link with the underlying Khepera, which is only employed for sending motor commands, reading wheel speeds and proximity sensors. The cameras (see § IV), instead, are directly connected to KEvoPic, avoiding the transfer of vision stream via the Khepera main processor.

B. B & PEvoPic – microcontroller boards for aerial robots

The boards designed for the aerial robots, BEvoPic (Blimp Evolutionary PIC) and PEvoPic (Plane Evolutionary PIC), share the same layout. PEvoPic is simply a subpart of BEvoPic, which in addition features Blimp-specific components (three motor drivers and a multiplexer for optional bumpers). The common part of the two boards comprises the microcontroller, the step-up converter, and two extension connectors, one for the vision system and another for the radio module (Fig.3). Both electronic circuits have been designed to minimize weight. For instance, PEvoPic weighs only 4.2 g, including the wireless module.

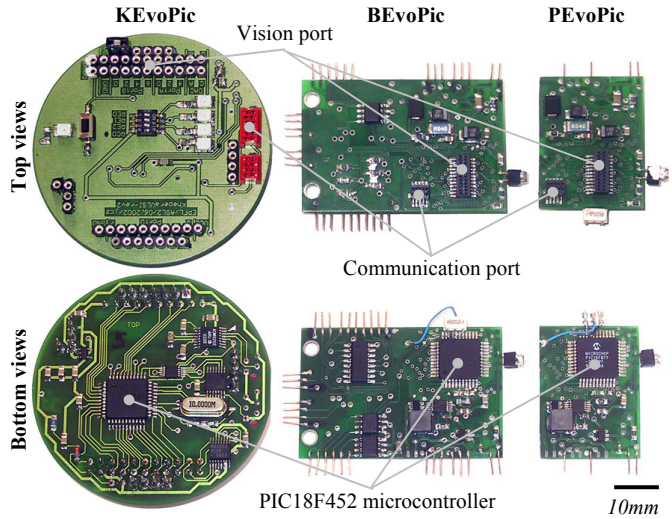

Fig.3 Microcontroller boards for the Khepera (left), the Blimp (center), and the Plane (right).

C. Microcontroller

The Microchip™ PIC18F452 microcontroller (Table III) has been selected as the core processor of our boards for different reasons. First, with only 20 mW average consumption at 20 MHz, it has extremely low power consumption. It also supports low voltage (3V) power supply. Then it is available in very small packaging (44-pin Thin Quad Flat Package) and has therefore a minimal weight. Furthermore, it features a number of integrated hardware peripherals like, e.g., USART (Universal Synchronous Asynchronous Receiver Transmitter), MSSP (Master Synchronous Serial Port, in particular I2C), and ADCs (Analog to Digital Converters) allowing different types of interfaces with the robot sensors and actuators.

TABLE III

PIC18F452 MICROCONTROLLER OVERVIEW

Operating frequency	20 MHz
Typical consumption	20 mW
Execution time for 1 instruction	200 ns
Data memory (RAM)	1536 Bytes
Instruction set	75
Program memory	16 kWords
Integrated peripherals	USART (RS232), MSSP (I2C), 8 ADCs

The PIC18F452 is an 8-bit microcontroller with reduced instruction set and without floating point arithmetic. However, our implementation of evolvable spiking circuits meets those constraints by relying only on basic instructions such as AND, OR, NOT, and bit shift. In particular, it is able to process 8 spikes in parallel by storing them in one byte and using basic, 1-cycle, byte-wide instructions of the processor [4]. It is also to notice that the microcontroller can be programmed as well in assembler as in C-language (using, e.g., CCS PICC compiler), which enhances the code readability, portability, and modularity.

D. Communication

Along with the boards we developed a communication system to enable efficient data logging and supervision from a computer. At the hardware level we use RS232 whenever it is possible to have a cable between the robot and the computer (which is always the case with KEvoPic). For wireless connection, we switch to Bluetooth. For sake of flexibility, at the software level, a unified packet-based communication protocol has been developed, which is RS232 compatible and can easily be wrapped in Bluetooth packets.

BEvoPic (Fig.4) and PEvoPic have an extension connector supporting either a RS232 cable or a Bluetooth module[1]. When Bluetooth is used, the PIC controls the module via its serial port, at the HCI (Host Controller Interface) level and another similar module must be connected to the serial port of the computer. The maximum communication range between robot and desktop modules is about 30 meters. The power consumption is 100 to 150mW per module in transmit mode.

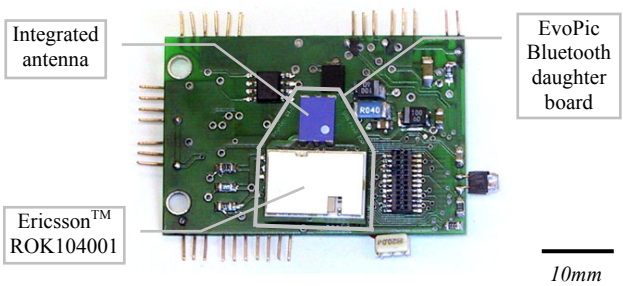

Fig.4 BEvoPic with its radio module plugged in.

The advantages of using Bluetooth technology are first the benefit of current development toward low power and small modules using a standardized protocol; second a good robustness to electromagnetic noise thanks to frequency hopping and automatic packet retransmission on errors. The host microcontroller is not required to take care of RF encoding or error detection and recovery. It just implements the standard Bluetooth HCI protocol.

IV. EMBEDDED VISION SENSORS

We are experimenting with four different visual sensors based on CMOS (Complementary Metal Oxide Semiconductor) technology. A major advantage of CMOS over CCD (Charge Coupled Device) cameras is the ability to integrate additional circuitry on the same die as the sensor itself. This makes it possible to integrate analog to digital converters or other functionalities like windowing. Furthermore, CMOS imagers offer lower power dissipation, and smaller system size than CCD sensors. These advantages are at the expense of image quality but this criterion is of less importance in our case. The main characteristics of the four vision sensors are summarized in Table IV. Except the K213, which was used in initial experiments [3], all of them are compatible with the three boards described above.

TABLE IV

COMPARISON OF THE DIFFERENT VISUAL SENSORS

	K213	TSL3301	OV7645	SMDM1
Type	1D	1D	2D	1D
Optic flow output	no	no	no	yes
Number of pixels	64	102	640x480	32
Field of view (°)	36	70 or 130	50	50
Max frame rate (Hz)	20	1000	12.5	> 1000
Min weight (g)	40	~ 8	~ 1	~ 1
Interface	Khepera	Serial	Parallel	Analog

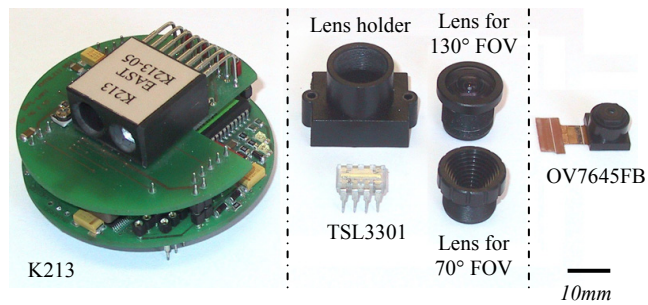

Fig.5 K213 turret (left), TSL3301 and optics (center), OV7645FB (right).

A. K213 – the standard Khepera extension turret

The K213 (Fig.5, left) is a commercially available extension turret for the Khepera with its own processor and proprietary bus. It features one line of 64 pixels with a 36° FOV (Field Of View), and has been used only before the EvoPic boards where developed. In addition to a quite tight FOV, fixed optics, and relatively heavy packaging, its main drawback is the very long integration time (min 50ms, depending on the average light intensity) that produces blurred images when the robot moves rapidly. Fig.6 highlights this problem by plotting sub-sampled images (only one pixel every four) produced by the K213 along time, while the robot is rotating[2] on the spot in the centre of the arena (Fig.2). When the robot rotates slowly (left) the shift of the patterns and the patterns themselves can be recognized easily. But as soon as the speed [3] increases (center and right), the spatial pattern blurs out and there is no structure anymore allowing for visual motion cues detection.

B. TSL3301 – a 102 pixels linear camera

In order to compensate for the limitations of the K213, we developed a module based on a TAOS™ TSL3301 CMOS linear vision sensor with 102 pixels. This module is composed of a lens, a lens holder and the sensor itself (Fig.5, center). The TSL3301 features an integrated analog to digital converter and a 2-wire serial bus allowing a straight forward connection to the EvoPic boards.

[1] See http://asl.epfl.ch/~jzuffere/Bluetooth.html for a list of Bluetooth modules and some references.

[2] Optic flow due to rotation is generally higher than the one resulting from translations. Moreover, it is independent from distance to objects (contrasts), in the case of pure rotations [5].

[3] A turning rate of 200 deg/s corresponds to the maximum value allowed during the preliminary experiments [3].

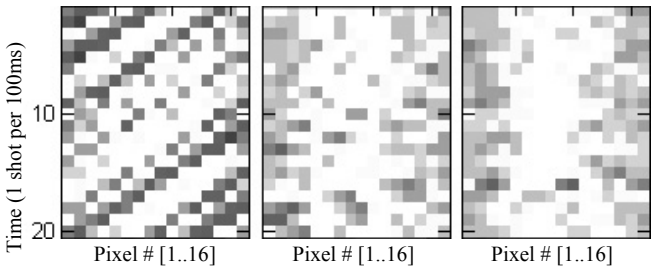

Fig.6 Sequences of images from the K213. Each line of a box represents a snapshot of 16 pixels measured every 100 ms as the robot rotates on the spot. The first sequence is measured at 30 deg/s (left), the second at100 deg/s (center), and the third at 200 deg/s (right).

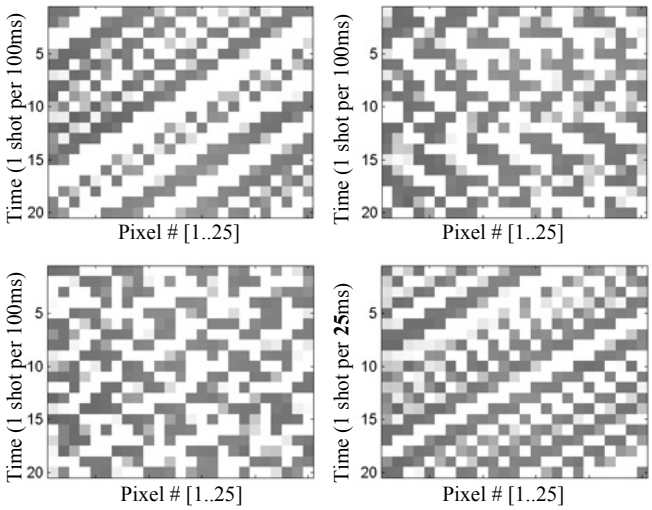

Fig.7 Sequences of images from the TSL3301 connected to KEvoPic. Each line of a box represents a snapshot of 25 pixels as the robot rotates on the spot. The bottom right box displays images grabbed every 25 ms as the three others are every 100 ms. The first sequence is measured at 30 deg/s (top left), the second at 100 deg/s (top right), and the two lasts at 200 deg/s (bottom).

With this chip, the image acquisition time is 1 to 10ms, depending on optic aperture. This module was plugged on the KEvoPic board (Fig.1, left) and tested in the same conditions as above with the K213. The three first graphs of Fig.7 show that the images remain sharp at higher rotating speeds (this is due to the shorter light integration time). In addition, the last graph (bottom right) shows that by increasing the acquisition rate (one image every 25ms, which was impossible with the K213) the spatial pattern remains stable even at high speed and therefore can be used for motion detection.

Another attractive feature of the TSL3301 is the large light sensitive area (8.75mm long) allowing very large FOV by using short focal lenses (K213 did not allow modifying the FOV since the lens was glued in a plastic box on top of the turret). For example, with only one camera oriented in the forward direction, large FOV allows for gaining better information from the lateral optic flow patterns, which are particularly relevant for ego-motion estimation in case of translations.

C. OV7645FB – a miniature VGA camera

A drawback of the TSL3301 module lies in its weight (about 8 g depending on optics and support). Considering the Plane's payload of 10g, different solutions must be considered. Another problem is that a linear array of photoreceptors is not sufficient to control a more than 3-DOF robot. Therefore we looked for a compact CMOS vision sensor with integrated optics and selected the OV7645FB from Omnivision™ (Fig.5, right). This module is sufficiently small (10x9x7mm) and lightweight (0.7g) to be mounted on the Plane. The core chip is a VGA color camera, in which acquisition can be restricted to a sub-region of the whole image (windowing). An 8-bit parallel bus is used for transferring pixel values while a serial port allows for adjustment of camera parameters such as brightness, contrast, gain, windowing, etc.

Fig.8 2D sub-sampled image (35x24 pixels) from OV7645FB camera. The camera is wired to PEvoPic via a flex connector supporting. The microcontroller is programmed to grab the frame, store the 840 pixels in its data memory before sending them upon request to a desktop computer.

Despite the remarkable capabilities of this camera module for such a small package, the OV7645FB tends to output too much information (more than 300000 pixels) for the microcontroller, whose data memory has only 1536 Bytes. Although a large part of these pixels could be ignored in case we need only a low resolution image (Fig.8), the camera still needs to scan every pixel internally. Therefore, it is very difficult to obtain high frame rates while maintaining a sufficiently slow pixel clock for the microcontroller to be able to read and store the pixels from the parallel port. The maximum frame rate obtained so far is 12.5 Hz with a pixel clock frequency (given by the camera) close to the maximum instruction frequency of the PIC. This frame rate is quite reasonable if one think in terms of distance traveled by the fastest robot we are interested in, i.e., the Plane flying at 1.5 m/s will get a new image every 12 cm.

D. SMDM1 – a custom-designed camera

In parallel, another visual sensor named SMDM1 (Smart Motion Detector Module, first version) has been developed jointly with the Institute of Neuroinformatics in Zurich and is currently under testing. The core chip is based on aVLSI (analog Very Large Scale Integration) technology and is custom-designed for this project. This technology allows implementation of processing circuitry next to the

phototransistors on the same die. In our case, the chip is based on a circuit developed by Kramer et al. [6] implementing analog computation of optic-flow. This first version has one line of 32 photoreceptors connected to 31 motion detectors (returning an estimation of optic flow between two adjacent photoreceptors). It has only three analog outputs retrieving respectively light intensity, rightward and leftward motion. An internal scanner allows to select the pixel and motion detector, which must be connected to the analog output and thus to the microcontroller ADCs.

The main advantage of this aVLSI module is to reduce information size and load of the microcontroller by pre-processing the image already in the sensor itself. Moreover, the need for high frame rate in order to estimate optic flow disappears since even if the spatial patterns are uncorrelated between two successive images, the optic flow given directly by the chip is always consistent, even if sampled at low frequency. To answer the question of the limitations due to a linear pixel arrangement, a future version, the SMDM2 is already planned with other photoreceptors arrangements, e.g., with two lines, horizontal and vertical.

V. CONCLUSION

In this paper we have described an incremental approach towards the development of autonomous indoor flyers that use only vision to navigate in textured environments. Our initial experiments on evolution of spiking neural controllers with a wheeled robot [3] and a blimp [12] indicate that this control methodology is suitable to generate efficient control systems for this problem. However, in those experiments the entire evolutionary and neural algorithms were implemented on an external desktop computer, which accessed the robotic platforms for reading sensory information and sending motor commands. Since the long term goal is to have the neural system running on board a fully autonomous robot, in this paper we described the development of three compatible electronic boards that will allow us to evolve the spiking neural controllers on the robots. These boards and their vision sensors have been designed by keeping in mind the increasing weight and energy constraints as we move from wheeled to winged robots. The communication capabilities of the boards allow us to monitor the dynamics of the control system and also give us the choice of running part of the algorithm (the evolutionary process) on an external computer when we need memory storage for data logging of the evolving populations.

A number of compact vision systems, some of which are still under development, have been described. Early tests indicate that these chips can provide good contrast and maintain spatial structure at navigation speeds required by an indoor flying robot. The direct connection of the chosen vision systems on the electronic boards allows drastically reducing the input-output cycle (from 100 ms when running the entire algorithm on an external desktop computer to approximately 20 ms), making the detection and exploitation of the motion cues possible. Estimation of optic flow is expected to be further improved by the use of dedicated aVLSI chips currently under development.

VI. ACKNOWLEDGEMENTS

We would like to especially thank Jean-Daniel Nicoud (http://didel.com) and Cyril Halter who significantly contributed to build the flying platforms and the Adaptive Intelligence Laboratory in Utrecht for fruitful collaboration on software development. We are indebted to Dr. Jörg Kramer, Dr. Shih-Chii Liu for the ongoing collaboration on the design of custom aVLSI chips, which is supported by the Institute for Neuromorphic Engineering (INE). We equally thank Yannick Fournier for his work on the OV7645FB camera and Portescap S.A., who provided a number of electrical miniature motors both for the Blimp and the Plane. This project is supported by the Swiss National Science Foundation, grant nr. 620-58049.

VII. REFERENCES

[1] M. Egelhaaf, and A. Borst, “A Look into the Cockpit of the Fly: Visual Orientation, Algorithms, and Identified Neurons”, *The Journal of Neuroscience*, 13(11), 1993, pp. 4563-4574.

[2] R.S. Fearing, K.H. Chiang, M.H. Dickinson, D.L. Pick, M. Sitti, and J. Yan, “Wing Transmission for a Micromechanical Flying Insect”, in *Proc. of 2000 IEEE Int. Conf. Robotics and Automation*.

[3] D. Floreano, and C. Mattiussi, “Evolution of Spiking Neural Controllers for Autonomous Vision-based Robots”, in T. Gomi (ed.), *Evolutionary Robotics IV*, 2001.

[4] D. Floreano, N. Schoeni, G. Caprari, and J. Blynel, “Evolutionary Bits'n'Spikes”, in R. K. Standish, M. A. Bedau and H. A. Abbass (ed.), *Artificial Life VIII*: Proc. of the 8th Int. Conf. on Artificial Life, MIT Press, 2002.

[5] J.J. Koendrink, A.J. van Doorn, “Facts on Optic Flow” *Biological Cybernetics*, 56, 1987, pp- 247-254.

[6] J. Kramer, R. Sarpeshkar, and C. Koch, “An analog VLSI velocity sensor”, in *Proceedings of 1995 IEEE Int. Symp. Circuits and Systems*, pp. 413-416.

[7] A. Kroo, “The Mesicopter: A Meso-Scale Flight Vehicle”, web page, 1999, http://aero.stanford.edu/mesicopter

[8] W. Maas, and C. Bishop, *Pulsed Neural Networks*, Cambridge, MA: MIT Press, 1999.

[9] F. Mondada, E. Franzi, and P. Ienne, “Mobile robot miniaturization: A tool for investigation in control algorithms”, in *Proc. of the Third Int. Symposium on Experimental Robotics*, Springer Verlag, pp501-513, 1993.

[10] J.D. Nicoud, and J.C. Zufferey, “Toward Indoor Flying Robots”, in *Proc. of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 787-792.

[11] S. Nolfi, D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*, Cambridge, MA: The MIT Press, 2001.

[12] J.C. Zufferey, D. Floreano, D., M. van Leeuwen, and T. Merenda, “Evolving Vision-based Flying Robots”, in *Proc. of the 2nd Int. Workshop on Biologically Motivated Computer Vision*, LNCS, Springer-Verlag, 2002.