Proceedings of the International Multiconference on Computer Science and Information Technology pp. 225–236
ISSN 1896-7094

Manifestation of selective attention in Sigma-if Neural Network

Maciej Huk[1]

Wroclaw University of Technology
Institute of Applied Informatics
Wroclaw, Poland
Maciej.Huk@pwr.wroc.pl, http://www.iis.pwr.wroc.pl/~huk

Abstract. Artificial neural networks are a very well known biologically-inspired machine learning technique. This technique has been widely applied in many domains, such as real-time signal filtering, modelling and synthesis, process control and classification (e.g. of images, diseases or star spectra). Other examples of use of neural networks include generation of rules for expert systems and knowledge discovery. However, effective analysis of multidimensional data sets still causes particular difficulties. It has been shown that processing too many data features is costly and has adverse effects on the classification properties of resulting models. Thus, while development of techniques for selecting features from very large data sets is not trivial, it is nonetheless very important. It can be observed that analogous problems are effectively solved by human low-level distributed selective attention mechanisms. For this reason, building a general (e.g. neural) model of such functionality would provide great benefits for the machine learning domain. This work addresses selective attention functionality found in the recently developed Sigma-if neural network. Experiments show how this selective attention model can reduce data acquisition and processing costs as well as the probability of classification errors.

1 Introduction

Artificial neural networks are one of the most valuable biologically-inspired machine learning techniques. They play a significant role in many data processing applications - from real-time signal filtering, modelling and synthesis, through process control to classification (e.g. of images, diseases or star spectra). Other important examples of neural network use include generation of rules for expert systems and knowledge discovery in data [1–4]. Typical examples of artificial neural networks are multilayer perceptrons (MLP), radial basis function networks (RBF) and asynchronous spiking neural networks. A wider description of artificial neural network types can be found in many recent surveys [5–8].

In turn, biological inspirations of artificial neural networks suggest that such systems can effectively realise feature selection tasks easily observed in humans and other primates [9, 10]. Artificial models of this phenomenon could reduce problems with proper data selection and integration still observed in most machine learning systems. This could be of great benefit to all domains in which effective analysis of multidimensional data sets is crucial [11–13]. The most important question in this field is what basic neural mechanisms implement the functionality called selective attention [14–16].

Despite the great variety of known neural network models, very few realize any aspect of selective attention functionality [15, 17, 18]. This is probably due to the fact that such models do not copy some important processes which evidently take place during input signal aggregation in real neurons [17, 19–25]. In this context it is very interesting that selective attention functionality, which seems to effectively mimic low-level attentional processes observed in humans, was found in the recently developed simple generalization of the well-known MLP network called Sigma-if [4, 26].

The remainder of this paper is organized as follows. Section 2 presents the concept and basic properties of selective attention systems as well as main results of to-date research in the field of selective attention modelling. Section 3 describes the Sigma-if neural network model, which realizes selective attention functionality. Section 4 illustrates how selective attention manifests itself in properties of the Sigma-if neural network on the example of selected UCI Machine Learning benchmark data sets. Finally, conclusions are presented in Section 5.

2 Previous work

In nature, selective attention is a mechanism which provides living organisms with the possibility to sift incoming data, to extract information which is most important at a given moment and which should be processed in detail [14, 27]. This mechanism is necessary due to limited processing capabilities of the nervous system which does not allow rapid analysis of the whole scene of visual and other senses [17, 20]. Selective attention can be viewed as a strategy of dynamical input space selection for gaining predefined goals by an organism interacting with a very complicated environment.

Potential benefits from adopting the selective attention mechanisms in domains such as data analysis and feature selection have led many researchers to search for appropriate models of this feature of the human brain [9, 10, 15, 16, 20, 28, 29]. The first widely known result was Broadbent's bottleneck model [30], but even after improvements by Treisman [31] it was found to be too inaccurate. Another important development was the late selection model by Deutsch and Norman [32] and its two-step version by Snyder and Posner [33]. These models were founded on strong evidence, through research of Lewis [34], Corteen [35], Neely [36] and Pynte [37]. Although late selection models were computationally very costly, they showed that fast, low-level automatic data selection is crucial for achieving effectiveness of slow contextual data analysis realized by higher organizational levels of neural structures [38–41]. On this basis, the most important achievement of neuropsychological studies arose - namely, the development of earlier works by Noton and Stark on saccadic eye movements [42–44]. Element of automatic, cyclic attentional field selection, according to learned scan paths, turned out to be common in almost all input channels of a human brain [45, 46].

Many models of scan path selection strategies were subsequently proposed - this included hierarchical neural routing and shifting circuits [10, 16, 47] as well as oscillating and adaptive neural networks [29, 48, 49]. They made it possible to build specialised tools, e.g. for face recognition, object tracking or indicating interesting objects for observation by space probes [50–53]. However, none of these solutions yielded a general model of low-level selective attention. In fact, most of them were very complicated and realised only centralised or predefined (i.e. not learned or evolved) feature selection strategies [15, 17, 18]. The degree of their specialisation and neurobiologically unfounded solutions induced researchers to look for basic selective attention mechanisms on the level of single synapses and neurons [54–57].

Neuropsychological and neurobiological studies of selective attention phenomena have led researchers to the conclusion that low-level selective attention functionality is realized in a distributed manner by serial processes that are carried out by one to three layers of neurons [19, 17, 20]. Simultaneously, selective attention at higher levels of brain structure organization seems to emerge as an effect of synergy between elementary structures at lower levels (e.g. neurons or groups of neurons). Unfortunately, the exact mechanisms that lie at the base of selective attention are still unknown.

This is why networks that use higher-order neuron models, such as Sigma-Pi [58–60], Power Unit [61] or Clusteron [62], realize only a very limited set of attentional mechanisms. Thus it can be very interesting that selective attention functionality which seems to effectively mimic low-level attentional processes observed in humans, was found in a recently-developed simple generalization of the well-known MLP network called Sigma-if [4, 26, 63].

3 The Sigma-if neural network

The proposed Sigma-if neural network is a type of fully connected, synchronous multilayer perceptron neural network (MLP) which possess selective attention abilities [26, 63]. Such a neural network does not use separate centralized attention guidance modules. Its ability to realize selective attention functionality emerges as an effect of synergy between its hidden, Sigma-if neurons. Each Sigma-if neuron is a special direct generalization of a sigmoidal perceptron which implements basic selective attention functionality via input connections grouping and stepwise conditional input signal accumulation. This is due to the new neuron's aggregation function [5, 26].

3.1 The Sigma-if neuron

The Sigma-if neuron, in contrast to a typical perceptron, aggregates input signals for the given data vector not in one but in a series of given K steps according to the corresponding state graph. Its input connections are divided into K discrete subsets during the training process. Subsequently, when the neuron's aggregation function value is computed, in every k-th step of this process, the subset of input signals X_k is taken from the environment and processed to determine the current value of the partial activation level φ_k of the neuron. The process continues until the value of φ_k exceeds a given aggregation threshold φ^*. When that condition is met, signals which were not analyzed are ignored, and φ_k is considered the input value for the neuron's activation function F. As a result, signal level information coding and even the use of a non-local activation function (e.g. sigmoid) do not degrade the neuron's selective attention abilities. A sample scheme of a state graph for a Sigma-if neuron is presented in Fig. 1.

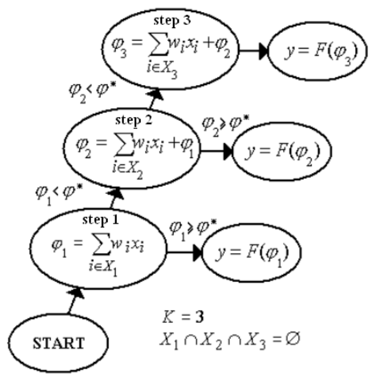

Fig. 1. Sample process of a three-step input signal aggregation in the Sigma-if neuron

Formally speaking, M dendrites of the Sigma-if neuron are divided into K distinct groups, by complementing each i-th input connection with an additional integer parameter $\theta_i \in \{0,1,..,K-1\}$, determining membership in one of the groups. This allows us to divide the process of signal accumulation into K steps, where K is a function of the neuron's grouping vector $\theta^T = [\theta_1, \theta_2, ..., \theta_M]$:

$$K(\theta) = \max_{i=1}^{M}(\theta_i). \tag{1}$$

During each step k (from 0 to K-1), the neuron accumulates data belonging to one selected group, such that

$$\theta_i = k. \tag{2}$$

Within each k-th group, partial activation $\Delta\varphi(k)$ is determined as a weighted sum of input signals and the appropriate Kronecker's delta:

$$\Delta\varphi_k(w, x, \theta) = \sum_{i=1}^{M} w_i x_i \delta(k, \theta_i), \tag{3}$$

where w_i and x_i are coefficients of the neuron's weight vector w and an input vector x. This process is repeated until the activation derived from respective groups exceeds a preselected aggregation threshold φ^*. It can be described by the following recursive formula (vectors w, x and θ are omitted for clarity):

$$\varphi_k = \begin{cases} \Delta\varphi_k H(\varphi^* - \varphi_{k-1}) + \varphi_{k-1} : k \geq 0 \\ 0 \qquad\qquad\qquad\qquad\quad : k < 0 \end{cases} \tag{4}$$

where H is Heaviside's function. This sum is then treated as the neuronal activation value. The input from remaining (heretofore unconsidered) groups is neglected. Thus, the proposed form of the aggregation function A is:

$$A(w, x, \theta) = \varphi_K(w, x, \theta). \tag{5}$$

In the final stages of determining the output value Y of the neuron, function (5) serves as a parameter of the nonlinear threshold (e.g. sigmoidal) function F:

$$Y(w, x, \theta) = F(A(w, x, \theta)). \tag{6}$$

It is worth noting that the described model assumes that the state graph used during signal aggregation is always a simple directed path of nonterminal nodes corresponding with the neural activation accumulation procedure. In a general case, the Sigma-if neuron, besides the weights vector w, includes one continuous valued parameter for aggregation threshold φ^*, and an additional connections grouping vector θ with only one nominal valued coefficient for each neuronal input connection.

3.2 Sigma-if network training

In comparison to MLP neural network training, searching for a globally optimal set of Sigma-if network parameters would be very computationally challenging. This is due to the noncontinuous character of Sigma-if neuron grouping vectors.

While there is no quick and effective method for global searching of network weights and grouping vectors, the proposed solution assumes that at each Sigma-if neuron, coefficients of the grouping vector θ are in fact direct functions of that weight vector. As a result, network connection weights are established by the well known error backpropagation algorithm, but for every ω training epoch, actual grouping vectors are computed. This reflects the application of the self-consistency idea widely used in physics.

In this work, the grouping vector computation procedure (as well as the predefined value of the aggregation threshold φ^*) is common for all Sigma-if neurons. It simply divides input connections into a given number of groups, according to "the greater the connection weight, the smaller the connection group number" principle. Such a search problem reduction leads us to very interesting results, and to practical elimination of the additional θ parameters vector. However, during further analysis of the general Sigma-if model, it is still very helpful to use the grouping vector concept.

4 Properties of the Sigma-if neural network

During examination of the proposed model properties, Sigma-if neural networks were compared to MLP networks with the same architectures. All networks were fully connected and had one hidden layer with the number of neurons for which the MLP network gained highest test data classification accuracy. As the sigmoidal perceptron is a special case of a Sigma-if neuron, MLP networks were in fact simulated by Sigma-if networks with the number of inputs groups K of all Sigma-if neurons set to one. This was consistent with the additional assumption that K and aggregation threshold φ^* values of all hidden Sigma-if neurons in the given network were equal. In all cases, standard input signal coding was used, and answers of the network were computed in the winner-takes-all manner.

Beside classification accuracies for training (u) and test (γ) data, properties such as network data processing time (τ), as well as hidden connections and network input activity (*hca* and *nia* respectively) were considered. The data processing time τ for all trained networks was measured on the same computer (P4 2.6GHz) to check relative data processing costs for MLP and Sigma-if networks. Regardless of the very precise time measurement procedure used (following measurement of real processor core frequency, the number of processor clock cycles was counted, starting from the end of loading data into network inputs and finishing with the end of the network output value generation), actual timings on other hardware setups may vary considerably.

Hidden connection activity (*hca*) and network input activity (*nia*) were used to representing the percentage ratio of the number of hidden and input connections used during data processing, compared to all of the network's hidden and input connections respectively. These parameters allowed us to check if hidden Sigma-if neurons used their selective attention ability in practice. For the completeness of analysis, for each given problem and trained network, the percent of all inputs used to classify all test vectors (*niu*) was calculated. This procedure was important in order to determine if selective attention functionality is also realized at the level of the whole Sigma-if network. All presented values for all considered UCI Machine Learning Repository benchmark problems were calculated as averaged outcomes of ten independent 10-fold cross validations.

The conducted experiments can be divided into two groups: examination of Sigma-if network classification properties and verification of the hypotheses that Sigma-if networks realize selective attention at the level of single neurons and of the whole network. In both cases all of the considered properties were analyzed in relation to the number of Sigma-if neuron inputs groups K. This was because the K value has the highest influence on the properties of the proposed network. Other parameters, such as the aggregation threshold φ^* and the grouping vector actualization interval ω, were set to 0.6 and 25 respectively, following preliminary tests. It is also important to note that parameters such as *hca*, *nia* and *niu* are informative only for Sigma-if networks; for MLP networks ($K = 1$) they are, by default, equal to 100%.

The obtained results indicate that increasing the number of Sigma-if neuron input groups K to more than one results in an increase of test data classification accuracy γ as well as in simultaneous decrease of data processing time τ. The drawback here is a decrease of training data classification accuracy u. A typical example of such a dependency can be observed for the HeartC problem, which is presented in Fig. 2. The observed decrease in training data classification ac-

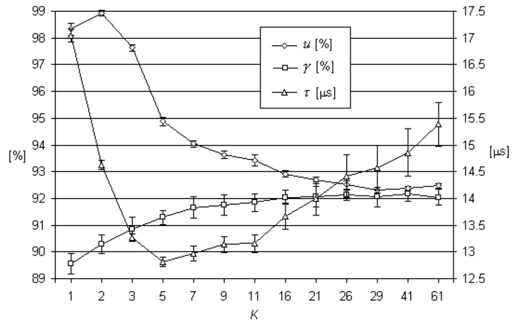

Fig. 2. The time of Sigma-if network output signal generation (τ), the classification accuracy of training (u) and test (γ) data for the HeartC problem versus the number of hidden neuron input connections groups K (networks architecture: 28 inputs, 10 hidden neurons, 5 outputs)

curacy u was most probably due to the fact that Sigma-if neurons process data in the same way after and during training. It is harder to learn when the neuron's input space is changed every ω epoch. However, and more importantly, the obtained increase in test data classification accuracy γ is a result of rejecting redundant or noisy signals from processed data and the consequence of effective reduction of problem dimensionality. This thesis is confirmed by the observed simultaneous decrease in data processing time τ. Reduction of τ can only be caused by reduction of the network's hidden and input connections activities. However, regardless of the reasons, these results shows that the Sigma-if neural network has better classification properties than MLP.

The visible increase of HeartC data processing time τ for K above 7 inputs groups is the effect of a linear increase of time cost, connected with the existence of additional instructions for grouping vector θ information processing. This factor can be easily seen for the number of groups K greater than the given number of network inputs. Without it, data processing time would semi-logarithmically decrease with rising K. This indicates the character of the changes of Sigma-if network hidden (hca) and input connection activities (nia) as a function of K, which can be observed in Fig. 3 (for the Sonar problem) and in Fig. 4 (for the Votes problem).

It can be easily seen that in the case of the Sonar problem the increase in K causes an increase of the test data classification accuracy, with a corresponding decrease of the training data classification accuracy. These changes are accompa-

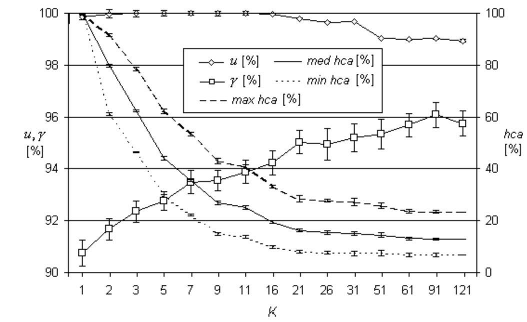

Fig. 3. The Sigma-if network hidden connection activity (hca), the classification accuracy of training (u) and test (γ) data for the Sonar problem versus the number of hidden neuron input connections groups K (network architecture: 60 inputs, 30 hidden neurons, 2 outputs)

nied with a much stronger reduction of hidden connection activities. The shape of the *hca(K)* function confirms earlier conclusions that the data processing time reduction is connected with Sigma-if neurons' selective attention abilities. All this is clear evidence that Sigma-if neurons use selective attention, and that this can reduce the generalization error level as well as data processing costs.

The last example concerns how Sigma-if selective attention abilities manifest themselves on the level of the whole network. The analysis of results for the Votes problem (Fig. 4) shows that when a significant decrease of network input activity (nia) occurs, one can expect a simultaneous reduction in the number of Sigma-if network inputs used to classify data (niu) without a notable decrease of classification accuracy in comparison to the MLP network. While the Sigma-if

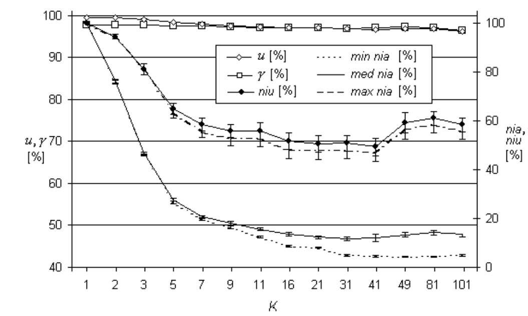

Fig. 4. The Sigma-if network input activity (nia), the number of network input used (niu), and the classification accuracy of training (u) and test (γ) data for the Votes problem versus the number of hidden neuron input connections groups K (network architecture: 48 inputs, 2 hidden neurons, 2 outputs)

network has no specialized or separate attention guiding unit, all such activities can emerge only as an effect of synergy between individual neurons. Thus, the observed selective attention behaviour of the proposed network, treated as a black box, is a significant indication that the Sigma-if model effectively mimics low level attentional processes observed in nature. This can, in turn, make the model an interesting tool for feature selection and other data processing purposes.

5 Conclusion

In this paper the Sigma-if neuron model and benefits from using the Sigma-if neural network instead of the MLP network were considered. The model's selective attention ability for medium-size test problems manifests itself in an increase of classification accuracy and in a simultaneous decrease of data processing costs. Additional reduction of the number of network inputs used to classify data hints at the possibility of further reduction of data collection costs, during as well as after the training process.

The Sigma-if selective attention feature also introduces new possibilities in the area of analyzing the network decision process via its input activity interpretation. This can point at features of given data sets that are most important for classification, and help identify features that are irrelevant, redundant or contaminated by noise.

What is more important, the Sigma-if neural network seems to be a promising model of low-level human selective attention. This functionality is realized in a distributed manner by sequential input signal accumulation, carried out by Sigma-if neurons, and emerges as an effect of synergy among the network's hidden neurons.

All this makes the Sigma-if neural network a very useful tool for the data analysis domain, despite the fact that the proposed method requires further tests on benchmark and real-life data. New Sigma-if network training methods as well as new selective attention neurons with aggregation functions other than those presented in this paper will be the subject of further research.

References

1. Anthony, M., Bartlett, P.: Neural networks learning: Theoretical foundations. 1st ed. edn. Cambridge University PressCambridge University Press, Cambridge (1999)
2. Srinivasan, T., Chandrasekhar, A., Seshadri, J., Siddharth, J.: Knowledge discovery in clinical databases with neural network evidence combination. In: Second IEEE International Conference on Intelligent Sensing and Information Processing - ICISIP 2005, Chennai, India, IEEE Press, (2005) 512–518
3. LiMin, F.: Knowledge discovery based on neural networks. Commun. ACM **42** (1999) 47–50

4. Huk, M.: The sigma-if neural network as a method of dynamic selection of decision subspaces for medical reasoning systems. Journal of Medical Informatics & Technologies **7** (2004) 65–73
5. Duch, W., Jankowski, N.: Survey of neural transfer functions. Neural Computing Surveys **2** (1999) 163–212
6. Duch, W., Adamczak, R., Diercksen, G.: Neural networks in non-euclidean metric spaces. Neural Processing Letters **10** (1999) 201–210
7. Chowdhury, F., Wahi, P., Raina, R., Kaminedi, S.: A survey of neural networks applications in automatic control. In: Southeastern Symposium on System Theory. (2001) 349–353
8. Zhao, Y., Gao, J., Yang, X.: A survey of neural network ensembles. In: ICNN&B'05. International Conference on Neural Networks and Brain. Volume 1. (2005) 438–442
9. Desimone, R., Duncan, J.: Neural mechanisms of selective visual-attention. Annual Review of Neuroscience **18** (1995) 193–222
10. Olshausen, B., Anderson, C., Van Essen, D.: A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. J. Neurosc. **13** (1993) 4700–4719
11. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: Knowledge discovery and data mining: Towards a unifying framework. In: 2nd Intl. Conf. on Knowledge Discovery and Data Mining (KDD96), Portland, Oregon (1996) 82–88
12. Yu, H., Yang, J., Han, J.: Classifying large data sets using svm with hierarchical clusters. In: 2003 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'03), Wash., D.C. (2003)
13. Fayyad, U., Djorgovski, S., Weir, N.: Automating the analysis and cataloging of sky surveys. In: Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press (1996) 471–493
14. Broadbent, D.: Task combination and selective intake of information. Acta Psychologica **50** (1982) 253–290
15. Heinke, D., Humphreys, G.: Computational Models of Visual Selective Attention: A Review. In: Connectionist Models in Psychology. Psychology Press (2002)
16. Koch, C., Ullman, S.: Shifts in selective visual attention: towards the underlying neural circuitry. Hum. Neurobiology **4** (1985) 219–227
17. Tsotsos, J., Culhane, S., Cutzu, S.: From foundational principles to a hierarchical selection circuit for attention. In: Visual Attention and Cortical Circuits. MIT Press, Cambridge, MA (2001) 285–306
18. Privitera, C., Stark, L.: Algorithms for defining visual region-of-interest: Comparison with eye fixations. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 970–982
19. Spratling, M., Hayes, G.: Learning synaptic clusters for nonlinear dendritic processing. Neural Processing Letters **11** (2000) 17–27
20. VanRullen, R., Koch, C.: Visual selective behavior can be triggered by a feedforward process. Journal of Cognitive Neuroscience **15** (2003) 209–217
21. Anderson, R., Essich, G., Siegal, R.: Encoding of spatial location by posterior parietal neurons. Science **230** (1985) 456–458
22. Karlholm, J.: Associative memories with short-range, higher order couplings. Neural Networks **6** (1993) 409–421
23. Stuart, G., Spruston, N.: Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. Journal of Neuroscience **18** (1980) 3501–3510
24. Larkum, M., Zhu, J., Sakmann, B.: A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature **398** (1999) 338–341

25. Körding, K., König, P.: Neurons with two sites of synaptic integration learn invariant representations. Neural Computation **13** (2001)) 2823–2849
26. Huk, M., Kwasnicka, H.: The concept and properties of sigma-if neural network. In Ribeiro, B., Albrecht, R., Dobnikar, A., Pearson, D., Steele, N., eds.: Adaptive and Natural Computing Algorithms, Springer Computer Science (2005) 13–17
27. Treisman, A.: Contextual cues in selective listening. Quarterly Journal of Experimental Psychology **12** (1960) 242–248
28. Xiao-Ping, Z.: Thresholding neural network for adaptive noise reduction. IEEE Transactions on Neural Networks **12** (2001) 567–584
29. Houghton, G., Tipper, S.: Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. Brain and Cognition **30** (1996) 20–43
30. Broadbent, D.: Perception and Communication. Pergamon Press, Oxford (1958)
31. Treisman, A.: Selective attention in man. British Medical Bulletin **20** (1964) 12–16
32. Norman, D.: Toward a theory of memory and attention. Psychological Review **75** (1968) 522–536
33. Posner, M., Snyder, C.: Facilitation and inhibition in the processing of signals. In: Attention and Performance. Academic Press, London (1975)
34. Lewis, J.: Semantic processing of unattended messages using dichotic listening. Journal of Experimental Psychology **85** (1970) 220–227
35. Corteen, R., Wood, B.: Autonomic responses to shock-associated words in an unattended channel. Journal of Experimental Psychology **94** (1972) 308–313
36. Neely, J.: Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. Journal of Experimental Psychology: General **106** (1977) 226–254
37. Pynte, J., Do, P., Scampa, P.: Lexical decisions during the reading of sentences containing polysemous words. In: Preparatory States and Processes. Erlbaum, Hillsdale NJ (1984)
38. Sperling, G.: A unified theory of attention and signal detection. In: Varieties of attention. Academic Press, New York (1984) 103–181
39. Dosher, B., Sperling, G.: Strategy optimization in human information processing. In: Handbook of perception and human performance: Sensory processes and perception. Volume 1. John Wiley and Sons, New York (1986) 1–65
40. Neumann, O.: Beyond capacity: A functional view of attention. In: Perspectives on perception and action. Lawrence Erlbaum Associates, Hillsdale, NJ (1987)
41. Posner, M., S.E., P.: The attention system of the human brain. Ann. Rev. Neuro. **13** (1990) 25–42
42. Noton, D., Stark, L.: Scanpaths in saccadic eye movements while viewing and recognizing patterns. Vision Research **11** (1971) 929–942
43. Findlay, J.: Global visual processing for saccadic eye movements. Vision Research **22** (1982) 1033–1046
44. Chelazzi, L., Miller, E., Duncan, J., Desimone, R.: A neural basis for visual search in inferior temporal cortex. Nature **363** (1993) 345–347
45. Tsal, Y.: Movements of attention across the visual field. Journal of Experimental Psychology: Human Perception and Performance **9** (1983) 523–530
46. Kastner, S., De Weerd, P., Desimone, R., Ungerleider, L.: Mechanisms of directed attention in the human extrastriate cortex as revealed by functional mri. Science **282** (1998) 108–111
47. Anderson, C., Van Essen, D.: Shifter circuits: a computational strategy for dynamic aspects of visual processing. Proc. Natl. Academy Sci. USA **84** (1987) 6297–6301

48. Niebur, E., Koch, C., Rosin, C.: An oscillation-based model for the neural basis of attention. Vision Research **33** (1993) 2789–2802
49. Tsotsos, J., Culhane, S., Wai, W., Lai, Y., Davis, N., Nuflo, F.: Modeling visual attention via selective tuning. Artificial Intelligence **78** (1995) 507–547
50. Yamada, K., Cottrell, G.: A model of scan paths applied to face recognition. In: 17th Ann. Cog. Sci. Conf., Pittsburgh (1995)
51. Rybak, I., Gusakova, V., Golovan, A., Podladchikova, L., Shevtsova, N.: A model of attention-guided visual perception and recognition. Vision Research **38** (1998) 2387–2400
52. Hager, G., Toyama, K.: Incremental focus of attention for robust visual tracking. International Journal of Computer Vision **35** (1999) 45–63
53. Privitera, C., Azzariti, M., Stark, L.: Locating regions-of-interest for the mars rover expedition. JRS **21** (2000) 3327–3347
54. Mel, B.: Why have dendrites? A computational perspective. In: Dendrities. Oxford University Press (1999) 271–289
55. Baldassi, S., Verghese, P.: Comparing integration rules in visual search. Journal of Vision **2** (2002) 559–570
56. Shimozaki, S., Eckstein, M., Abbey, C.: Comparison of two weighted integration models for the cueing task: linear and likelihood. Journal of Vision **3** (2003) 209–229
57. Hamker, F.: Predictions of a model of spatial attention using sum- and max-pooling functions. Neurocomputing **56** (2004) 329–343
58. Feldman, J., Ballard, D.: Connectionist models and their properties. Cognitive Science **6** (1982) 205–254
59. Rumelhart, D., Hinton, G., McClelland, J.: A general framework for parallel distributed processing. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations. Volume 1. The MIT Press (1986) 45–76
60. Mel, B.: The sigma-pi column: a model of associative learning in cerebral cortex. Technical report, CNS Memo 6, Computation and Neural Systems Program, California Institute of Technology (1990)
61. Durbin, R., Rumelhart, D.: Product units: A computationally powerful and biologically plausible extension to backpropagation networks. Neural Computation **1** (1990) 133–142
62. Mel, B.: The clusteron: toward a simple abstraction for a complex neuron. In: Advances in Neural Information Processing Systems. Volume 4. Morgan Kaufmann (1992) 35–42
63. Huk, M.: Sigma-if neural network as a use of selective attention technique in classification and knowledge discovery problems solving. Annales UMCS Informatica AI **5** (2006) 121–131