Gradient, STEM, and Regression Models for Motion Perception: Relationships and Extensions

Eero P. Simoncelli
and
Edward H. Adelson

MIT Media Laboratory
Cambridge, MA

Outline

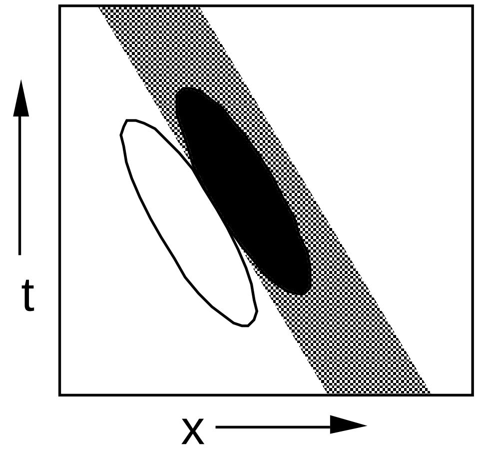
• Three seemingly unrelated standard techniques for estimating motion.

• Similarity of these techniques.

• Failure of these techniques to detect or represent multiple motions.

• A simple extension that detects multiple motions.

Spatio-temporal Energy Models



- Concept: motion is orientation in "space-time".

  f_r, f_l, f_s

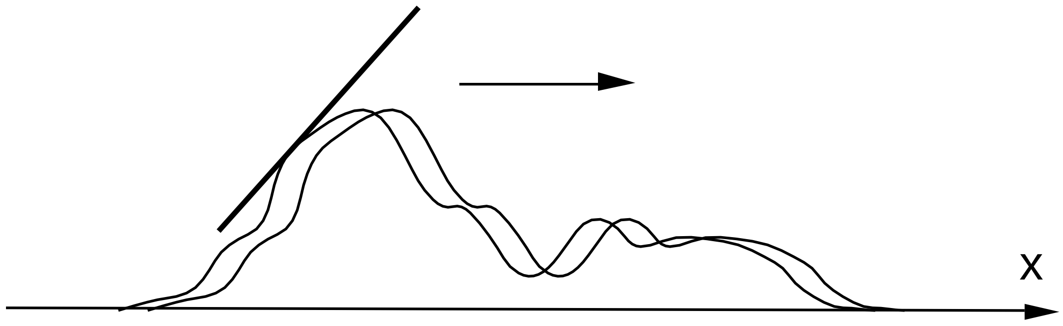
- Measurements: convolution with tuned filters.

  $R = f_r * I$, $L = f_l * I$, $S = f_s * I$

- Computation of opponent "motion energy" from quadratic combinations of filter outputs:

  $E = R^2 - L^2$

Gradient techniques



• Concept: intensity conservation. Image approximated by a translating ramp.

$$I_x v + I_t = 0$$

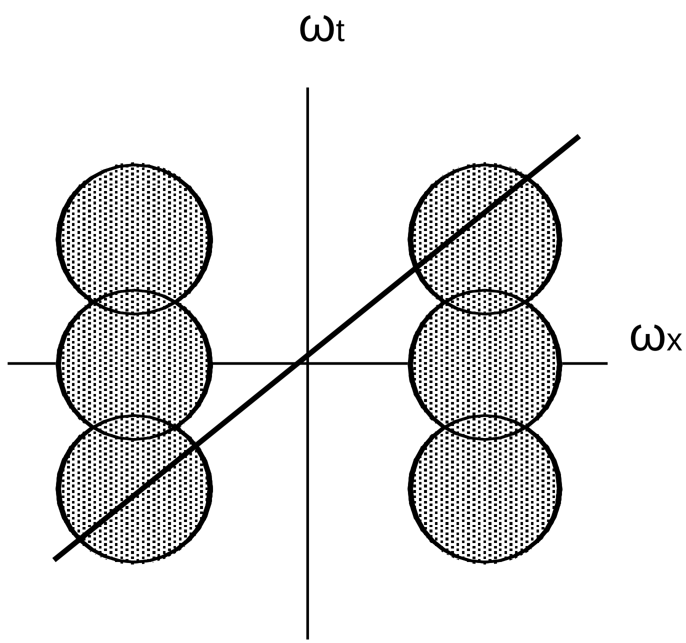
• Measurements: partial derivatives.

$$I_x = dI/dx, \qquad I_t = dI/dt$$

• Least squares velocity: computed by blurring quadratic combinations of measurements.

$$v = \Sigma I_x I_t \; / \; \Sigma I_x^2$$

Spatio-temporal Frequency Regression



• Concept: Fourier transform of a translating image lies on a line.

• Measurements: use Gabor-like filters to estimate spectral content.

• Velocity: slope of best-fitting line.

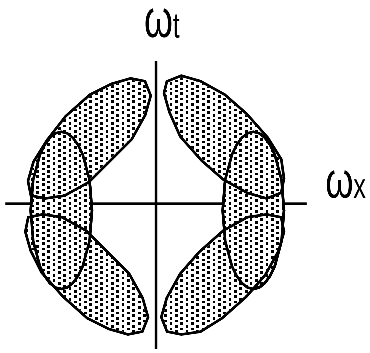
STEM: common form

• Choose filters that are directional derivatives of prefilter, g:

$f_r = g_x + g_t$

$f_l = g_x - g_t$

$f_s = g_x$



• Now can compute v from opponent energy:

$$R^2 - L^2 = (g_x * I)(g_t * I) = I_x I_t$$

$$v = \Sigma (R^2 - L^2) / \Sigma S^2$$

Gradient: common form

• Derivatives are computed by convolving with two filters:

$I_x = g_x * I, \qquad I_t = g_t * I$

where g is a (lowpass) interpolation prefilter.

• As with STEM, R, S, and L may be computed from space-time derivatives:

$$R^2 - L^2 = I_x I_t$$

$$v = \Sigma (R^2 - L^2) \; / \; \Sigma S^2$$

Regression: common form

- Compute linear regression on the prefiltered image spectrum:

$$E(v) = \Sigma\, (v\, \omega_x + \omega_t)^2 \; |\tilde{I}\,|^2$$

$$= \Sigma\, |\, v\, \omega_x \tilde{I} + \omega_t \tilde{I}\,|^2$$

- Use Parseval's Theorem to switch to space-time domain:

$$E(v) = \Sigma\, (v\, I_x + I_t)^2$$

$$v_{min} = \Sigma\, I_x I_t \;/\; \Sigma\, I_x^2$$

- As with gradient solution, can write as opponent energy computation:

$$v = \Sigma\, (R^2 - L^2) \;/\; \Sigma\, S^2$$

All Three Techniques . . .

• ... are fundamentally the same when considered as linear least-squares velocity estimators with suitable choice of filters.

• ... are based on local linear measurements (convolutions) followed by non-linear (quadratic) velocity computation.

• ... can be extended to compute physiologically plausible distributed velocity representations (ARVO-90).

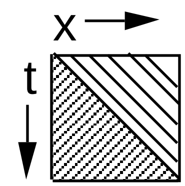
• ... are designed to compute *single* motions:

Information about multiple motions is discarded at the measurement stage!

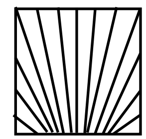
Multiple Motions

• Three Example Cases:

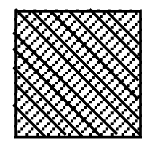
- Occluding contours.



- Non-translational flow: (e.g. expansion, contraction, rotation).



- Transparency.



• Comments:

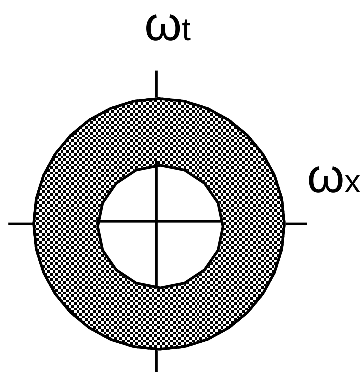
- Occur frequently in natural scenes.

- Provide important information about depth ordering, approach/withdrawal, material properties, etc.

Detecting Multiple Motions

• One additional linear measurement:

$$A = f_{total} * I$$



• To determine if there are multiple motions, compare total energy to derivative energy:

$$E_{total} = A^2$$

$$E_{derivative} = I_x^2 + I_t^2 = R^2 + L^2$$

Results

• One-dimensional test image:

[Image Here]

• Multiple motion detector output:

[Image Here]

Conclusions

• Gradient, Spatio-temporal Energy, and Regression models for early motion processing are very similar (identical with proper choice of parameters).

• All three of these techniques (in their common forms) cannot detect or represent multiple motions.

• There are simple extensions to these techniques which allow detection of multiple motions.