

A Behavioral Model of Sensory Alignment in the Superficial and Deep Layers of the Superior Colliculus

Matthew C. Casey and Athanasios Pavlou

Abstract—The ability to combine sensory information is an important attribute of the brain. Multisensory integration in natural systems suggests that a similar approach in artificial systems may be important. Multisensory integration is exemplified in mammals by the superior colliculus (SC), which combines visual, auditory and somatosensory stimuli to shift gaze. However, although we have a good understanding of the overall architecture of the SC, as yet we do not fully understand the process of integration. While a number of computational models of the SC have been developed, there has not been a larger scale implementation that can help determine how the senses are aligned and integrated across the superficial and deep layers of the SC. In this paper we describe a prototype implementation of the mammalian SC consisting of self-organizing maps linked by Hebbian connections, modeling visual and auditory processing in the superficial and deep layers. The model is trained on artificial auditory and visual stimuli, with testing demonstrating the formation of appropriate spatial representations, which compare well with biological data. Subsequently, we train the model on multisensory stimuli, testing to see if the unisensory maps can be combined. The results show the successful alignment of sensory maps to form a multisensory representation. We conclude that, while simple, the model lends itself to further exploration of integration, which may give insight into whether such modeling is of benefit computationally.

I. INTRODUCTION

THE ability to fuse, process and act upon sensory information is an important attribute of humans and animals. Traditionally, low level processing of sensory information, such as for vision, audition and touch, were thought to occur in isolation from other senses. However, there is now a wealth of evidence suggesting that even low level processing is multisensory [1,2], with perhaps unisensory processing being the rarity [3]. This change in our understanding of natural cognitive systems has implications for artificial systems. First, since computational techniques are an established tool for exploring models of cognition, these must also be able to integrate sensory stimuli in order to prove effective. Second, such an integrative approach may help us overcome the limitations of existing computational paradigms (cf. speaker independent speech recognition [4]).

Multisensory integration in mammalian brains is exemplified by the superior colliculus (SC) [5]. While the SC was originally thought to process only visual stimuli, studies have demonstrated that it combines visual, auditory and somatosensory topographic maps into a multisensory representation [1]. The integration of these stimuli then causes gaze shifts (the combination of head movements and eye saccades) [5]. In the SC, stronger reactions are obtained to multisensory stimuli compared to just unisensory stimuli. This *multisensory enhancement* (and indeed *suppression*) occurs in deep layers of the SC, and is controlled by descending afferents from the cortex (anterior ectosylvian sulcus and the lateral suprasylvian sulcus) [2]. However, although the general architecture of the SC is well established, as yet we do not fully understand the process of integration and the role of cortical feedback [6].

While a multisensory model of the SC may help answer this question of how integration is influenced in natural systems, such a model may also help to establish new ways in which artificial sensory inputs can be combined in a computational system. Previous computational models of the SC have focused on understanding the outputs of the SC for saccades (for example [7]). Such models have been used to improve our understanding of multisensory enhancement, suppression and saccades, but these models necessarily ignore the wider behavioral aspects of the SC, such as sensory alignment in both the superficial and deep layers.

In this paper, we describe a computational model of the SC that integrates representations of visual and auditory stimuli, simulating both the superficial and deep layers, and aligns these unisensory representations. While this is a simple behavioral model constructed using a number of assumptions about the inputs, it demonstrates how different sensory maps may be integrated through a process of learning in order to translate sensory coordinate spaces into a single multisensory representation. Here, in order to explore whether integration can be learnt, we make the simplifying assumption that integration is affected through associative learning without cortical feedback. While this hypothesis is perhaps a limitation of the model, we demonstrate however that an emergent property of this integration is a simplified form of multisensory enhancement, albeit not as capable as some previous dynamic models. This early prototype provides insight into how larger behavioral models of brain function can be developed, and at the same time successfully simulates behavior in the SC that compares well to biology.

Manuscript received December 14, 2007.

M. C. Casey is with the Department of Computing, University of Surrey, Guildford, Surrey, GU2 7XH, UK (corresponding author phone: +44 (0) 1483 689635; fax: +44 (0) 1483 686051; e-mail: m.casey@surrey.ac.uk).

A. Pavlou is with the Department of Computing, University of Surrey, Guildford, Surrey, GU2 7XH, UK (e-mail: a.pavlou@surrey.ac.uk).

In section II of this paper, we review past computational models of the SC to motivate our method. In section III we describe the model. In section IV we evaluate the model’s outputs compared with the biological evidence. In section V we conclude and summarize the insights gained.

II. MODELING THE SUPERIOR COLLICULUS

The SC has been characterized as a laminated structure with superficial and deep layers [5]. The superficial layers receive visual stimuli directly from the retina and visual cortex. The deep layers receive auditory and somatosensory stimuli, but respond to multisensory signals to produce a motor output via so called *burst* and *buildup* neurons. Of particular interest is that the different layers of the SC form mutually aligned spatial sensory maps for the different modalities, which appears crucial for multisensory integration in that the alignment allows the maps to be combined and matched with appropriate motor outputs for eye, head and body orienting. The process of alignment of sensory and motor maps occurs during development of the SC to form a coordinate transformation, with in particular, vision being used to calibrate auditory responses (cf. [8] for a model of the inferior colliculus).

There have been several models of the SC developed that focus on replicating the functional neuronal saccadic output of the SC. For example, Grossberg et al [7] used Adaptive Resonance Theory (ART) to model the burst and buildup responses of the deep SC. Of note is that their model learnt the transformation between auditory, visual and motor responses using associative learning, demonstrating both multisensory enhancement and suppression. Similar dynamic models of the physiology of the deep SC have been developed to help determine how the SC controls movement. Here, models have been used to simulate parallel pathways between the SC and cerebellum [9], the competitive combination of exogenous (sensory) and endogenous (voluntary) information to simulate response times for saccade initiation [10] and antisaccades [11], and ensemble-coding that adds intended saccade trajectory information to each output spike of the SC [12].

In contrast to these physiologically motivated models, computational paradigms have also been used to explore how neural information is encoded for multisensory enhancement and suppression, using Bayesian and perceptron models [13]. However, despite being of interest because of the more abstract, computational approach, these models have still focused on the deep SC. Unlike such past models, in this paper we simulate the behavior of both the superficial and deep layers of the SC to explore sensory alignment and integration. We take motivation from [7], in which the associations between representations were learnt, but add the formation of sensory topographic maps. For ease of development, we use simple and well understood rate-coded algorithms [14,15], rather than the more complex neural dynamic models. We hope to gain insight into

whether such techniques will prove beneficial for larger-scale implementations of brain function.

III. A BEHAVIORAL MODEL OF SENSORY ALIGNMENT

Our model consists of sensory topographic maps that are aligned and linked together to form a multisensory space. We simplify our approach in two ways. First, we consider only visual and auditory stimuli within an approximate azimuth and elevation range for humans, but do not consider touch. Here, we use an auditory space that surrounds the head completely (elevation [-180, 180], azimuth [-90, 90]), at a fixed depth. The visual space is a subset of this (elevation [-65, 65], azimuth [-90, 90]), allowing us to evaluate both single modality stimuli locations (auditory only), multimodal locations (auditory and visual) as well as coincident (same location multisensory) and non-coincident (different location multisensory) inputs.

Second, our model of the biology is simplified by using rate-coded neural network models. While rate-coded models are highly abstract, they provide a useful prototyping tool for computational modeling with some degree of plausibility. For example, Hebbian learning [15] is an established method that is both biologically supported and implemented in a range of rate-coded algorithms. In particular, Kohonen’s Self-organizing Map (SOM) [14] has become an established method for forming topographic maps. In our model we use two SOMs to form topographic maps of the auditory and visual spaces. We then link these together using Hebbian connections to align the visual and auditory representations (Fig. 1).

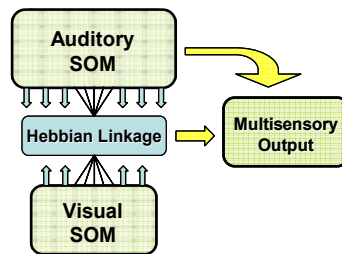


Fig. 1. Block diagram of the SC model. The model consists of a SOM for both the visual and auditory senses, connected together via Hebbian links. The linkage acts as a coordinate translator of the visual SOM to the auditory SOM. These are then combined to form the multisensory map.

A. Input Representation

As input to our model we use a simplified representation of visual and auditory space that corresponds to a spatial stimulus in each modality. Although there is limited understanding of the actual input to the SC, it is clear that the SC translates inputs into spatial topographic maps in order to determine the co-ordinates for any gaze shift. For our model we use Gaussian activity patterns (see for example [16]) to achieve this for both the auditory and visual modalities (Fig. 2). Each pattern is centered at regular discrete intervals to give the value of an input x at elevation i and azimuth j as:

$$x_{ij} = \lambda e^{-\left(\frac{i^2 - j^2}{\sigma^2}\right)} \quad (1)$$

where λ is the amplitude, and σ is the bandwidth.

The use of this representation offers a flexible and simple way of testing feature detection [17]. At the same time it allows us to define receptive field refinement by adjusting the amplitude and bandwidth of the Gaussian pattern. For both auditory and visual inputs the amplitude and bandwidth vary between two pairs of values corresponding to a dense (fovea) and a non-dense region (peripheral). For the denser firing regions, greater amplitude with smaller width is used to capture the refined representation of the stimuli occurring within these spatial windows (Fig. 2).

B. Sensory Maps

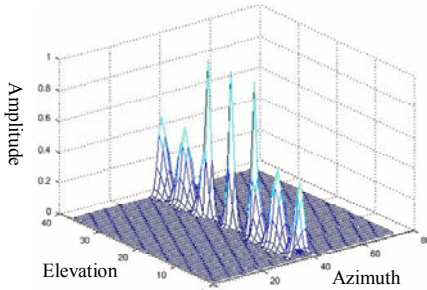


Fig. 2. Sample auditory input with 7 individual patterns shown. Each Gaussian pattern represents a single input for the auditory SOM. The central 3 peaks are inside the dense region, and hence are less wide and with a higher peak than the remaining patterns outside this region.

The topographic organization of stimuli is a property met in brain structures across a number of species [1]. Kohonen's SOM [14] is a biologically motivated algorithm that builds upon earlier studies that examined spatial ordering in sets of feature-sensitive cells [18]. We use SOM to form our auditory and visual topographic maps, using the standard weight update rule. The output of a map is calculated as the normalized inverse distance between an input vector x^{aud} and each of the weight vectors w_n^{aud} , in this case for the auditory map, for each neuron n , with $1 \leq n \leq N^{aud}$, so that:

$$u_n^{aud} = \frac{1}{\|x^{aud} - w_n^{aud}\|} \quad (2)$$

$$y_n^{aud} = \frac{u_n^{aud} - \alpha^{aud}}{\beta^{aud} - \alpha^{aud}} \quad (3)$$

In this way we highlight similarly responding areas of the map for the Hebbian links. The normalization parameters are chosen to be the minimum α^{aud} and maximum β^{aud} values of u_n^{aud} obtained from the map's unisensory training for calibration. The same equations hold for the visual map with the appropriate change of superscripts.

C. Multisensory Translation

Our output multisensory space has a spatial range that corresponds to the superset of elevation and azimuth values

for the sensory inputs, which in this case corresponds to the auditory space. Therefore, to form a multisensory representation, we combine the outputs of the auditory SOM with that of the translated outputs of the visual SOM. The translation is achieved by learning the association between the output of the auditory and visual maps for coincident stimuli at the same spatial location (a multisensory stimulus). To form the association between the map responses we use Hebbian learning [15]. Here each neuron in the visual map is connected to each neuron in the auditory map.

The output from the links are the normalized weighted summations of the outputs from the visual map y^{vis} multiplied by the corresponding connection weight w_{ni}^{link} :

$$u_n^{link} = \sum_{i=1}^{N^{vis}} w_{ni} y_i^{vis} \quad (4)$$

During training, the link weights are updated using the activity product rule:

$$w'_{ni} = w_{ni} + \eta y_i^{vis} y_n^{aud} \quad (5)$$

using learning rate η . These are then normalized to have unit magnitude to prevent exponential growth:

$$w_{ni} = \frac{w'_{ni}}{\|w'_{ni}\|} \quad (6)$$

Once trained, the output from the links are normalized in the same way for the sensory map outputs, using the minimum α^{link} and maximum β^{link} values of u_n^{link} obtained from the multisensory training data.

$$y_n^{link} = \frac{u_n^{link} - \alpha^{link}}{\beta^{link} - \alpha^{link}} \quad (7)$$

Finally, the multisensory output is formed from the summation of the normalized auditory map and visual link outputs. Normalization of the values is used here to equally bias the contribution of each modality:

$$y^{ms} = y^{aud} + y^{link} \quad (8)$$

IV. EXPERIMENTS AND EVALUATION

To evaluate the model, we explore the ability of the SOMs to represent distinct sensory inputs spatially, and then how each of these representations can be combined using Hebbian learning to give a multisensory space. All experiments were carried out using Matlab (version 7.3.0.298) and the SOM Toolbox [19]¹.

Unisensory training took place using two independently generated data sets for each of the SOMs. The data sets used Gaussian pattern activations centered in random locations (uniformly distributed) within the auditory and visual spaces. The selection of training data used varied depending upon the input space, with additional data

¹ Matlab source and experimental data files for this work can be found at <http://www.cs.surrey.ac.uk/BIMA/People/M.Casey/software.html>.

selected explicitly for each of the dense regions to ensure good coverage of stimuli. For a given input, discrete values for the Gaussian were calculated using a uniform elevation and azimuth step size.

For multisensory training of the links, datasets were generated for both the auditory and visual space using a single set of randomly selected centers (uniformly distributed). The centers were selected from the range of coordinates from the visual space to ensure that coincident stimuli were presented to both the auditory and visual maps to allow an association to form between the outputs.

For both unisensory and multisensory training phases, randomly selected testing data were generated using the same specifications as for the training data. Details of the training and testing data sets are presented in Table I.

TABLE I
TRAINING AND TESTING DATA

		Auditory	Visual
Input Space			
Non-dense area	Elevation	[-90, -30]; (30, 90]	[-65, -15]; (15, 65]
	Azimuth	[-180, 180]	[-90, -15]; (15, 90]
	Gaussian	$\lambda = 0.5; \sigma = 10$	$\lambda = 0.5; \sigma = 10$
Dense area	Area	33%	1%
	Elevation	[-30, 30]	[-15, 15]
	Azimuth	[-180, 180]	[-15, 15]
	Gaussian	$\lambda = 1; \sigma = 5$	$\lambda = 1; \sigma = 5$
Discrete step size	5	5	
Input dimension	2701 (37 by 73)	999 (27 by 37)	
Unisensory training and testing			
Whole area examples	1675	810	
Additional dense examples	825 (33%)	90 (10%)	
Total examples	2500	900	
Multisensory training and testing			
Whole area examples	1000	1000	
Additional dense examples	125 (11%)	125 (11%)	
Total examples	1125	1125	

A. Unisensory maps

The sizes of the auditory and visual SOMs were selected to be computationally efficient, although being approximately in proportion with the corresponding input space (Table II). After training each SOM individually on the training data for 1000 epochs, both maps were tested on the testing data. The results show that both SOMs organized the inputs based upon spatial location, and in particular, larger regions of the maps correspond to the dense areas.

TABLE II
SOM PARAMETERS

		Auditory	Visual
Size		20 by 15 $N^{aud} = 300$	10 by 10 $N^{vis} = 100$
Lattice		Rectangular	Rectangular
Training epochs		1000	1000
Learning rate	Type	Inverse	Inverse
	Initial rate	0.5	0.5
Neighborhood	Type	Gaussian	Gaussian
	Initial radius	20	10
	Final radius	1	1

Fig. 3(a) shows the auditory SOM U-matrix output (grayscale background), overlaid with the best matching units for different azimuth strips. On the lower part of the

map, clusters of all inputs centered within an azimuth of [30, 64.9] and [65, 90] (dark blue and blue) were observed. At the center of the map, only clusters of azimuth strips [-0.9, 29] and [-29.9, -1] were found. These clusters (green, dark green) comprised approximately 55% of the auditory SOM and included all inputs centered within the dense area, which is only 33% of the total input space, hence showing an increased representation within the map corresponding with increased acuity. Finally, the last two clusters corresponded to azimuth strips [-64.9, -30] and [-90, 65].

Fig. 3(b) shows the same visualization for the visual

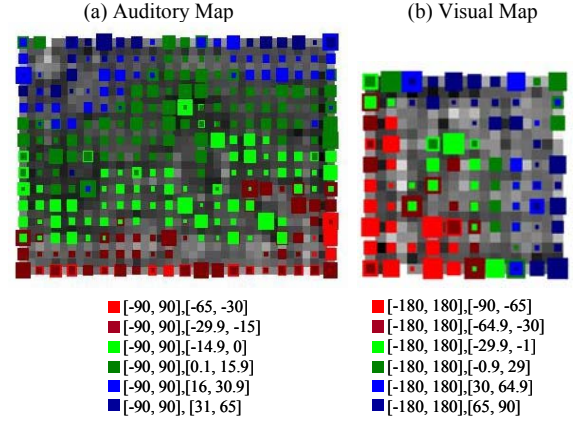


Fig. 3. U-matrix visualization of the (a) auditory and (b) visual SOMs overlaid with best matching units. Testing data were visualized by selecting azimuth strips that gradually cover the entire input space. The best matching units are colored depending on their corresponding inputs' centers. Both maps have learnt to preserve the spatial relationships of their inputs. Note that the representation of the dense areas comprise a significant portion of the maps: 2500 hits for the auditory SOM cover 55% of the map, whereas 900 hits for the visual SOM cover 26% of the map.

SOM, which shows the same pattern of organization. Note here that the dense area represents over 26% of the visual SOM, compared to just 1% of the visual space.

The first step of our model's training has successfully organized the visual and auditory maps based on the inputs' spatial relationships. In particular, the map has yielded expanded coverage of inputs centered within the dense areas. These results are in accordance with biological findings showing a magnification of the central areas on visual and auditory processing in the SC [1]. Furthermore, the preservation of spatial similarity in each modality implies that they can be aligned through a simple linear translation, such as a Hebbian association.

B. Multisensory Integration

Having developed the two sensory maps to process unisensory stimuli, we then combined them in the SC model to train the Hebbian links on multisensory stimuli. Prior to training, the outputs from both maps were normalized using the unisensory training minimum and maximum responses ($\alpha^{aud}=0.43, \beta^{aud}=1.54, \alpha^{vis}=0.37, \beta^{vis}=1.97$). The links were then trained for 100 epochs with a learning rate $\eta = 0.1$ on the multisensory training data. The links were then normalized ($\alpha^{link}=0.42, \beta^{link}=2.98$).

To determine if the links had established a

correspondence between visual and auditory map units, we examined the coincidence of auditory unit hits via the Hebbian links for each input. The effectiveness of the training was examined by recording the number of coincident hits between the maximum visual link output and the best matching unit of the auditory SOM. A direct hit is recorded when the link output matches exactly the auditory output. We also explored how near each of the hits were by

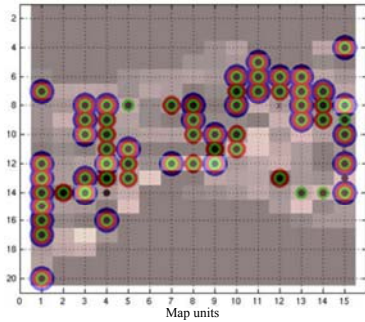


Fig. 4. Auditory best matching units overlaid with visual Hebbian link outputs. Blue circles signify that the maximal link output corresponds to the required auditory unit (0 radius). Red circles denote that the maximal output (for a different input) is located within a radius of 1 unit in the auditory map, green a radius of 2 units, black a radius of 3 units and an X a radius of 4 units – for example, unit (12, 8).

recording the coincidence between the maximum link output and units within a defined radius of the auditory best matching unit.

Fig. 4 shows the auditory best matching units overlaid with the Hebbian link outputs. We can see from these testing results that the links learnt to approximately associate auditory and visual stimuli map locations. Here, the Hebbian link outputs covered approximately 35% of the auditory SOM for a radius of 0 units, rising to 46% for a radius of 4 units. As expected, for spatial locations outside of the visual space (essentially a zero input to the visual map), the linkage could not translate this to an auditory unit, even though this would be within the auditory space. When using an input centered within the visual space then the Hebbian linkage activation area approximated that of the auditory activation. Direct unit to unit translation was achieved with 27.28% of the testing data. Although this percentage of direct hits appears low, if we look at the coincidence of link outputs to units within a defined radius of the auditory best matching unit, the percentage significantly increases. For a radius of 1 unit, the coincidence rises to 61.51% and for 2 units this rises to 70.66%. Overall therefore, the links learn to associate between units with the most hits, as seen by the high coincidence compared to the low coverage of the map.

The Hebbian links therefore successfully translate the visual spatial representation into an auditory representation, aligning the different sensory spatial coordinate systems. With this translation in place, we can now form the multisensory space as the addition of the auditory output and the visual link output. Fig. 5 shows the outputs from the model for (a, b, c) a non-coincident auditory and visual

input, and (d, e, f) a coincident auditory and visual input.

When selecting an input at the edge of the auditory space (outside the visual field), the link output is correspondingly low and is the same as the average link output (0.28). The resulting multisensory response (0.57) was therefore driven most by the auditory response. When using a multisensory input centered within the visual space, the link activation area approximated that of the auditory activation resulting in an enhanced response (0.74). Maximum multisensory enhancement was achieved only in the case of direct hits.

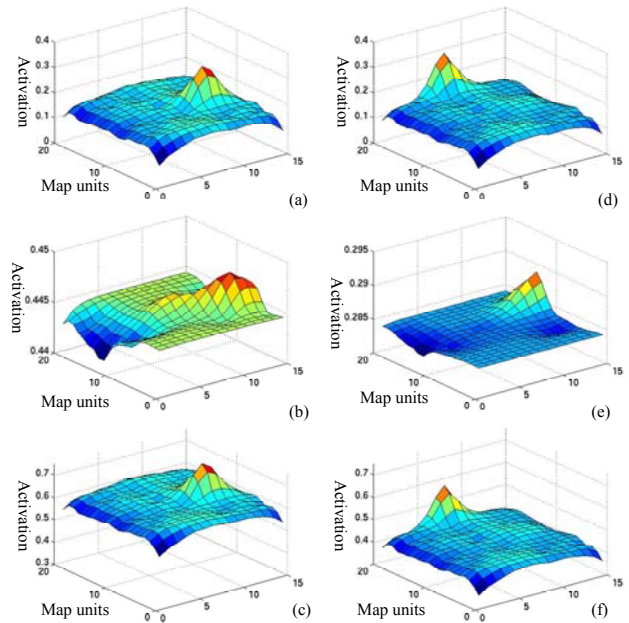


Fig. 5. Multisensory responses to non-coincident (a, b, c) and coincident inputs (d, e, f). Each plot shows the value of the outputs from the auditory SOM (a, d), visual links (b, e) and multisensory outputs (c, f).

(a, b, c) Auditory and visual stimulus with centre (elevation -9.6, azimuth 161.1) and multisensory output (0.57) at coordinate (20, 8). The link (b) and auditory response (a) at that location is 0.28 and 0.28, respectively. The location of the maximum link response (0.29) is at (14, 15).

(d, e, f) Auditory and visual stimuli with centre (elevation 18.3, azimuth -7.6) and multisensory output (0.74) at coordinate (10, 11). The link (e) and auditory response (d) on that location is 0.45 and 0.29, respectively. The location of the maximum link response (0.45) is at (8, 13).

The increase in output seen with coincident multisensory stimuli is commensurate with multisensory enhancement, in which a higher activation is achieved [20]. However, while our model shows that we have successfully integrated sensory responses, and in particular translated the visual representation into an auditory representation, it uses a too simplistic combination method to show multisensory suppression when only one sensory stimulus is presented (seen by the 0.57 response), especially since current research has started to show that the enhancement response is logarithmic in nature [6]. While our model does not show this type of response, it does successfully align the sensory maps and combine them to show enhancement. Further work can be used to refine the combination, including varying unimodal stimuli intensity and combination method (for example multiplicative or logarithmic).

V. CONCLUSION

In this paper we have presented a simple behavioral model of the superficial and deep layers of the mammalian SC. Our model comprises a visual (superficial) and auditory (deep) topographic map, each behaving in a way that is comparable to the biology, with the preservation of spatial similarity and the increased representation of denser regions (fovea). We then combined these unisensory representations using associative learning to test whether this is a viable method for modeling integration. Through the presentation of coincident multisensory stimuli, the visual map is aligned with the auditory map. Once aligned, the representation shows a simple form of multisensory enhancement (deep), which is a consequence of the additive combination of (translated) sensory signals.

With this work we set out to answer two questions. First, can we use a computational model to understand better the process of sensory alignment and multisensory integration within the SC? Using simple techniques we have successfully shown that Hebbian learning can be used to align two spatial representations through a process of multisensory training (although not three senses). Not only does this show that this is possible with a simple model, it also demonstrates how cortical feedback may play a role in the integration. Although we have not explicitly included cortical control, implicitly we have selected to associate multisensory responses together during training. In this way our training regimen has simulated cortical selectivity, although our approach is not sufficiently sophisticated to provide further insight. Future work will address this issue by explicitly modeling cortical feedback and enhancing the integration.

Second, we chose to implement a fuller model of the SC in order to gain insight into whether such modeling is of benefit computationally. While this has obviously not provided a step change in computational capabilities, what this has shown is that we can learn to associate sensory stimuli. Although our choice of input representation is simple, they can be replaced by simple real-world artificial sensory inputs, which can be input to the model at discrete time intervals. While this excludes more complex inputs, such as images, videos or sound without further preprocessing, it does mean that this model can be embedded into a prototype robot in order to get it to react to multisensory stimuli, much like the cats used in the biological experiments [1]. The insight is therefore that we should incrementally increase the size and complexity of such models, and embed them into the real-world where possible to find the step change.

ACKNOWLEDGEMENT

We would like to thank Jim Austin, Jim Bednar, Alan Murray, Leslie Smith, Barry Stein and Stefan Wermter for early discussions on this work, as well as the three anonymous reviewers for their constructive comments.

REFERENCES

- [1] B.E.Stein and M.A.Meredith, *The Merging of the Senses*. Cambridge, MA.: A Bradford Book, MIT Press, 1993.
- [2] B.E.Stein, W.Jiang and T.R.Stanford, "Multisensory Integration in Single Neurons of the Midbrain," G.A.Calvert, C.Spence, and B.E.Stein, Ed. Cambridge, MA.: A Bradford Book, MIT Press, 2004, pp. 243-264.
- [3] A.A.Ghazanfar and C.E.Schroeder, "Is Neocortex Essentially Multisensory," *Trends in Cognitive Sciences*, vol. 10, pp. 278-285, 2006.
- [4] M.Denham and T.Tarassenko, "Sensory Processing," Foresight Cognitive Systems Project, London, 2003.
- [5] A.J.King, "The Superior Colliculus," *Current Biology*, vol. 14, pp. R335-R338, 2004.
- [6] B.A.Rowland, T.R.Stanford and B.E.Stein, "A Model of the Neural Mechanisms Underlying Multisensory Integration in the Superior Colliculus," *Perception*, vol. 36, pp. 1431-1443, 2007.
- [7] S.Grossberg, K.Roberts, M.Aguilar and D.Bullock, "A Neuronal Model of Multimodal Adaptive Saccadic Eye Movement Control by Superior Colliculus," *Journal of Neuroscience*, vol. 17, pp. 9706-9725, 1997.
- [8] A.Haessly, J.Sirosh and R.Miikkulainen, "A Model of Visually Guided Plasticity of the Auditory Spatial Map in the Barn Owl," in *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, 1995, pp. 154-158.
- [9] C.Quaia, P.Lefèvre and L.M.Optican, "Model of the Control of Saccades by Superior Colliculus," *Journal of Neurophysiology*, vol. 82, pp. 999-1018, 1999.
- [10] T.P.Trappenberg, M.C.Dorris, D.P.Munoz and R.M.Klein, "A Model of Saccade Initiation Based on the Competitive Integration of Exogenous and Endogenous Signals in the Superior Colliculus," *Journal of Cognitive Neuroscience*, vol. 13, pp. 256-271, 2001.
- [11] V.Cutsuridis, N.Smyrnis, I.Evdokimidis and S.Perantonis, "A Neural Model of Decision-making by the Superior Colliculus in an Antisaccade Task," *Neural Networks*, vol. 20, pp. 690-704, 2007.
- [12] H.H.L.M.Goossens and A.J.Van Opstal, "Dynamic Ensemble Coding of Saccades in the Monkey Superior Colliculus," *Journal of Neurophysiology*, vol. 95, pp. 2326-2341, 2006.
- [13] P.E.Patton and T.J.Anastasio, "Modeling Cross-Modal Enhancement and Modality-Specific Suppression in Multisensory Neurons," *Neural Computation*, vol. 15, pp. 783-810, 2003.
- [14] T.Kohonen, "Self-Organized Formation of Topologically Correct Feature Maps," *Biological Cybernetics*, vol. 43, pp. 59-69, 1982.
- [15] D.O.Hebb, *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons, 1949.
- [16] R.Miikkulainen, J.A.Bednar, Y.Choe and J.Sirosh, *Computational Maps in the Visual Cortex*. New York: Springer Science+Business Media, 2005.
- [17] J.L.McClelland, A.Thomas, B.D.McCandliss and J.A.Fiez, "Understanding Failures of Learning: Hebbian Learning, Competition for Representational Space and Some Preliminary Experimental Data," J.Reggia, E.Ruppin, and D.Glanzman, Ed. Oxford: Elsevier, 1999, pp. 75-80.
- [18] C.von der Malsburg, "Self-organization of Orientation Sensitive Cells in the Striate Cortex," *Kybernetik*, vol. 14, pp. 85-100, 1973.
- [19] J.Vesanto, J.Himberg, E.Alhoniemi and J.Parhankangas, "Self-Organizing Map in Matlab: the SOM Toolbox," http://www.cis.hut.fi/projects/somtoolbox/package/papers/toolbox2pa_per.pdf, 2000.
- [20] M.T.Wallace, M.A.Meredith and B.E.Stein, "Multisensory Integration in the Superior Colliculus of the Alert Cat," *Journal of Neurophysiology*, vol. 80, pp. 1006-1010, 1998.