Edge detection and contour integration in human and machine vision

Keith May

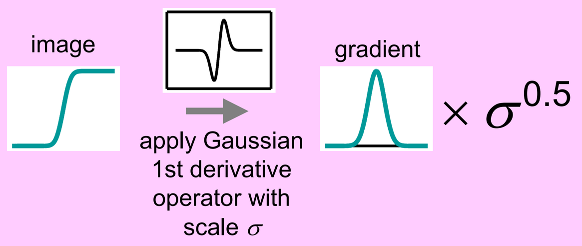
www.keithmay.org

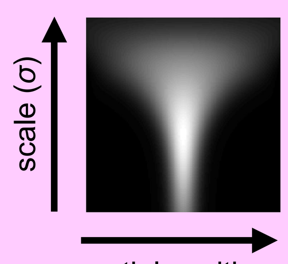
UCL Department of Computer Science

Scale-space edge detection algorithms in human and machine vision

1) Lindeberg (1998): Machine vision algorithm

- Multiple channels with different scales, σ





- Detect edges by looking for peaks in scale-space
- Position of peak along “spatial position” dimension gives the spatial position of the edge
- Position of peak along “scale” dimension gives the blur of the edge

Reference

Lindeberg, T. (1998). Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, *30*, 117–154

2) McIlhagga (2011): Optimal linear edge detection filter

- Corrected errors in Canny's (1986) derivation
- Modelled image noise as white noise, with flat power spectrum, n_0^2
- Modelled surrounding edges as brown noise, with power spectrum, C^2/ω^2, where ω is spatial frequency
- Decompose filter into whitening filter, $W(\omega)$, followed by detection filter $K(\omega)$

$$F(\omega) = W(\omega)K(\omega) \quad \text{where} \quad W(\omega) = \frac{i\omega}{\sqrt{C^2 + n_0^2\omega^2}}$$

- For high-contrast Gaussian edges, optimal $K(\omega)$ is whitened Gaussian edge:

$$K(\omega) = W(\omega)\left(\frac{-i}{\omega}\right)Gauss(\omega,\sigma)$$

- So optimal filter given by

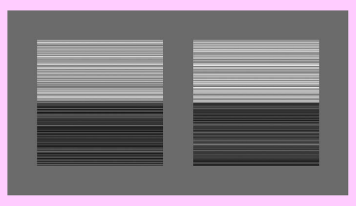
$$F(\omega) = \frac{i\omega}{C^2 + n_0^2\omega^2}Gauss(\omega,\sigma)$$

- For low image noise ($n_0 \approx 0$), $F(\omega)$ is a Gaussian 1st derivative filter, and McIlhagga's optimal edge detection algorithm is identical to Lindeberg's
- Unlike in Lindeberg's algorithm, McIlhagga's filters adapt to changes in image noise or surrounding image clutter, so they remain optimal

Reference

McIlhagga, W. (2011). The Canny edge detector revisited. *International Journal of Computer Vision*, *91*, 251–261

3) McIlhagga & May (2012): Blur detection by humans



- Observers see a sharp edge next to a Gaussian-blurred edge, both with added noise
- Asked which is the blurred edge
- Simulations of the task with McIlhagga's (2011) optimal algorithm explain human performance with remarkable accuracy on a trial-by-trial basis

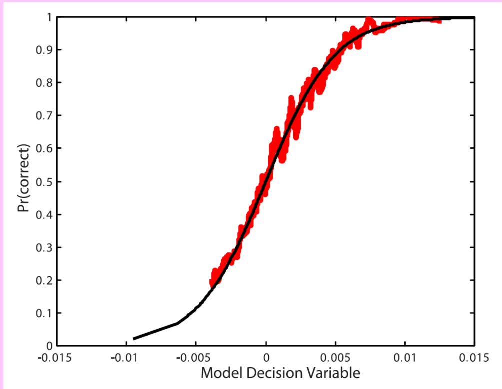


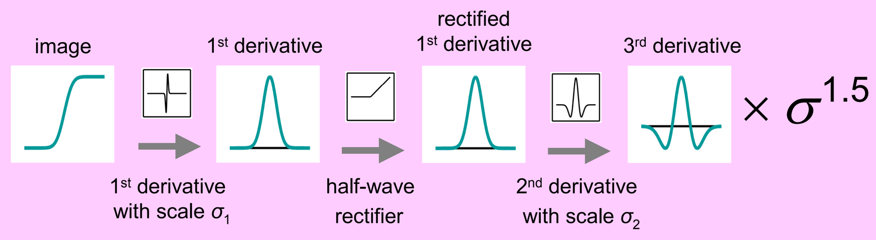
Figure 5. How well the optimal edge detector model accounts for observer KAM's probability correct. The *x* axis plots the decision variable for the optimal model, and the black curve gives the model's probability correct as a function of the decision variable. The red jagged line shows the human observer's probability correct, as a function of the model decision variable. This was calculated as follows. For each value of *x*, we selected trials in which the model decision variable was near *x*, and then calculated the observer's probability correct within that set of trials.

Reference

McIlhagga, W.H. & May, K.A. (2012). Optimal edge filters explain human blur detection. *Journal of Vision*, *12*(10):9, 1–13

4) Georgeson, May, Freeman & Hesse (2007): Blur matching by humans

- Observers see a Gaussian-blurred edge and a blurred edge with a non-Gaussian profile
- They adjust the Gaussian edge until it looks as blurred as the other edge
- Both stimuli are noise-free, so McIlhagga's optimal algorithm reduces to Lindeberg's algorithm
- But Georgeson et al. found that performance was best explained by a scale-space algorithm similar to Lindeberg's but with a nonlinear operation in each channel (the N_3^+ model)



σ is the channel scale $= \sqrt{\sigma_1^2+\sigma_2^2}$

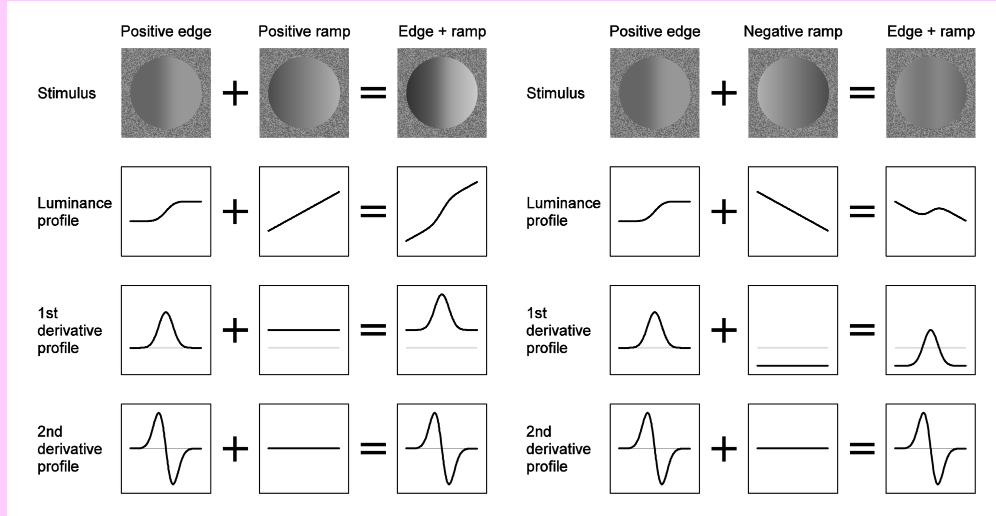
- The N_3^+ model correctly predicts blur matches by humans for a large collection of different edge profiles
- Why does McIlhagga's optimal model fail to predict the data in this study? His filters are the optimal *linear* filters: Maybe in low noise conditions, Georgeson et al.'s nonlinear filter is better than the best linear filter

Reference

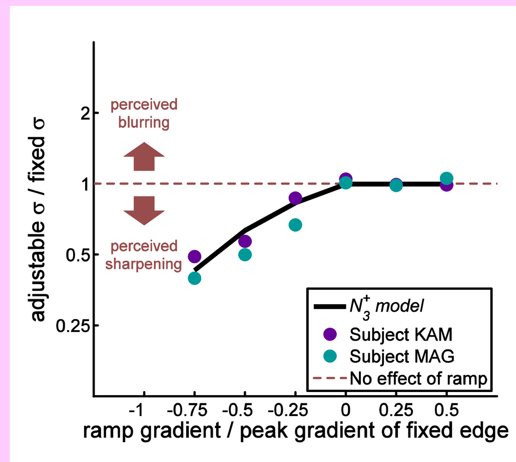
Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale–space analysis of edge and blur coding in human vision. *Journal of Vision*, *7*(13):7, 1–21

5) May & Georgeson (2007): Test of 2nd-derivative based edge detection in human vision

- Many edge detection models in biological vision begin by filtering the image with a 2nd-derivative operator (Marr & Hildreth, 1980; Watt & Morgan, 1985; Georgeson, 1992; Kingdom & Moulden, 1992)
- These models predict that adding a linear ramp to an edge should not change its appearance



- But adding a ramp with opposite polarity to the edge makes the edge look much sharper
- This rules out any model that starts off by applying a 2nd derivative operator to the image
- Georgeson et al.'s (2007) N_3^+ model correctly predicts the effect of adding the ramp



References

Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London B*, *249*, 235–245

Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale–space analysis of edge and blur coding in human vision. *Journal of Vision*, 7(13):7, 1–21

Kingdom, F. & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, *32*, 1565–1582

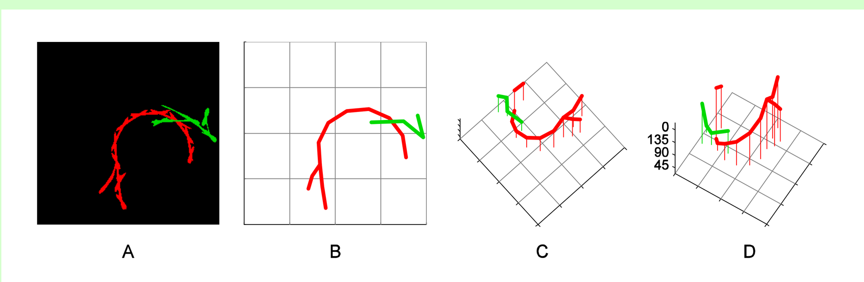
Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, *207*, 187–217

May, K. A. & Georgeson, M. A. (2007). Added luminance ramp alters perceived edge blur and contrast: A critical test for derivative-based models of edge coding. *Vision Research*, *47*, 1721–1731

Watt, R. J. & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, *25*, 1661–1674

May & Hess (2008): Filter-rectify-filter algorithm for contour integration

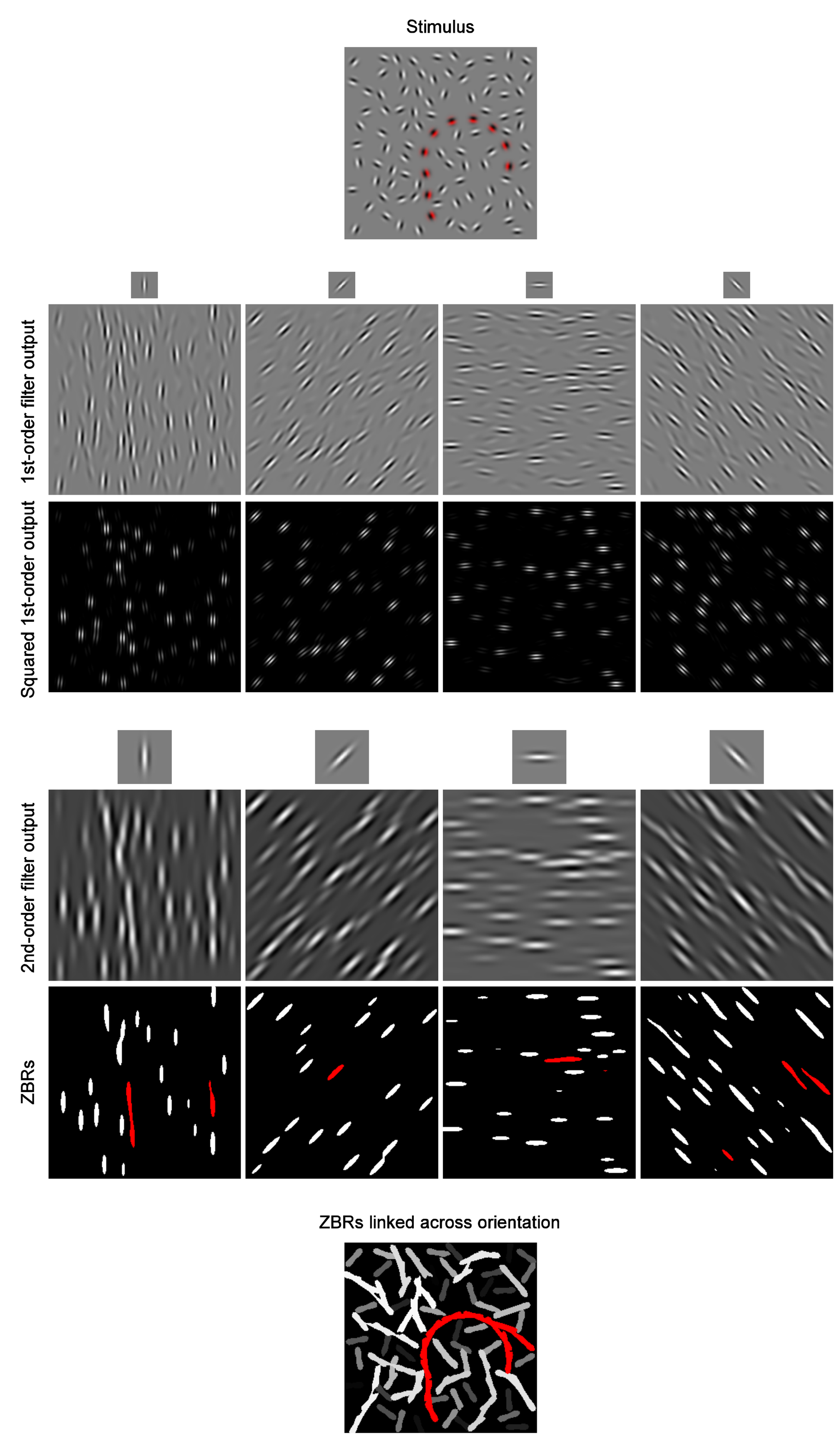
- Filter-rectify-filter at a range of orientations
- If 1st- and 2nd-stage filters are parallel, the algorithm detects “snakes”
- If 1st- and 2nd-stage filters are orthogonal, the algorithm detects “ladders”
- Applying a threshold to the 2nd-stage filter output generates zero-bounded response distributions (ZBRs) that extend across space and orientation, tracing out the contours
- 3D representation allows contours to overlap spatially without joining up

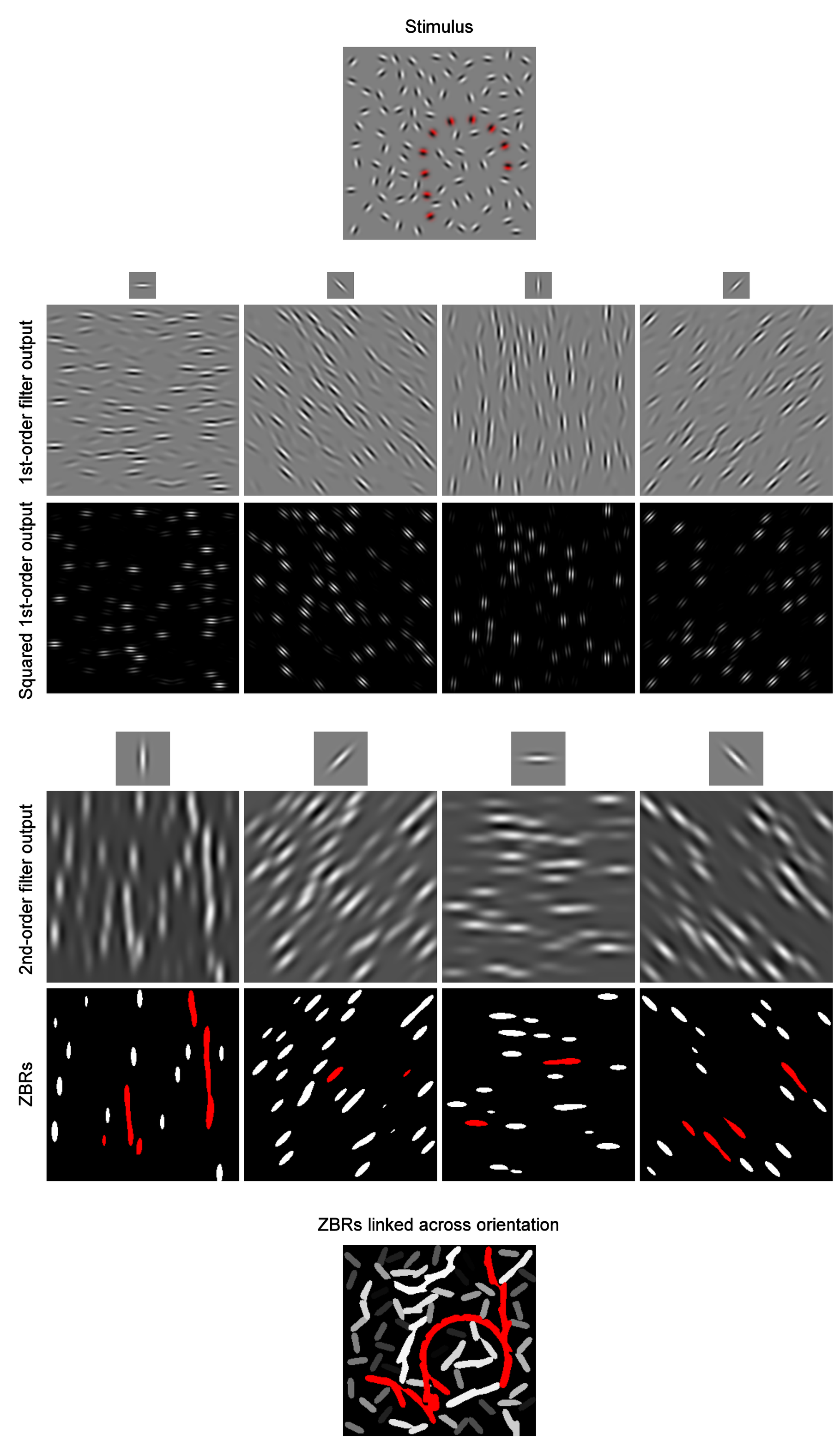


- Parameters of the model are the filter parameters and the threshold
- Scale of 1st-stage filter should match stimulus elements
- Scale of 2nd-stage filter should match spacing between the elements
- With one set of physiologically plausible parameters, the model can account for human performance on 176 experimental conditions in which the following contour parameters were varied: contour curvature, element orientation jitter, element orientation bandwidth properties (Hansen, May & Hess, under review)

Snake detection

Ladder detection





References

Hansen, B.C., May, K.A. & Hess, R.F. (under review). One “shape” fits all: The orientation bandwidth of contour integration

May, K.A. & Hess, R.F. (2008). Effects of element separation and carrier wavelength on detection of snakes and ladders: Implications for models of contour integration. *Journal of Vision*, *8*(13):4, 1–23