waterfall SOLUTIONS

UNIVERSITY OF SURREY

Small-scale Anomaly Detection in Panoramic Imaging using Neural Models of Low-level Vision

Matthew Casey[*], Duncan Hickman[#], Athanasios Pavlou[*], James Sadler[#]

[*]University of Surrey
[#]Waterfall Solutions Ltd

26 April 2011

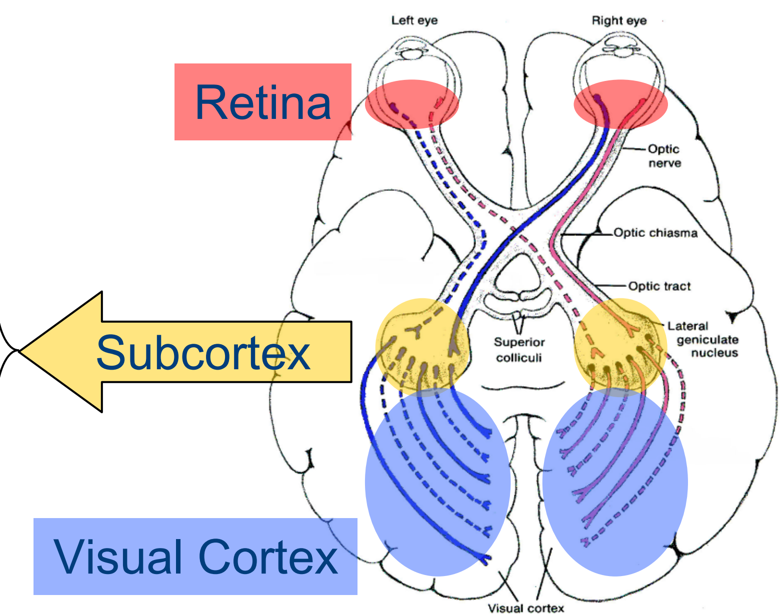
www.surrey.ac.uk/computing
www.waterfallsolutions.co.uk

Opportunity

- Models of human senses ready for exploitation
  - New models of low-level vision
  - Rapid, adaptive techniques working on crude stimuli
  - Innovative neural network models
- Exploitation
  - Can we improve anomaly detection in imaging?
  - Real-time, small-scale anomalies in CCTV
  - Panoramic imaging monitored by operators
- Partnership
  - Panoramic imaging: Waterfall Solutions
  - Models of vision: University of Surrey
  - EPSRC: Knowledge Transfer Account (EP/H500189/1)

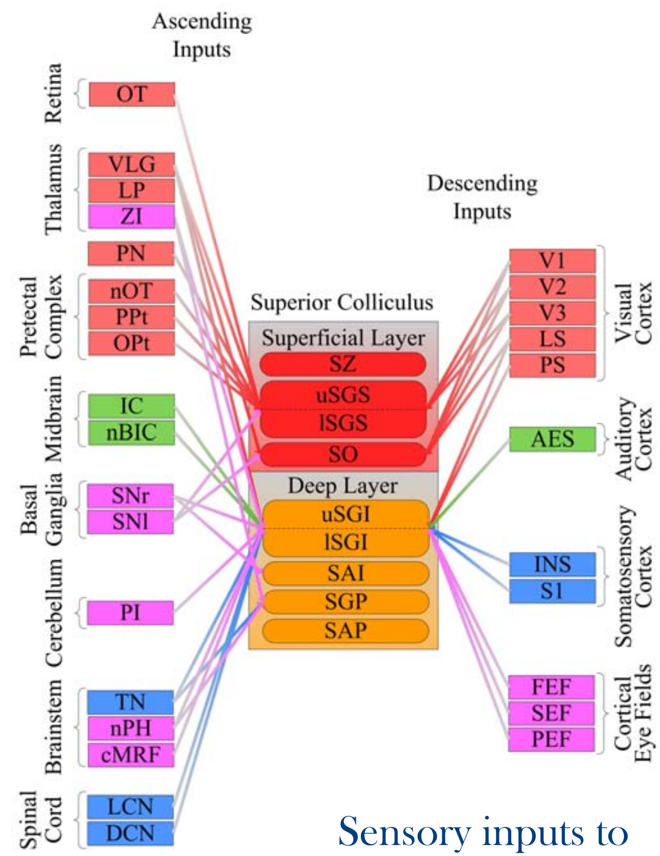
Human Low-level Vision

- How do humans rapidly react to threats?
  - Subcortical sensory processing in the brain
  - Prepares our bodies to react before conscious perception
- Superior colliculus [1,2]
  - Controls gaze shifts [6]
  - Rapid localization
  - For spatial and temporal events [3]
  - Fuses with sounds and touch [2]
- Amygdala [7]
  - Rapid discrimination of threats
  - Learns to react
  - Prepares the body to react



Superior Colliculus

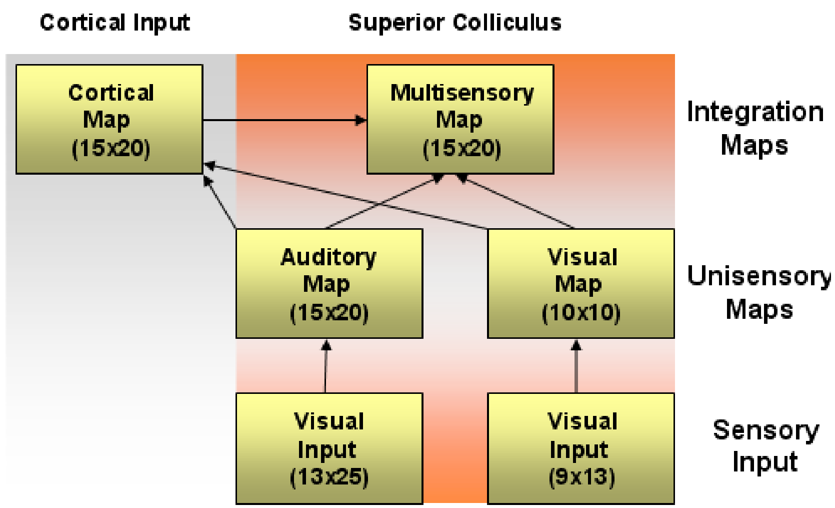
- Layered structure in the midbrain [1]
  - Sensory topographic maps [6]
  - Eye-centered alignment
- Superficial layer
  - Visual input
  - Representation is binocular
  - Sensitive to motion [11-13]
  - Sensitive to contrast [14,15]
- E.g. layering in frogs [3] (optic tectum)
  - Spatial contrast at sharp edges
  - Convex-shaped dark objects (spot the fly)
  - Temporal contrast of moving edges
  - Sudden reduction in illumination (predator's shadow)



Sensory inputs to the SC [6]

Model Specification

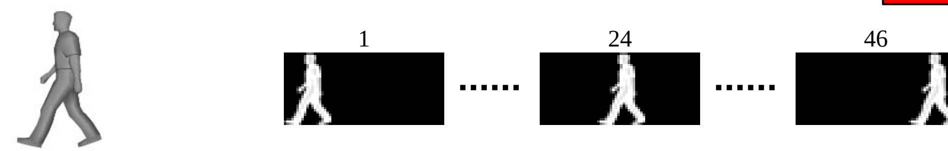
- Previous neuroscience model of the SC [8]
  - Rate-coded neural network
  - Simple layering
  - Weights learnt by competition
- Achieved
  - Localization on abstract data
  - Multisensory integration
- Adaptation
  - Could learn to localize different patterns
- Binocular, rapid localization of anomalies
  - Ideal study on anomaly detection in imaging: small-scale movement



Neural network model [8]

Pilot Experiment

- Can the map used in the SC model detect walking people?
  - 768 by 576 pixels images: grayscale intensity values as input to the map
  - Mosaic of 12 by 18 maps, each with 64 by 2 neurons
  - Trained on a small walking silhouette (from a 3d model)



a) 3-dimensional template

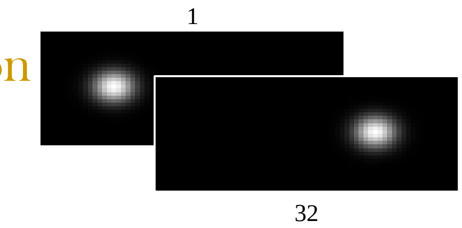
b) Sequence of training images used

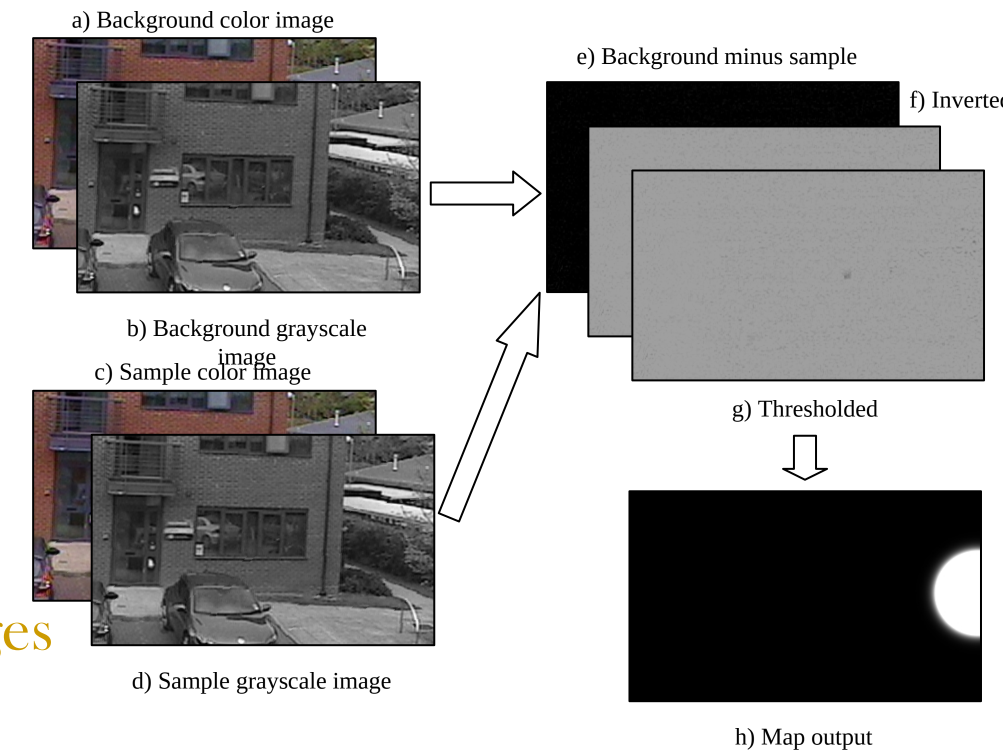
- Localized person (red circle)
  - When obscured by trees 10%
  - When in clear view 24%
  - Total detection rate of 21%; false alarms: 0
- But low detection rate, if no false alarms
  - Input relied upon pixel intensity (dark vs. light clothes)
  - Silhouette was too specific, only detects walking
  - No detection when on boundary between maps

Promising approach, but too specific

Generic Small-scale Detection

- Overcome issues in pilot
  - Detect change in the image using background subtraction
  - More generic training pattern for small-scale anomalies
- Pre-processing crudely approximates retina
  - Simple and rapid
  - But subject to noise
- Steps:
  - a), b) grayscale background
  - c), d) grayscale sample
  - e) subtraction
  - f) inverted (change is high)
  - g) threshold
  - h) topographic map
- Applied to panoramic images





www.surrey.ac.uk/computing
www.waterfallsolutions.co.uk

UNIVERSITY OF SURREY

waterfall SOLUTIONS

Panoramic Imaging

- Real-time combination of three cameras
  - Cylindrical warp of overlapping images [22]: 100° horizontal field of view
- Scenes designed to include small-scale or obscured targets
  - Includes 'large-scale' movement of people and cars
  - But also small-scale (normally undetectable) movement: head pop
- Mosaic of maps
  - 1344 by 580 pixels images, split into 64 by 32 pixels for each map
  - Mosaic of 21 by 18 maps, each with 64 by 2 neurons

Look for the head pop

UNIVERSITY OF SURREY

waterfall SOLUTIONS

Panoramic Detection Experiments

- Designed to test the capability of the technique
  - Movement partially obscured which is visible in as little as 3 pixels

Sequence	Description	Frames	Detect Obscured	Detect Clear	False Alarms
1	Person walking behind a tree and parked motorbike	70	56%	100%	0
2	Person crouched behind a car, stands up to be half visible, moves, crouches	56	63%	*	0
3	Person standing out of view behind a corner who then pops their head out	32	84%	*	3
4	Person standing out of view behind a corner who then pops their head out	61	100%	*	28

* Target always obscured

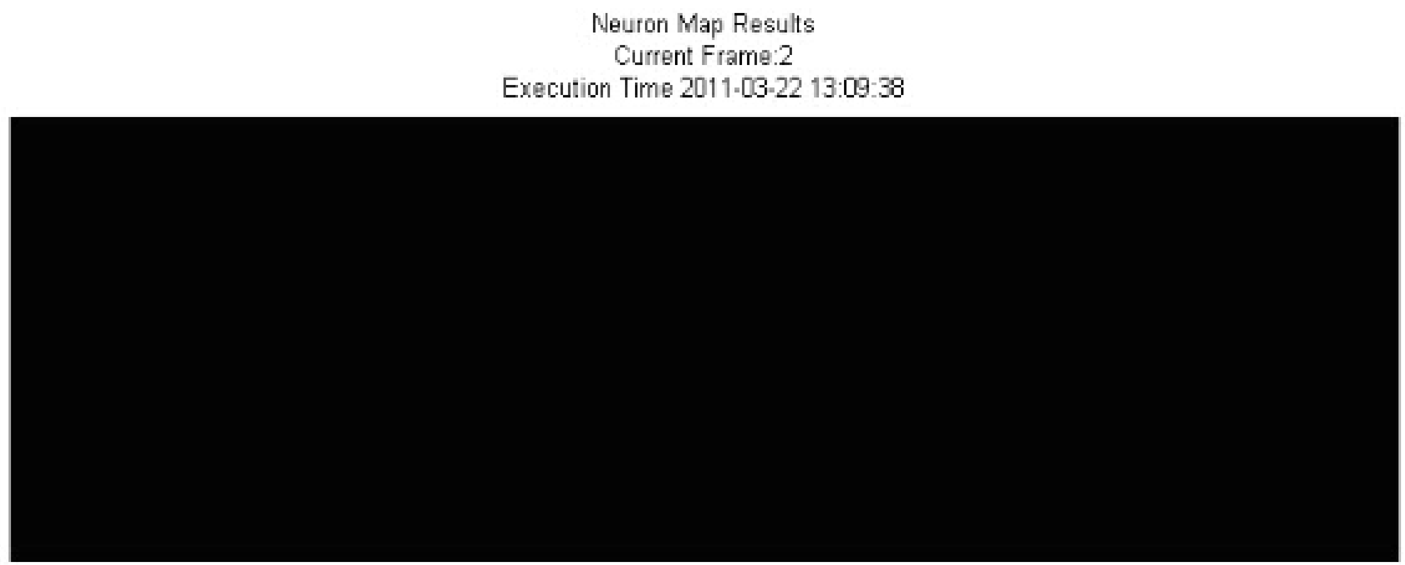
On these constrained sequences: good detection rate, low false alarms

Something More Challenging

- Sequences so far are not realistic
  - Picked to test ability, not to test for issues
  - Longer sequence used which includes significant noise
- Low light scene with various anomalies
  - Person walking towards camera, obscured, hiding and full view
  - Person starts visible (head) at 1 pixel, grows to (full view) 68 pixels

Look for person hiding →

Example


Neuron Map Results
Current Frame:2
Execution Time 2011-03-22 13:09:38

Results

- Rapid detection of obscured and small-scale movement
  - Beyond the capability of conventional techniques
  - But has a high false alarm rate

Sequence	Description	Frames	Detect Obscured	Detect Clear	False Alarms
5	Person behind a parked car (head 1 pixel) Walks towards the camera Passes behind signs; hides in a bush Also a fluttering flag Low light conditions with moving clouds	821	Person: 27%	Person: 41% Flag: 46%	859

Maximum 4 false alarms per frame

Demonstrated: small-scale detection, but high false alarms

Summary

- Biologically-inspired technique
  - Human low-level vision
  - Sensitive to small and large-scale motion and contrast
  - Neuroscience model using neural networks
- Applied to panoramic imaging
  - Demonstrated generic capability to detect movement
  - Particularly small-scale or obscured movement
  - Can be tuned to shape
- Limitations
  - Simple pre-processing (background subtraction)
  - Generic pattern leads to high false alarms
  - May be overcome by improved biological model
  - Temporal contrast detection built into the retina
- Opportunities
  - Improved biological model
  - Link to audio signals [23]

Selected References

[1] King, A. J., "The Superior Colliculus," Current Biology, 14(9) R335-R338 (2004).

[2] Stein, B. E. and Meredith, M. A., [The Merging of the Senses], A Bradford Book, MIT Press, Cambridge, MA., (1993).

[3] Lettvin, J. Y., Maturana, H. R., Mcculloch, W. S. and Pitts, W. H., "What the Frog's Eye Tells the Frog's Brain," Proceedings of the IRE, 47(11) 1940-1951 (1959).

[6] May, P. J., [The Mammalian Superior Colliculus: Laminar Structure and Connections], Buttner-Ennever, J.A. (Ed), Progress in Brain Research: Neuroanatomy of the Oculomotor System, Volume 151, Elsevier, 321-378 (2006).

[7] Öhman, A., Flykt, A. and Esteves, F., "Emotion Drives Attention: Detecting the Snake in the Grass," Journal of Experimental Psychology: General, 130(3) 466-478 (2001).

[8] Pavlou, A. and Casey, M. C., "Simulating the Effects of Cortical Feedback in the Superior Colliculus with Topographic Maps," Proceedings of the International Joint Conference on Neural Networks (IJCNN) 2010, (2010).

[11] Sterling, P. and Wickelgren, B. G., "Visual Receptive Fields in the Superior Colliculus of the Cat," Journal of Neurophysiology, 32(1) 1-15 (1969).

[12] Rauschecker, J. P. and Harris, L. R., "Auditory and Visual Neurons in the Cat's Superior Colliculus Selective for the Direction of Apparent Motion Stimuli," Brain Research, 490(1) 56-63 (1989).

[13] Wallace, M. T., McHaffie, J. G. and Stein, B. E., "Visual Response Properties and Visuotopic Representation in the Newborn Monkey Superior Colliculus," Journal of Neurophysiology, 78(5) 2732-2741 (1997).

[14] Prévost, F., Lepore, F. and Guillemot, J. P., "Spatio-temporal Receptive Field Properties of Cells in the Rat Superior Colliculus," Brain Research, 1142 80-91 (2007).

[15] Schneider, K. A. and Kastner, S., "Visual Responses of the Human Superior Colliculus: A High-Resolution Functional Magnetic Resonance Imaging Study," Journal of Neurophysiology, 94(4) 2491-2503 (2005).

[22] Sadler, J. R. E., Davis, J. and Hickman, D. L., "A Compact Wide-area Surveillance System for Defence and Security Applications," Proceedings of SPIE Defense, Security, and Sensing Conference 2011 on Enhanced and Synthetic Vision, 8042B (2011).

[23] Casey, M. C., Pavlou, A. and Timotheou, A., "Mind the (Computational) Gap," Proceedings of the UK Workshop on Computational Intelligence (UKCI 2010), (2010).

Thank you

Questions?

Department of Computing
University of Surrey
Guildford
Surrey
GU2 7XH, UK

Waterfall Solutions Ltd
1 & 2 Saxton
Parklands
Guildford
Surrey
GU2 9JX, UK

www.surrey.ac.uk/computing
www.waterfallsolutions.co.uk