Philipp Wolfrum

Information Routing, Correspondence Finding, and Object Recognition in the Brain

Studies in Computational Intelligence, Volume 316

Editor-in-Chief

Prof. Janusz Kacprzyk Systems Research Institute Polish Academy of Sciences ul. Newelska 6 01-447 Warsaw Poland *E-mail:* kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 294. Manuel F.M. Barros, Jorge M.C. Guilherme, and Nuno C.G. Horta Analog Circuits and Systems Optimization based on Evolutionary Computation Techniques, 2010 ISBN 978-3-642-12345-0

Vol. 295. Roger Lee (Ed.) Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2010 ISBN 978-3-642-13264-3

Vol. 296. Roger Lee (Ed.) Software Engineering Research, Management and Applications, 2010 ISBN 978-3-642-13272-8

Vol. 297. Tania Tronco (Ed.) New Network Architectures, 2010 ISBN 978-3-642-13246-9

Vol. 298. Adam Wierzbicki Trust and Fairness in Open, Distributed Systems, 2010 ISBN 978-3-642-13450-0

Vol. 299. Vassil Sgurev, Mincho Hadjiski, and Janusz Kacprzyk (Eds.) Intelligent Systems: From Theory to Practice, 2010 ISBN 978-3-642-13427-2

Vol. 300. Baoding Liu (Ed.) Uncertainty Theory, 2010 ISBN 978-3-642-13958-1

Vol. 301. Giuliano Armano, Marco de Gemmis, Giovanni Semeraro, and Eloisa Vargiu (Eds.) Intelligent Information Access, 2010 ISBN 978-3-642-13999-4

Vol. 302. Bijaya Ketan Panigrahi, Ajith Abraham, and Swagatam Das (Eds.) *Computational Intelligence in Power Engineering*, 2010 ISBN 978-3-642-14012-9

Vol. 303. Joachim Diederich, Cengiz Gunay, and James M. Hogan Recruitment Learning, 2010 ISBN 978-3-642-14027-3

Vol. 304. Anthony Finn and Lakhmi C. Jain (Eds.) Innovations in Defence Support Systems, 2010 ISBN 978-3-642-14083-9 Vol. 305. Stefania Montani and Lakhmi C. Jain (Eds.) Successful Case-Based Reasoning Applications-1, 2010 ISBN 978-3-642-14077-8

Vol. 306. Tru Hoang Cao Conceptual Graphs and Fuzzy Logic, 2010 ISBN 978-3-642-14086-0

Vol. 307. Anupam Shukla, Ritu Tiwari, and Rahul Kala Towards Hybrid and Adaptive Computing, 2010 ISBN 978-3-642-14343-4

Vol. 308. Roger Nkambou, Jacqueline Bourdeau, and Riichiro Mizoguchi (Eds.) *Advances in Intelligent Tutoring Systems*, 2010 ISBN 978-3-642-14362-5

Vol. 309. Isabelle Bichindaritz, Lakhmi C. Jain, Sachin Vaidya, and Ashlesha Jain (Eds.) *Computational Intelligence in Healthcare 4*, 2010 ISBN 978-3-642-14463-9

Vol. 310. Dipti Srinivasan and Lakhmi C. Jain (Eds.) Innovations in Multi-Agent Systems and Applications – 1, 2010

ISBN 978-3-642-14434-9

Vol. 311. Juan D. Velásquez and Lakhmi C. Jain (Eds.) Advanced Techniques in Web Intelligence, 2010 ISBN 978-3-642-14460-8

Vol. 312. Patricia Melin, Janusz Kacprzyk, and Witold Pedrycz (Eds.) Soft Computing for Recognition based on Biometrics, 2010 ISBN 978-3-642-15110-1

Vol. 313. Imre J. Rudas, János Fodor, and Janusz Kacprzyk (Eds.) *Computational Intelligence in Engineering*, 2010 ISBN 978-3-642-15219-1

Vol. 314. Lorenzo Magnani, Walter Carnielli, and Claudio Pizzi (Eds.) *Model-Based Reasoning in Science and Technology*, 2010 ISBN 978-3-642-15222-1

Vol. 315. Mohammad Essaaidi, Michele Malgeri, and Costin Badica (Eds.) *Intelligent Distributed Computing IV*, 2010 ISBN 978-3-642-15210-8

Vol. 316. Philipp Wolfrum Information Routing, Correspondence Finding, and Object Recognition in the Brain, 2010 ISBN 978-3-642-15253-5 Philipp Wolfrum

Information Routing, Correspondence Finding, and Object Recognition in the Brain



Dr. Philipp Wolfrum Siemens AG Corporate Technology Otto-Hahn-Ring 6 81739 Munich Germany E-mail: wolfrum@fias.uni-frankfurt.de

ISBN 978-3-642-15253-5

e-ISBN 978-3-642-15254-2

DOI 10.1007/978-3-642-15254-2

Studies in Computational Intelligence

ISSN 1860-949X

Library of Congress Control Number: 2010934961

© 2010 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

 $9\ 8\ 7\ 6\ 5\ 4\ 3\ 2\ 1$

springer.com

Acknowledgements

This book would not have been possible without the people I have had the honor of working with over the past years. Major parts of the book originate from my doctoral studies at the Frankfurt Institute for Advanced Studies. My first thanks therefore go to my academic advisor, Christoph von der Malsburg. His scientific ideas have inspired my research and had a strong influence on the concepts developed in this book. Discussions with Geoff Goodhill, Steve Grossberg, Bruno Olshausen, Rudolf Mester, Jochen Triesch, and Alan Yuille, among others, have shaped and widened my thinking. I'm sure you will find the traces of our discussions in here! I also thank all colleagues at FIAS for good company and the truly interdisciplinary interaction we had. Special thanks go to the members of our group, Urs Bergmann, Jenia Jitsev, Jörg Lücke, and Junmei Zhu. Last but not least I owe thanks to my parents, who planted curiosity and a desire for understanding in me that ultimately lead to this book.

Foreword

At first sight, this book is about face recognition in the brain. Its more lasting value, however, lies in the paradigmatic way in which this particular problem is treated. From the basic ideas that are worked out here in concrete detail, it is a natural and simple next step to at least imagine, if not realize in model form, much more general structures and processes, thus helping to bridge the still tremendous chasm between mind and brain. It is the purpose of this foreword to point out these generic traits.

For centuries, thinking about the brain has been dominated by the most complex mechanistic devices of the time, clockwork, communicating hydraulic tubes or, today, the computer. The computer, taken as incarnation of the Universal Turing Machine, can implement any conceivable process, so that also a functional brain can surely be simulated on it, an idea that, beginning in the fifties of the last century, has been seducing scientists to create "artificial intelligence" in the computer. As a result we now have an information technology that displays many functional capabilities formerly regarded as the exclusive domain of the mind. As fascinating as this is, doting on "intelligent machines" is systematically diverting our attention away from the true problems of understanding the working of the brain.

"Intelligence in the machine" is, of course, in truth canned human intelligence. The computer itself is without a clue, behaving like a blind man guided by a distant voice step by step over a tight-rope. It is humans that are exploiting their own insight into complex processes to endow the computer with algorithms, reaction patterns that cover those and only those situations the programmer has thought of ahead of time. From the technological point of view this would be fine, if the complexity of it all weren't beginning to grow out of hand. Little noticed by the general public, for a decade or two there has been growing awareness among the experts of a "software crisis" that is bugging information technology. Large software projects have an alarmingly high likelihood of complete failure, and if not failure, of time and cost overruns and an abundance of functional deficits. The venture of artificial intelligence has completely underestimated the complexity of such processes as language or motion control or vision, and it is evident by now that no reasonable effort can generate such capabilities directly in algorithmic fashion. Realizing this, instead of running the computer with borrowed human intelligence we should rather concentrate on the organization processes that create functional structures in the child's brain (or indeed in the programmer's mind), implement them on our computers and let them do the rest of the work.

The basis for the programmer's ability to create algorithms is creative infrastructure comprising world knowledge, sample structures, methodology as well as the ability to project goals, diagnose mistakes or interpret the symbols handled in the machine. It is this infrastructure that enables humans to formulate abstract goals and to translate these abstract goals into concrete process descriptions. Understanding this infrastructure will enable us to understand mind and brain, and it will be the basis for another profound technological revolution.

As misleading as the computer is when using human intelligence as a shortcut to problem solving, as instrumental can it be as a tool to model the processes going on in the brain. Once we understand creative mind processes, the computer, being Turing-universal, will be able to emulate them. Information technology may then retract from the details in the machine, to concentrate fully on the formulation of our human goals and intentions. This will not only liberate humans from programming chores, setting free their creativity in terms of goal formulation, but will also liberate the computing medium from its leash and let it deal with challenges as they arise, meeting them by reference to the abstract goals we set instead of to distant human foresight.

Trying to model mind processes in the computer is a powerful tool to study the fundamental issues of the neural and cognitive sciences. It is a false expectation that experiments can force functional ideas on us. Experimentation needs guidance by hypotheses and beliefs. Now, it is a sad observation that many highly convincing functional ideas, some of them the inspiration of extensive experimental paradigms, turn out to be flawed when tried on the computer. It therefore is a wise strategy to develop functional computer models with generic brain features, and to demonstrate their functional viability "experimentally". They will powerfully influence thinking about the brain and, given time, will inspire neuroscientific experiments, just as Rosenblatt's perceptron once inspired Hubel and Wiesel's experimental paradigm.

This book, essentially the PhD thesis of its author, is focused on one such function, the invariant representation and recognition of visual structure. It brings together a number of threads that have been developed over decades in my lab. The model that is developed in these pages offers conceptual solutions to a number of fundamental issues regarding brain function. Among these are, first and foremost, the four cognitive architecture questions — what are the data structures of the brain on the fast time scale of brain state and on the slow time scale of memory; and, what are the mechanisms of organization, again on both time scales, brain state and memory? To this day, the neurosciences are dominated by a certain set of answers to these questions, beginning with a very specific idea regarding the data structure of brain state. According to this, neurons are taken as elementary symbols (blue light in position x, an edge of orientation ϕ in position y, muscle fiber z to be twitched etc.), and the brain's state in a given moment is thought to be completely described by the vector of neural activities. Further, brain states are organized by exchange of excitation and inhibition between neurons, regulated such that they evolve slowly from quasi-stationary state to quasi-stationary state, each lasting for a tenth of a fifth of a second. The data structure of memory has the form of strengths of neural connections (synaptic weights), and the mechanism of organization of memory is constituted by different forms of synaptic plasticity.

Modelling of brain functions, a very active field in recent decades, has almost exclusively been based on this conceptual framework, and although a great wealth of results has sprung forth from this effort, one cannot help to suspect that progress is held back by some conceptual roadblock. In the mid-eighties the enthusiasm of the field of artificial neural networks reached the sky, claiming that in a few years' time information technology would be revolutionized and the brain understood, but although activity levels in neural theory and such related fields as machine learning, Bayesian estimation or robotics are higher than ever today, those great aspirations have not materialized, no one daring to come forth with predictions when they will. In some quarters, one might even sense a feeling of depression and rising doubt back-engineering of the brain could be possible in a few decades, given that evolution had taken so long to generate it.

There is, however, reason to be optimistic. Cerebral cortex, our cognitive organ, is estimated to contain 5×10^{10} neurons, and each neuron to have 10^4 synaptic connections. It would take about 10^{16} bits (or one PetaByte) to describe a list of these connections. In contrast, there are only about 10^9 bits of information in the human genome (only a small part of which can be assumed to code the brain's wiring diagram). Also a lifetime of learning is not enough to close the gap. Generously counting 30 years of growing up, which amounts to 3×10^8 seconds of waking time, and generously assuming the brain to absorb 300 bits of sensory information each second, only 10^{11} bits of information could come from that source. This leaves us short by a factor of 10^5 of being able to account for the information content of the list. The conclusion of this little calculation is that to an overwhelming extent the connectivity patterns of the brain are dominated by regularity. (Random connectivity is no alternative, in spite of the neuroanatomist's impression under the microscope, and in spite of some of it being invoked by certain theorists, as coordination between different brain structures would be severely hampered by ubiquitous independent random decisions.) Part of this regularity is due to spatial organization, neurons connecting mainly within a relatively small neighborhood. If each synapse has only to choose from 10^n neurons in its environment (with n likely to be between 4 and 5) instead of all neurons in cortex, the information requirement is reduced from 10^{16} bits to $n \times 10^{15}$ bits, still leaving a tremendous gap. The tremendous remaining amount of structural regularity is the hunting ground of brain theory.

The most promising approach to closing the information gap is trying to understand the process of development of the nervous system. Studying this is a very active field, and has been for over a hundred years. In distinction to our human technology, where the manufacture of artefacts starts with a blueprint of the whole, and parts are shaped and assembled by outside forces, life constructs its organisms inside out, the movers being the component molecules and cells, the final Gestalt emerging gradually in the process in a progression from coarse to fine. This development is all driven by the behavioral repertoire of individual cells, inherited, refined and differentiated in a continuous line from our single-celled ancestors. Cells divide and multiply, differentiate, shape themselves, move, adhere selectively to each other, put out extensions and communicate with each other by sending and receiving molecular, mechanical and electrical signals in ever-richer patterns. Thus, growth of the embryo, of the brain, of the detailed connectivity of the nervous system, is a complex dance of cells, choreographed from moment to moment by the molecular network of genetic regulation. One way of looking at it is to see the organism as a party joyfully celebrated by its population of cells.

In distinction to human artefacts, which are totally passive during design and assembly and start to perform only after being finished, the body, nervous system and molecular signalling web of the organism are functioning from the outset, and this function is essential for shaping development and even design. Development progresses from simple to complex in a process of gradual differentiation, with more and more detail emerging in stages. The enormous number of degrees of freedom that are generated in this process are dealt with as they arise, system dynamics being regulated such that it is of contracting nature, an aspect that Waddington called channeling, such that most variables tend towards set points, annihilating initial and any accidental deviations. Only a very small set of variables is unstable, with a tendency to grow, leaving never more than a few choices, and it is at these branching points that the genes exert their essential influence. In a breath-taking development over the last two decades molecular biology has discovered a set of genes that to a large extent is common to all animals, from worm to human. The experts speak of the ontogenetic toolkit, the piano on which evolution lighthandedly plays animal tunes.

When it comes to the nervous system, after development of the neural tube and its various bulbar swellings, cells differentiate into neurons and they begin to grow processes with which to connect up with each other, starting the game of network self-organization that will continue into adulthood. Initially, cells reach out with their processes into their neighborhood, but then they send them out to more distant targets, guided by molecular signals that themselves are created by other cells. These signals form complex patterns that are shaped by the growing network and in turn are shaping the network. As result of this feedback, networks converge towards patterns that are stabilized by the very signals they put out themselves. Over the last three decades, some prime examples for this process have been studied intensively, prominent among them the establishment of neighborhood-preserving, retinotopic, fiber projections between eye and brain.

Network self-organization, probably the most pervasive and fundamental process by which Life gives itself shape from molecular to societal levels, is in some respects akin to crystallization. A crystal is shaped by the preferences that individual molecules have concerning their immediate neighborhood. Just as these local preferences conspire in a crystal to create long-range order, local preferred network arrangements conspire to form globally ordered networks. Local preferences may differ in detail, depending on molecular signals and genetic influences, but dominant among them are likely to be a tendency towards sparsity (relatively few incoming and outgoing connections per neuron) and cooperativity between alternate pathways (such as the pathways $A \rightarrow C$ and $A \rightarrow B \rightarrow C$, or the pathways $A \rightarrow B \rightarrow C$ and $A \rightarrow B' \rightarrow C$). During network self-organization, spatial neighborhood relationships, like within retina or within cortex, will play a potent influence, evident in the relative abundance of short-range connections. However, a neuron's neighbors in the network sense are all those other neurons to whom it is connected — and these may be at any distance in the spatial sense. Many aspects and details of self-organized networks therefore may not be evident to anatomists looking through the microscope, the reason why they speak of random connectivity. They may be random in the spatial sense, but they aren't in the sense of connectivity.

It is a very interesting and important question how the attractor states of network self-organization, the connectivity patterns that dominate the structure of the nervous system, may be characterized in terms of their global structure. There may be no general answer to this question, or at least not any time soon. However, it is possible to characterize an important class of attractor networks. Imagine a two- or three-dimensional space populated by neurons with short range connections in this space. Such a network will produce signal ensembles that are dominated by short-range correlations, which in turn will favor short-range connections. It is therefore natural to expect that such "topological networks" are attractor states. In spite of its simple construction, the universe of such networks is complex and interesting, as many different topological networks can be created on the same total set of neurons, each individual network having as support a different subset of all neurons, supporting a combinatorially rich family of networks.

Network self-organization takes place on many temporal scales, the evolutionary growth of gene-regulatory networks being the slowest (although it may happen under the biologist's eye when studying bacterial evolution). Much faster are the processes of network self-organization during ontogenesis and learning. A taste of this can be found in Chapter 5 of this book, where the ontogenesis of the so-called switchyard circuits is described as a process of network self-organization.

At the heart of this book, however, is the hypothesis that network selforganization takes place also on the psychological time scale of about a tenth of a second, and that rapidly switching connections form an integral part of the data structure of brain state, in importance on a par with neural activity switching, the classical neural data structure. According to this view, the permanent, physical connections formed by slow synaptic plasticity during ontogenesis or learning can be switched off or on again, in a matter of milliseconds. The activity state of a connection means two things, on the one hand, an active connection is conducting and transmits signals, on the other it expresses the statement that the two connected neurons are related to each other, that they form a composite symbol.

A symbol system that wants to be expressive has to exploit combinatorics, building up more complex symbols from more elementary ones. In order to be able to do so, the system has to have some kind of glue, by which elements can be attached to each other. All known symbol systems have this capability. Thus, by arranging elements spatially on paper we form characters out of strokes, words out of characters etc., or mathematical expressions out of elementary symbols. Neurons in the brain cannot move about to form composite symbols, and there correspondingly has to be another means of building up higher structures. This, the hypothesis goes, is done with the help of switching synapses.

In this book, two applications of this idea play an important role. To build up the representation of a piece of visual structure, textural elements in their two-dimensional spatial arrangement are attached to each other by activated links. Having this means of expression is particularly important in memory, where novel representations are to be fluently composed out of pre-existing structural pieces under the guidance by retinal input. The other application is the center-piece of a recognition mechanism based on point-topoint correspondences between incoming images and stored object models, as modelled in detail in Chapter 3.

If switching connections are an important aspect of the data structure of brain state there must be a mechanism to organize and control these variables. This can be achieved by a rapid, reversible kind of Hebbian plasticity: the synapse between two neurons with signals that are correlated on a fine time scale is activated (or stabilized in its active resting state), whereas the synapse is switched off temporarily if the two neurons are both active but uncorrelated in their activity. On this basis, a process of network self-organization can take place, by which, for instance, the set of links that establish the correspondence between image and model can be activated or the representation of a complex scene can be glued together in a structured way.

The *de novo* organization of a complex network with many connections is a complicated affair, and if it was to be based entirely on signal correlationcontrolled synaptic plasticity it would be relatively slow, for two reasons. On the one hand, neural signals have low bandwidth and the temporal resolution of signal coincidences is surely not better than one millisecond. When many connections are to be sorted into the active or passive state, it takes a large number of time slices of that width to make that distinction for them all. On the other hand, as the signal correlations themselves are to be shaped by the network, they are confused during the early stages of network organization and a longish process of gradual differentiation is required to reach the organized state. Thus, organized network structure is a valuable resource which, once formed, should be stored for quick later retrieval.

Such storage is indeed possible by finding and storing associations between sets of synapses that during signal correlation-controlled network selforganization are switched on simultaneously. This novel process, modelled in (Bergmann and von der Malsburg n.d.), is analogous to the formation of neural assemblies by an associative mechanism of Hebbian type, and one might speak of synaptic assemblies. To the extent that the synapses to be associated are contacting different target neurons a specific communication medium between synapses is required. This medium is probably constituted by a particular class of cells, which would correspond to the control units proposed by Charles Anderson (Olshausen et al. 1993), modelled as associative mechanism in (Lücke and von der Malsburg 2006) and further developed and used intensively in Chapter 3 of this book. Several possibilities for physiological realization are discussed in Chapter 4. An idea to be considered is that control units are to be identified with astrocytes (Möller et al. 2007).

The ideas expressed in this foreword may or may not be convincing to you, feeling that intuitively appealing arguments are too easily offered, and armchair modelling goes wrong too often. It therefore is important to take pains and work out models in all concreteness — even if experiment will later prove many or most of the details wrong — so that the functional viability of the underlying ideas can be tested on the computer. This is the significance of the work presented here. For the sake of concreteness it had to concentrate on some specific application, face recognition, and had to limit itself to mere sketches of the processes going on in the brain or thought to be going on in the brain. The lasting value of this work is to be found, however, in it generic traits. The most fundamental among these is the proposal that dynamic links, rapidly switching synapses, are an integral part of the data structure of brain state. Given that for each cortical neuron there are an estimated 10,000 synapses, the classical view, which restricts attention to one activity variable per neuron, would be ignoring 99.99% of all the information needed to describe that state. What this book is advocating is a profound paradigm shift, opening the door to vistas not even imaginable within the classical view.

One of these is the ability of active networks to represent the structure of mental objects. The reality that we experience is in our head, of course, and is a construction of our mind, although a construction guided by sensory input. This reality of the present scene is to a very large extent ephemeral, no two scenes ever being identical in detail. The representation of our inner reality is composed by the brain as a collage of structural pieces collected in the past. Dynamic links are the glue with which to put pieces together in a structured way.

Sets of active neurons, the classical data structure, have no inherent structure, structure that would allow, for instance, to judge whether two such sets stand in any kind of meaningful relationship (other than having been associated with each other in the past, indicated by associatively strengthened mutual permanent connections). By contrast, two well-organized dynamic graphs may or may not be composable to a larger well-organized dynamic graph again. In the example given in this book, a stored model network may or may not be homomorphic to an active network in primary visual cortex. Both networks have two-dimensional topological structure, and they are homomorphic if there is a topological mapping between them connecting cells labeled with similar features. This correspondence can be discovered by the system the first time the two dynamic nets are activated. The basis for this is previous training of control structures that connect small sub-networks (small sets of neighboring feature-specific neurons) here and there, sub-networks that are re-used again and again as part of different larger networks. Whereas in this book links are allowed to be dynamic only between stored model and image, when flexible models are to be constructed in short-term memory as a collage of stored subnetworks, dynamic links will be required to glue them together, and structural relationships must be defined to decide whether two subnetworks fit together or not, even if they are to be combined for the first time.

This book intends to demonstrate the viability of a set of qualitative ideas as model for the brain. Important further steps are necessary to complete the paradigm shift. One is neuroscientific validation of the basic aspects of the theory, among them the rapid reversible switching of synapses under signal control, the existence and structure of control units, and the generic nature of self-organized dynamic networks. A second, corresponding to T. Kuhn's puzzle solving, is the massive amount of work required to interpret the elements of cognitive science in terms of the machinery advocated here. And finally, it will be necessary to develop the mathematical apparatus that alone can turn what in retrospect will one day look like tinkering into a coherent theory of the working of brain and mind.

Frankfurt, April 2010

Christoph von der Malsburg

References

- Bergmann, U., von der Malsburg, C.: Self-organization of topographic bilinear networks for invariant recognition. Neural Computation (Under revision)
- Lücke, J., von der Malsburg, C.: Rapid correspondence finding in networks of cortical columns. In: Kollias, S.D., Stafylopatis, A., Duch, W., Oja, E. (eds.) ICANN 2006. LNCS, vol. 4131, pp. 668–677. Springer, Heidelberg (2006)
- Möller, C., Lücke, J., Zhu, J., Faustmann, P.M., von der Malsburg, C.: Glial cells for information routing? Cognitive Systems Research 8(1), 28–35 (2007)
- Olshausen, B.A., Anderson, C.H., van Essen, D.C.: A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. Journal of Neuroscience 13(11), 4700–4719 (1993)

Contents

Бас	skground and Concepts
2.1	The Primate Visual System
	2.1.1 Topographic Organization of the Visual Areas
2.2	Approaches to Invariant Object Recognition
	2.2.1 Two Concepts for Achieving Invariance
	2.2.2 State of the Art in Object Recognition
2.3	Computational and Biological Plausibility of the Two
	Approaches
	2.3.1 Computational Arguments
	2.3.2 Experimental Evidence
2.4	Proposal for Dynamic Routing as a Principle of Brain
	Function
Refe	erences
Refe	erences
Refe A (Correspondence-Based Neural Model for Face
Refe A (Rec	Correspondence-Based Neural Model for Face
Refe A C Rec 3.1	Correspondence-Based Neural Model for Face Cognition
Refe A (Ref 3.1 3.2	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition
Refe A (Ref 3.1 3.2 3.3	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns
Refe A (Ref 3.1 3.2 3.3	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background
Refe A (Rec 3.1 3.2 3.3	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column
Refe A (Rec 3.1 3.2 3.3 3.4	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column The Network
Refe A (Ref 3.1 3.2 3.3 3.4	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column The Network 3.4.1 Input Layer
Refe A (Rec 3.1 3.2 3.3 3.4	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column The Network 3.4.1 Input Layer 3.4.2 Assembly Layer
Refe A (Ref 3.1 3.2 3.3 3.4	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column The Network 3.4.1 Input Layer 3.4.2 Assembly Layer 3.4.3 Gallery Layer
Refe A (Rec 3.1 3.2 3.3 3.4 3.4	Correspondence-Based Neural Model for Face cognition Correspondence Finding Face Recognition The Basic Computational Units: Cortical Columns 3.3.1 Neurobiological Background 3.3.2 A Model of the Cortical Column The Network. 3.4.1 Input Layer 3.4.2 Assembly Layer 3.4.3 Gallery Layer Results

		3.5.2 Position Invariance	51
		3.5.3 Tests on Standard Databases	53
		3.5.4 Attention Experiments	57
	3.6	Discussion	60
	Refe	erences	63
4	Swi	tchyards—Routing Structures in the Brain	69
•	4 1	Multistage Routing	69
	4.2	Physiological Background of Dynamic Routing	70
	4.3	Optimized Architectures for Bouting	73
		4.3.1 Routing between Two Regions of the Same Size	73
		4.3.2 Routing Circuit with Different Sizes of Input and	
		Output Laver	78
	4.4	Interpretation of Results	80
		4.4.1 Difference to Sorting Networks	80
		4.4.2 Physiological Interpretation	82
	4.5	Discussion	85
	Refe	erences	87
_			
5	Ont	cogenesis of Switchyards	91
	5.1	Ontogenetic Plasticity Mechanisms in the Brain	92
	5.2	A Model for the Growth of Routing Networks	94
	5.3	Kesuits	97
		5.3.1 Robustness to Disturbances	99
	54	0.5.2 Growth of Three-Dimensional Networks	105
	0.4 5.5	Conclusion	100
	D.D Dofe		107
	neie	ences	108
6	Put	ting the Pieces Together: Recognition with	
	\mathbf{Swi}	tchyards	109
	6.1	Matching of Two Patterns	110
	6.2	Recognition from a Gallery of Patterns	117
	6.3	Conclusion	121
	Refe	erences	122
7	\mathbf{Dis}	cussion and Outlook	123
	Refe	erences	126
Α	Ap	pendix	127
	A.1	Self-normalization Properties of Columnar Dynamics	127
	A.2	Gabor Transform	129
Inc	lev		131
Inc	UA .	• • • • • • • • • • • • • • • • • • • •	101

List of Figures

2.1 Processing in the retina 2.2 Retina to cortex 2.3 Ventral and dorsal stream 2.4 Receptive fields in V1 2.5 Illusory contours 2.6 Retinotopy 2.7 Problem of invariance in machine vision. 2.8 The perceptron 2.9 Challenges for a vision system 3.1 The correspondence problem 3.2 Cortical areas involved in face recognition 3.2 Cortical areas involved in face recognition 3.3 Columnar organization of cortex 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A		
2.2 Retina to cortex. 2.3 Ventral and dorsal stream 2.4 Receptive fields in V1 2.5 Illusory contours 2.6 Retinotopy 2.7 Problem of invariance in machine vision. 2.8 The perceptron 2.9 Challenges for a vision system 3.1 The correspondence problem 3.2 Cortical areas involved in face recognition 3.3 Columnar organization of cortex 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database	2.1	Processing in the retina
 2.3 Ventral and dorsal stream	2.2	Retina to cortex
2.4 Receptive fields in V1 2.5 Illusory contours 2.6 Retinotopy 2.7 Problem of invariance in machine vision. 2.8 The perceptron 2.9 Challenges for a vision system 2.9 Challenges for a vision system 3.1 The correspondence problem 3.2 Cortical areas involved in face recognition 3.2 Cortical areas involved in face recognition 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for	2.3	Ventral and dorsal stream
2.5Illusory contours2.6Retinotopy2.7Problem of invariance in machine vision.2.8The perceptron2.9Challenges for a vision system3.1The correspondence problem3.2Cortical areas involved in face recognition3.3Columnar organization of cortex3.4Timecourse of unit activities3.5Principal layout of the face recognition system3.6Faces represented by a grid or a face graph3.7Architecture of the network3.8Information processing by control units3.9Information flow in the network3.11Interaction among control units3.12Matching process3.14Position invariance3.15A sample of 30 faces from the FERET database3.16Different facial expressions in the AR database3.17Cumulative match scores for the AR database3.18Cumulative match scores for the AR database3.19Spatial attention experiments	2.4	Receptive fields in V1
2.6 Retinotopy 2.7 Problem of invariance in machine vision. 2.8 The perceptron 2.9 Challenges for a vision system 2.9 Challenges for a vision system 3.1 The correspondence problem 3.2 Cortical areas involved in face recognition 3.2 Cortical areas involved in face recognition 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the AR database 3.18 Cumulative match scores for the AR database <td< td=""><td>2.5</td><td>Illusory contours</td></td<>	2.5	Illusory contours
 2.7 Problem of invariance in machine vision. 2.8 The perceptron . 2.9 Challenges for a vision system . 3.1 The correspondence problem . 3.2 Cortical areas involved in face recognition . 3.3 Columnar organization of cortex . 3.4 Timecourse of unit activities . 3.5 Principal layout of the face recognition system . 3.6 Faces represented by a grid or a face graph . 3.7 Architecture of the network . 3.8 Information processing by control units . 3.9 Information flow in the network . 3.10 Average face graph . 3.11 Interaction among control units . 3.12 Matching process between Input and Input Assembly . 3.13 Recognition process. 3.14 Position invariance . 3.15 A sample of 30 faces from the FERET database . 3.16 Different facial expressions in the AR database . 3.17 Cumulative match scores for the FERET database . 3.18 Cumulative match scores for the AR database . 3.19 Spatial attention experiments . 3.20 Object search experiments . 	2.6	Retinotopy
2.8The perceptron2.9Challenges for a vision system3.1The correspondence problem3.2Cortical areas involved in face recognition3.3Columnar organization of cortex3.4Timecourse of unit activities3.5Principal layout of the face recognition system3.6Faces represented by a grid or a face graph3.7Architecture of the network3.8Information processing by control units3.9Information flow in the network3.10Average face graph3.11Interaction among control units3.12Matching process between Input and Input Assembly3.13Recognition process3.14Position invariance3.15A sample of 30 faces from the FERET database3.16Different facial expressions in the AR database3.17Cumulative match scores for the FERET database3.18Cumulative match scores for the AR database3.19Spatial attention experiments3.20Object search experiments	2.7	Problem of invariance in machine vision
 2.9 Challenges for a vision system	2.8	The perceptron
3.1 The correspondence problem . 3.2 Cortical areas involved in face recognition . 3.3 Columnar organization of cortex . 3.4 Timecourse of unit activities . 3.5 Principal layout of the face recognition system . 3.6 Faces represented by a grid or a face graph . 3.7 Architecture of the network . 3.8 Information processing by control units . 3.9 Information flow in the network . 3.10 Average face graph . 3.11 Interaction among control units . 3.12 Matching process between Input and Input Assembly . 3.13 Recognition process . 3.14 Position invariance . 3.15 A sample of 30 faces from the FERET database . 3.16 Different facial expressions in the AR database . 3.17 Cumulative match scores for the FERET database . 3.18 Cumulative match scores for the AR database . 3.19 Spatial attention experiments . 3.20 Object search experiments .	2.9	Challenges for a vision system
 3.1 The correspondence problem		
 3.2 Cortical areas involved in face recognition 3.3 Columnar organization of cortex 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.1	The correspondence problem
 3.3 Columnar organization of cortex 3.4 Timecourse of unit activities 3.5 Principal layout of the face recognition system 3.6 Faces represented by a grid or a face graph 3.7 Architecture of the network 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process. 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.18 Cumulative match scores for the FERET database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.2	Cortical areas involved in face recognition
3.4Timecourse of unit activities3.5Principal layout of the face recognition system3.6Faces represented by a grid or a face graph3.7Architecture of the network3.8Information processing by control units3.9Information flow in the network3.10Average face graph3.11Interaction among control units3.12Matching process between Input and Input Assembly3.13Recognition process3.14Position invariance3.15A sample of 30 faces from the FERET database3.16Different facial expressions in the AR database3.18Cumulative match scores for the AR database3.19Spatial attention experiments3.20Object search experiments	3.3	Columnar organization of cortex
 3.5 Principal layout of the face recognition system	3.4	Timecourse of unit activities
 3.6 Faces represented by a grid or a face graph	3.5	Principal layout of the face recognition system
 3.7 Architecture of the network	3.6	Faces represented by a grid or a face graph
 3.8 Information processing by control units 3.9 Information flow in the network 3.10 Average face graph 3.11 Interaction among control units 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.7	Architecture of the network
 3.9 Information flow in the network	3.8	Information processing by control units
 3.10 Average face graph	3.9	Information flow in the network
 3.11 Interaction among control units	3.10	Average face graph
 3.12 Matching process between Input and Input Assembly 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.11	Interaction among control units
 3.13 Recognition process 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments	3.12	Matching process between Input and Input Assembly
 3.14 Position invariance 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.13	Recognition process
 3.15 A sample of 30 faces from the FERET database 3.16 Different facial expressions in the AR database 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments	3.14	Position invariance
 3.16 Different facial expressions in the AR database	3.15	A sample of 30 faces from the FERET database
 3.17 Cumulative match scores for the FERET database 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments	3.16	Different facial expressions in the AR database
 3.18 Cumulative match scores for the AR database 3.19 Spatial attention experiments 3.20 Object search experiments 	3.17	Cumulative match scores for the FERET database
3.19 Spatial attention experiments	3.18	Cumulative match scores for the AR database
3.20 Object search experiments	3.19	Spatial attention experiments
	3.20	Object search experiments

3.21	Activity of the Gallery Assembly after priming of female faces	60
3.22	Result of a priming experiment	61
4.1	One- and two-dimensional routing networks	72
4.2	Number of required units as a function of intermediate layers	7
4.3	Prefactors c and \tilde{c}	7
4.4	Possible forms of tapered networks	78
4.5	Routing network with linear decrease of layer size	78
4.6	Number of possible conflicts as a function of distance of input nodes	8'
17	Dependence of network size on parameter α	8/
4.7	Dependence of network size on parameter α	04
5.1	Axonal growth cone	95
5.2	Switchyard architecture	94
5.3	Term G for alignment of coordinate systems	90
5.4	Role of the marker similarity term	9'
5.5	Snapshots of the growth process	98
5.6	Results for layer size $n = 27$	100
5.7	Results for layer size 125 and noisy initial conditions	10
5.8	Noise robustness of the ontogenetic mechanism	102
5.9	Growth of three-dimensional networks	10^{4}
5.10	Results with and without wraparound boundary	
	conditions	10
6.1	Challenge of matching patterns via a Switchyard	11(
6.2	Information flow in the dynamic Switchyard	111
6.3	Typical pattern used for matching experiments	112
6.4	Principle of matching in a Switchyard	11:
6.5	Snapshots of the matching process	114
6.6	Information flow in the full recognition system	117
6.7	Patterns stored in the gallery	118
6.8	Matching and recognition process in the full system	119
6.9	Signals propagated down the Gallery Assembly stream	120
6.10	Random dot stereogram	122
7.1	Extension to recognizing several categories	125