2011 IEEE Int'l Conference on Robotics and Biomimetics (ROBIO)
http://www.cds.caltech.edu/~murray/papers/fm11-robio.html

A hovercraft robot that uses insect-inspired visual autocorrelation for motion control in a corridor

Sawyer B. Fuller and Richard M. Murray

***Abstract*— In this paper we are concerned with the challenge of flight control of computationally-constrained micro-aerial vehicles that must rely primarily on vision to navigate confined spaces. We turn to insects for inspiration. We demonstrate that it is possible to control a robot with inertial, flight-like dynamics in the plane using insect-inspired visual autocorrelators or "elementary motion detectors" (EMDs) to detect patterns of visual optic flow. The controller, which requires minimal computation, receives visual information from a small omnidirectional array of visual sensors and computes thrust outputs for a fan pair to stabilize motion along the centerline of a corridor. To design the controller, we provide a frequency-domain analysis of the response of an array of correlators to a flat moving wall. The model incorporates the effects of motion parallax and perspective and provides a means for computing appropriate inter-sensor angular spacing and visual blurring. The controller estimates the state of robot motion by decomposing the correlator response into harmonics, an analogous operation to that performed by tangential cells in the fly. This work constitutes the first-known demonstration of control of non-kinematic inertial dynamics using purely correlators.**

I. INTRODUCTION

Autonomous flight by an insect-sized robot will require feats of miniaturization on multiple fronts: actuation, system integration, power use, and even computation. For instance, it may be necessary to replace electric motors and ball bearings with piezo actuators and flexure joints because of scaling considerations [1]. In this work we are concerned with the sensors and feedback control of such a vehicle, aiming to insure it flies stably and keeps collisions to a minimum.

Miniaturization may require dispensing with traditional sensors and looking toward biology for inspiration. Larger unmanned aerial vehicles (UAVs) use radar, lidar, and the global positioning system (GPS) for self-localization. But for a tiny robot that may need to fly in enclosed or cluttered environments, these sophisticted sensors may be unavailable. The GPS signal will be compromised or provide insufficient precision or bandwidth [2]. And emissive sensors may use too much power or be too heavy. In nature, the fly is a virtuoso flyer among insects that performs under these constraints [3]. To perform their aerial feats, flies carry large omnidirectional eyes that sense patterns as well as visual motion [3]. Considering that our small robot and the fly operate under similar constraints, we aim here to take inspiration from the fly in the design of our robot. Despite its reliance on biomaterials, Nature's solution is nonetheless quite adept: flies can land on inverted surfaces and right themselves within tens of wingbeats after tumbling surprise take-offs, and can navigate complicated enclosed spaces like kitchens and often avoid flyswatters [3].

This work supported by the Institute for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office and a National Science Foundation Fellowship to S. Fuller

S. Fuller is with the department of Bioengineering and R. Murray is with the Faculty of Control and Dynamical Systems, California Institute of Technology, 1200 East California Blvd., Pasadena, CA 91125, USA {minster@caltech.edu, murray@cds.caltech.edu}

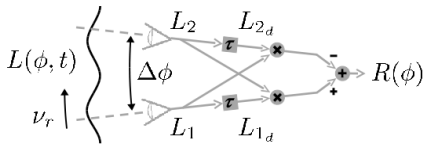

Fig. 1. The correlator model for visual optic flow estimation in insects. A pair of visual elements aimed at different angles observes luminance signal L moving across the retina with retinal velocity ν_r. The element τ is a delay or time-lagging low-pass filter and $\times$ is a multiplication or "correlation." For a sinusoid of a given spatial frequency, the correlator response R increases linearly with ν_r for ν_r sufficiently small.

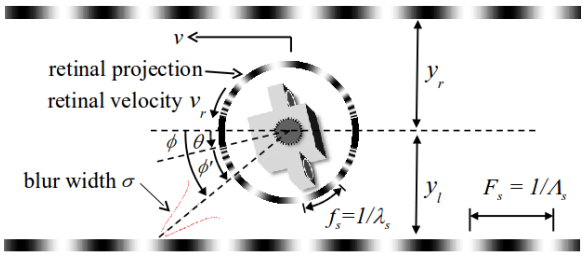

Fig. 2. A fan-actuated robot with hovercraft dynamics carries an omnidirectional array of correlators and navigates a corridor. The goal is to move along the axis, avoiding the walls. A representation of the projection of the visual environment onto the retina surrounds the robot.

Insects rely heavily on the pattern of optic flow, that is, the rate of visual motion across the retina (measured in radians per second), for flight control. In a variety of insects, including flies, bees, and beetles, visual flow is detected by a computation that is modelled faithfully by an autocorrelation scheme in which outputs from neighboring pairs of visual sensors are delayed and multiplied [4], [5], [6] (Figure 1). Each pair is known as an "elementary motion detector," or EMD. The time-mean response of an EMD pair has the same sign and, below saturation, is proportional to the magnitude of the optic flow ν_r.

It is not yet understood the underlying neural mechanism that performs the delay-and-correlate operation in insects [3], or why correlation is used rather than a different method such as the gradient method [6]. Correlator response is strongly dependent on both spatial frequency and contrast in the image [5]. For naturalistic imagery, this leads to a low signal-

to-noise ratio that requires averaging over space and time to obtain a lower-noise estimate [7]. Nevertheless, insects use correlators in spite of these non-idealities. It may be that autocorrelation-based control is easy to bootstrap [8] or that performing a division or matrix inversion as required by the gradient scheme is too neurally expensive [6]. In any case, an advantage is that the correlation computation is minimal, and the output may be sufficient to carry out the feedback controls tasks required by the fly.

In this work we are concerned with whether the output of correlators can be employed as the primary source of information for the controller of a robot with non-kinematic, flight-like dynamics. Correlators are appealing from the perspective of controlling small flying vehicles because they require minimum computation. They eschew serial, digital computation (for example, comparing lists of visual features [9]) in favor of simple operations like addition and multiplication that could be performed in parallel on low-power analog silicon. We are concerned with motion in a corridor (Figure 2) because it is the essential behavior required for navigating between obstacles in a cluttered environment. A reflexive corridor following controller could operate in real-time, freeing a high-level controller to pursue long range goals such as searching or path planning.

Previous work reported controlling a wheeled robot using a different motion detection system based on time-of-flight [10], but this robot had static kinematic motion that did not admit sway so was much easier to control. Later results using this optic flow algorithm controlled altitude [11] and corridor/wall following [12], but rely on an unspecified mechanism for attitude (yaw or pitch) control. Another wheeled robot used image interpolation for corridor following [13]. Other controllers have been formulated around pure optic flow and used a more compute-intensive gradient algorithm to estimate it [14], [15]. Another flying robot used a gradient-based algorithm to sense visual expansion to perform sudden turns in a square arena inspired by the body-saccades observed in flies [16]. In the domain of simulation alone, forward velocity regulation [17] and simplified lateral position control [18] have been demonstrated in a visually-realistic fruit fly simulator, but these simulations relied on artificial constraints of certain degrees of freedom.

In this work we report that it is possible to stabilize the motion of a dynamic fan-actuated hovercraft robot in the 2D plane along the centerline of a corridor using fly-like visual correlators (Figure 2). Contributions include a frequency-domain model for the response of an array of correlators moving relative to a large flat textured surface. In addition, we propose decomposing correlator response into square harmonics for improved performance, rather than the sinusoidal harmonics of [14].

II. FREQUENCY-DOMAIN ANALYSIS OF CORRELATORS

A. Correlator response to panoramic image motion

To analyze correlators, we first consider the case of luminance readings coming from panoramic image motion as would be induced by self-rotation. Since by Fourier decomposition an arbitrary image can be represented by a sum of sinusoids of different frequencies and amplitudes, we start with an analysis of a single, arbitrary sinusoid.

A single correlator consists of two luminance sensors oriented at slightly different body-centric angles separated by an angle $\Delta\phi$ (Figure 1). Suppose a sinusoid luminance signal L with spatial frequency f_s (cycles/rad) moves at ν_r rad/sec in front of the retina. Each sensor reads

$$L(\phi, t) = C_0 \cos(2\pi f_s \phi + 2\pi f_s \nu_r t), \tag{1}$$

where t is time. The correlator response is $R = L_2 L_{1_d} - L_1 L_{2_d}$ where the subscript d indicates a delayed or filtered version of the luminance signal. If a pure delay is used, the correlator response can oscillate between positive and negative with increasing ν_r. Accordingly, we preferred low-pass filter $\frac{1}{\tau s+1}$ as the delay element because it never goes negative. Assuming zero-mean input, it can be shown that this form of the correlator asymptotically reaches a steady-state (constant in time) response [7]

$$R(t) = \frac{1}{2\pi\tau} C_0^2 \frac{f_t}{f_t^2 + 1/(2\pi\tau)^2} \sin(2\pi f_s \Delta\phi), \tag{2}$$

where

$$f_t = f_s \nu_r \tag{3}$$

is the temporal frequency of the sinusoid's oscillation as it moves. Initial transients die off with an exponential time constant τ.

B. Decomposing correlator response

The correlator response to a sinusoid (2) can be decomposed into a product of factors

$$R = \frac{1}{2\pi\tau} C^2 T A, \tag{4}$$

where

$$C = C_0 \tag{5}$$

is a contrast factor that depends only on the amplitude of the luminance input L,

$$T = \frac{f_t}{f_t^2 + 1/(2\pi\tau)^2} \tag{6}$$

is a temporal factor that depends only on the temporal frequency f_t of the sinusoid, and

$$A = \sin 2\pi f_s \Delta\phi \tag{7}$$

is an aliasing factor that depends only on the product of spatial frequency f_s and angular separation between the pair of luminance sensors $\Delta\phi$.

C. Incorporating the effect of spatial blurring

We model the luminance sensors as having a Gaussian sensitivity profile, blurring the image. This is effectively a spatial-frequency-dependent attenuation of C. Each luminance sensor has an angle-dependent blurring function

$$G(\phi) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\phi^2}{2\sigma^2}\right),$$

where σ is proportional to the width of the blurring function. If the original luminance signal is L_0 (1), then by convolving it with G we get the blurred lumincance signal

$$L = G \circ L_0.$$

To find the resultant amplitude attenuation, we turn to the frequency domain. Because it contains only one frequency, the Fourier transform $\hat{L}_0$ of L_0 is a pair of delta functions at $\pm f_s$. Using the property that convolution in the spatial domain is equivalent to multiplication in the frequency domain, we take the Fourier transform of G,

$$\hat{G}(f_s) = \exp\left(-\frac{1}{2}(2\pi f_s \sigma)^2\right),$$

where f_s is the spatial frequency, and thus $\hat{L}(f_s) = \hat{L}_0(f_s)\hat{G}(f_s)$. The two delta functions of $\hat{L}_0$ are scaled by $\hat{G}(f_s)$. When the inverse transform is applied on $\hat{L}(f_s)$, a sinusoid is recovered, but in general with a change in amplitude and phase according to $\hat{G}(f_s)$. Because $\hat{G}$ is a real-valued function, the phase is unchanged and the effective amplitude of the blurred image is thus

$$C(f_s) = \hat{G}(f_s)C_0, \tag{8}$$

where C_0 was the amplitude of the original luminance sinusoid (1). At low frequencies where $f_s \ll 1/(2\pi\sigma)$, $C \approx C_0$, and at high frequencies $C \ll C_0$.

D. Incorporating motion parallax and perspective

We would like to extend the equation for R (2) to a moving flat surface of infinite extent. To do so, we need only consider how know the spatial and temporal frequencies (f_s and f_t) project onto the retina and change as a function of angle ϕ and state of the vehicle $\mathbf{q}$.

The vehicle is moving at a velocity v (m/sec) near the midpoint between two walls pattered with sinusoids with spatial frequency F_s cycles/m (Figure 2). The walls are separated by $2y_d$ where y_d is the desired distance the robot wants to keep from the walls. The distance to the left and right walls are y_l and y_r respectively, with $y_l + y_r = 2y_d$. The fly's position $\tilde{y}$ is its distance from the centerline, giving $y_l = y_d + \tilde{y}$ and $y_r = y_d - \tilde{y}$.

The first matter is to find an expression for how the spatial frequency F_s on the walls is projected onto the retina as a (spatially varying) spatial frequency f_s. By multiplying the effect of changing distance to the wall $f_s = \frac{y}{\sin\phi}F_s$ by the effect of changing the angle of the wall $f_s = \frac{1}{\sin\phi}F_s$, the spatial frequency f_s projected onto the retina is

$$f_s(\phi) = \frac{F_s y}{\sin^2\phi}, \tag{9}$$

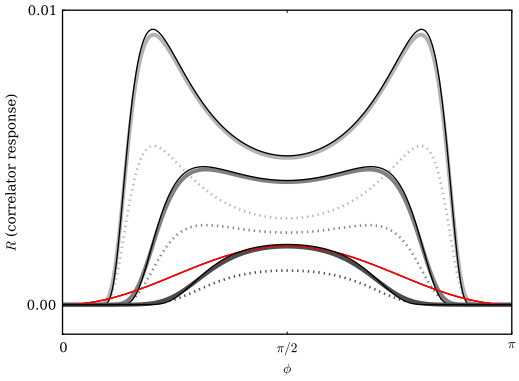

Fig. 3. Simulated and analytic model correlator response R for different visual sensor blurring widths σ. As σ increases, the response diminishes and simplifies, until it begins to resemble the red line, a scaled version of the retinal velocity or "optic flow" ν_r. Thick lines are simulated at 5 kHz, black lines are from (11-13), and dashed lines are simulated result at 60 Hz with a zero-order hold and are qualitatively the same. The vehicle is situated 1.5 m from the wall moving at v=0.25 m/s. The red line is scaled for easier comparison.

where y is the distance to the wall seen by that sensor and ϕ is the angle relative to the axis of the corridor. (Or alternately, equate frequency $\omega_s = 2\pi f_s$ to rate of phase change $\omega_s = \frac{d\varphi}{d\phi} = \frac{d\varphi}{dx}\frac{dx}{d\phi}$ where $x = y\tan(\phi - \pi/2)$ is the linear distance along the wall from the vehicle and $\frac{d\varphi}{dx} = 2\pi F_s$.)

Second, the retinal velocity or "optic flow" ν_r in the corridor is [14]

$$\nu_r = -\dot{\theta} + \frac{1}{r}\left(v\sin\phi - \dot{y}\cos\phi\right), \tag{10}$$

where $r = \frac{y}{\sin\phi}$ is the distance to the wall at angle ϕ. θ is the angle of the fly in the counter-clockwise direction relative to the axis of the corridor.

To find the correlator response in the corridor, we need only substitute the corresponding f_s, $f_t = \nu_r f_s$, and $C(f_s)$ into (5-7), to arrive at

$$C = C_0 \exp\left(-\frac{1}{2}\left(\frac{2\pi F_s y\sigma}{\sin^2\phi}\right)^2\right) \tag{11}$$

$$T = \frac{F_s\left(v - \frac{\dot{\theta}y}{\sin^2\phi} - \dot{y}\cot\phi\right)}{F_s^2\left(v - \frac{\dot{\theta}y}{\sin^2\phi} - \dot{y}\cot\phi\right)^2 + \frac{1}{(2\pi\tau)^2}} \tag{12}$$

$$A = \sin\left(\frac{2\pi F_s y\Delta\phi}{\sin^2\phi}\right) \tag{13}$$

for the correlators facing the left wall of the corridor. For the right wall, substitute y_r for instead of y_l and negate any appearance of v. In the rotated frame of the robot, substitute $\phi' - \theta = \phi$.

The analytic form is faithful to the response under full simulation (Figure 3). One limitation is that the model assumes both visual sensors of the correlator observe the same local spatial frequency f_s, but in fact f_s is continually varying across the retina. This is not a significant problem

because rapid changes in f_s coincide with high f_s, which are blurred out.

III. A CONTROLLER THAT USES CORRELATORS TO APPROXIMATE RETINAL VELOCITY

The control task is to provide thrust force commands to a pair of fans operating in unison (thrust u_1) or differentially (torque u_2) on a hovercraft robotic testbed. To avoid impacting the walls, the controller uses omnidirectional visual imagery to stabilize its motion along the middle of the corridor (for details of the robotic implementation, see Section V). The vehicle rolls on uni-directional roller balls which allow motion in all directions like a hovercraft with small linear b and rotational c damping coefficients. The dynamics of the vehicle in the moving coordinate frame of the vehicle are thus modelled as

$$\begin{aligned} m\dot{v} &= -bv + u_1 \cos\theta \\ m\ddot{y} &= -b\dot{y} + u_1 \sin\theta \\ J\ddot{\theta} &= -c\dot{\theta} + u_2, \end{aligned}$$

where m is mass and J is the rotational moment of inertia about the center of mass. The vehicle is underactuated because it cannot generate lateral force directly, but lateral dynamics are controllable because they are coupled to forward dynamics by the θ term.

This inertial (non-kinematic) control problem is similar to that encountered by the fly, with its aerial dynamics and pair of independently-controlled wings. An approach for flight control proposed by Humbert et. al. [14] decomposes the retinal velocity function $\nu_r(\phi)$, also known as "optic flow," into sinusoid basis functions (Fourier harmonics), emulating the lobula plate tangential cells (LPTCs) of insects [19]. It can be shown that within a straight corridor, the first few harmonics correspond implicitly to the state variables of the vehicle, $\{v, y, \theta, \dot{\theta}\}$. Linearizing about an operating point of baseline motion along the center of the corridor and using the first few harmonics as the outputs of the system, the vehicle's state is both observable and controllable by the fan pair.

A. Tuning σ and $\Delta\phi$ angles for the environment to approximate retinal velocity ν_r

While correlators do not measure pure retinal velocity, their response rises monotonically with it under certain conditions. Figure 3 shows that the correlator response have a complicated shape for low σ, but as it is increased (increasing blurring), its response resembles a Gaussian function and also the retinal velocity ν_r. It is possible tune the inter-sensor angle $\Delta\phi$ and the width of the Gaussian blurring kernel σ together to insure that response appears Gaussian and does not alias to a negative response.

There are two criteria to satisfy. The first is to insure the spatial aliasing term A does not go negative in the operating regime. For this we require that A be positive for a correlator facing laterally ($\phi = \pm\pi/2$) no matter where the robot is in the corridor $0 < y < 2y_d$ where y_d is the desired distance to be maintained from each wall. This can be insured if the argument to the sin function $2\pi F_s y \Delta\phi \leq \pi$. This puts an upper limit on $\Delta\phi$, and since R is strongest for larger $\Delta\phi$ (thereby minimizing the effects of noise) we choose the upper limit

$$\Delta\phi = \frac{1}{4F_s y_d}. \tag{14}$$

The second criterion is to insure that attenuated contrast factor C^2 falls off sufficiently fast with increasing y that when A does go negative at $y > 2y_d$, the blurring effect in C has attenuated the response to near zero. Under this condition, as the vehicle moves still further from the wall the correlators' response will remain near zero because all of the scenery is blurred away. This is preferable to having aliasing cause the response to go negative. Since the contrast factor dies off as $e^{-(2\pi F_s \sigma y)^2}$, we set the blurring width σ such that at $y = 2y_d$ the signal attenuation has diminished by three standard deviations, or

$$\sigma = \frac{3}{4\sqrt{2}\pi F_s y_d}. \tag{15}$$

Lastly, scale the correlator response so that it matches the retinal velocity at $\phi = \pi/2$ by the factor

$$\alpha = \frac{\nu_r}{R} = \frac{2\pi\tau\left(F_s^2 v_d^2 + 1/(2\pi\tau)^2\right)}{C_0 F_s y_d \exp\left(-\frac{1}{2}\left(2\pi F_s y_d \sigma\right)^2\right)\sin 2\pi F_s \Delta\phi},$$

where v_d is the desired forward velocity operating point for the robot.

The foregoing insure that the correlator response will resemble the pure retinal velocity across ϕ near the desired operating point of forward motion along the axis of the corridor, enabling the use of the Humbert controller. Figure 3 shows the correlator response function $R(\phi)$ for a few different values of σ, showing that the model is faithful to a full correlator simulation.

B. Implementation in simulation

A simulation environment was written in MATLAB to render the scene from the perspective of the robot. Visual updates were received at 60 Hz, equal to the servo rate of the physical robot's controller software. The visual environment consisted of 2880 visual rays, and each of the 64 visual sensors was simulated with a local Gaussian pool across a 3-σ neighborhood of the sensor's orientation, blurring the image. The discrete-time low-pass filters in the correlators were also updated at 60 Hz, and dynamics were integrated with a zero-order hold. To replicate the actual robot, the following parameters were used: $F_s = 1$ cycle/m, corridor width $2y_d$ = 2.5 m and length 12 m, v_d = .25 m/s, σ = .14 rad from (15), and $\Delta\phi \approx .2$ rad from (14) (achieved by using second-nearest neighbors when constructing correlators), $m = 6$ kg, $b = 4.45$ N-s/m, $J = .06$ kg-m^2, and $c = .06$ N-m-s.

Because of the difference between correlators and pure retinal velocity, it was necessary to select different gains than those reported in [14]. Following the same naming convention, the lateral gains of $K_{a_0} = -0.1$, $K_{a_1} = 0.8$, and

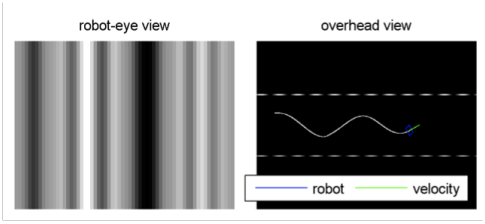

Fig. 4. Simulation of robot using Humbert controller (sinusoid harmonics) and tuned blur width σ and correlator distance $\Delta\phi$. On the left is the 360° visual environment used for state estimation by the robot and on the right the robot is shown by a blue rectangle with its velocity vector a green line. The robot was stable with this controller, but barely.

$K_{a_2} = 4$ were found to give stable performance as shown in Figure 4. However, because the difficulty of decomposing the response R (11-13) into sinusoid harmonics analytically, it was impossible to provide a satisfactory stability analysis. This led to the design of the more-tractable controller in Section IV.

IV. A CONTROLLER DESIGNED FOR CORRELATORS

The limitation with applying the Humbert controller to the correlator response is that correlators do not behave like the retinal velocity. In areas of low visual contrast (such as where it is blurred out at angles near the axis of the corridor) there is no correlator response, leading to an ambiguity between $\dot{\theta}$ and y (Section V). In addition, because of the complexity of the correlator response R, taking sinusoid harmonics of this function is analytically intractable, making a stability analysis impractical.

Our approach is to instead seek an approximation R' of the correlator response that captures its salient characteristics. With the approximation, it becomes possible to analytically take square harmonics which are also functionally equivalent to the wide-field integration performed by the tangential cells in flies [19], sidestepping the difficulty of taking inner products with sinusoids. To disambiguate y and $\dot{\theta}$, an estimate $\hat{\dot{\theta}}$ must be made from a non-visual source such as a gyroscope. With that information, as well as the centroid of the correlator response to estimate θ, by taking first-order Taylor expansions of the sum and difference of the harmonics, it is possible to explicitly estimate the state.

A. Approximations to correlator model

We start with an approximation to the correlator model, simplifying (11-13) by making certain assumptions that hold for the conditions of the robot described in Section V.

1) Aliasing factor (13): A attains an infinite frequency in the vicinity of $\phi = \{0, \pi\}$. However, the correlator response is also heavily attenuated by Gaussian blurring C factor there, meaning that the correlator response at these angles is nearly zero and can be ignored. In addition, the Gaussian blurring "bump" around $\phi = \pm\pi/2$ in C is much narrower across ϕ than the corresponding "bump" in A, so we may approximate A as being constant across ϕ.

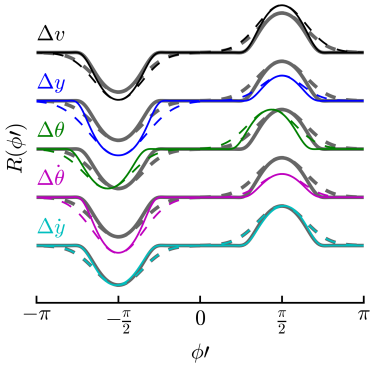

Fig. 5. Correlator response to state perturbations as a function of body-frame angle ϕ'. The baseline response $R(\mathbf{q}_d, \phi')$ is shown in grey and perturbation responses are shown as solid ($R(\mathbf{q}, \phi')$) or dashed (approximation $R'(\mathbf{q}, \phi')$) lines. Changes in $\dot{\theta}$ are essentially indistinguishable from changes in y, necessitating an externally-derived estimate $\hat{\dot{\theta}}$. Changes in $\dot{y}$ have essentially no effect on R. The perturbation magnitudes are $\Delta v = 0.05$ m/s, $\Delta y = 0.2$ m, $\Delta\theta = .2$ rad, $\Delta\dot{\theta} = 0.08$ rad/s, and $\Delta\dot{y} = 0.05$ m/s.

2) Temporal factor (12): T rises to a "saturated" maximum at $f_t = 1/2\pi\tau$ and then drops off, mimicking the temporal frequency peak observed in studies on insect behavior. If, however, the temporal frequency is far below saturation, $f_t \ll 1/2\pi\tau$, then the f_t^2 term in the denominator can be neglected. We also assume assume $\dot{y}$ is small and neglect it. Though it is amplified by the large factor $\cot\phi$ near $\phi = \{0, \pi\}$, temporal frequency saturation negates the effect of this term as $\cot\phi$ grows large, as does the Gaussian blurring. In simulation, it was found to have little effect on the correlator response (Figure 5).

3) Perspective approximation: The quantity $1/\sin^2(\phi)$ appears repeatedly. To approximate it, substitute $\alpha = \phi \pm \frac{\pi}{2}$ and perform a Taylor expansion around $\alpha = 0$ (the same approximation holds in both cases) to find

$$\left.\frac{1}{\sin^2\phi}\right|_{\phi=\pm\frac{\pi}{2}} = \frac{1}{\sin^2\left(\alpha+\frac{\pi}{2}\right)} = \frac{1}{\cos^2\alpha} \approx 1+\alpha^2+\frac{3}{2}\alpha^4+\ldots$$

Applying the approximations to equations (11-13) and truncating the Taylor expansion to 2nd order terms, we arrive at the tractable approximation R' whose factors are

$$C' = C_0 \exp\left[-\tfrac{1}{2}\left(2\pi F_s y\sigma\right)^2\left(1+2\alpha^2\right)\right] \approx C \quad (16)$$

$$T' = \left(2\pi\tau\right)^2 F_s\left(v - \dot{\theta}y\left(1+\alpha^2\right)\right) \approx T \quad (17)$$

$$A' = \sin(2\pi F_s y\Delta\phi) \approx A. \quad (18)$$

Full analytic and approximate responses under different perturbations are shown in Figure 5, showing reasonable agreement across the perturbations of interest.

B. Finding state variables by square harmonics

There is a finite minimum number of correlator pairs at different directions ϕ needed to make the system observable. But we seek an explicit inverse relation between correlator response and state variables with a significant effect on the correlator response for ease of constructing a controller. In

analogy to the fly, suppose we have two lobula plate tangential cells, call them L_L and L_R that take the mean response over the left or right hemisphere, respectively. Because this integration range is much larger than the width of each $R(\phi)$ Gaussian, the integration range may be extended out to infinity without significantly affecting the result, assuming the sensor blur width $\sigma \ll \pi/2$. Thus,

$$L_L = \int_0^\pi R(\phi)d\phi \approx \int_{-\infty}^{\infty} R'(\phi)d\phi = \int_{-\infty}^{\infty} R'(\alpha)d\alpha = L'_L.$$

By approximating this π-width square harmonic as an infinite-domain integral, we need only evaluate integrals of the form $\int x^2 e^{-ax^2}dx$ and $\int e^{-ax^2}$ for which closed-form limits exist. This avoids taking analytically-intractable inner products with sinusoids. And because the limits of integration are much wider than the width of these Gaussian-like functions, for small θ the orientation of the robot may be ignored here. Integrating,

$$L'_L = \gamma e^{-\sigma_s^2 y_l^2}\left(\frac{v}{\sigma_s y_l} - \frac{\dot\theta}{\sigma_s} - \frac{\dot\theta}{4\sigma_s^3 y_l^2}\right)\sin(2\pi F_s y_l \Delta\phi), \tag{19}$$

where we have defined $\sigma_s = 2\pi F_s \sigma$ and the constant $\gamma = 2\pi\tau C_0^2 F_s\sqrt{\pi/2}$ for compactness. For L'_R, negate v and substitute y_r for y_l.

By taking the sum and difference of the two L's (compare to the zeroth cosine and first sine harmonic),

$$\Sigma = L_L + L_R \tag{20}$$
$$\Delta = L_L - L_R \tag{21}$$

as well as the centroid of the bumps, and using internal knowledge of $\dot\theta$ gathered from e.g. gyroscopes, the full state can be extracted as follows.

For a chosen desired operating point $\mathbf{q}_d = \{y_d, v_d, \theta_d = 0, \dot\theta_d = 0\}$ there is a corresponding Σ_d and Δ_d. Shifting the origin to the desired operating point, $\tilde v = v - v_d$, $\tilde y = y_l - y_d = y_d - y_r$, $\tilde\Sigma = \Sigma - \Sigma_d$, $\tilde\Delta = \Delta - \Delta_d$, the shifted coordinates can be interpreted as error to be driven to zero.

First, $\hat{\dot\theta}$ is estimated from gyroscopes, a reasonable proposition for both flies and robots. Then, $\hat\theta$ is found by taking the centroid of the two opposite-hemisphere "bumps" from the perspective of the vehicle, that is, using ϕ' instead of ϕ,

$$\hat\theta = \frac{\langle R(\phi'), |\phi'| - \pi/2\rangle}{\overline{R(\phi')}} \approx \frac{\langle \hat R(\phi'), |\phi'| - \pi/2\rangle}{\Sigma'/2\pi}. \tag{22}$$

where R is the true reading and $\hat R$ is the response read in from sensors. The divisor Σ' is known from the model and thus can be inverted beforehand to be a multiplicative scaling factor, avoiding a division operation. This inner product is similar to the first cosine harmonic, but has a larger domain.

The other states are found by taking Taylor expansions of $\tilde\Sigma$ and $\tilde\Delta$ about the origin. For Σ,

$$\tilde\Sigma = \left.\frac{\partial\Sigma}{\partial\tilde y}\right|_{\tilde y=0}\tilde y + \left.\frac{\partial\Sigma}{\partial\tilde v}\right|_{\tilde v=0}\tilde v + \left.\frac{\partial\Sigma}{\partial\tilde\theta}\right|_{\tilde\theta=0}\tilde\theta + \left.\frac{\partial\Sigma}{\partial\tilde{\dot\theta}}\right|_{\tilde{\dot\theta}=0}\tilde{\dot\theta} + \dots \tag{23}$$

where the ellipsis denotes higher-order terms. Using the approximation Σ', an analytic form can be found for each of the derivatives. Both $\frac{\partial\Sigma'}{\partial\tilde v} = 0$ because the v term changes sign between L_R and L_L (19) and $\frac{\partial\Sigma'}{\partial\tilde\theta} = 0$ because by construction Σ does not depend on θ. Because we have an estimate $\hat{\dot\theta}$ from gyros, we can rearrange (23) to get the estimate

$$\hat{\tilde y} = \left(\hat{\tilde\Sigma} - \hat{\dot\theta}\frac{\partial\Sigma'}{\partial\dot\theta}\right)\Big/\frac{\partial\Sigma'}{\partial y}, \tag{24}$$

where $\hat{\tilde\Sigma}$ is read from the sensors. Both derivatives of Σ' are evaluated at $\mathbf{q}_d$ and can be found from (19) and (20) and can be easily calculated with symbolic software such as Sympy or Mathematica. As in (22), the divisor can be inverted beforehand. We can then calculate

$$\hat{\tilde v} = \left(\hat{\tilde\Delta} - \hat{\tilde y}\frac{\partial\Delta'}{\partial y} - \hat{\dot\theta}\frac{\partial\Delta'}{\partial\dot\theta}\right)\Big/\frac{\partial\Delta'}{\partial v}. \tag{25}$$

The controller was implemented by using an outer loop to set a desired orientation θ_d according to the estimate of the lateral distance from the center, $\theta_d = 0.25\hat{\tilde y}$ to steer the vehicle to the center. A slightly-damped inner loop regulated $\hat\theta$ with the torque command $\tau = K_\theta(s+10)\hat\theta_e$ where $\hat\theta_e = \theta_d - \hat\theta$ is the θ error and $K_\theta/J = 3$. This controller was implemented in simulation and was much more stable. Though we can offer no proof of stability in this report, we note that this controller (with $2\times$ larger blur width) is able to stabilize in a corridor patterned with a texture taken from a real photograph (Figure 6), suggesting its ability to control motion under more real-world circumstances.

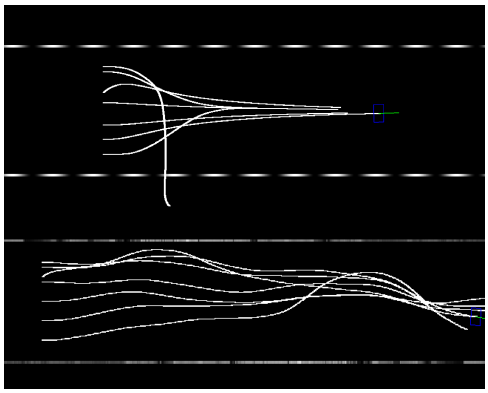

Fig. 6. The second controller has a larger basin of attraction, extending nearly to the walls of the corridor (too close and it overcompensates) (top). It is also able to stabilize motion in a corridor patterned with imagery taken from a real photograph (bottom).

V. ROBOTIC IMPLEMENTATION

We constructed a custom infrared omnidirectional visual sensor that emulated the eyes of the fly to test the controllers outside of simulation. The sensor was mounted on a 20×35×20 cm tall 3 kg robot platform that rolled on low-friction rollerballs and was actuated by fans powerd by 12 V NiMH battery [20]. Computation was performed by an onboard laptop running the RHexLib library on top of the real-time QNX operating system. Fan forces were calibrated beforehand and interpolated from a look-up table to achieve desired forces. An overhead vision system could

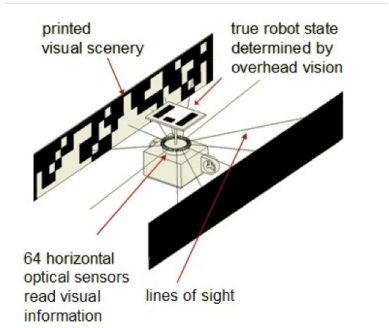

Fig. 7. Diagram of the fan-actuated hovercraft robot in its environment.

track position and pose of the robot for analysis afterward (Figure 7). The parameters of the vehicle and dimensions of the arena were what were used in the simulation and are given in Section III-B.

The custom, 64-element visual array used a short segment of heat-shrink tubing attached to each infrared diode so that its field of view was narrowed to a blur width of $\sigma \approx .077$ rad or $4.4°$. The output of the array was sequentially read by multiplexing through two layers of 8-input analog multiplexors, amplified by operational amplifier, and read into the 8-bit analog-to-digital converter of a PIC microcontroller. The microcontroller read the entire array of luminance readings and communicated the result to the host laptop at 60 Hz over its parallel port. A spatial discrete Gaussian blurring was performed in software with $\sigma = .11$ using the property that the convolutions of two Gaussians with standard deviations σ_1 and σ_2 is a Gaussian with standard deviation $\sqrt{\sigma_1^2 + \sigma_2^2}$ to arrive at the desired $\sigma = .14$ rad. Parallel walls were constructed and illuminated by DC incandescent lights powered by a large power supply to minimize 60 Hz line interference. The infrared sensors could detect printed black vs. white on paper if it was printed by laser, but not by other printing technologies such as ink-jet printing [21]. To normalize the luminance to zero, the mean was taken of all sensors at the beginning of each trial and subtracted out. The robot and its environment are shown in Figure 8.

Correlator response and velocity estimation of the robot are shown in Figure 9. Representative trajectories captured by the overhead vision system are shown in Figure 10. The robot could navigate the corridor, but not always consistently. Clumps of lint and cracks in the floor were large disturbances that were hard to compensate for.

A significant problem was that if it got too close to the walls it would turn into them rather than away. This can be explained as follows. The first few sinusoid harmonics of the retinal velocity ν_r give enough information to observe the vehicle state. But the zeroth cosine harmonic (the mean $\overline{\nu_r}$) is $a_0 = -\sqrt{2}\dot{\theta} + \frac{v_d}{\sqrt{2}y_d^2}y$ and is a function of two state variables. This is not a problem using pure optic flow because they can be separated using information from the second cosine harmonic $a_2 = \frac{v_0}{2y_d^2}y$ also has information on y alone. However, using correlators, the two states cannot be

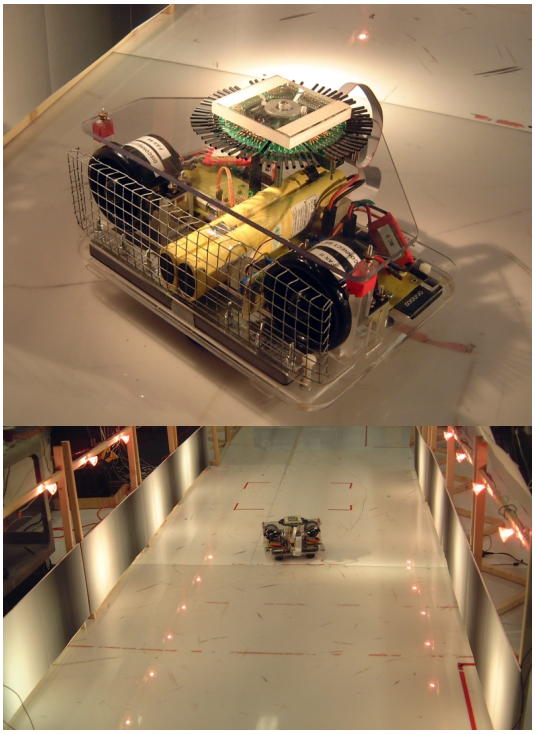

Fig. 8. Fan-actuated hovercraft robot with a 64-element circular omnidirectional array of infrared light sensors (top). The view of each infrared sensor was constricted by a segment of heat-shrink tubing to increase visual accuity and eliminate spurious sources of light. Computation was performed by a subnotebook laptop at the base of the robot and forces were generated by two ducted fans (black) powered by a NiMH battery (yellow). The "hat" used by the overhead tracking system attached by velcro to the translucent piece of acrylic at the top and was removed so that the visual sensor was visible. The vehicle is shown in its environment, patterned walls lit by DC incandescent lighting and a low-friction floor.

disambiguated because there is no response near $\phi = \{\pi, 0\}$ because of blurring of the high spatial frequencies. Hence, being leftward of the centerline of the corridor (stronger correlator response on left hemisphere) is indistinguishable from a rightward rotation $\dot{\theta} < 0$ (larger response on left hemisphere) (see Figure 5). These two conditions require opposite torque responses to reach the desired $y = y_d, \dot{\theta} = 0$. As soon as the robot turned away from the wall arising from a large detected error in y, its large rotation rate $\dot{\theta}$ would immediately eliminate the perceived error and, overcompensating, the robot would turn back into the wall. For this reason, the improved controller of Section IV was devised, but implementation of this controller on the robot we leave for future work.

VI. CONCLUSIONS AND FUTURE WORK

This paper reports progress on the visual control of micro-aerial vehicles using a small number of visual sensors. The controllers rely only on simple-to implement neurologically plausible multiplies and adds and were computationally efficient, requiring less than $6n$ multiply-accumulate operations per step, where n is the number of omnidirectional pixels.

The frequency-domain correlator analysis may extend to naturalistic scenery with known power spectra (for example,

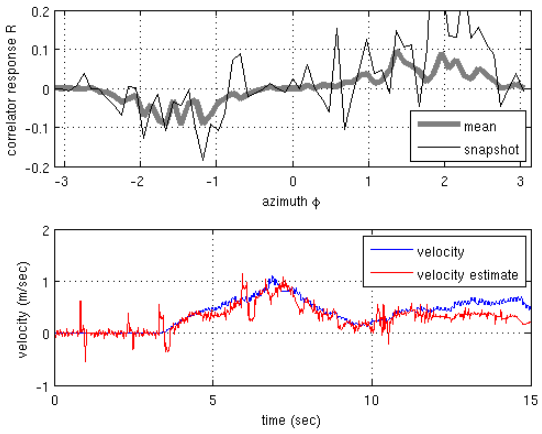

Fig. 9. Correlator response (top) during a run in which the robot was pulled along the corridor by a cord. Mean was taken during seconds 4-7. Variations in mean correlator response arise from variability in visual sensors and inter-sensor spacing. Comparison (bottom) of velocity measured by overhead vision system to estimate derived from sin harmonic of correlator response (filtered with a 10-element box filter).

$1/f$) following [7]. It may be extended to a flat surface projected onto a 3-dimensional hemisphere by incorporating a longitude-like lateral angle β between the poles of expansion and contraction, substituting $y/\cos\beta$ for y. The sensitivity to frequency or contrast exhibited by correlators may be mitigated by compressive nonlinearities [7] or by local contrast estimation [8]. A further question is how many sensors are needed to reach a desired variance in the state estimate.

This work is related to [8] and [22] in that both works aim to stabilize around "snapshot" desired visual response, but differs in that this work seeks a desired correlator snapshot corresponding to a continuous state of motion rather than a fixed pose. Future work may build on both, demonstrating an ability to learn or "bootstrap" the weighting kernels for corridor following.

VII. ACKNOWLEDGEMENTS

The authors would like to acknowledge Andrew Straw for insightful discussion regarding correlators and optic flow estimation.

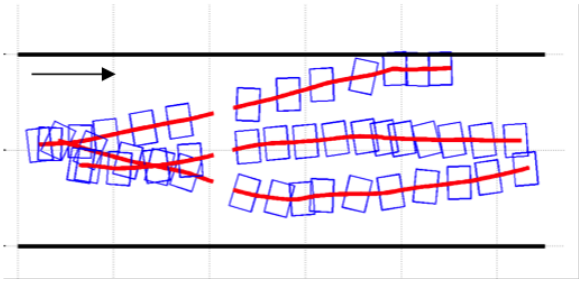

Fig. 10. Trajectories (left to right) of the robot in the corridor captured by the overhead vision system. Gaps in the trajectories coincide with areas where tracking was not available.

REFERENCES

[1] Michael Karpelson, John P. Whitney, Gu-Yeon Wei, and Robert J. Wood. Energetics of flapping-wing robotic insects: Towards autonomous hovering flight. In *International Conference on Intelligent Robots and Systems*, 2010.

[2] Antoine Beyeler, Jean-Christophe Zufferey, and Dario Floreano. Vision-based control of near-obstacle flight. *Autonomous Robots*, 27:201–219, 2009. 10.1007/s10514-009-9139-6.

[3] Graham K. Taylor and Holger G Krapp. Sensory systems and flight stability: What do insects measure and why? *Advances in Insect Physiology*, 34:231–316, 2008.

[4] Martin Egelhaaf, Alexander Borst, and Werner Reichardt. Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *Journal of the Optical Society of America A*, 6(7):1070–1087, Jul 1989.

[5] E. Buchner. *Photoreception and Vision in Invertebrates*, chapter Behavioral Analysis of Spatial Vision in Insects, pages 561–621. Plenum, 1984.

[6] Mandyam V. Srinivasan. Honeybees as a model for the study of visually guided flight, navigation, and biologically inspired robotics. *Physiological Reviews*, 91(2):413–460, 2011.

[7] Ron O. Dror, David C. O'Carroll, and Simon B. Laughlin. Accuracy of velocity estimation by Reichardt Correlators. *Journal of the Optical Society of America A*, 18(2):241–252, February 2001.

[8] Andrea Censi, Shuo Han, Sawyer B. Fuller, and Richard M. Murray. A bio-plausible design for visual attitude stabilization. In *IEEE Conference on Decision and Control*, 2009.

[9] Yang Cheng, Mark Maimone, and Larry Matthies. Visual odometry on the mars exploration rovers - a tool to ensure accurate driving and science imaging. *Robotics and Automation Magazine*, 13(2):54 – 62, June 2006.

[10] Nicolas Franceschini, J. M Pichon, C Blanes, and J. M. Brady. From insect vision to robot vision. *Philosophical Transactions: Biological Sciences*, 337:283–294, 1991.

[11] Nicolas Franceschini, Franck Ruffier, and Julien Serres. A bio-inspired flying robot sheds light on insect piloting abilities. *Current Biology*, 17:1–7, 2007.

[12] J. Serres, D. Dray, F. Ruffier, and N. Franceschini. A vision-based autopilot for a miniature air vehicle: joint speed control and lateral obstacle avoidance. *Autonomou*, 25:103–122, 2008.

[13] K. Weber, S. Venkatesh, and M. V. Srinivasan. *From Living Eyes to Seeing Machines*, chapter Insect inspired behaviors for the autonomous control of mobile robots, pages 226–248. Oxford University Press, 1997.

[14] J. Sean Humbert, Richard M. Murray, and Michael H. Dickinson. Sensorimotor convergence in visual navigation and flight control systems. In *16th International Federation of Automatic Control World Congress, Prague, Czech Republic*, 2005.

[15] James Sean Humbert and Andrew Maxwell Hyslop. Bioinspired visuomotor convergence. *IEEE Transactions on Robotics*, 26:121–130, 2010.

[16] J.-C. Zufferey and D. Floreano. Fly-inspired visual steering of an ultralight indoor aircraft. *IEEE Transactions on Robotics*, 22(1):137–146, Feb. 2006.

[17] Michael Epstein, Stephen Waydo, Sawyer Fuller, Andrew D. Straw, William B. Dickson, Michael H. Dickinson, and Richard M. Murray. Biologically inspired feedback design for *Drosophila* flight. In *American Control Conference*, pages 3395–3401, July 2007.

[18] William B. Dickson, Andrew D. Straw, and Michael H. Dickinson. Integrative model of drosophila flight. *Journal of the American Institute of Aeronautics and Astronautics*, 46(9):2150–2164, September 2008.

[19] Holger G. Krapp and Roland Hengstenberg. Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, 384:463–466, 1996.

[20] Timothy Chung, Lars Cremean, William B. Dunbar, Zhipu Jin, Eric Klavins, David Moore, Abhishek Tiwari, Dave van Gogh, and Stephen Waydo. A platform for cooperative and control of multiple vehicles: the caltech multi-vehicle wireless testbed. In *Conference on Cooperative Control and Optimization*, 2002.

[21] Sawyer B. Fuller, Eric J. Wilhelm, and Joseph M. Jacobson. Ink-jet printed nanoparticle microelectromechanical systems. *IEEE Journal of Microelectromechanical Systems*, 11(1):54–60, Feb. 2002.

[22] Shuo Han, Andrea Censi, Andrew D. Straw, and Richard M. Murray. A bio-plausible design for visual pose stabilization. In *International Conference on Intelligent Robots and Systems*, 2010.