Space-time representations of surfaces in motion

David Young

School of Cognitive and Computing Sciences, University of Sussex

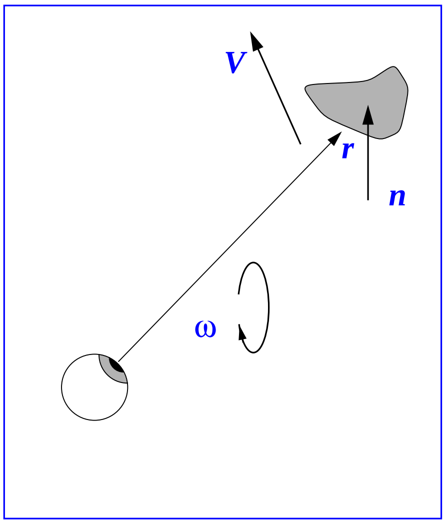
The problem

To understand the visual information provided by a single small patch of a planar textured surface in motion.

- This is a classical problem in computational vision. The formal framework is well established (e.g. in the references below).
- Assumption: a careful analysis will help in understanding more complex visual information gathering. Successful interaction with surface patches can be seen as the basis of locomotion and many interceptive acts such as catching and grasping.
- Approach: to describe what the visual information tells us in terms that relate to possible interactions between the observer and the surface.
- Restriction: we deal with the *instantaneous* state of the surface — its position and velocities — not its trajectory extended in time.

Koenderink, J.J. (1986) Optic flow. *Vision Res.* **26**, 161-180.

Koenderink, J.J. & van Doorn, A.J. (1992) Second-order optic flow. *J. Opt Soc. Am.*, **9**, 530-538.



Localisation

Here, *localisation* of the surface is taken to mean specifying that subspace of the full parameter space that is consistent with the perceptual information.

Physical description

The *physical* relationship between the observer and the patch can be described using the following set of variables. We take a frame of reference attached to the eye and define:

$\boldsymbol{r}$: the position of the surface patch;

$\boldsymbol{V}$: the velocity of the patch;

ω: the angular velocity (rate of rotation) of the patch about the observer — equal to the inverse of the angular velocity of the eye relative to the patch;

$\boldsymbol{n}$: the inverse surface normal of the patch (slant and tilt).

Each of these is a 3-vector, so we seem to need 12 scalar parameters. However, $\boldsymbol{n}$ is a unit vector so only 11 of them are independent.

Tracking - an important special case

If the eye is tracking the patch, only one component of $(\boldsymbol{V} + \omega \times \boldsymbol{r})$ is non-zero, and only one component of $\boldsymbol{r}$ is non-zero, so we have only 7 physical parameters. (The non-zero components are those along the line of sight.)

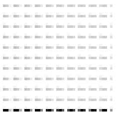
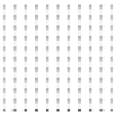
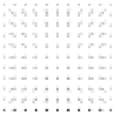
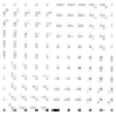
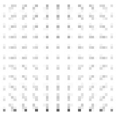
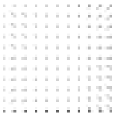
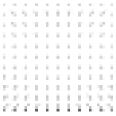
Optical description

The *optical* relationship between the observer and the patch can be described in terms of the *optic flow*.

- We assume that the surface has sufficient visual texture for the optic flow to correspond to measurable image motion.
- We restrict attention to the *instantaneous* optic flow — that is the rate of change of direction of texture elements — and ignore the extended history of the flow.

Optic flow parameterisation

The optic flow produced by the surface patch can be *parameterised*: it is a linear combination of the patterns below (plus a relatively small third-order component). The parameters are the coefficients giving the contribution of each pattern.

Zero order flow	Translation v_1 	Translation v_2 		
First order flow	Dilation D 	Rotation R 	Shear S_1 	Shear S_2 
Second order flow	Second order Q_1 	Second order Q_2 	(This is not *general* second order flow, only the two components produced by a planar surface patch.)	

Tracking and zero-order flow

The zero-order flow is affected by pan and tilt eye movements. It provides the obvious feedback signal for tracking — i.e. it is normally nulled out by gaze control.

This also means that it gives no useful information unless an independent motor frame of reference can be established, relative to which eye rotation can be estimated. Two possibilities are:

- a visual frame, established by observing other, preferably distant, elements of the environment;
- a mechanical frame, established by proprioceptive information about gaze direction relative to the body, and maybe physical contact with surfaces.

Frames of reference and higher-order flow

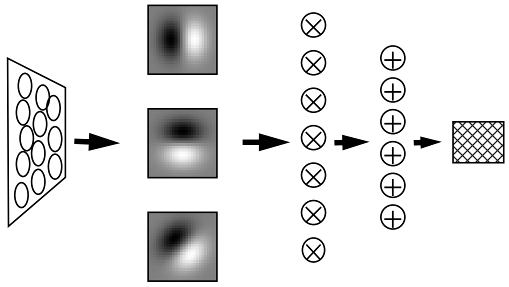
Higher-order flow is unaffected by pan and tilt eye movements. Thus it can provide useful information even if no independent frame of reference is available.

This is the case explored in the present work.

Estimating the parameters

Estimating the parameters of a flow field for a smooth textured surface is easier and more reliable than estimating individual flow vectors. First order is more accurate than second order.

Many approaches have been investigated — a common scheme is:



Action description

The relationship between the observer and the surface may also be specified using parameters which relate to how the observer might *interact* with the surface. An example would be their distance apart, in the context of a grasping task.

Possibly stretching Gibson's terminology, we refer to these parameters as *affordances*.

We propose the following as affordances in the context of locomotor tasks in 3 dimensions. (An *immediacy* is the inverse of a predicted time to contact.)

i_P: the immediacy of the plane containing the surface patch;

i_A: the immediacy of the plane normal to the line of sight and intersecting the surface patch;

c: course angle, specifying the orientation in the image of the projection of $\boldsymbol{V}$;

t: tilt, specifying the orientation in the image of the projection of $\boldsymbol{n}$;

s: slant, the angle between $\boldsymbol{n}$ and the line of sight;

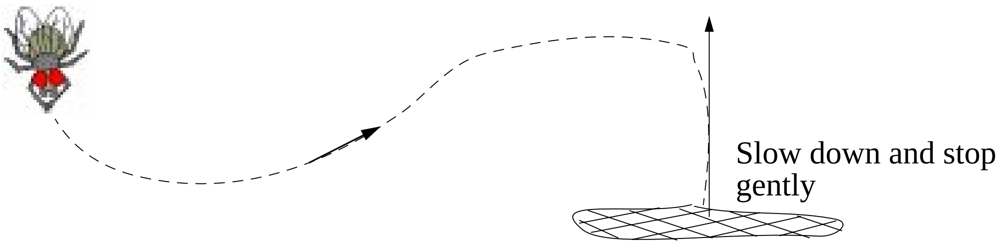
m: miss, the angle between $\boldsymbol{V}$ and the line of sight (not independent of rest).

There are, of course, many other candidates for sets of affordances.

See http://www.cogs.susx.ac.uk/users/davidy/ecvp2000/presentation.html.

Why is it reasonable to call these parameters "affordances"?

Computational experiments show that knowledge of i_P, i_A, c and t is sufficient to support a simple strategy for a form of docking: making a gentle controlled approach along the surface normal after starting with random position and velocity, all in 3 dimensions.



How are these descriptions related?

The physical description specifies the optic flow parameters. (That is, if we know the physical description, we can calculate the optic flow.)

The optic flow **almost** specifies the physical description. We need in addition to make a choice between up to 6 discrete solutions, together with one independent scalar measurement to resolve the *speed-scale* ambiguity. (A patch with given flow may be far and fast or close and slow.) However, computation of the physical parameters from the optic flow is complex, difficult and involves all components of the flow.

The physical description specifies the affordances.

The affordances **do not** specify the physical description.

The first order optic flow **almost** specifies the four affordances i_A, i_P, c, t. We need in addition independent information about one unknown parameter, such as *spin* (the component of ω along the line of sight), or one of the affordances, and a way of choosing between up to 4 discrete solutions.

If second order optic flow is included, then in addition slant s and miss m are specified.

Flow information and parameter subspaces

A summary of how various kinds of information contribute to restricting the parameter space of the system. To some extent this restates the results given in the references on p. 1, but that work does not relate the geometrical findings to the control of action.

Measurements available	No. of undetermined parameters	Undetermined parameters/ ambiguities	No. of determined parameters	Determined parameters	Possibilities for control of action
None	11	All	0	None	None
Position of patch in image	9	All except direction	2	Direction of patch in visual frame of ref.	Orientation towards surface
Position + 0-order flow	7	Many	4	Projection of $\boldsymbol{V} + \omega \times \boldsymbol{r}$	Tracking eye movement
Position + 0-order flow + independent frame of reference giving pan, tilt components of ω	5	Speed-scale; spin; miss; $\boldsymbol{n}$	6	Above + projection of $\boldsymbol{V}$ hence course angle c	Motion towards predicted nearest approach — e.g. some catching strategies
Position + 1st order flow	5	Various combinations	6	Various combinations of $\boldsymbol{V}$, $\boldsymbol{n}$, ω	No clear additional operations
Position + 1st order flow + independent measure of e.g. spin	4	Speed–scale; slant–miss; pan, tilt rates	7	Direction, affordances i_A, i_P, c, t, spin	Docking and other manoeuvres involving surface normal
Position + 1st order flow + 2nd order flow	3	Speed-scale; pan, tilt rates	8	Above + s and m	Tasks requiring explicit knowledge of slant or predicted collision point
Position + all components of flow up to 2nd order	1	Speed-scale	10	Above + remaining components of ω	Tasks requiring explicit control of rate of rotation

Experimental questions

In general, how do biological visual systems exploit this information?

Specifically:

- There is a trade-off between accuracy of flow parameter measurement and the number of useful parameters determined. Do biological visual systems use second order flow? Do they use shear?
- Does this depend on the task?
- Additional information — e.g. spin — can sometimes be replaced by an assumption — e.g. the spin is close to 0 because cyclotorsional movements are small. Do biological visual systems rely on any such assumptions, and if so, what?

Experiments

Simulated flow fields can give the observer the perception of a surface in motion. Under these conditions a subject can be asked to perform tasks that use properties of the surface.

For example:

- The spin-zero assumption for first order flow use can be tested by looking at estimates of course angle c and tilt t.
- The contribution of second order flow can be tested by looking at whether slant s or miss m can be estimated.
- Second order flow can be omitted, exaggerated or brought into conflict with first order flow.

The mathematical details

We choose a co-ordinate system centred on the point of observation with the 3-axis along the line of sight, and the 1- and 2-axes in the plane of projection. The vector $\boldsymbol{x}$ is position in the spherical image, measured in radians of visual angle. The optic flow field $\boldsymbol{u}$ is expressed up to second order in position thus:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} D + S_1 & S_2 - R \\ S_2 + R & D - S_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} Q_1 & 0 & Q_2 \\ 0 & Q_2 & Q_1 \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{bmatrix}$$

Defining $d = \boldsymbol{n} \cdot \boldsymbol{r}$ (the perpendicular distance from the observer to the plane of the surface patch), the optic flow parameters are determined by the physical parameters according to:

$$v_1 = V_1 n_3 / d + \omega_2 \qquad D = (V_1 n_1 + V_2 n_2 - 2 V_3 n_3)/(2d) \qquad Q_1 = (V_3 n_1 + V_1 n_3)/d$$

$$v_2 = V_2 n_3 / d - \omega_1 \qquad R = (V_2 n_1 - V_1 n_2)/(2d) + \omega_3 \qquad Q_2 = (V_3 n_2 + V_2 n_3)/d$$

$$S_1 = (V_1 n_1 - V_2 n_2)/(2d)$$

$$S_2 = (V_1 n_2 + V_2 n_1)/(2d)$$

The affordances are related to the physical variables by

$$i_A = (V_1 n_1 + V_2 n_2 + V_3 n_3)/d$$

$$i_P = V_3 n_3 / d$$

$$t = \operatorname{atan}(n_1 / n_2)$$

$$c = \operatorname{atan}(V_1 / V_2)$$

$$s = \operatorname{atan}(\sqrt{n_1^2 + n_2^2} / n_3)$$

$$m = \operatorname{atan}(\sqrt{V_1^2 + V_2^2} / V_3)$$

The relationships between optic flow and affordances given at http://www.cogs.susx.ac.uk/users/davidy/ecvp2000/presentation.html follow.