Mind the (Computational) Gap

UKCI 2010

Matthew Casey
Athanasios Pavlou
Anthony Timotheou
10 September 2010

www.surrey.ac.uk/computing

Motivation

- "Are we there yet?" [1]
  - Cristianini's view on the state of computational intelligence (CI)
  - We have not achieved Turing's "intelligent machinery" dream [2]
  - But some impressive "data driven AI" [1:467]
  - Advocating a broader reference for intelligence
- Yet neuroscience has seen some impressive advances
  - Large-scale simulations of the human brain [3,4]
  - Mostly focused on the cortex (like Turing)
  - Can we still learn from computational neuroscience (CN)?
- Do large-scale CN models hold the key to "intelligence"?
  - There appears to be a computational gap between CN & CI
  - How can we apply these models and plug this gap?

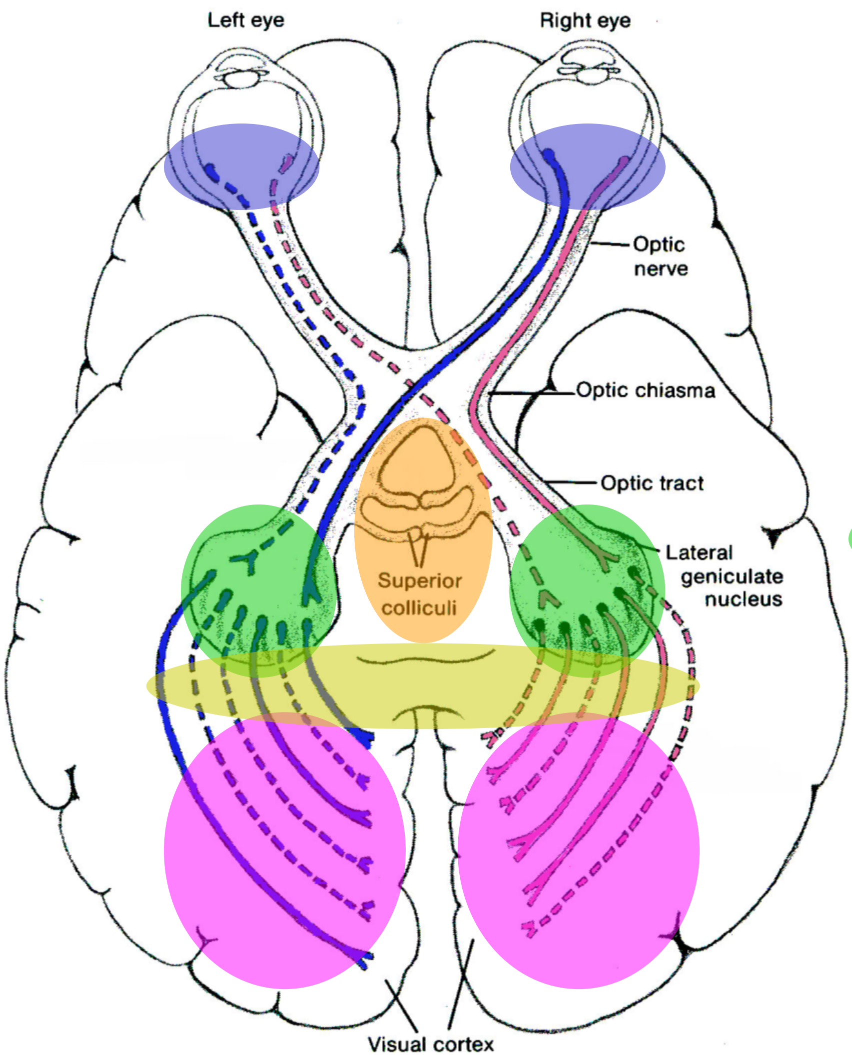
Aim

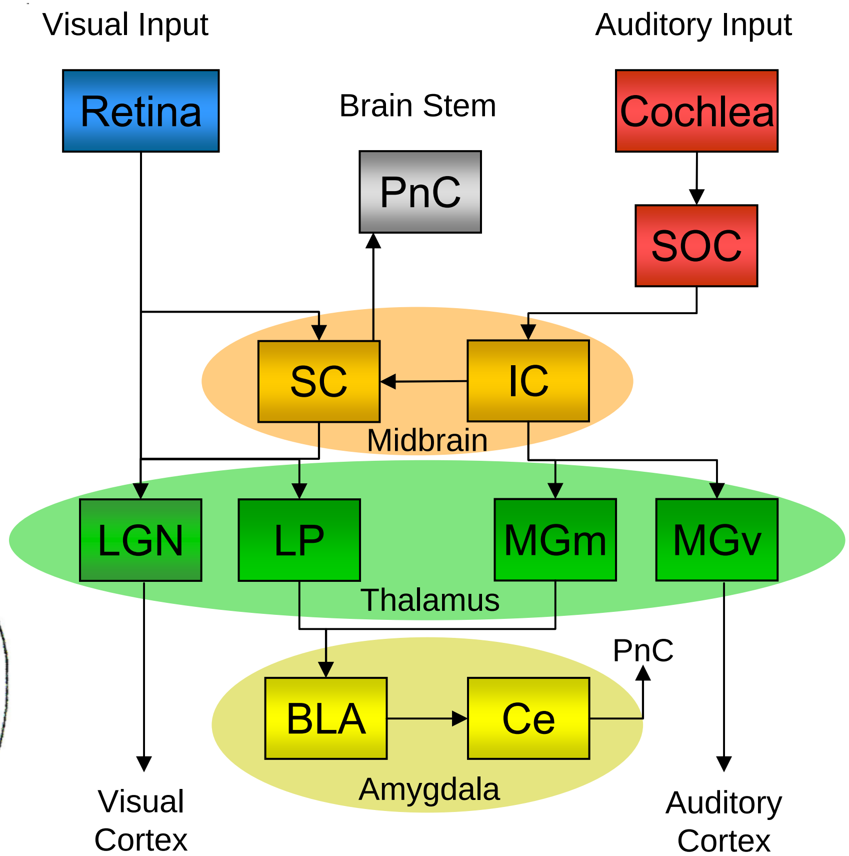
- Large-scale models are impressive
  - But not connected with the real-world
- Can we develop large-scale models connected to the real world?
  - Useful for CN: behavioural models
  - Useful for CI: applying models of intelligence
- What do we need?
  - To focus on how to provide cortical models with input/output
- Focus on function below the cortex
  - Sensorimotor function: subcortical processing

Proof-of-concept

- Use an existing CN model and apply it to real-time inputs
  - Extract computational principles
  - Apply models to an example CI problem
- Which brain structures?
  - Subcortical visual processing
  - Link between senses and cortical processing
  - Extensive studies of key structures
  - Good example: superior colliculus
  - Using existing CN models (cf. [9,10])

Audio-Visual Subcortical Function

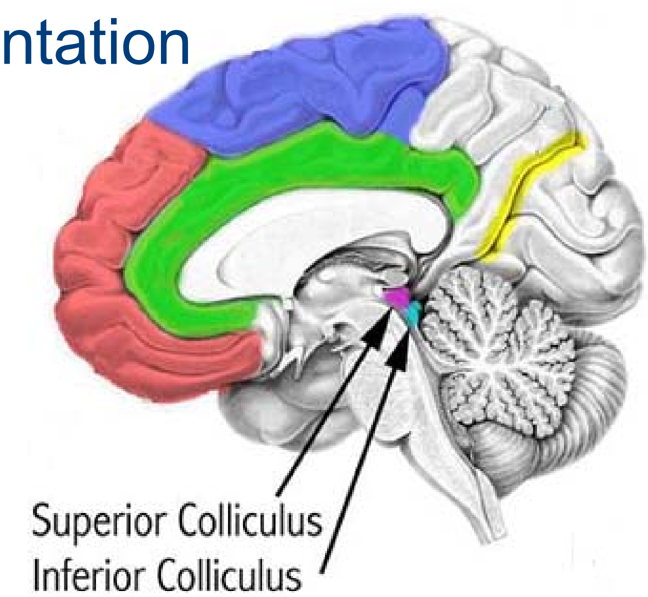




Constructed from [11,14,15,18]

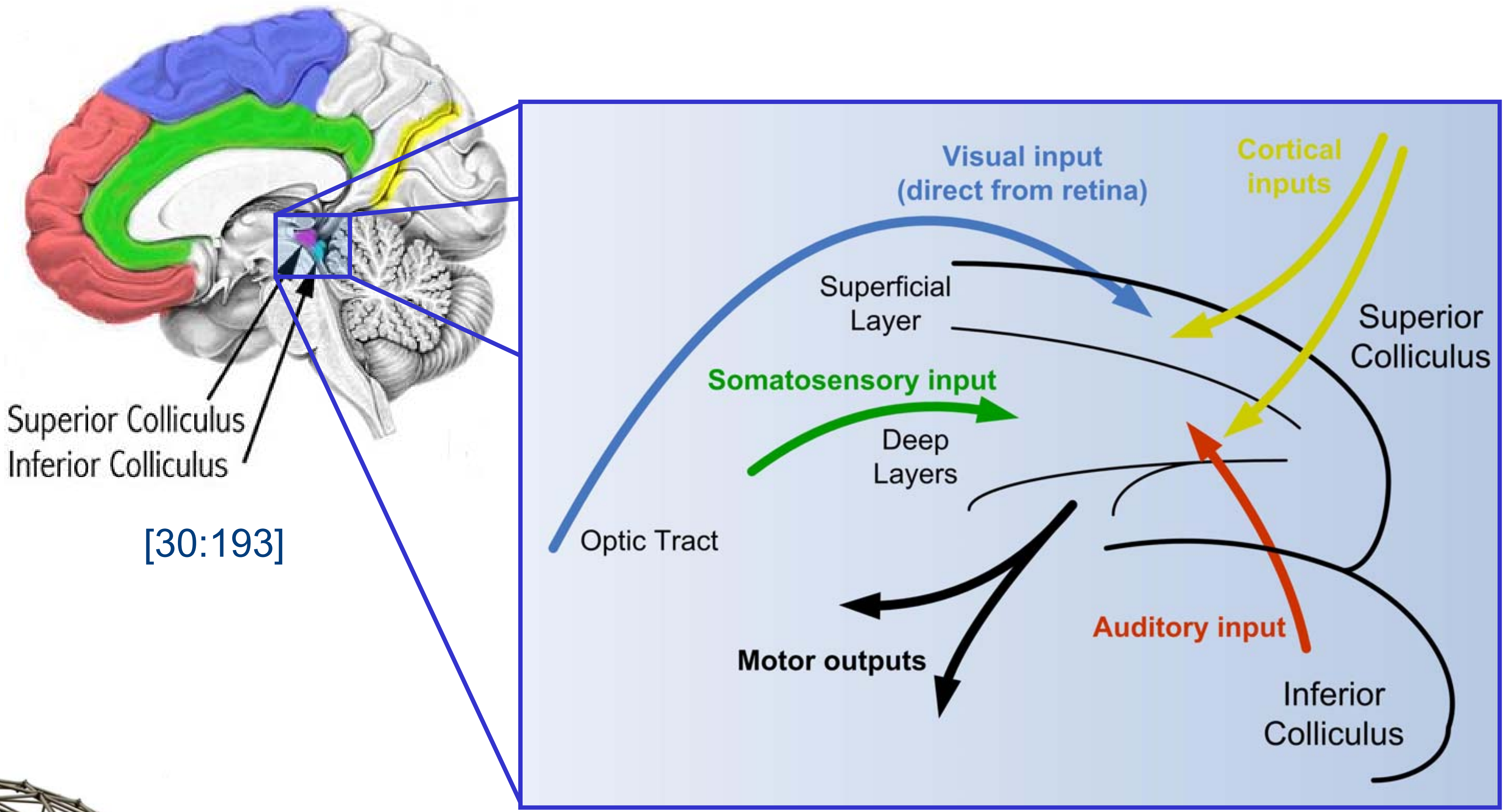
Superior Colliculus

- Laminated midbrain structure [13,21]
  - Combines visual, auditory and somatosensory stimuli
  - Sensory alignment of topographic maps
  - Develops a multisensory representation
- Causes gaze shifts
  - Prioritises to multimodal stimuli
  - Enhancement and suppression
- Integration is moderated
  - Cortical feedback [23]
  - Unisensory co-operation?



[30:193]

Superior Colliculus

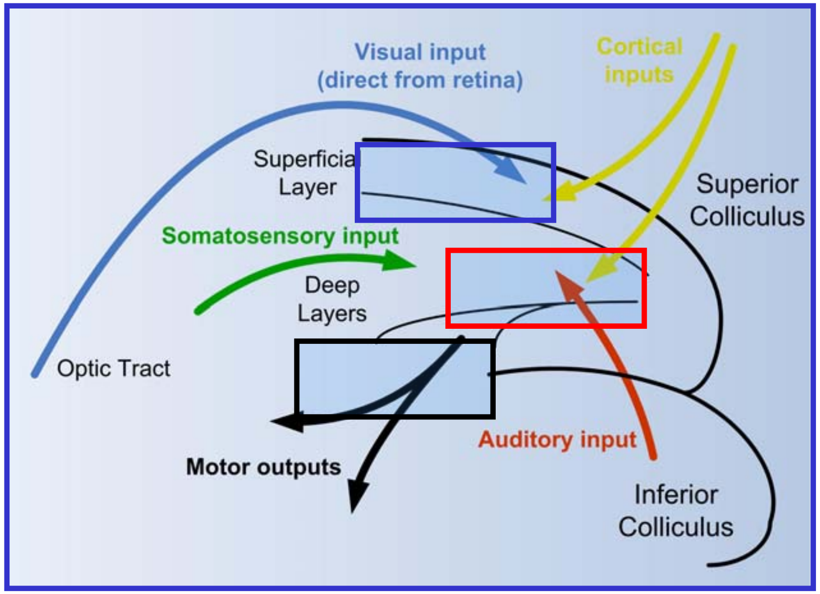


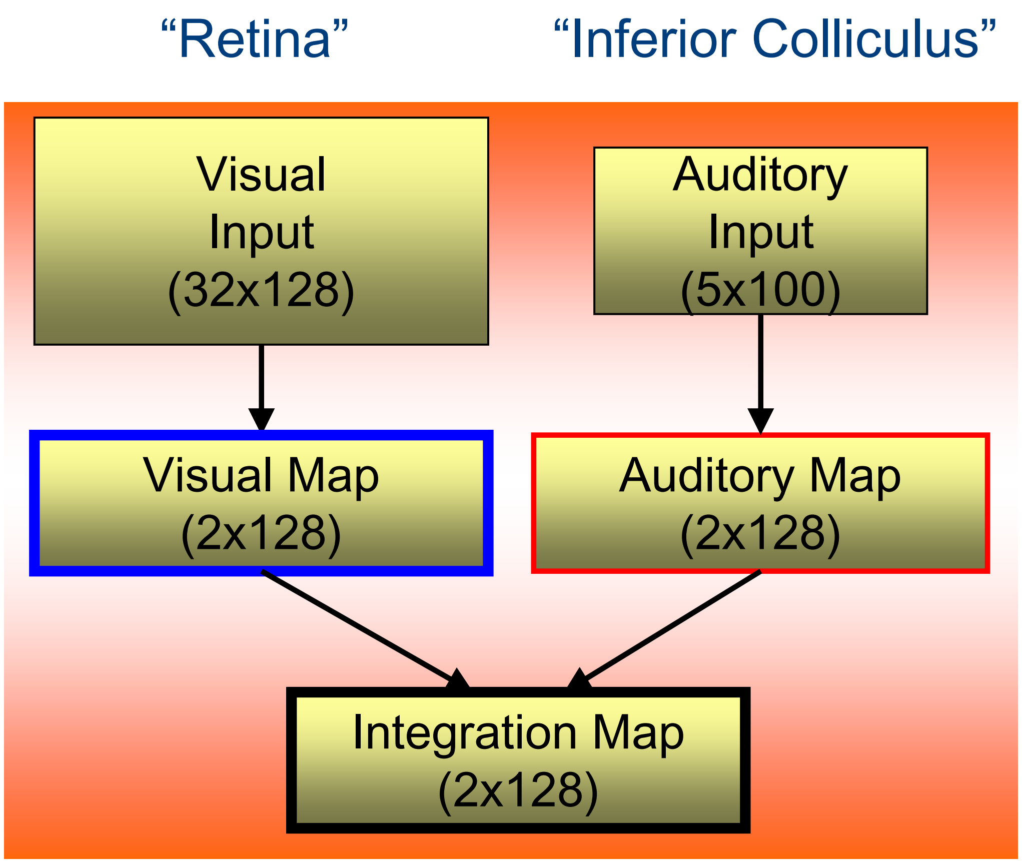
[30:193]

Computational Principles

Subcortical structures	Example
1. Develop	• Self-organisation
2. Adapt	• Through conditioning
3. Discriminate	• Albeit on crude stimuli
4. Operate rapid	• Express saccades 80ms [6]
5. Direct input/output	• Optic tract, plus others [19]
6. Multisensory	• Early stages of processing [21]
7. Essential for cortex	• Sensory routing to cortex, plus cortex-to-cortex [12,22,23]

Behaviour Model of the SC





Based upon our previous CN models [9,10]

Rate-coded Neural Model

- Topographic maps
  - Spatial representation [9]
  - Develops organisation
- Competitive learning
  - Hebbian association [18]
  - Maps trained in layers
- Integration
  - Map outputs as inputs
  - Unisensory → multisensory

Neuron output

$$u_{ij} = \sum_{k=1}^{m} x_k w_{kij}(t),$$

$$y_{ij} = \begin{cases} f(u_{ij}) & \text{if } \|c_{ij} - c_{win}\| < h(t) \\ f(u_{ij} - y_{win}) & \text{otherwise} \end{cases}$$

$$f(u) = \begin{cases} 1 & u \geq 1 \\ u & 0 < u < 1 \\ 0 & u \leq 0 \end{cases},$$

Weight Update

$$w'_{kij}(t+1) = w_{kij}(t) + \epsilon(t) x_k y_{ij}$$

$$w_{kij}(t+1) = \frac{w'_{kij}(t+1)}{\sum_{l=1}^{m} w'_{lij}(t+1)}$$

Neighbourhood and learning rate

$$h(t) = r_{min} + (r_{max} - r_{min}) e^{-\left(\frac{(t/t_e)^2}{2r_s^2}\right)}$$

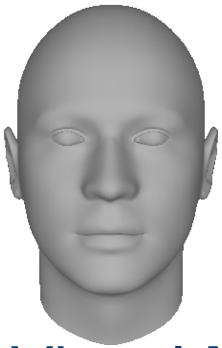
$$\epsilon(t) = l_{min} + (l_{max} - l_{min}) e^{-\left(\frac{(t/t_e)}{2l_s^2}\right)}$$

Speaker Localisation

- We are attempting to show that CN models can be connected to real-world inputs to apply their principles
- As a toy example we select speaker localisation
  - To locate a coincident facial and sound stimulus
  - Not aiming for state-of-the-art, but generic, adaptive solutions
- This will demonstrate:

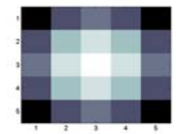
1. Develop	• Internal spatial representations
2. Adapt	
3. Discriminate	• Face silhouettes
4. Operate rapidly	• Real-time inputs (25 frames per second)
5. Direct input/output	• Sensory inputs, location output
6. Multisensory	• Video and audio
7. Essential for cortex	

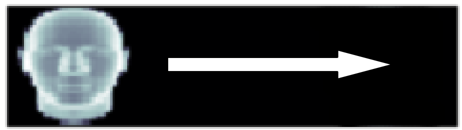
Training: Artificial Stimulus



FaceGen [27]

$$x_{ij} = \lambda e^{-\left(\frac{(i-c)^2+(j-d)^2}{2\sigma^2}\right)}$$



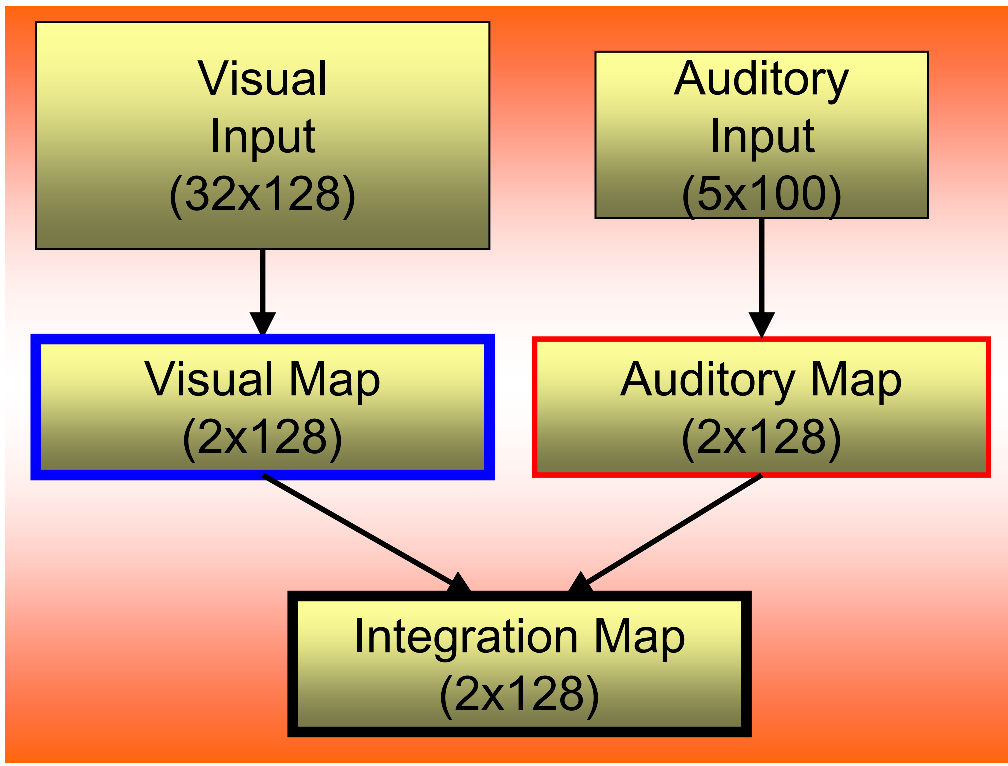


Visual Map

96 frames

Generic face

400 epochs



Auditory Map

96 frames

Gaussian blob

400 epochs

Develops a spatial representation of visual space, discriminates face silhouettes

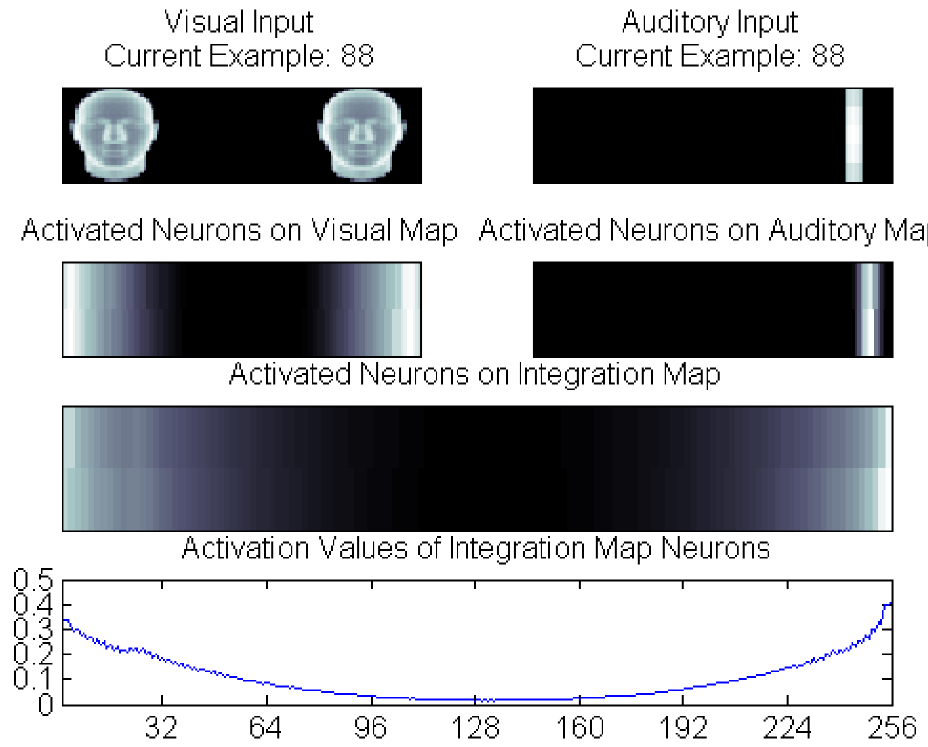
Develops a spatial representation of auditory space

Integration of visual and auditory locations

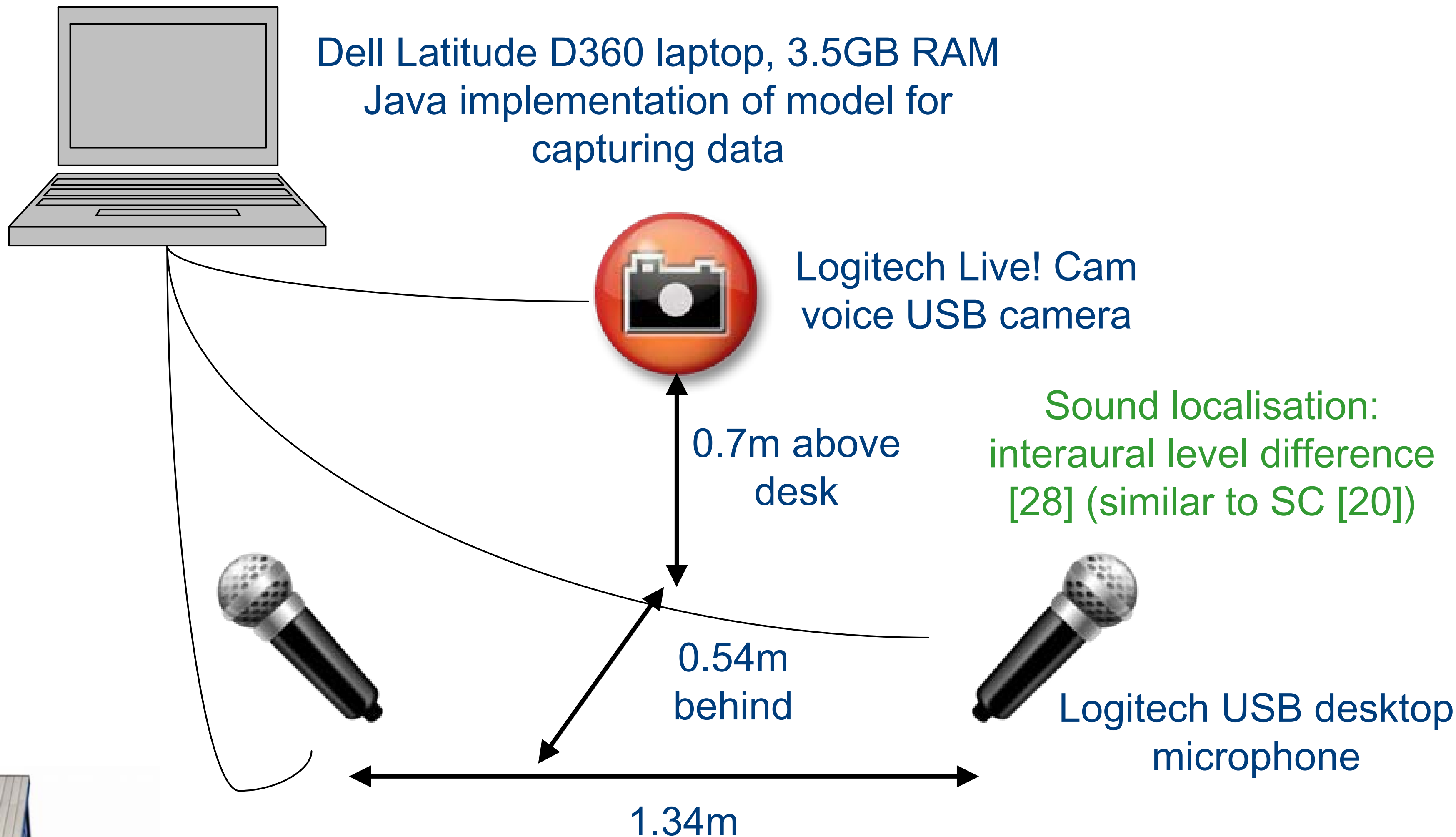
UNIVERSITY OF SURREY

Model Development

- Learnt spatial representations
  - Discriminates and locates (crude) faces
  - Locates abstract sound stimuli
  - Integrates locations
  - (Parameters [9])
- When coincident
  - Higher integrated activation



Experimental Set-up

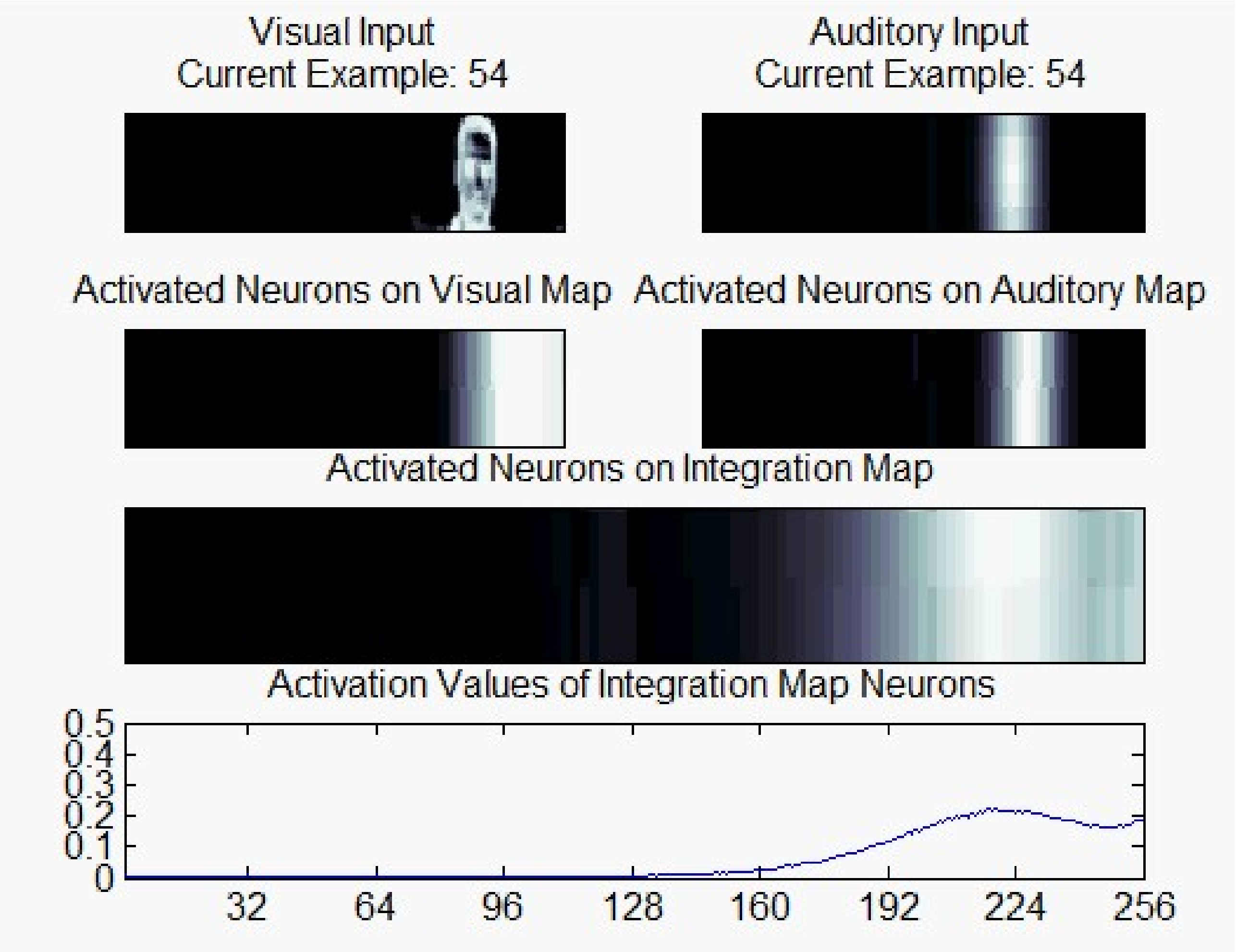


Independent sound source (pure tone) located 0.42m in front of microphones

Experiments

- Data
  - 325 example frames (13 seconds at 25 fps)
  - Various conditions: no faces, one face enters, another face enters, first leaves, last leaves, plus moving sound source
- Responses
  - Faces and sound localised at 25 fps
  - Integration response of 0.22 or above when coincident
- However
  - Crude discrimination of faces
  - Sound needs to be sufficiently loud and sustained

Example

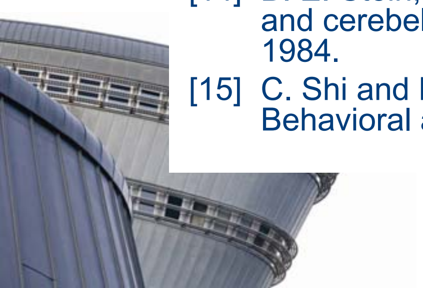

Visual Input
Current Example: 54
Auditory Input
Current Example: 54
Activated Neurons on Visual Map
Activated Neurons on Auditory Map
Activated Neurons on Integration Map
Activation Values of Integration Map Neurons
0.5
0.4
0.3
0.2
0.1
0
32
64
96
128
160
192
224
256

Conclusion

- Can we develop large-scale real-world CN models?
  - To achieve this, we need CN models that can be connected to real-world inputs
- We demonstrated
  - How an existing CN model can be applied to real-time inputs
  - Simple speaker localisation to show crude discrimination
  - But generic and fast pattern recognition
- Limitations
  - Model uses less plausible rate-coded neurons: spiking needed
  - Need to increase the scale: link to cortical models
  - Simple localisation, hardware constrained, but functional
  - Ad-hoc computational principles identified
- Yet this shows how CN models can be applied to CI

References

[1] N. Cristianini, “Are we there yet?” Neural Networks, vol. 23, no. 4, pp. 466–470, 2010.
[2] A. M. Turing, “Intelligent machinery,” National Physical Laboratory, Tech. Rep., 1948. [Online]. Available: http://www.alanturing.net/turing archive/archive/l/l32/L32-001.html
[3] H. Markram, “The blue brain project,” Nature Reviews Neuroscience, vol. 7, no. 2, pp. 153–160, 2006.
[4] E. Izhikevich and G. M. Edelman, “Large-scale model of mammalian thalamocortical systems,” Proceedings of the National Academy of Sciences of the USA, vol. 105, no. 9, pp. 3593–3598, 2008.
[5] I. R. Cohen and D. Harel, “Two views of a biology-computer science alliance,” in Proceedings of the 2009 Workshop on Complex Systems Modelling and Simulation (CoSMoS 2009), S. Stepney, P. Welch, P. S. Andrews, and J. Timmis, Eds. Luniver Press, 2009, pp. 1–8.
[6] A. K. Moschovakis, C. A. Scudder, and S. M. Highstein, “The microscopic anatomy and physiology of the mammalian saccadic system,” Progress in Neurobiology, vol. 50, no. 2-3, pp. 133–254, 1996.
[7] J. Qu, X. Zhou, H. Zhu, G. Cheng, W. S. Ashwell, and F. Lu, “Development of the human superior colliculus and the retinocollicular projection,” Experimental Eye Research, vol. 82, no. 2, pp. 300–310, 2006.
[8] A. G. Constantin, H. Wang, and J. D. Crawford, “Role of superior colliculus in adaptive eye-head coordination during gaze shifts,” J Neurophysiol, vol. 92, no. 4, pp. 2168–2184, 2004.
[9] A. Pavlou and M. C. Casey, “Simulating the effects of cortical feedback in the superior colliculus with topographic maps,” in Proceedings of the International Joint Conference on Neural Networks (IJCNN) 2010. IEEE, 2010.
[10] ——, “Identifying emotions using topographic conditioning maps,” in Advances in Neuro-Information Processing: Proceedings of the 15th International Conference on Neuro-Information Processing, Lecture Notes in Computer Science 5506, M. Koeppen, N. Kasabov, and G. Coghill, Eds. Springer-Verlag, 2009, pp. 40–47.
[11] T. Pasternak, J. W. Bisley, and D. J. Calkins, “Visual processing in the primate brain,” in Biological Psychology, ser. Handbook of Psychology, M. Gallagher and R. J. Nelson, Eds. John Wiley and Sons, Inc., 2003, vol. 3, chapter 6, pp. 139–185.
[12] S. M. Sherman and R. W. Guillery, “The visual relays in the thalamus,” in The Visual Neurosciences, L. M. Chalupa and J. S. Werner, Eds. A Bradford Book, The MIT Press, 2004, vol. 1, chapter 35, pp. 565–591.
[13] A. J. King, “The superior colliculus,” Current Biology, vol. 14, no. 9, pp. R335–R338, 2004.
[14] B. E. Stein, R. F. Spencer, and S. B. Edwards, “Efferent projections of the neonatal cat superior colliculus: Facial and cerebellum-related brainstem structures,” The Journal of Comparative Neurology, vol. 230, no. 1, pp. 47–54, 1984.
[15] C. Shi and M. Davis, “Visual pathways involved in fear conditioning measured with fear-potentiated startle: Behavioral and anatomic studies,” The Journal of Neuroscience, vol. 21, no. 24, pp. 9844–9855, 2001.



References

[16] C. Casanova, “The visual functions of the pulvinar,” in The Visual Neurosciences, L. M. Chalupa and J. S. Werner, Eds. A Bradford Book, The MIT Press, 2004, vol. 1, chapter 36, pp. 592–608.

[17] L. Pessoa, “To what extent are emotional visual stimuli processed without attention and awareness?” Current Opinion in Neurobiology, vol. 15, pp. 188–196, 2005.

[18] J. L. Armony, D. Servan-Schreiber, L. M. Romanski, J. D. Cohen, and J. E. LeDoux, “Stimulus generalization of fear responses: Effects of auditory cortex lesions in a computational model and in rats,” Cerebral Cortex, vol. 7, no. 2, pp. 157–165, 1997.

[19] J. A. Gottfried, “Smell: Central nervous processing,” in Taste and Smell: An Update (Advances in Oto-Rhino-Laryngology), T. Hummel and A. Welge-Lssen, Eds. S. Karger AG, 2006, pp. 44–69.

[20] B. Delgutte, P. X. Joris, R. Y. Litovsky, and T. C. Yin, “Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus,” J Neurophysiol, vol. 81, no. 6, pp. 2833–2851, 1999.

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

[22] D. G. Amaral, H. Behniea, and J. L. Kelly, “Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey,” Neuroscience, vol. 118, no. 4, pp. 1099–1120, 2003.

[23] J. C. Alvarado, T. R. Stanford, B. A. Rowland, J. W. Vaughan, and B. E. Stein, “Multisensory integration in the superior colliculus requires synergy among corticocollicular inputs,” The Journal of Neuroscience, vol. 29, no. 20, pp. 6580–6592, 2009.

[24] R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh, Computational Maps in the Visual Cortex. New York: Springer Science+Business Media, 2005.

[25] T. Kohonen, “Self-organized formation of topologically correct feature maps,” Biological Cybernetics, vol. 43, pp. 59–69, 1982.

[26] K. Hyun-Don, C. Jong-Suk, and K. Munsang, “Speaker localization among multi-faces in noisy environment by audio-visual integration,” 2006, pp. 1305–1310.

[27] Singular Inversions Inc. (2010, Jun) Facegen - 3d human faces. [Online]. Available: http://www.facegen.com/

[28] S. T. Birchfield and R. Gangishetty, “Acoustic localization by interaural level difference,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005), vol. 4, 2005, pp. iv/1109–iv/1112.

[29] S. Stepney, R. E. Smith, J. Timmis, A. M. Tyrrell, M. J. Neal, and A. N. W. Hone, “Conceptual frameworks for artificial immune systems,” International Journal of Unconventional Computing, vol. 1, no. 3, pp. 315–338, 2005.

[30] G. A. Calvert and T. Thesen. Multisensory integration: Methodological approaches and emerging principles in the human brain. Journal of Physiology - Paris, 98(1-3):191-205, 2004.

UNIVERSITY OF SURREY

Thank you

Questions?

T: +44 (0)1483 689635
F: +44 (0)1483 686051
m.casey@surrey.ac.uk