

# Multimodal FeedForward Self-organizing Maps

Andrew P. Papliński<sup>1</sup> and Lennart Gustafsson<sup>2</sup>

<sup>1</sup> Clayton School of Information Technology,  
Monash University, Victoria 3800, Australia  
`app@csse.monash.edu.au`

<sup>2</sup> Computer Science and Electrical Engineering,  
Luleå University of Technology, S-971 87 Luleå, Sweden  
`Lennart.Gustafsson@luth.se`

**Abstract.** We introduce a novel system of interconnected Self-Organizing Maps that can be used to build feedforward and recurrent networks of maps. Prime application of interconnected maps is in modelling systems that operate with multimodal data as for example in visual and auditory cortices and multimodal association areas in cortex. A detailed example of animal categorization in which the feedforward network of self-organizing maps is employed is presented. In the example we operate with 18-dimensional data projected up on the 19-dimensional hyper-sphere so that the “dot-product” learning law can be used. One potential benefit of the multimodal map is that it allows a rich structure of parallel unimodal processing with many maps involved, followed by convergence into multimodal maps. More complex stimuli can therefore be processed without a growing map size.

## 1 Introduction and Motivation

We present a generalization of Kohonen self-organizing maps [1] that allows us to build feedforward and feedback structures consisting of interconnected self-organizing maps. This work has originated from our attempt to model deficits in learning, caused by attention abnormalities in autism [2, 3, 4, 5, 6, 7], therefore is inclined towards modelling functions of brain. However, the results presented are more general and can be applied to interpretation of any multimodal data.

Self-organizing maps (SOMs) and their applications are very popular topic to study. The search of *IEEEExplore* shows that since 2000 75 papers has been published in IEEE periodicals and another 528 in IEEE conference proceedings. A significant number of papers have been also published outside IEEE. Most of the publications are devoted to a variety of applications using a variant of the basic Kohonen algorithm for a single SOM. In [1] Kohonen discussed possible variations of SOMs that include: maps of varying topology that has been studied, for example, in [8, 9], and tree-structured SOMs to improve the winner search procedure, e.g. [10]. The systems of interconnected SOMs that our paper is preoccupied with have not been to our knowledge studied yet and is designated “for future research” in [1] p.194.

There are also self-organizing systems of four hierarchically organized maps, *VisNet* and *VisNet2* [11, 12]. These systems, designed to capture some essential characteristics of the visual system, achieve learning through variants of Hebb's learning rule and ordering of the maps by lateral connections. These systems, while in important aspects more biological than the Kohonen maps presented in this paper, are not multimodal.

### 1.1 An Introduction to the Modular Architecture of Mammalian Sensory Processing

Organization of the mammalian neocortex is modular and hierarchical. A columnar structure, consisting of minicolumns and macrocolumns, comprising in the order of fifty to a hundred and a number of thousand neurons respectively, is ubiquitous [13]. On a higher aggregational scale sensory processing of different modalities is done in visual cortex, auditory cortex etc. with a number of areas, hierarchically arranged, in each (the human visual cortex has upwards of thirty areas), see e.g. [14]. These areas are primary areas (visual area V1, auditory area A1, etc.) and several levels of higher unimodal association areas. The primary areas are organized retinotopically, tonotopically etc., each neuron having a small receptive field. Higher association areas process information of a more specific kind, e.g. face recognition, each neuron having a large receptive field. For an introduction to the organization of neocortex see [15].

The stimuli presented to our nervous system are not restricted to one modality however, they are generally bimodal or multimodal. We see and hear, we see and feel, we smell and taste. If we hear an automobile crash it will catch our attention and guide our vision, through a saccade or a head movement or both, to the scene of the accident. While lower level processing of stimuli is largely unimodal, multimodal percepts are formed when the results of the unimodal processing are combined in multimodal areas in, e.g. the orbitofrontal cortex, see e.g. [16]. There are different kinds of connections between and within areas. Between areas in a hierarchy there are bottom up connections but also top down connections. There are lateral connections between areas on the same level in a hierarchy and within areas there are recurrent feedback connections. The areas may be seen as parallel processors, interchanging processing results.

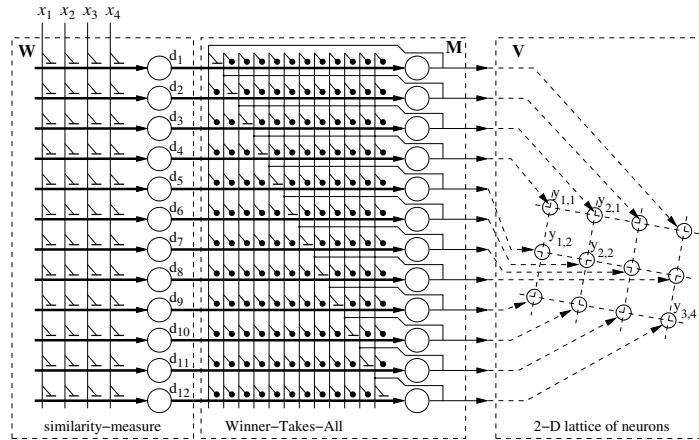
The modular architecture of cortex may be formed from such biological restrictions as upper limits of a neuron's connectivity. It may also offer a processing speed advantageous to process sensory stimuli in a highly modular way with processing in all modules running in parallel. The hierarchical structure of cortex may even offer a specific advantage to humans. There is evidence that synaptogenesis is heterochronous in human neocortex as opposed to synaptogenesis in rhesus monkeys ([17, 18]). In humans the maximum density of synapses is reached earlier in sensory cortices (earlier in visual cortex than in auditory cortex) than in prefrontal cortex, in contrast to the development in rhesus monkeys where synaptogenesis is concurrent in all neocortical areas. Successive organization of areas as they appear in the sensory processing hierarchy may well be advantageous.

## 2 The Structure of the Feedforward Self-organizing Maps

Neural networks have been inspired by the possibility of achieving information processing in ways that resemble those of biological neural systems. Even though some neural network architectures and learning rules, such as the multilevel perceptrons and the backpropagation error correcting learning algorithms have developed into independent signal processing systems, others such as pattern associators based on Hebbian learning [19] and self-organizing networks [1] show considerable similarities with biological neural systems. It has been shown that pattern associators may well simulate the multimodal sensory processing in cortex [16].

The purpose of this paper is to show that the Kohonen networks, contrary to popular belief, lend themselves very well to building networks in a modular and multimodal way. Simulation of such modular networks consisting of interconnected Kohonen networks are suitable for execution on parallel processors. Also, a large number of interconnected Kohonen networks may serve as a simulation model for biological sensory processing.

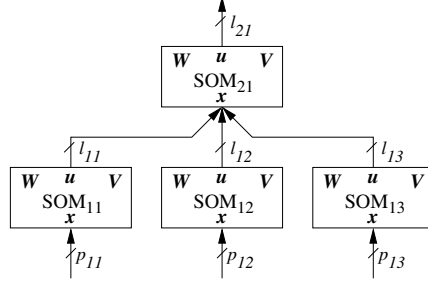
Kohonen Self-Organizing Maps (SOMs) [1] are well-recognized and efficient tools for mapping multidimensional stimuli onto a low dimensionality (typically 2) neuronal lattice. Example of such a map is given in Fig. 1.



**Fig. 1.** A SOM mapping a 4-dimensional stimuli onto a 2-dimensional  $3 \times 4$  neuronal lattice

For our purposes we can identify a similarity layer, a competitive layer based on lateral inhibition and local self-excitatory feedback and a mapping mechanism to the output layer. We can say that for a given matrix of neuronal weights,  $W$ , the  $n$ th stimulus  $\mathbf{x}(n)$  is mapped into a position of the **winner**  $\mathbf{v}(n)$ , that is, the neuron with the weight vector most similar to the stimulus.

$$\mathbf{v}(n) = g(\mathbf{x}(n); W) ; \mathbf{v} \in \mathbb{R}^l \quad (1)$$



**Fig. 2.** A simple feedforward structure of self-organizing maps built from self-organizing blocks.

where  $l$  is a dimensionality of the neuronal grid. The self-organizing building block can now be used to form feedforward and feedback structures. In this paper we concentrate on the feedforward networks as in example in Fig. 2.

The bottom level maps,  $SOM_{1,k}$ , received sensory stimuli  $\mathbf{x}_{1,k}$  of different dimensionality  $p_{1,k}$  representing different modalities of afferent signals. Dimensionality of each map is, in general, equal to  $l_{1,k}$ , but is, most typically, equal to 2 for the ease of visualisation. The higher level map,  $SOM_{21}$ , received the re-coded stimuli from each lower level map of the total dimensionality  $p_{21} = \sum_{k=1}^K l_{1,k}$ , e.g. 6 for the example of Fig. 2. Each individual map, e.g. the one from Fig. 1, consists of  $M = m_1 \times m_2 \dots \times m_l$  neurons, where  $m_i$  is the map size in the  $i$ th dimension. In a feedforward structure as in Fig. 2 learning takes place concurrently for each SOM, according to the well-known Kohonen learning law.

In the practical example presented below we work with normalised stimuli and activity data, therefore we use the simple “dot-product” learning law [1, 5]. In this case the update of the weight vector for the  $j$ th neuron is described by the following expression:

$$\Delta \mathbf{w}_j = \eta \cdot \Lambda_j \cdot (\mathbf{x}^T - d_j \cdot \mathbf{w}_j) ; \quad d_j = \mathbf{w}_j \cdot \mathbf{x} \quad (2)$$

where  $\Lambda_j$  is a neighbourhood function, Gaussian in our case, centred on the position of the winning neuron, and  $d_j$  is the post-synaptic activity of the  $j$ th neuron. It is easy to show that for the above learning law, if the stimuli  $\mathbf{x}$  are on a unity hyper-sphere, the resulting weight vectors  $\mathbf{w}$  are located on, or close to such a sphere.

The higher level map learns from the combined centres of activity produced by the lower level maps. During testing, a multimodal stimulus is applied and for each  $i, k$  map on every level, we can record the position of the winning neuron,  $\mathbf{v}_{i,k}$ .

### 3 Example of Mapping Multi-modal Stimuli

As an illustrative example we consider a well-known problem of categorization of animals. Our animal kingdom that we have also used in our work on modelling

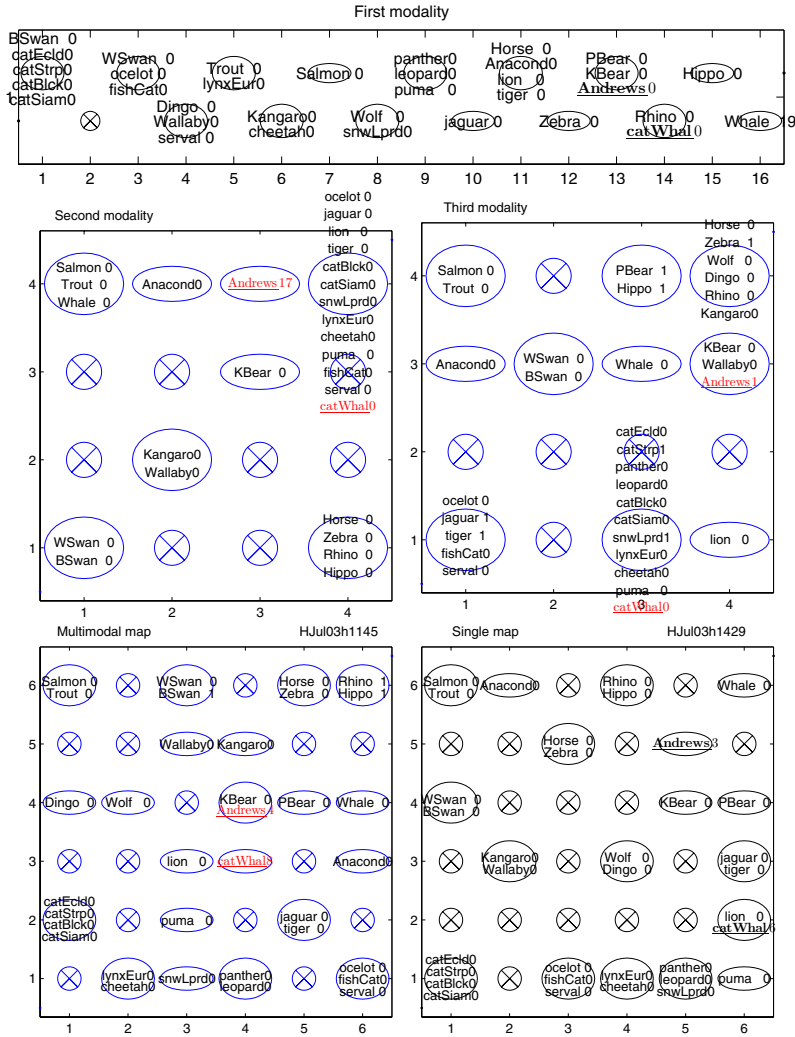
autism [5, 4, 3] consists of 32 animals each characterized by 18-dimensional data. Half of the animals are variety of cats. The animals are listed below sorted according to their weight:

Grey whale, Hippopotamus, White rhinoceros, Kodiak bear, Polar bear, Grevy's zebra, Przewalski's horse, Tiger, Lion, Anaconda, Jaguar, Puma or cougar, Panther, Leopard, Snow Leopard, Canis lupus (Wolf), Atlantic salmon, Cheetah, Grey western kangaroo, Eurasian lynx, Rainbow trout, Dingo, Swamp wallaby, Serval, Ocelot, Fishing cat, Mute (white) swan, Black swan, Domestic cat (even coloured), Domestic cat (striped), Domestic cat (black), Siamese cat.

The features chosen to characterize these animals are as follows:  $x_1 = \log(\text{weight})$ ;  $x_2 \in \{1, 2, 3\}$  – food (herbivores, omnivores, carnivores);  $x_3, x_4, x_5, x_6$  (binary) – locomotion (fins, wings, two legs, four legs);  $x_7$  (binary) – equipped with hooves (perissodactyls) or cloven hooves (artiodactyls);  $x_8$  (binary) – equipped with claws;  $x_9$  (binary) – equipped with other feet;  $x_{10} \in \{1, 2, 3\}$  – cover (fur, feathers and scales);  $x_{11}$  (binary) – colour black;  $x_{12}$  (binary) – colour white;  $x_{13}$  (binary) – even coloured;  $x_{14}$  (binary) – spotted;  $x_{15} \in \{0, 1/4\}$  (binary) – striped;  $x_{16} \in \{2, 4\}$  – facial feature (short faced, long faced);  $x_{17} \in \{1, 2, 3\}$  – aquatic;  $x_{18} \in \{1, 2, 3\}$  – social behaviour (single living, pair living, group living). In addition, for testing of the generalization of the multi-map structure we use two unusual animals, one being a domestic cat weighing two tonnes ('catWhale'), the other, *Andrewsarchus mongoliensis*, is an extinct cloven-hoofed 1-tonne carnivour.

The features by which we describe our animals are of course multimodal and some, like behaviour, represent semantic memory data rather than immediate sensory stimuli. This example does not imply that a biological neural system would treat the total data in three primary unimodal and one multimodal map. Rather the example illustrates the functioning of a very small multiple map system. The 18-dimensional vector representing an animal is split into three sub-vectors, representing three modalities:  $\mathbf{x}_{1,1} = x_1$  (weight),  $\mathbf{x}_{1,2} = (x_2 \dots x_9)$  (food and locomotion),  $\mathbf{x}_{1,3} = (x_{10} \dots x_{18})$  (coloration, facial features, behaviour). Single modality stimuli of dimensionality  $p_{11} = 1, p_{12} = 8, p_{13} = 9$ , excite three first-level maps as shown in Fig. 2. In the example below we use 16 neurons in each sensory map organized on a  $1 \times 16$  grid for the (1,1) map and  $4 \times 4$  grids for (1,2) and (2,3) maps. The second level map has 36 neurons organized on a  $6 \times 6$  grid. The result of learning can look like in Fig. 3. The numbers that follows the animal's name is a relative distance between the animal and the closest neuron. In well-developed maps most of these distances are expected to be zero.

The organization of the three unimodal maps justifies some comments. The **first modality** map, representing weights, is a one-dimensional map with the animals simply arranged in descending weight order from left to right. The neuron which is the best fit for a particular animal has an oval shape with the name of the animal written in. Since there are thirty-two animals and the map was provided with only sixteen neurons, in some cases several animals have to share one neuron for its best fit. The **second modality** map represents food and locomotion. Therefore it is natural that salmon, trout and whale share one neuron, whereas almost all land predators share one neuron. The other animals can be



**Fig. 3.** Categorization of animals with the feedforward multimodal maps and a conventional single map

commented in similar ways. The **third modality** map represents coloration, facial features and behaviour. This map results in more discrimination than the second modality map. The cats which forms a tight knit group in the second modality are here split onto three neurons, cats which appreciate water, cats which do not and the lion which lives in larger groups. The Kodiak bear and the wallaby might seem an odd pair, but they both have very elongated faces, they are both even colored and they both are solitary animals. The kangaroo lives in groups and is therefore grouped with zebras, wolves and other social animals. The clustering of animals seen in the third modality map may at first seem absurd but it quite correctly reflects the features that were chosen for this map.

In the multimodal map we see that the animals are represented in a way which we find natural. Some animals like the whale and the anaconda are solitary in the multimodal map in all experiments with these data, simply because no other animals in our animal kingdom resemble them. Other animals form pairs and then they are very similar, as described by the features we chose. The dingo and the wolf are solitary in this map while in other experiments they appear as a pair. The same is true for the wallaby and the kangaroo, and the Kodiak bear and the polar bear. The cats are more often than not grouped with other cats, usually distinguished from other groups of cats mainly by weight. The four house cats which is one species with different coloration are grouped together here and this is the case for all experiments with maps of this size. The lion often has its own neuron, in some experiments it shares one neuron with the tiger and the jaguar. There are quite a number of neurons which are not best fits for any animal. This is in the nature of Kohonen maps. The two animals that were presented to the map without any prior learning, the fictitious catwhale and the extinct *Andrewsarchus* are represented in a reasonable way. The catwhale is close to the lion which is a good compromise for an animal with the conflicting features of a house cat except for the weight of a rhino. The *Andrewsarchus* is grouped with the Kodiak bear, which is reasonable, given the features by which it was described.

For comparison, the “single map” in figure 3 has been generated using a traditional unimodal Kohonen algorithm. This map has the same characteristics as the multimodal map. The detailed results are different but different experiments yield slightly different results and it is not possible to tell by the result if the map was generated as a multimodal map or as a single map.

## 4 Conclusion

The multimodal maps discussed in the paper have many potential benefits over their single map counterparts. One potential benefit of the multimodal map is that it allows a rich structure of parallel unimodal processing with many maps involved, followed by convergence into multimodal maps. More complex stimuli can therefore be processed without a growing map size. A second potential benefit to be explored is that this modular multimodal structure may be used to simulate biological neural systems.

**Acknowledgment.** This work has been supported by Monash University Small Grant, Faculty of Information Technology (Australia) and by the Department of Computer Science and Electrical Engineering, Luleå University of Technology (Sweden).

## References

1. Kohonen, T.: *Self-Organising Maps*. 3rd edn. Springer-Verlag, Berlin (2001)
2. Gustafsson, L., Papliński, A.P.: Self-organization of an artificial neural network subjected to attention shift impairments and novelty avoidance: Implications for the development of autism. *J. Autism and Dev. Disorder* **34** (2004) 189–198

3. Papliński, A.P., Gustafsson, L.: An attempt in modelling early intervention in autism using neural networks. In: Proc. Int. Joint Conf. Neural Networks, Budapest, Hungary (2004) 101–108
4. Gustafsson, L., Papliński, A.P.: Neural network modelling of autism. In Casanova, M.F., ed.: Recent developments in autism research. Nova Science Publishers, Inc., Hauppauge, New York (2005) 100–134 in press.
5. Papliński, A.P., Gustafsson, L.: Detailed learning in narrow fields – towards a neural network model of autism. In Kaynak, O., Alpaydin, E., Xu, L., eds.: Lect. Notes in Comp. Sci. Volume 2714., Springer (2003) 830–838
6. Gustafsson, L., Papliński, A.P.: Preoccupation with a restricted pattern of interest in modelling autistic learning. In Pallade, V., Howlett, R.J., Jain, L., eds.: Lecture Notes in Artificial Intelligence. Volume 2774, Part II., Springer (2003) 1122–1129
7. Papliński, A.P., Gustafsson, L.: An attempt in modelling autism using self-organizing maps. In: Proc. 9th Intern. Conf. Neural Information Processing, Singapore (2002) 301–304
8. Alahakoon, D., Halgamuge, S.K., Srinivasan, B.: Dynamic self-organizing maps with controlled growth for knowledge discovery. *IEEE Trans. Neural Networks* **11** (2000) 601–614
9. Milano, M., Koumoutsakos, P., Schmidhuber, J.: Self-organizing nets for optimization. *IEEE Trans. Neural Networks* **15** (2004) 758–765
10. Xu, P., Chang, C.H., Papliński, A.: Self-organizing topological tree for on-line vector quantization and data clustering. *IEEE Tran. System, Man and Cybernetics, Part B: Cybernetics* **35** (2005) 515–526
11. Wallis, G., Rolls, E.: Invariant face and object recognition in the visual system. *Progress in Neurobiology* **51** (1997) 167–194
12. Rolls, E., Milward, T.: A model of invariant object recognition in the visual system: learning rules, activation function, lateral inhibition, and information-based performance measures. *Neural Computation* **51** (2000) 2547–2572
13. Mountcastle, V.B.: The columnar organization of the neocortex. *Brain* **120** (1997) 701–722
14. Felleman, D.J., van Essen, D.C.: Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* **1** (1991) 1–47
15. Kandel, E.R., Schwartz, J.H., Jessel, T.M., eds.: Principles of Neural Science. 4th edn. McGraw Hill, New York (2000)
16. Rolls, E.T.: Multisensory neuronal convergence of taste, somatosensory, visual, and auditory inputs. In Calvert, G., Spencer, C., Stein, B.E., eds.: The Handbook of multisensory processes. MIT Press (2004) 311–331
17. Huttenlocher, P.R., Dabholkar, A.S.: Regional differences in synaptogenesis in human cerebral cortex. *J. Comparative Neurology* **387** (1997) 161–178
18. Rakic, P., Bourgeois, J.P., Goldman-Rakic, P.S.: Synaptic development of the cerebral cortex: implications for learning, memory, and mental illness. In van Pelt, J., Comer, M.A., Uylings, H.B.M., Lopes da Silva, F.H., eds.: Prog. Brain Research. Elsevier Sci. (1994) 227–243
19. Hopfield, J.: Neural networks and physical systems with emergent collective computational properties. *Proc. Nat. Academy of Sci. USA* **79** (1982) 2554–2588