ARTICLE IN PRESS

Computers in Biology and Medicine ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Eye movements and smart technology

Christian Freksa, Sven Bertel*

Cognitive Systems Research Group & SFB/TR 8 Spatial Cognition, Universität Bremen, Cartesium 3.47, 28359 Bremen, Germany

Received 29 July 2005; received in revised form 23 November 2006; accepted 13 December 2006

Abstract

From a cognitive science/artificial intelligence perspective, this paper identifies the scanpath concept as the instantiation of a general sequencing principle that permeates the organization of spatial scene knowledge throughout the levels of mental processing. As such, it helps create methodologies to open up windows onto higher-level cognitive processes, particularly by relating shifts of visual focus to shifts of attention in mental reasoning. The paper argues that these methodologies form a robust basis for smart applications that employ eye movements to assess and to assist in diagrammatic problem solving.

Keywords: Mental imagery; Neighborhood; Diagrammatic problem solving; Scanpath; Spatial structure; Visual attention; Attentional focus

1. Introduction

Eye movement technology has come a long way. In its beginning, technologies to register eye movements were needed: copper coils were attached to the eye balls to record induced electric currents in a magnetic field when the eyes were moved and cornea reflections of infrared light sources were measured by photo cells to compute eye positions. Determining eye position through image analysis from a camera image was first proposed and realized in the 1970s, but cameras were large, computer memory was expensive, and processing times were too long for camera-based real-time eye tracking in those days.

Due to the technological issues involved in tracking eye positions, much of the early eye movement studies were more determined by technological boundary conditions than by the cognitive issues of interest to the researchers. For example, the head movements of the participants in eye movement experiments had to be strictly constrained through chin and head rests and through bite bars that restricted not only the head movements but also the comfort and relaxation of the participants. The quality of equipment calibration competed against the quality of the participants' response as calibration required full attention, took a long time, and often tired out the participants.

This situation has changed considerably: modern camera-based mobile eye recording equipment functions with by and large unconstrained head movements, is frequently portable, and can be worn and carried around comfortably. Off-the-shelf digital video processing equipment allows for affordable real-time processing of the cameras' output, tracking of the pupils, and mapping onto a visual scene. The use of eye movement data as a method in academic research or in applications for industries and services has increased manifold over the last decades; in fact, it has increased to such a degree that eye movement tracking and data analysis have been embraced by growing open-hardware and open-source communities (e.g. [1]).

We now can go beyond simple eye movement recording tasks and employ advanced, real-time technologies for new tasks that make even better use of the sophistication of the oculomotor system. In particular, we can use eye movement data as an input for a robust and dynamic modeling of attentional shifts, including shifts that occur during mental problem solving and in particular with regard to spatial or diagrammatic reasoning problems. From this, dynamic models of control of focus in (spatial) problem solving come into close reach, thereby opening up exciting new perspectives on human–computer collaborative reasoning in domains such as architectural design or spatial configuration.

* Corresponding author. Tel.: +49 421 218 64235; fax: 49 421 218 98 64235.
E-mail address: bertel@informatik.uni-bremen.de (S. Bertel).

0010-4825/$ - see front matter
doi:10.1016/j.compbiomed.2006.12.009

2. Spatial structure and visual attention

Grasping the physical environment with our senses and giving it meaning provides a great challenge to cognitive systems—both natural and artificial. The entities in the environment and their relations span a huge space too complex to exhaustively search by computational approaches. This complexity is due to a high dimensionality of the conceptual space which we obtain when we relate each entity to every other entity; the problem is compounded when we take into account the wide range of scales and granularities at which we are interested in entities in the environment ranging from subatomic structures to astronomic scales.

In other words, there is no way that cognitive agents—specifically humans and autonomous robots—can understand the world around them by exhaustively analyzing the interrelationships of the entities involved; they must be extremely selective to come to grasp with the world. Cognitive agents and their perceptual and conceptual mechanisms can establish certain relationships between perceptual entities easily and quickly while others require a lot of effort and cannot be used in real-time recognition processes.

For this reason, it is natural that the concepts and meaningful structures that cognitive systems establish and employ closely follow their perceptual and conceptual abilities [2] and such is in fact a central idea of studying embodied cognitive systems (cf. [3]). Conversely, perceptual and conceptual abilities have developed through evolutionary processes that provide support for establishing connections between related entities in the world.

2.1. Spatial structure

Perception makes use of and is biased by spatial structures in the physical environment extensively: if each entity in the world was equally related to each other entity, cognitive systems had to deal with a space whose dimensionality equals the number of entities in the world minus one. Physical space restricts the number of dimensions to three. This essentially means that not all entities can have the same distance to one another; some entities are closer to one another than others. Thus, by restricting the problem space to few dimensions it will be much less complex—at the expense of accessibility of more distant concepts.

As embodied cognitive agents are at the same time physical agents, they are also subject to the same restrictions of physical space with regards to their physical movements: some locations in space will be closer than others. Additional restrictions apply: due to gravity and/or other physical constraints, agents cannot move equally easily in all three dimensions, thus, not all geometrically equidistant locations are equally reachable. For perceptual space other restrictions apply: vision, for example, can overcome egocentric distance to a certain degree, but it is confined to the two true spatial dimensions of the visual field; as a consequence, information about the third dimension (i.e. distance) can only be construed indirectly from the two-dimensional information. Thus, vision is partially constrained by spatial structure and partially overcomes the spatial restrictions.

2.2. Visual attention

Let us consider the two spatial dimensions that are maintained spatially on a retina. Spatial structure—specifically spatial neighborhoods—can be exploited by the homomorphic 'campotopic' organization of the retina: neighboring locations in the visual field are mapped to neighboring locations on the retina. As the retina is 'implemented' through physical neural networks, it is also subject to spatial constraints. In particular, neighboring neurons can interact more easily than distant neurons and neighboring receptive fields are usually stronger connected than distant ones. For many visual tasks, the spatial organization of the retina as well as of subsequent early cortical areas is very helpful; think of how the 'hole' in the visual image that is created by the retinal blind spot gets filled in by extrapolating neighboring information to give the complete picture (e.g. [4]). Also, a variety of recognition tasks (e.g. edge detection, movement detection) can be performed at these early processing levels. For other tasks, local neighborhood structures are too restrictive.

This is exactly the level where eye movements set in. Eye movements also act locally—but at a different scale than processes on the retinal organization. 'Strategic' attention-shifting eye movements depend on retinal information and are capable of abbreviating tedious local propagation of information at lower processing levels. Viewed in this way, eye movements generate spatial neighborhoods at a coarser level of granularity; these neighborhoods are induced by higher-level connections between larger-scale entities. As cognitive agents are equipped with world and domain knowledge, they do not have to analyze all details of their visual input; in fact, there is no way that visual systems could, for reasons of complexity. Strategic selection of specific locations to support (or refute) a hypothesis is enough to build up a coherent image and understanding of a visual scene.

3. Visual recognition strategies

Visual recognition strategies use spatial structure in two orthogonal ways: (1) region growing and boundary detection by relating visual input on a given level of granularity (horizontal processing) and (2) aggregation/refinement by relating visual input across various levels of granularity (vertical processing). While the first type of processes is a data-driven interaction between visual input and neural structures, the second type relies heavily on internal structures, e.g. [5], including those retrieved from memory. Together, the two types represent an interchange between bottom–up and top–down visual processing. Their interplay results in a sequence of eye movements with which a subject scans a visual scene and organizes features in sequential order [6].

The fact that visual recognition is partly controlled by higher-level processes suggests that high-level knowledge can be strategically employed for active search [7] and for problem solving, as will be discussed later on. One has to keep in mind, however, that as with any feedback system bottom–up and top–down processing do not constitute two separate streams of activities; they are codependent systems that interact in many ways.

4. Scene analysis and image construction

The spatial ordering of a scene's content that is induced by the scanpath is not just a phenomenon of early levels of visual perception. Instead, on a representation-theoretic level, the underlying sequentialization of information can be regarded as one of many principles of spatial organization that transcend the perception of physical space and serve to structure knowledge throughout a number of cognitive (e.g. memory) subsystems, cf. [8]. The concept 'scanpath' instantiates the sequentializing principle in the domain of visual perception; on the other hand, this principle stems from the dominance of linear ordering in visual scene analysis.

On a practical note, where physical space restricts the number of dimensions to three, sequencing the salient features of the space linearly further reduces representational complexity. In mental reasoning, linear representations can facilitate emergence of goal-directed behavior from distributed processes (i.e. by providing a 'thread' for mental processes).

4.1. Mental representation

The spatial ordering of visual information can be traced to later stages of data-driven cognitive processing; for example, the sequence of eye movements has been suggested to be part of the mental representation of a visual scene or configuration, cf. [7]. As such, it is stored in long-term memory along with other perceived information; it also serves as a spatial key to retrieving memory content.

4.2. Spatial index in image construction

It is important to underscore that data-driven and top–down streams of visual processing rely in part on highly overlapping cognitive subsystems, such as for the shifting and zooming of the spatial focus of attention [9], and that the subsystems' actual states and outputs constitute combined effects of inputs from both streams. Eye movements can be seen as resulting from some of these outputs, and more, they are reflective of spatial shifts of focus in underlying attentional processes.

That being said, it becomes clear that the concerned output necessarily depends on a relation between inputs from both streams. Where, for example, input from top–down has to stand against a constant broadband input through visual perception its overall effect on eye movements has to be less than it is in the absence of perceptual input. The latter situation characterizes the case of mental imagery.

Imagery provides good examples of the pervasiveness of the scanpath/sequentialization concept. Even in the absence of visual perception, effects of spatial ordering of a scene can be found. Brandt and Stark [10] show that the scanpath during mental imagery reflects the content of the imagined scene. What is more, Laeng and Teodorescu suggest that eye movements during mental imagery are not epiphenomenal but play a functional role for imagery processes [11] further argue that scanpaths may provide a spatial index to the parts of a mental image, a position that is also embraced by Mast and Kosslyn [12]. Under this notion, again, the scanpath can be viewed as a part of mental representations that is abstracted from the actual eye movements, and that plays a functional role in general relative spatial indexing.

This abstract conception of eye movement patterns is further substantiated by the finding that the movement patterns generated under imagery conditions may reflect the spatial relations in a scene even when the initial scene stimulus was non-visual in nature but instead purely verbal [13]. Possibly, in such cases, subjects were inspecting a mental scene they had previously constructed while listening to the verbal description and were thereby partially re-enacting the scanpaths memorized during the original construction process. In that respect, one may think of the spatial index provided by a scanpath as abstract (i.e. non-modal) rather than representing concrete oculomotor patterns.

4.3. Eye movements and attentional control

The focus of attention is similarly governed: as a matter of fact, shifts of attentional focus are often related to 'a moving of the mind's eye'. A substantial body of research associates attentional processes with processes in eye movement control (and vice versa); it is based on broad psychological, functional anatomical or neural evidence. Accordingly, members from both sets of processes—attention and eye movement control—are seen as rather tightly related and interdependent and it has been suggested that attentional shifts may be essentially oculomotor in nature [14–16]. Thus, the spatial index provided by a scanpath typically also tells a story about underlying shifts of attention.

In particular, under normal conditions, attention and eye movements are synchronized and attentional and visual foci coincide on a common visual target (*overt attention*). During a fixation the two can be dissociated and, during an eye movement, they can be moved even to opposite directions [17] (*covert attention*). Yet, this dissociation does not seem to be complete (as can be demonstrated with spatial cueing, cf. [15]). Accordingly, models that predict or explain eye movements induced by a scanpath structure cannot be (completely) segregated from models that predict or explain shifts of attention.

4.4. Scene analysis and image construction combined

Pictorial representations (including diagrams, sketches, pictures, etc.) have been attributed with particular perceptual, cognitive, representational, and computational advantages over sentential representations [18]. This applies particularly to tasks that can capitalize on the spatio-analogical properties of these forms of representation. Specifically, diagrams have

Please cite this article as: C. Freksa, S. Bertel, Eye movements and smart technology, Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2006.12.009

been related to visual mental images, either by representational or by computational metaphors. Cognitive mechanisms involved in the inspection of diagrams and those involved in the construction and inspection of mental images are found to interface at later [19] and earlier stages of mental processing [20,21]. Such conclusions complement the findings on overlapping cognitive subsystems in both capacities, for example regarding attentional focus.

They also help explain why human reasoning, when it is based on mental images and external diagrams combined, is as effective as it is: mental and external representations are interfaced by the significant overlap in the mental processing of either one of them. Thus, their different representational and computational strengths complement each other rather than that their weaknesses would add up.

5. Mental/visual construction

While visual recognition aims at identifying entities and features that are really out there in the visual field, visual problem solving also has to deal with hypothetical views that are supplied by mental construction. The boundary between visual and mental input to attention control can be shifted even further towards the mental side: imagery-driven processes can control eye movements for creative construction tasks.

Visual imagery in working memory is subjected to more severe capacity constraints than vision; therefore imagery-based construction reaches a complexity threshold beyond which the consistency of a mental construction cannot be assured. At this point, we can reverse the relation between visual input and mental representation we have maintained in problem solving: whereas mental imagery assisted to fill gaps in the visual domain, we now can employ the visual domain to fill gaps in the mental realm. We do this by externalizing mental images on a visual medium and then employ vision processes to provide feedback about consistency and other properties of the mental construction.

6. Diagrammatic problem solving

We have seen that the organizational principle of linearization, as it is instantiated in the scanpath concept, among others, permeates the levels of mental processing. Along various lines of exploration, eye movements have been shown to be reflective of attentional shifts as well as of underlying functional organizations of mental representations and processes, both under conditions that involve visual perception or mental imagery. The combination of these two thoughts leads us to postulate that some telling relation holds between eye movements on a visual scene on the one hand and attentional shifts on mental representations of this scene, on the other. With it, we should be able to define robust (i.e. partial) mappings from eye movements to manipulations on higher level, cognitive concepts which have been abstracted from a scene, and vice versa. Additional support for our claim comes from research into insight problem solving with diagrammatic problems where significant differences in eye movement patterns have been found between successful and unsuccessful problem solvers [22,23].

Clearly, one should not expect to find homomorphism between attentional shifts in the oculomotor system and moving foci in executive control of working memory. There are just too many abstraction levels inbetween and other factors and components involved to make a simple and direct mapping generally possible.

However, many tasks require close coupling between mental imagery and visual perception. For example, designers make extensive use of diagrammatic representations to visualize imagined configurations. With these tasks, attentional shifts are well coupled across abstraction levels, so that eye movement records taken during the problem solving can effectively help generate hypotheses as to what happens on higher (i.e. problem solving) levels at a given point in time. We thus propose to start investigating the interrelations between scanpaths and shifts of focus for well-defined spatial or diagrammatic problems and then expand the scope from there on.

6.1. Relating eye movements and spatial problem solving

With respect to mental reasoning about spatial configuration problems, we further suggest that eye tracking during the mental construction phase of a solution model in a diagrammatic reasoning task can be employed to robustly assess a reasoner's individual preferences. Fig. 1 shows an example of a spatial configuration task used in a recent eye tracking study. In it, participants were asked to mentally reconfigure geometric arrangements of matchsticks to fit a specific, different pattern which had been verbally described to them. All problems were underspecified, thus permitting a multitude of correct solution models. The construction process was carried out entirely mentally (i.e., without the aid of sketches or other tools) while the reasoner had the problem in view and his eye movements were tracked. After each problem, the reasoner indicated his solution model (if any had been discovered). In the case of the example problem in Fig. 1, four different valid solution models exist, as

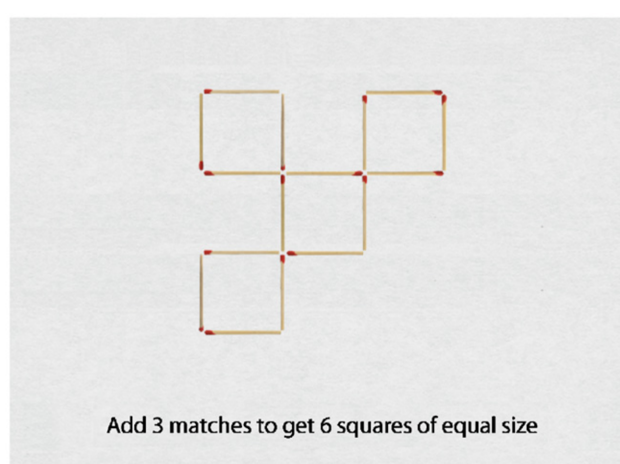

Fig. 1. An example of a spatial matchstick configuration task.

Please cite this article as: C. Freksa, S. Bertel, Eye movements and smart technology, Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2006.12.009

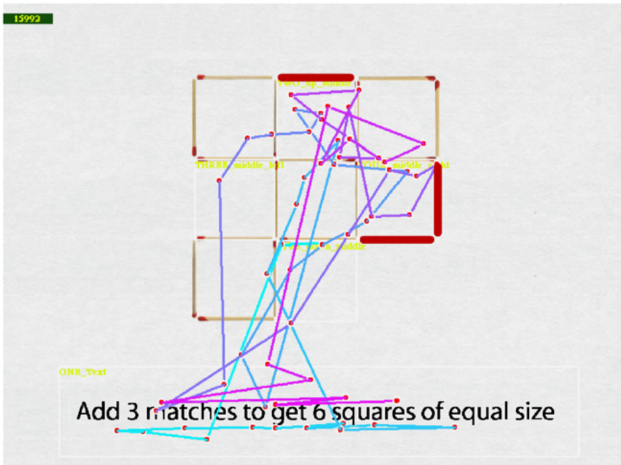

Fig. 2. Eye movements of one subject as shown by fixations and actually constructed solution (bold).

there are two ways each to create another square by adding two of the three matches, and an additional sixth square by adding the third match.

As a first step of the analysis, the eye movements that occur under problem solving are analyzed post hoc for various quantitative and qualitative measures, leading to the postulation of actions, behaviors, and problem solving episodes. Based on an evaluation of these concepts, a hypothesis is then generated as to the constructed solution model, and consequently checked against the solution provided by the reasoner.

One frequent finding was that the constructed solution model was in fact a function of distribution of relative fixation frequencies over diagram parts. Similarly, relative time spent on relation parts carried a meaning (compare Fig. 2; matches were added where the subject looked more often as well as longer). In addition, the constructed solution was often found to be related to the relative frequencies of transitions and/or sequences of fixations where typical sequences differed for different solution models. This finding nicely complements the reports by Grant and Spivey [22] that transition statistics between significant regions differ for successful and unsuccessful problem solvers.

Similar to findings of Knoblich et al. [23], we could observe distinct phases in the solving of matchstick problems: early phases were more concerned with a general inspection and understanding of the problem, while later phases (especially in later parts of a trial) were more indicative of the eventual solution, suggesting that this was when the solution's construction took place. Our data further suggests that it could be in earlier phases when different aspects of the problem and possible solutions get mainly explored.

Generally, there seem to exist important inter-individual differences in the reasoning methods and strategies that subjects applied in solving the different matchstick problems. For example, with problems for which subjects were required to take away a certain number of matches, most subjects either spent much more time on those parts that were eventually taken away than on those parts that stayed, or vice versa. The differences are interesting both from an informational as well as from a cognitive modeling point of view, as they suggest that smart (i.e. assistive) computational systems should react differently for different reasoners in such spatial layout tasks.

6.2. *Smart assistants based on eye movement analysis*

An analysis of eye movements during problem solving, however, does not have to end with generating post hoc hypotheses. As a second step, we propose that the developed classification and evaluation routines can be applied *during* problem solving. The aim is to establish reasoning and interaction schemes that can relate eye movement data points to their medium-duration reasoning contexts. Implemented in a computational reasoner and applied to collaborative human–computer reasoning scenarios the proposed routines could generate anytime behavior in which the currently best hypothesis of the eventual result is, respectively, used to tailor the behavior of a computational assistant in the task.

Applications for such joint reasoning can be found in many human–computer collaboration scenarios, in particular in spatial reasoning or sketch-/plan-based architectural design as for such collaboration to be satisfactory the computational side has to dynamically adapt to changes in the human reasoner's attentional focus as well as to the problem solving decisions that he takes along the way [24]. It is with design tasks in particular that one can expect a good acceptance of new computational tools among practitioners [25] as only part of a task will usually be complex and cognitively demanding while other parts require repetitive and mechanical labor. It is for reasons of implicitness of knowledge, style or esthetics that a fully computational treatment is precluded. Resulting are settings which exhibit an exceptionally strong need for good (i.e. cognitively adequate) human–computer collaboration. Tools that gather data on scanpaths during problem solving through the registering of eye movements and then based on this data try to predict a human reasoner's current cognitive states and attention could be a first approach.

7. Conclusion

The research discussed here relies on the hypothesis that attentional shifts in the human reasoner provide indications as to his mental processes and representations. Cognition thus guides attention, at least within the topics conversed here. Eye movements are seen as reflecting attentional shifts. The linear organization of features in the visual scene as scanpaths possess corresponding concepts on the various levels of cognitive processing. We argue that the combination of these properties helps to propagate attentional foci across different levels of processing.

Today, eye movement recording technology has advanced to a point where it can be used not only to register naturally occurring eye movements but to provide feedback about intentionally induced eye movements in connection with computer-driven visualization. In the context of human–computer

Please cite this article as: C. Freksa, S. Bertel, Eye movements and smart technology, Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2006.12.009

collaboration scenarios, the discussed mental and visual construction processes have good potential to be extended towards true dynamic procedures which serve to coordinate concurrent problem solving activities of more than one agent. It is in combination with other human–computer interface technologies that eye movement recording will prove useful in yet more domains.

Acknowledgment

This work is supported by the German Research Foundation (DFG) in the framework of the project *R1*-[*ImageSpace*] of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition: Reasoning, Action, Interaction.

C. Freksa acknowledges countless stimulating discussions and inspiration on the cognitive control of eye movements with Larry Stark over a span of 30 years. The authors also acknowledge fruitful discussions and cooperation with Mary Hegarty and members of her spatial thinking lab at UCSB.

References

[1] D. Li, D. Winfield, D.J. Parkhurst, Starburst: a hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches, Proceedings of the IEEE Vision for Human-Computer Interaction Workshop at CVPR, 2005, pp. 1–8.

[2] G.R. Lockhead, J.R. Pomerantz, The Perception of Structure. American Psychological Association, 1991.

[3] M. Wilson, Six views of embodied cognition, Psychon. Bull. Rev. 9 (2002) 625–636.

[4] P. de Weerd, R. Gattass, R. Desimone, L.G. Ungerleider, Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma, Nature 377 (1995) 731–734.

[5] L. Stark, S. Ellis, H. Inoue, C. Freksa, Z. Portnoy, J. Zeevi, Cognitive models direct scanpath eye movements: evidence obtained by means of computer processing of perceptual scanpath eye movements, in: XIIth International Conference on Medical & Biological Engineering, Jerusalem, 1979.

[6] D. Noton, L. Stark, Eye movements and visual perception, Sci. Am. 224 (6) (1971).

[7] L.W. Stark, Y.S. Choi, Experimental metaphysics: the scanpath as an epistemological mechanism, in: W.H. Zangemeister, H.S. Stiehl, C. Freksa (Eds.), Visual Attention and Cognition, Elsevier, Amsterdam, 1996.

[8] D. Engel, S. Bertel, T. Barkowsky, Spatial principles in control of focus in reasoning with mental representations, images, and diagrams, in: C. Freksa, et al. (Eds.), Spatial Cognition IV. Reasoning, Action, and Interaction, Lecture Notes in Computer Science, Springer, Berlin, 2005, pp. 181–203.

[9] A.C. Nobre, J.T. Coull, P. Maquet, C.D. Frith, R. Vandenberghe, M.M. Mesulam, Orienting attention to locations in perceptual versus mental representations, J. Cognitive Neurosci. 16 (3) (2004) 363–373.

[10] S.A. Brandt, L.W. Stark, Spontaneous eye movements during visual imagery reflect the content of the visual scene, J. Cognitive Neurosci. 9 (1997) 27–38.

[11] B. Laeng, D. Teodorescu, Eye scanpaths during visual imagery reenact those of perception of the same visual scene, Cognitive Sci. 26 (2002) 207–231.

[12] F.W. Mast, S.M. Kosslyn, Visual mental images can be ambiguous: insights from individual differences in spatial transformation abilities, Cognition 86 (2002) 57–70.

[13] R. Johansson, J. Holsanova, J., and K. Holmqvist (2005). What Do Eye Movements Reveal About Mental Imagery? Evidence From Visual And Verbal Elicitations, in: B.G. Bara, L. Barsalou, M. Bucciarelli (Eds.), Proceedings of the 27th Annual Conference of the Cognitive Science Society, Erlbaum, Mahwah, NJ, pp. 1054–1059.

[14] M. Corbetta, Frontoparietal cortical networks for directing attention and the eye to visual locations: identical, independent, or overlapping neural systems?, Proc. Natl. Acad. Sci. USA 95 (1998) 831–838.

[15] M. Shepherd, J.M. Findlay, R.J. Hockey, The relationship between eye movements and spatial attention, Q. J. Exp. Psychol. 38A (1986) 475–491.

[16] T. Moore, M. Fallah, Control of eye movements and spatial attention, Proc. Natl. Acad. Sci. USA 98 (2001) 1273–1276.

[17] M.I. Posner, Orienting of attention, Q. J. Exp. Psychol. 32 (1980) 3–25.

[18] J.H. Larkin, H.A. Simon, Why a diagram is (sometimes) worth ten thousand words, Cognitive Sci. 11 (1987) 65–99.

[19] S.M. Kosslyn, A.L. Sussman, Roles of imagery in perception: or, there is no such thing as immaculate perception, in: M.S. Gazzaniga (Ed.), The Cognitive Neurosciences, MIT Press, Cambridge, MA, 1995, pp. 1035–1042.

[20] A. Ishai, D. Sagi, Common mechanisms of visual imagery and perception, Sci. 268 (1995) 1772–1774.

[21] A. Ishai, D. Sagi, Visual imagery facilitates visual perception: psychophysical evidence, J. Cognitive Neurosci. 9 (4) (1997) 476–489.

[22] E. Grant, M. Spivey, Eye movements and problem solving: guiding attention guides thought, Psychol. Sci. 14 (5) (2003) 462–466.

[23] G. Knoblich, S. Ohlsson, G.E. Raney, An eye movement study of insight problem solving, Mem. Cognit. 29 (7) (2001) 1000–1009.

[24] S. Bertel, Show me how you act on a diagram and I'll tell you what you think (or: spatial structures as organizing schemes in collaborative human–computer reasoning), in: Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance. Papers from the 2005 AAAI Spring Symposium (SS-05-06), American Association for Artificial Intelligence, Menlo Park, CA, 2005.

[25] K. Meniru, H. Rivard, C. Bédard, Specifications for computer-aided conceptual building design, Des. Stud. 24 (2003) 51–71.

Christian Freksa received a Ph.D. in computer science/artificial intelligence/ cognitive science from UC Berkeley. He was a postdoctoral Thyssen fellow at the Max-Planck-Institute for Psychiatry in Munich and a research associate and assistant professor in the informatics department at TU Munich before becoming a professor for knowledge-based systems at the University of Hamburg. Christian Freksa currently holds the chair for cognitive systems at the informatics department of the University of Bremen. He is the director of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition at the Universities of Bremen and Freiburg. Christian Freksa's research interests are in the representation and processing of lean, coarse, fuzzy, and multimodal knowledge in terms of qualitative relations in spatial and temporal domains.

Sven Bertel studied informatics and cognitive sciences at the University of Hamburg and Indiana University, Bloomington. He received a diploma in informatics from the University of Hamburg with a thesis on supporting history-based web navigation by means of basic spatial concepts. Sven currently works as a research and teaching assistant in the Cognitive Systems Research Group at the informatics department of the University of Bremen, and in the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition. His research interests include computational models and architectures for reasoning about spatial configurations, the role of diagrams and mental imagery in architectural design processes, human–computer collaborative reasoning and the role of eye movements in diagrammatic problem solving.