Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Visual Navigation with a 2-pixel Camera—Possibilities and Limitations [1]

John Baillieul * **and Feiyang Kang**

* *The authors are with the Departments of Mechanical Engineering, Electrical and Computer Engineering, and the Division of Systems Engineering at Boston University, Boston, MA 02115. John B. may be reached at* johnb@bu.edu.

Abstract: Borrowing terminology from fluid mechanics, the concepts of *Eulerian* and *Lagrangian optical flow sensing* are introduced. Eulerian optical flow sensing assumes that each photoreceptor in the camera or eye can instantaneously detect feature image points and their velocities on the retina. If this assumption is satisfied, even a two pixel imaging system can provide a moving agent with information about its movement along a corridor that is precise enough to be used as a robust and accurate steering signal. Implementing Eulerian optical flow sensing poses significant challenges, however. Lagrangian optical flow, on the other hand, tracks feature image points as they move on the retina. This form of visual sensing is the basis for many standard computer vision implementations, including Lukas-Kanade and Horn-Schunck. Lagrangian optical flow has its own challenges, not least of which is that it is badly confounded by rotational components of motion. Combined steering and sensing strategies for mitigating the effects of rotational motions are considered.

Keywords: Eulerian optical flow, Lagrangian optical flow

1. INTRODUCTION

Depth cues coming from both binocular vision as well as from optical flow are central to animal navigation, and recent advances in light-weight high-speed computer hardware have made it possible to implement optical flow calculations (e.g. Lucas et al. (1981) and Horn and Schunck (1981)) on mobile robots. Challenges remain, however, in implementing bio-inspired robot control based on optical flow due to the complexity of extracting reliable steering signals from a moving camera. See, e.g. Márquez (2012), Corvese (2018), and Seebacher (2015).

Factors that confound perception based on visual cues include

- Depth discontinuities associated with obstacle boundaries that may be difficult to distinguish from noise in optical flow,
- Moving objects in the field of view that produce localized optical flow that is inconsistent with the optical flow that is generated by self-motion,
- Ephemeral persistence of features within the field of view (FoV),
- Flow indeterminacy due to very sparse optical sensor data in part of the FoV,
- Rotational movement of the optical sensor relative to the features being observed.

In order for autonomous mobile robots to realize full autonomy of movement it will be essential to develop a deeper understanding of the ways in which animal sensory-neuro-motor systems deal with such confounding factors. This very brief note treats visual confounding due to the presence of a rotational component in a robot vehicles's motion.

2. IDEALIZED OPTICAL FLOW

Over many decades, a great deal of research has been devoted to measuring and mitigating the effects of rotational disturbances Royden (1997). While a number of mechanism for determining heading based on optical flow have been proposed—see, for example, Layton and Fajen (2016)—satisfactory solutions to reliable navigation have remained elusive. It is against this backdrop that we hope to obtain insights from the idealized models proposed in what follows. Specifically, we shall consider optical flow based steering laws using the perceived quantity *time-to-transit* along the lines introduced in Sebesta and Baillieul (2012), Kong et al. (2013), Seebacher (2015), Corvese (2018), and Baillieul (2019).

Recall that if a vehicle is moving in a straight line and a feature point lies somewhere ahead of the vehicle—possibly to the left or right in the environment, we can draw a plane that passes through the feature point and is also perpendicular to the line of travel. The *time-to-transit* the feature is defined as the time it will take the vehicle to follow the straight line path at a constant speed until it reaches the plane—as illustrated in 2D in Fig. 1(a). As noted in Baillieul (2019), under the assumption that the vehicle travels at a constant velocity, the time-to-transit can be determined from the movement of feature images on the image plane. Let r denote the radial distance from

[1] Support from various sources including the Office of Naval Research grants N00014-10-1-0952, N00014-17-1-2075, and N00014-19-1-2571 is gratefully acknowledged.

Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

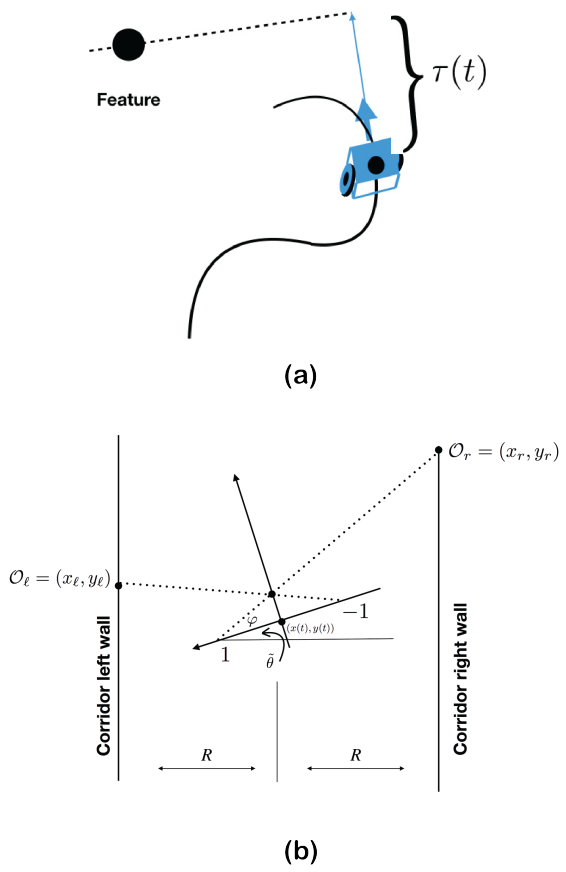

Fig. 1. Vehicle with kinematics (1). Here, the stylized image plane is compressed to one dimension and represented by a line segment on the vehicle y-axis. The vehicle x-axis is the direction of travel, $\tan\varphi = f$, the pinhole camera focal length, and $\theta = \tilde{\theta} + \frac{\pi}{2}$. **(a)** The geometry of time-to-transit, τ as discussed in (Baillieul (2019)). **(b)** The spatial features $\mathcal{O}_r$ and $\mathcal{O}_\ell$ that are registered at ± 1 respectively on the image plane (y-axis) have world frame coordinates $x_r = R$, $x_\ell = -R$, simple geometry shows that $y_\ell = y + f\sin(\theta) + \frac{(R+x+f\cos(\theta))(\cos(\theta)+f\sin(\theta))}{\sin(\theta)-f\cos(\theta)}$ and $y_r = y + f\sin(\theta) + \frac{(R-x-f\cos(\theta))(f\sin(\theta)-\cos(\theta))}{f\cos(\theta)+\sin(\theta)}$.

a feature's image point to the point where the optical axis intersects the image plane. Then the *time-to-transit* the feature point, denoted by, τ, is given by $r/\dot{r}$. Fig. 1(a) illustrates the concept, and details of how time-to-transit is visually perceived are given in Baillieul (2019).

Remark 1. Because τ can be determined by image motions in the image plane (or retina), it is not surprising that there is strong experimental evidence that it is something that animals can perceive. It must be noted, however, that τ as perceived as $r/\dot{r}$ is only equal to the actual time-to-transit when the velocity is constant, and this constrains our use of τ as a proxy for feature distance in our steering laws.

To fix ideas, consider a unicycle vehicle with kinematics

$$\begin{pmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\end{pmatrix} = \begin{pmatrix} v\cos\theta \\ v\sin\theta \\ u\end{pmatrix}, \tag{1}$$

where v is the forward speed in the direction of the body-frame x-axis, and u is the turning rate. For this simple vehicle, we examine the difference between a purely geometric definition of τ and the value of τ that is perceived by means of the movement of image points in the image plane.

Definition 1. Consider a feature point with coordinates (x_f, y_f) and a vehicle whose configuration evolves according to (1). Given the current configuration $(x(t), y(t), \theta(t))$ and speed $v(t)$, the *geometric time-to-transit* (*geometric* τ) is the time it would take the vehicle with its current speed $v(t)$ and heading $\theta(t)$ held constant to cross a line intersecting the feature and perpendicular to the current heading. (See Fig. 1(a).)

Remark 2. Geometric tau is given by the formula

$$\tau(t) = \frac{\cos\theta(t)(x_f - x(t)) + \sin\theta(t)(y_f - y(t))}{v}. \tag{2}$$

3. PERCEPTUAL ALIASING AND QUANTIZATION

It is easy to see that the value of geometric tau given by (2) is maximized if the vehicle is headed directly toward the feature point, and it is zero if the heading direction is perpendicular to the vector from the current vehicle position $(x(t), y(t))$ to the featue point (x_f, y_f). If the vehicle is moving along a curved path, as illustrated in Fig. 1(a), the geometric value of tau (2) will increase if the vehicle is turning toward the feature point and decrease if the vehicle is turning away. If the values of τ are computed by means of the movement of image points on the image plane for the same motions, the values of $\tau = r/\dot{r}$ as perceived in the image plane correspondingly increase or decrease, but they are significantly exagerated as illustrated in Fig. 2.

Time-to-transit is a proxy for depth or distance, and *geometric tau* is taken to be ground truth in visual estimation of distance. In light of the above remark, we must carefully determine the conditions under which perceived values of tau can be used for steering. To do this, it is useful to distinguish between optical flow sensing in which values of τ are registered instantaneously and continuously at a single or discrete set of fixed pixels and optical flow sensing in which feature images must be tracked as they move on the image plane. Borrowing terminology from fluid mechanics, we shall call optical flow tau-based perception in which the tau values are instantaneously sensed at a finite set of body-fixed photoreceptors *Eulerian optical flow sensing*. In Eulerian flow tau-sensing, no single feature point has any significance. The values of tau are instantaneously perceived, utilized in the steering law, and discarded in favor of a continuously updated stream of new values. This is clearly an idealization that will be difficult to realize with ordinary light cameras, but for depth cameras such as LIDARs, it is a useful concept. By contrast, we shall call optical flow sensing in which discrete environmental feature points give rise to continuously moving image points on the image plane or retina *Lagrangian optical flow sensing*. Cameras with standard optics can approximately

Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

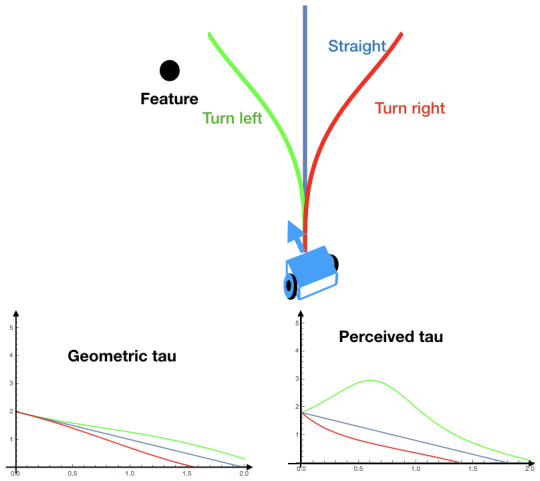

Fig. 2. When time-to-transit, τ is computed using pixel motions on the retina, values deviate from the corresponding geometric quantities. Here, the vehicle turns slightly toward the feature as it rises toward the top of the figure along the green path. While the geometry of this motion clearly increases τ relative to what it would have been traveling at the same speed along the straight (blue) path, the perceived value increases considerably more.

implement Eulerian flow sensing by using a restricted form of Lagrangian flow sensing under stringent conditions that include dense arrays of visible feature points, high densities of photoreceptors, and very modest rotational components of the motion.

In Baillieul (2019), the tau-based optical flow sensing was assumed to be Eulerian, and the following theorem provided a robust steering law for the vehicle (1).

Theorem 1. Consider a mobile camera moving along an infinitely long corridor with every point along both walls being a detectable feature that determines an accurate value of τ. Suppose the corridor has width $2R$ as depicted in Fig. 1**(b)**. Let $\tau_r = \tau(\mathcal{O}_r)$ and $\tau_\ell = \tau(\mathcal{O}_\ell)$ be the respective times to transit the two feature points whose images appear at points equidistant on either side of the optical axis (at ± 1). Then for any gain $k > 0$ there is an open neighborhood U of $(x, \theta) = (0, \frac{\pi}{2})$, $U \subset \{(x, \theta) : -R < x < R;\ \varphi < \theta < \pi - \varphi\}$ such that for all initial conditions (x_0, y_0, θ_0) with $(x_0, \theta_0) \in U$, the steering law

$$u(t) = k(\tau_\ell - \tau_r) \tag{3}$$

asymptotically guides the vehicle with kinematics (1) onto the center line between the corridor walls.

■

This is proved in Baillieul (2019). It provides a remarkably robust steering law that works well even when some of the stated assumptions are significantly relaxed. The requirement that the photoreceptors be symmetrically located on opposite sides of the optical axis is not needed, for instance, and we have the following.

Corollary 1. Consider a mobile camera moving along a corridor as in Theorem 1, and let $\tau_r = \tau(\mathcal{O}_r)$ and $\tau_\ell = \tau(\mathcal{O}_\ell)$ be the respective times to transit two feature points whose images register on opposite sides of the optical axis at $-\delta$ and ϵ respectively. Then for any gain $k > 0$ there is an open neighborhood $U \subset \{(x, \theta) : -R < x < R;\ \varphi < \theta < \pi - \varphi\}$ such that for all initial conditions (x_0, y_0, θ_0) with $(x_0, \theta_0) \in U$, the steering law

$$u(t) = k(\tau_\ell - \tau_r) \tag{4}$$

asymptotically guides the vehicle onto a line parallel to the corridor walls with the asymptotic limit $(x(t), \theta(t)) \to (R(\delta - \epsilon)/(\delta + \epsilon), \pi/2)$.

Proof The proof is essentially the same as for Theorem 1, except the geometry is changed such that the coordinates of the feature points are

$$\mathcal{O}_\ell = \begin{pmatrix} -R \\ y + f\sin(\theta) + \dfrac{(R + x + f\cos(\theta))(\delta\cos(\theta) + f\sin(\theta))}{\delta\sin(\theta) - f\cos(\theta)} \end{pmatrix}$$

and

$$\mathcal{O}_r = \begin{pmatrix} R \\ y + f\sin(\theta) + \dfrac{(R - x - f\cos(\theta))(f\sin(\theta) - \epsilon\cos(\theta))}{f\cos(\theta) + \epsilon\sin(\theta)} \end{pmatrix}.$$

■

Remark 3. Clearly if one photoreceptor is significantly closer than the other to the optical axis, the steady-state motion will be closer to the wall opposite to the photoreceptor.

Corollary 2. Consider a mobile camera as in Corollary 1 with the same $\tau_\ell = \tau(\mathcal{O}_\ell), \tau_r = \tau(\mathcal{O}_r)$ and photoreceptors at the same $(-\delta, \epsilon)$ locations in the body frame. For any gain $k > 0$, there is an open neighborhood $\subset \{(x, \theta) : -R < x < R;\ \varphi < \theta < \pi - \varphi\}$ such that for all initial conditions (x_0, y_0, θ_0) with $(x_0, \theta_0) \in U$, the steering law

$$u(t) = k(\delta\tau_\ell - \epsilon\tau_r) \tag{5}$$

asymptotically guides the vehicle onto a line parallel to the corridor walls with the asymptotic limit $(x(t), \theta(t)) \to ((\epsilon - \delta)/2, \pi/2)$.

■

Under the conditions assumed in the theorem and its corollaries, precise steering is possible. For standard light cameras, approximately comparable results can be achieved using Lagrangian tau-based flow sensing, provided there is an adequate density of features. The key is to segment motions into sequences of *sense - perceive - act* components. Toward validating this strategy, the following gives conditions under which piecewise constant curvature steering together with depth sensing aligns the vehicle on the centerline of the corridor as desired.

Theorem 2. Consider the planar vehicle (1) for which the steering law is of the sample-and-hold type:

$$u(t) = k[\tau_\ell(x(t_i), \theta(t_i)) - \tau_r(x(t_i), \theta(t_i))], \quad t_i \le t < t_{i+1}, \tag{6}$$

where the sampling instants $t_0 < t_1 < \dots$ are uniformly spaced with $t_{i+1} - t_i = h > 0$. Then for any sufficiently small sampling interval $h > 0$, there is a range of values of the gain $0 < k < k_{crit}$ such that the sampled control law (6) asymptotically guides the vehicle with kinematics (1) onto the center line between the corridor walls.

Proof: As in Theorem 1, we assume that the forward speed is constant ($v = 1$). We also assume a normalization of scales such that $f = 1$. It is again convenient to consider

Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

the angular coordinate $\phi = \theta - \pi/2$. In terms of this, we have

$$\dot{\phi} = k[\tau_\ell(x(t_i), \phi(t_i) + \pi/2) - \tau_r(x(t_i), \phi(t_i) + \pi/2)]$$

on the interval $t_i \leq t < t_{i+1}$. Given the explicit formulas for τ_ℓ and τ_r, and given that the right hand side of the above differential equation is constant, we have the following discrete time evolution

$$\phi(t_{i+1}) = \phi(t_i) + hk\frac{2\sin\phi(t_1)(R + \cos\phi(t_i)) - 2x(t_i)\cos\phi(t_i)}{\sin^2\phi(t_i) - \cos^2\phi(t_i)} \tag{7}$$

In other words, the discrete time evolution of the heading ϕ is given by iterating the x-dependent mapping

$$g(\phi) = \phi + hk\frac{2\sin\phi(R + \cos\phi) - 2x\cos\phi}{\sin^2\phi - \cos^2\phi.} \tag{8}$$

Differentiating, we obtain

$$g'(\phi) = 1 + \frac{2hk(-2 - 3R\cos\phi + R\cos 3\phi + 3x\sin\phi + x\sin 3\phi)}{\cos^2\phi - \sin^2\phi}. \tag{9}$$

The numerator is negative in the parameter range of interest, while the denominator is positive. Hence, we can choose k sufficiently small that g is a contraction on $-\pi/4 < \phi < \pi/4$ uniformly in x in the range $-R < x < R$. Thus the iterates of ϕ under the mapping (8) converge to 0, and because $\dot{x} = -\sin\phi$, this proves the theorem. ■

Remark 4. With this sample-and-hold Eulerian steering law, the next step is to show that the same path will be well approximated by a corresponding Lagrangian sample and hold result and that interleaving very short straight line segments with the constant curvature law (6) will also asymptotically align with the desired path. As noted in Hildreth (1992), for short straight paths, heading direction can be determined accurately with short two to three frame sequences.

The next step toward understanding path interpolation by alternating sequences of straight (for sensing) and curved (for steering) path segments is a heuristic model for calculating τ values along curved paths. Specifically, given a feature point (x_f, y_f), instead of computing $\tau = r/\dot{r}$ along a path $(x(t), y(t), \theta(t))$ of (1), we compute the *quasi-linear time-to-transit*

$$\tau^*(t) = r(x(t), y(t), \theta(t))/(\frac{\partial r}{\partial x}\dot{x} + \frac{\partial r}{\partial y}\dot{y}).$$

This value separates the component of image movement on the retina due to tangential speed along the path from the motion due to curving. Space does not allow a complete exploration of steering laws based on τ^*, but to summarize what's involved, the steps in comparing these with the Eulerian model are given in the algorithm below.

Conclusion. The stylized models considered above do not capture all essential aspects of visual navigation, but they are capable of highlighting purely geometric components of factors that confound visual perception of motion. Work is ongoing to understand how these factors are best dealt with in laboratory settings using state-of-the-art optical flow such as Ilg et al. (2017).

Algorithm 1 Determine steering signal from Lagrangian optical flow—*Sense - Perceive - Act*

1: Choose a time interval h that is compatible with the rate at which features enter and leave the field of view.
2: Form thee partition: $t_0 < t_1 < \ldots$ with $t_{j+1} - t_j = h$.
3: Define a segmentation protocol that groups feature images from each wall whose τ^* values and image locations determine the steering signal for the current time interval.
4: At each switching time t_j return to 3 and repeat.

REFERENCES

Baillieul, J. (2019). Perceptual control with large feature and actuator networks. In *58th IEEE Conference on Decision and Control (CDC)*, 3819–3826. IEEE.

Corvese, L. (2018). *Perceptual Aliasing in Vision Based Robot Navigation.* Thesis, Boston University. https://open.bu.edu/handle/2144/27453.

Gupta, A., Ingle, A., Velten, A., and Gupta, M. (2019). Photon-flooded single-photon 3d cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 6770–6779.

Hildreth, E.C. (1992). Recovering heading for visually-guided navigation. *Vision research*, 32(6), 1177–1192.

Horn, B.K. and Schunck, B.G. (1981). Determining optical flow. In *Techniques and Applications of Image Understanding*, volume 281, 319–331. International Society for Optics and Photonics.

Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., and Brox, T. (2017). Flownet 2.0: Evolution of optical flow estimation with deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2462–2470.

Kong, Z., Özcimder, K., Fuller, N., Greco, A., Theriault, D., Wu, Z., Kunz, T., Betke, M., and Baillieul, J. (2013). Optical flow sensing and the inverse perception problem for flying bats. In *52nd IEEE Conference on Decision and Control*, 1608–1615. IEEE.

Layton, O.W. and Fajen, B.R. (2016). Competitive dynamics in mstd: A mechanism for robust heading perception based on optic flow. *PLoS computational biology*, 12(6).

H.C. Longuet-Higgins and K. Prazdny, (1980). The interpretation of a moving retinal image. *Proc. R. Soc. Lond. B*, 208, 385-397.

Lucas, B.D., Kanade, T., et al. (1981). An iterative image registration technique with an application to stereo vision. DARPA Workshop.

Márquez-Valle, P., Gil, D., and Hernàndez-Sabaté, A., (2012). A complete confidence framework for optical flow. in *ECCV* 2012 Workshops/Demos, Part II, LNCS 7584, 124-133.

Royden, C.S. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *JOSA A*, 14(9), 2128–2143.

Sebesta, K. and Baillieul, J. (2012). Animal-inspired agile flight using optical flow sensing. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 3727–3734. IEEE.

Seebacher, J.P. (2015). *Motion control using optical flow of sparse image features.* Ph.D. thesis, Boston University. https://open.bu.edu/handle/2144/15191.