Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes

Sara Irina Fabrikant[1], Stacy Rebich-Hespanha[2], Natalia Andrienko[3], Gennady Andrienko[3] and Daniel R. Montello[2]

[1]Department of Geography, University of Zurich, Zurich, Switzerland, sara@geo.uzh.ch. [2]Department of Geography, University of California Santa Barbara, Santa Barbara, CA, USA. [3]Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Supported by eye-movement data collected during a controlled experiment on small-multiple map displays, a new concept coined inference affordance aimed at overcoming drawbacks of traditional empirical 'success' measures when evaluating static visual analytics displays and interactive visual analytics tools is proposed. Then, a novel visual analytics research methodology is outlined to quantify inference affordance, taking advantage of the well-known sequence alignment analyses techniques borrowed from bioinformatics. The presented visual analytics approach focuses on information reduction of large amounts of fine-grained eye-movement sequence data, including sequence categorisation and summarisation.

INTRODUCTION

Cognitive scientists have attempted to tackle the fundamental research question of how externalised visual representations (e.g., statistical graphs, organisational charts, maps, animations, etc.) interact with people's internal visualisation capabilities, and can facilitate inference and decision making (Scaife and Rogers, 1996; Simon and Larkin, 1987). Experimental research in psychology suggests that *static* graphics can facilitate comprehension, learning, memorisation, communication of complex phenomena, and inference from the depiction of *dynamic* processes (Hegarty, 1992; Hegarty and Sims, 1994).

The need to better understand the cognitive processes involved in using dynamic displays has become more important recently, paralleling the exponential growth of animation and dynamic graphics to which people are being exposed in their everyday life (e.g., virtual-globe viewers, game controllers, and weather animations on TV news). As with most rapid developments of new technologies, the theory and understanding of novel graphics technology and applications has lagged behind.

As real-time three-dimensional landscape fly-throughs and interactive map animations become ubiquitous with dissemination over the Internet, an important question that remains is how effective the potential increase of information density in these highly interactive visual forms really is for (spatial) knowledge construction and decision-making. We still know very little about how effective novel interactive graphical data depictions and visual analytics

tools are for knowledge discovery, learning, and sense-making of dynamic, multidimensional processes (Harrower and Fabrikant, 2008). Today, a pervasive theme underlying many current (geo)visualisation research challenges is the difficulty of effectively evaluating highly interactive visualisation tools and complex displays, and of identifying their potentially positive influence on exploratory data analysis, knowledge extraction, and learning (Harrower, 2007).

STATIC AND DYNAMIC DEPICTIONS OF PROCESSES AND EVENTS

Visual analytics is based on the intuition that highly interactive and dynamic depictions of complex and multivariate databases amplify human capabilities for inference and decision making, as they facilitate cognitive tasks such as pattern recognition, imagination, association, and analytical reasoning (Andrienko and Andrienko, 2007; Thomas and Cook, 2005).

This contention is supported by the congruence principle suggested by Tversky *et al.*, (2002). This principle states that well designed external representations such as graphic displays show a natural cognitive correspondence in structure and content with the desired structure and content of the internal (mental) representation (i.e., the appropriate analytical inference). For example, animations congruently depict the concept of time and change with

changing displays over time, so it seems obvious that humans will have less difficulty comprehending complex dynamic processes through well-designed dynamic displays.

However, in a series of publications surveying the cognitive research literature on animated graphics (that did not include map animations), Tversky and colleagues claim they failed to find benefits to animations for conveying dynamic processes (Bétrancourt *et al.*, 2000; Bétrancourt and Tversky, 2000; Morrison *et al.*, 2000; Morrison and Tversky, 2001; Tversky *et al.*, 2002). These cognitive scientists argue that studies reporting a superiority of animations over static displays had experimental design flaws. For example, additional interactivity for the animated sequences violated the information equivalence between static and animated graphics. However, as Krygier *et al.* (1997) suggest, interactivity has different intensities or modalities, with static graphics having the lowest interaction intensity — but not zero. Static graphics do afford mental (internal) interactivity. A sequence of static graphics (e.g., small multiples) can be seen as (mentally) interactive in the sense that people can proactively control with their eyes the viewing order of the static sequence; they can always go back to the beginning of the sequence, and they can choose to study the sequence at their own pace and in any order they wish. Generally, if a static map is presented on a piece of paper (or a small display), it can be rotated and/or folded, as travellers commonly do with maps, for instance. An animated sequence being slightly more (externally) inter-active when featuring a start, stop, and rewind button is less (internally) interactive, in the sense that the sequence must be passively viewed in a pre-defined order. This reduction in (internal) interactivity may add cognitive load onto a viewer's working memory, thus limiting the animation's potential for facilitating learning (Sweller, 1994). It seems that this particular hypothesis has not been adequately investigated in animation studies, especially not on dynamic map displays.

Results on (comparative) cartographic experiments are inconclusive, partly because it depends on how 'better' is defined and measured. In some experiments that compare map animations with static small-multiple displays, participants answer more quickly but not more accurately with animations (Koussoulakou and Kraak, 1992). In other experiments, they take longer and answer fewer questions more accurately (Cutler, 1998), or the time it takes to answer the question does not relate to accuracy at all (Griffin *et al.*, 2004). In a mostly qualitative map animation study, Slocum *et al.*, (2004) found that map animations and small multiples are best used for different tasks. The former are more useful for inspecting the overall trend in time-series data, the latter for comparisons of various stages at different time steps.

We argue in the next sections that the question of whether animations are superior to static maps is not only an ill-posed question but also an unanswerable one. As geovisualisation designers, we should instead be interested in finding out how highly interactive visual analytic displays work, identifying when they are successful and why (Fabrikant, 2005).

INFORMATION EQUIVALENCE VS. INFERENCE AFFORDANCE

There is a fundamental problem with these kinds of comparative studies. One the one hand, it seems obvious that well-designed animations need to be compared to well-designed static displays. Very often, however, the stimuli are not prepared by design experts, thus differences that are found might be attributed simply to bad design choices. To precisely identify differences in the measures of interest, the design of the animations and small multiples to be compared requires tight experimental control to the extent that it might make a comparison meaningless. Tversky and colleagues (citations above) argue that experimental studies reporting advantages of animation over static displays lacked equivalence between animated and static graphics in content or experimental procedures. For example, they argue that animations show more information than static graphics because only the coarse segments are portrayed in static graphics, whereas animations portray both the coarse and fine segments of change. This presupposes the notion that animations and static displays *can* be *informationally equivalent*, a term coined by Simon and Larkin (1987) to express the idea that all information encoded in one representation is also inferable from the other, and vice versa. In other words, can all information available in a map animation be used to build an informationally equivalent small-multiple map display (SMMD)? We argue that well-designed animations are inherently different from well-designed small multiples and *should* afford different kinds of information extraction, specifically amenable to the desired inference modes and to the specific knowledge construction tasks. Making an animation equivalent in information content to a small-multiple display in order to achieve good experimental control for comparisons may actually mean degrading its potential power for certain tasks. Animations are not simply a sequence of static small multiples (Harrower, 2003). Effective static displays depict configural information (i.e., states) or static snapshots (freeze frames) of events and processes. For example, in a static time series of choropleth maps, a seven-class solution might effectively display a complex pattern of change. In an animation using the same data, however, it might not be wise to portray maps with seven classes, as viewers would not be able to apprehend that much detail and keep it active in working memory when viewing a non-interactive animation. If the goal of the animation is to depict change, then good design should focus particularly on emphasising change most effectively, for example, by applying smooth transitions between display frames to avoid potential change blindness (Rensink et al., 1997), something that static displays are inherently unable to achieve (Fabrikant and Goldsberry, 2005).

With their term *computational equivalence*, Simon and Larkin (1987) suggest a much more useful concept for assessing the effectiveness of graphic representations, especially when comparing different visual-analytics displays that typically afford different modes of interactions for inference making. Two representations are said to be *computationally equivalent* when any inference that is easily and quickly drawn from the encoded information in one display can be easily and quickly drawn from the other
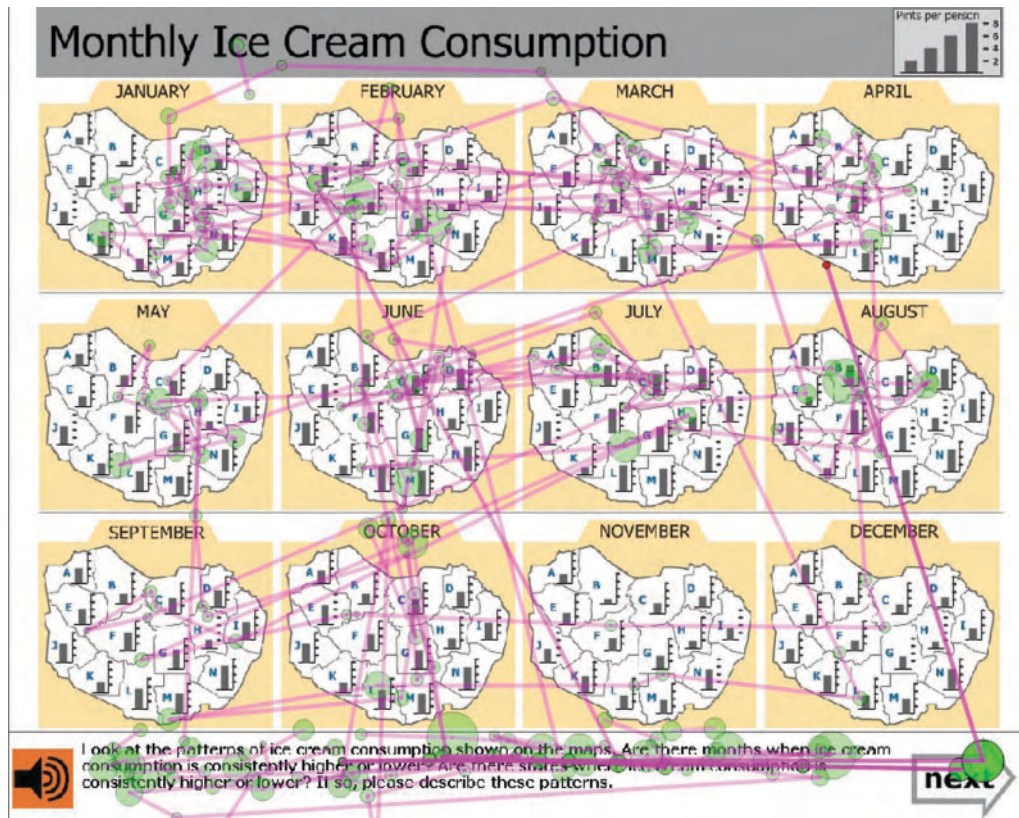
Figure 1. A gaze plot including eye fixations and saccades overlain onto a small-multiple map stimulus

(informationally equivalent) display and vice versa. Simon and Larkin (1987) do not specify what *easily* and *quickly* mean. They suggest that the advantages of graphics over text in general are computational, not because they contain more information, but because the presentation of the information can support extremely useful and efficient (computational) inference-making processes. This suggests that computational equivalence might be the more useful concept to use for comparison of complex graphics and interactive visual tools, with varying degrees and differing kinds of interaction affordances. We argue that computational equivalence is inherently linked to information equivalence and cannot be easily disentangled.

When comparing displays that afford different interaction modes, there seems to be a trade-off between informational equivalence and computational equivalence. To compare a non-interactive choropleth map animation in a fair way to a small-multiples display (e.g., with seven classes), the informational equivalence of the two displays has to be violated (i.e., the choropleth map classes must be reduced for the animation), because the limited interaction possibilities afforded by the animation leads to greater cognitive load, which affects its computational performance.

To better capture the effectiveness of a highly interactive and dynamic visual analytics display, we instead propose to use the concept of *inference affordance* that integrates both informational equivalence (amount and quality of content) and computational equivalence (quality and efficiency of inferences based on design). Effective visual analytics is not only about successfully extracting the content of the encoded data, but also about supporting different kinds of knowledge construction and inference-making processes through various cognitively adequate inference affordances.

What this discussion has not touched on so far is the complex issue of individual differences, including prior knowledge and training for visual-inference making. Elsewhere, it has been suggested that bottom-up (e.g., perceptual) and top-down (i.e., cognitive) processes are interlinked (Kriz and Hegarty, 2007). In other words, it does not just suffice to provide well designed graphics and visual tools and hope for success, but users also need to have an established base capacity for recognising and deciding which tool to select when, how, and for what aim and purpose (Lowe, 1999).

EYE-MOVEMENT ANALYSIS AND INFERENCE AFFORDANCE IN VISUAL ANALYTICS DISPLAYS

For over a century, psychologists and other researchers have recorded human eye movements, mostly on static displays, to learn how people read texts and view various static graphic displays, such as advertisements, works of art, charts, diagrams, etc. (Wade and Tatler, 2005). People move their eyes so that the fovea (the vision centre with highest acuity) is directed toward what they wish to attend to — to visually process at the highest possible detail (Rayner, 1992). Continual, ongoing eye movements are called saccades. Saccades are interrupted by eye fixations, phases where our eyes are relatively static, focusing on and attending to an object of interest.

The assumption that people's centre of visual attention is tightly linked with where they look during scene viewing has been recognised by a number of cognitive scientists who utilise eye-movement records to infer knowledge about the cognitive processes involved in various visual cognition tasks (Rayner, 1998). An advantage of eye-movement recordings, compared with traditional empirical data collection (e.g., questionnaires, interviews etc.), is that they provide relatively unobtrusive, real-time measures of (overt) visual and cognitive information processing behaviour (Henderson and Hollingworth, 1998).

Cartographers have utilised eye-movement recording as early as the 1970s to investigate how people look at static maps (Steinke, 1987). Cartographers were particularly interested in improving the design of their map products based on eye-movement research, thereby creating better and more user-friendly products (Montello, 2002). After increased interest in eye-movement studies with maps during the 1970s and early 1980s, the collection of eye-movement data in academic cartography has almost disappeared. Montello, (2002) suggested that one of the factors might have been that eye-movement analysis tended to be very expensive, and notoriously difficult to perform and analyze. Other critical voices argued that this kind of data collection did not tell mapmakers anything they did not already know, and thus did not warrant the extra effort and expense. Another reason for the limited success of eye-movement studies in cartography may have been that researchers tended to focus their studies on *where* people looked without getting at the *how* and *why* of map reading that generated the viewing pattern for particular map tasks (Brodersen *et al.*, 2002).

However, especially when evaluating visual analytics tools, where classic evaluation measures such as *accuracy of response* and *time to respond* might fall short (because of the entanglement of computational and informational equivalence), eye-movement behaviour analysis should provide additional insight into assessing the hard-to-measure concept of inference affordance proposed earlier.

SMALL-MULTIPLE DISPLAYS

Small-multiple displays (SMD), a graphic display type named and popularised by Tufte, (1983), had gained public attention for their potential to uncover complex dynamic processes at least since Muybridge introduced stop-action photography to study galloping horses in the late 19th century (Encyclopedia Britannica, 2008). Early on, cartographers achieved a high level of sophistication in representing complex, dynamic spatio-temporal reality through the power of abstraction, in the form of a series of static two-dimensional maps, which Bertin, (1967) calls the 'collection of maps with one map characteristic'. More recently, small multiples have resurfaced in highly interactive and dynamic visual analytics displays, allowing the user to reorder, brush, and otherwise manipulate the depicted spatio-temporal data on the fly (MacEachren *et al.*, 2003). The informational effectiveness of a static small-multiple display compared with an animation depends on using the appropriate number of small multiples and choosing the key events; that is, it depends on how many and which of the key events (macro steps) are selected to discretely represent the continuous and dynamic process (of micro steps). Well-designed small-multiple displays depict the most thematically relevant (pre-selected) key events and, unlike non-interactive animations, allow viewers to inspect the display at their own pace and viewing order. The inference affordance is directly related to the arrangement of the small multiples in the display, which in turn might be determined by the inference tasks the display should support.

EXPERIMENT

Utilising the eye-movement data collection method to track people's viewing behaviour, we investigated the role of inference affordance in static small multiple map displays (SMMD). The hypothesis at the outset is that if SMMDs and map animations are informationally equivalent, one would expect to find that viewers' knowledge gained from SMMDs would emphasise information about macro steps and the configurational aspects of the display (i.e., its visuo-spatial properties) more than on change (i.e., micro steps), as claimed by cognitive scientists in the work cited above. Moreover, in terms of computational equivalence, people's gazes would have to move sequentially from one map to the next in the SMMD, matching the sequential viewing order users are locked into in non-interactive animations, regardless of the knowledge-construction or inference-making tasks.

In this paper, we report on experimental results that were collected on SMMDs in isolation, without comparing the results to a map animation condition. As argued earlier, we believe the comparative 'what-is-better' question of SMMDs vs. animations to be unanswerable directly by means of a controlled experiment. The results reported in later sections will mostly focus on the computational aspects (inference events and process) of the inference affordance measure proposed earlier and specifically present a novel analysis approach to assess eye-movement behaviour for this purpose. We first present inference making patterns of individuals (exemplars) and then discuss methods for aggregation and summarisation. While we chose small multiple map displays as one typical static depiction method for representing a spatio-temporal process, the presented evaluation methodology is generic enough to be applicable to any spatial display (static or interactive) that may be produced to support spatio-temporal inference making.

In a controlled experiment, we first asked novice participants ($n=34$), tested individually, to study a series of small-multiple maps showing monthly ice cream consumption for an average year for different states in a fictitious country and then answer a number of questions about these maps (Figure 2).

The test questions required participants to make inferences varying in type and complexity; test question constituted a within-subject independent variable. For more complex inference questions we asked participants to explain their answers. Digital audio-recordings of participants' verbal statements permit joint analyses with
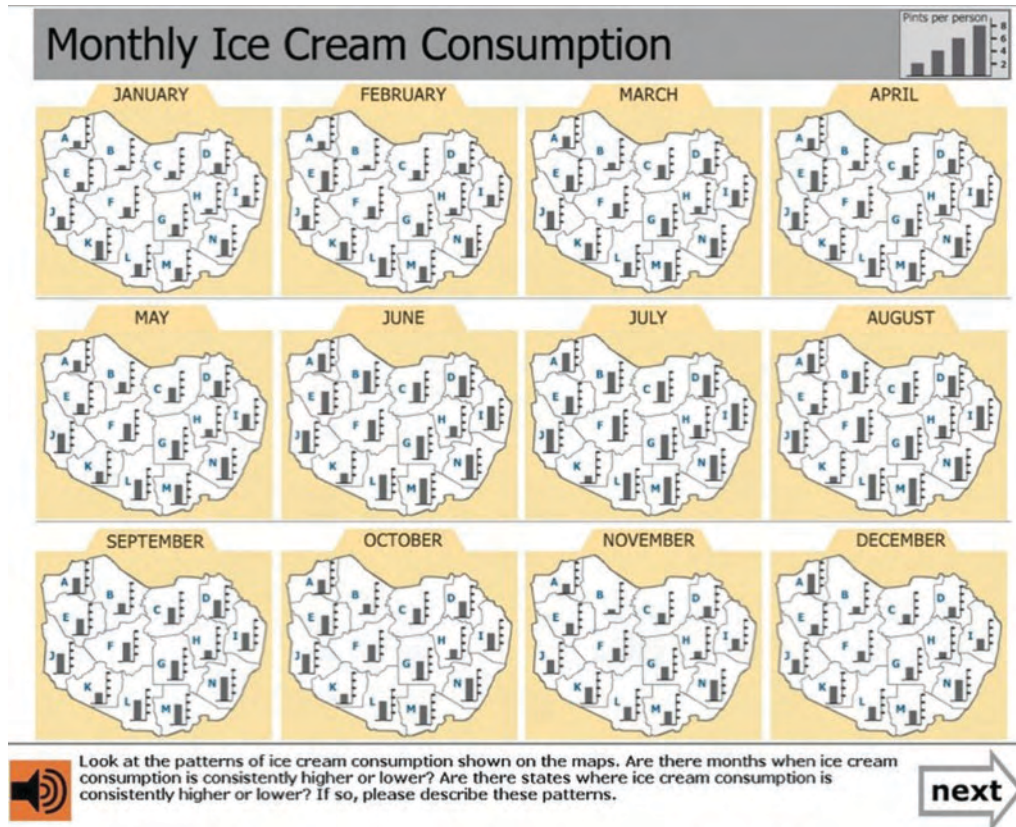
Figure 2. Sample small-multiple test stimulus with a general pattern detection question

the accuracy of their responses (inference quality measure) and their eye-movement recordings (inference process measure), all dependent variables.

Figure 3 shows a test participant's eye-movement patterns overlain on two identical SMMDs, but during two different inference-making tasks. The graduated circles show eye fixation durations (the larger the circle, the longer the fixation), and the connecting lines represent *saccades,* rapid eye movements between fixations. The passage of time is represented in both panels with transparency; that is, the more opaque the saccades and fixations are, the more recent.

In Figure 3a, the task is to gain an overall impression of the SMMD and verbally describe the patterns that are discovered during its visual exploration. In contrast, in Figure 3b, the task was to specifically compare two maps within the SMMD. When a map-use context requires a user to compare items in a time series (across time, space, or attribute), the non-interactive animations (locking a viewer into a pre-defined sequence) will always add cognitive load, as the viewer will have to wait and remember the relevant information until the respective comparative displays come into view. When animating the collected gaze tracks, one can clearly see that the viewer is not exploring the display in the implied sequence of the small-multiple arrangement, but going back and forth between the maps several times or jumping between different rows of maps. Ironically, this is one of our first success stories of the power of visual analytics: The interactive animation of eye-movement behaviour in the visual analytics tool we developed to

analyze eye movements turned out to be far superior for analyzing our collected data than the static gaze plot displays! To summarise: The SMMD allows the user to freely interact with the data in the viewing sequence they deem necessary for the task. This is one example of violating the computational equivalency of SMMDs and non-interactive animations in order to affect their informational equivalence.

Figure 4 depicts eye-movement behaviours during two magnitude-comparison tasks involving two maps at two different time steps in an SMMD. In Figure 4a, a user is asked to compare ice-cream consumption rates between the months of May and August, and in Figure 4b between the months of January and February. The gaze patterns reveals that only the information contained in those specific two maps is investigated to answer the test question. The remaining small maps are completely ignored. This suggests that for this particular task, non-interactive animations would indeed not be informationally equivalent to SMMDs, as they would force a user to see *much more* information than is relevant for the inference task. To maximise inference affordance, one could reduce the overload of presented information by offloading it, i.e., by making the animation interactive.

Moreover, these data further reveal that the design of the SMMD is an integral part of the inference affordance problem, which was not investigated in the cognitive work reviewed above. The particular design of the SMMD stimulus shown in Figure 4 seems to be ideal for detecting detailed change information between the adjacent months
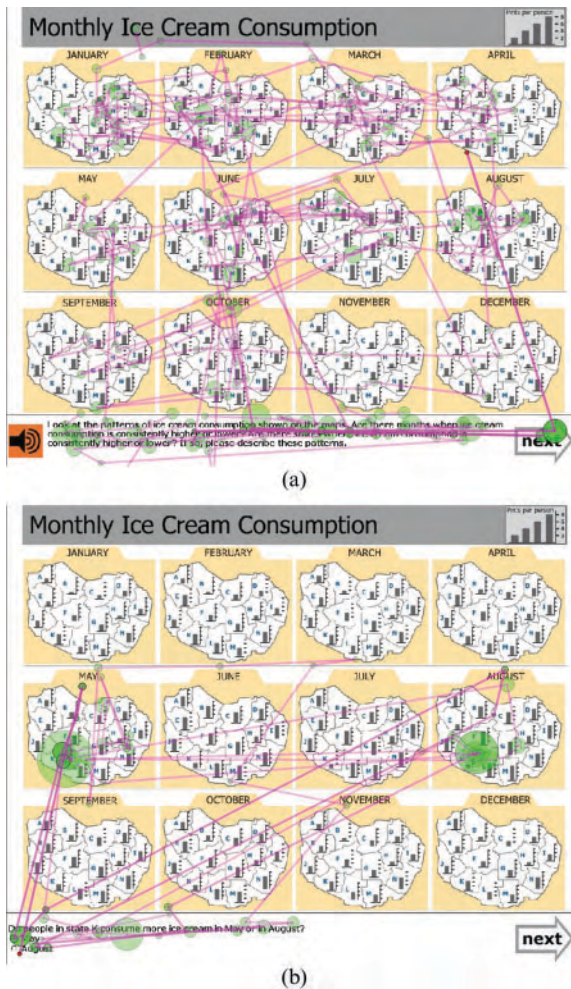
Figure 3. Task dependent viewing behaviour of two identical SMMD stimuli

Figure 4. Gaze plots for two different inference tasks affected by layout design

of January and February, but not between May and August, as these maps are far apart. In this case, adding interactivity to an SMMD might alleviate the reduced computational power produced by a suboptimal layout (e.g., by being able to move maps), as the arrangement of the SMMDs cannot be manipulated in the static version. A predefined layout might make this kind of inference task particularly difficult.

The significantly different viewing behaviours depicted suggest that small-multiple displays cannot generally be computationally or informationally equivalent to non-interactive animations; the computational and informational equivalence of displays certainly depends on the task, the information extraction goal, and the decision-making purpose.

VISUAL ANALYTICS OF EYE-MOVEMENT PATTERNS

Eye-movement research typically yields a tremendous amount of fine-grained behavioural data, both spatially and temporally, at very high levels of detail. For example, a 30-min recording will yield about 90 000 records, at a temporal resolution of 50 Hz (50 gaze points/seconds). Raw eye data are seldom used directly; they need to be
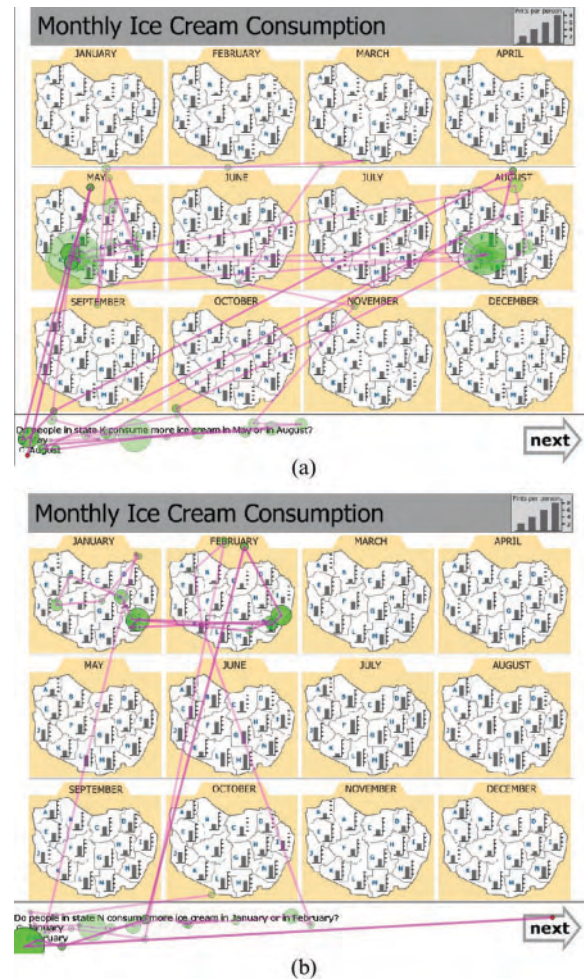
filtered based on a duration threshold, an empirical construct designed to better separate 'where people *look*' from where people cognitively 'process *seen* information'.

Data typically contained in an eye-movement record are depicted in Figure 5. A numeric identifier ('Map') links the eye record with a particular graphic stimulus. As stimuli are often randomised to avoid potential ordering biases, a second identifier ('Slide') indicates the order in which the stimuli have been seen. X- and Y-locations of the eye fixations are stored in display (screen) coordinates. Temporal information includes a time stamp released by a trigger event ('Start' in seconds) and a fixation duration ('Duration' in milliseconds). Additionally, investigators can identify areas of interest (AOI) in a stimulus that get recorded as 'interaction' events, as soon as the eyes have entered that particular AOI zone ('Top Zones' column). Other user interactions such as mouse or keyboard manipulations can be recorded as well and linked to gaze tracks. Based on available theory (Irwin, 2004; Henderson, 2007), only gaze points above 100 milliseconds have been retained for further analysis of the SMMD.

To analyze these large datasets, cross-fertilisation with GIScience/geovisualisation seems appropriate on several levels. Eye-movement software, and other related time-based observational data-analyses packages, typically do not

Slide	Map	Answer	X	Y	Start	Duration	Top Zones
12	1	x	245.4	842.4	19.75	0.25	QUESTION_01
12	1	x	201.7	827.3	20	0.265	QUESTION_01
12	1	x	263.5	815.3	20.265	0.172	QUESTION_01
12	1	x	366.5	813.5	20.437	0.297	QUESTION_01
12	1	x	282.2	825.3	20.734	0.328	QUESTION_01
12	1	x	342.3	816	21.062	0.125	QUESTION_01
12	1	x	401.6	817.7	21.187	0.141	QUESTION_01
12	1	x	476	817.8	21.328	0.297	QUESTION_01
12	1	x	596.1	815.8	21.625	0.25	QUESTION_01
12	1	x	682.8	804.7	21.875	0.187	QUESTION_01
12	1	x	778.8	801.1	22.062	0.219	NOVEMBER_01
12	1	x	882.4	811.6	22.281	0.219	QUESTION_01
12	1	x	800.4	823.9	22.5	0.843	QUESTION_01
12	1	x	919.2	826.5	23.343	0.422	QUESTION_01
12	1	x	964.4	820	23.765	0.532	QUESTION_01
12	1	x	1032.3	813.9	24.297	0.187	QUESTION_01
12	1	x	182.6	822.7	24.484	0.281	QUESTION_01
12	1	x	263.6	824.4	24.765	0.375	QUESTION_01
12	1	x	388.8	843.5	25.234	0.25	QUESTION_01
12	1	x	455.9	845.2	25.484	0.125	QUESTION_01
12	1	x	504.3	823.2	25.609	0.156	QUESTION_01
12	1	x	564.3	820.8	25.765	0.485	QUESTION_01
12	1	x	628.3	814.4	26.25	0.297	QUESTION_01
12	1	x	583.7	816.9	26.609	0.484	QUESTION_01
12	1	x	647.3	813.3	27.093	0.172	QUESTION_01
12	1	x	812.5	799.7	27.406	0.359	NOVEMBER_01

Figure 5. Extract of a processed eye-movement data set

include any spatial-analytical tools to analyze or summarise location-based data. Visual analytics methods are missing entirely. Herein lies a great opportunity for the GIScience/ geovis community to reach out to other disciplines and help in the analysis of eye-movement recordings. The amount and complexity of the collected eye-movement recordings required us to think carefully about how to make sense of the empirical data sets. For this reason, we developed a lightweight visual analytics interface (using Adobe Flash) that allows us quickly to visually explore the collected eye-movement data (play back, filter, visually summarise), gaining first insights on individual behaviours before running any hypothesis-testing analyses. Figure 6 below depicts the Flash-based graphical user interface of our *eyeview* software[1], developed as a proof-of-concept tool and described in Grossmann (2007).

The system allows one to load text-based eye-movement records, as shown in Figure 5 above, and filter data based on time, attribute, or location, including more advanced spatial analyses; the subset can then be displayed overlain on a graphic stimulus. The most useful feature of this system for this research simply turned out to be the play-back and sequencing function, which created animations of the eye-movement sequences!

SEQUENCE ANALYSIS (SA)

Visual analytics methods and data exploration tools for the effective depiction and analysis of time-referenced spatial data sets at high resolution have recently gained new attention (Laube and Purves, 2006; Andrienko and Andrienko, 2007). Location changes, order of events, smooth pursuits, etc., have become new foci of process-based research using spatio-temporal moving-objects databases of various kinds, and at different scales (i.e., moving humans over a year or moving eyeballs in milliseconds) (Laube *et al.*, 2007). Very large databases containing moving object behaviours are generated in abundance as a result of various tracking devices available today (i.e., LBS, GPS-enabled cell phones, eye trackers for market research and in psychology).

Sequence analysis (SA) is one promising approach to the analysis of process, event, and change, rather than the more traditional analysis of objects and their configurations, including location (Abbott, 1990). Depending on the research question and the collected sequence data, different kinds of SA methods are available. As for traditional statistical analysis, it is important first to distinguish continuous from categorical sequence data. Moreover, non-recurrent sequences of equal length (in which events cannot repeat in the sequence) or recurrent sequences with unequal lengths (containing sub-sequences with event repetitions) require different SA methods. One also needs to consider if states within a sequence are dependent on each other, or if whole sequences are dependent on each other.

For example, well-known Markov-type sequence analyses aim at modelling a process that reproduces a certain pattern (Hacisalihzade *et al.*, 1992). Markov analyses focus on internal sequence dependencies. These are modelled as a stochastic process by means of a 'step-by-step' computation, based on a transition probability matrix. There are several reasons why these kinds of models are not suitable
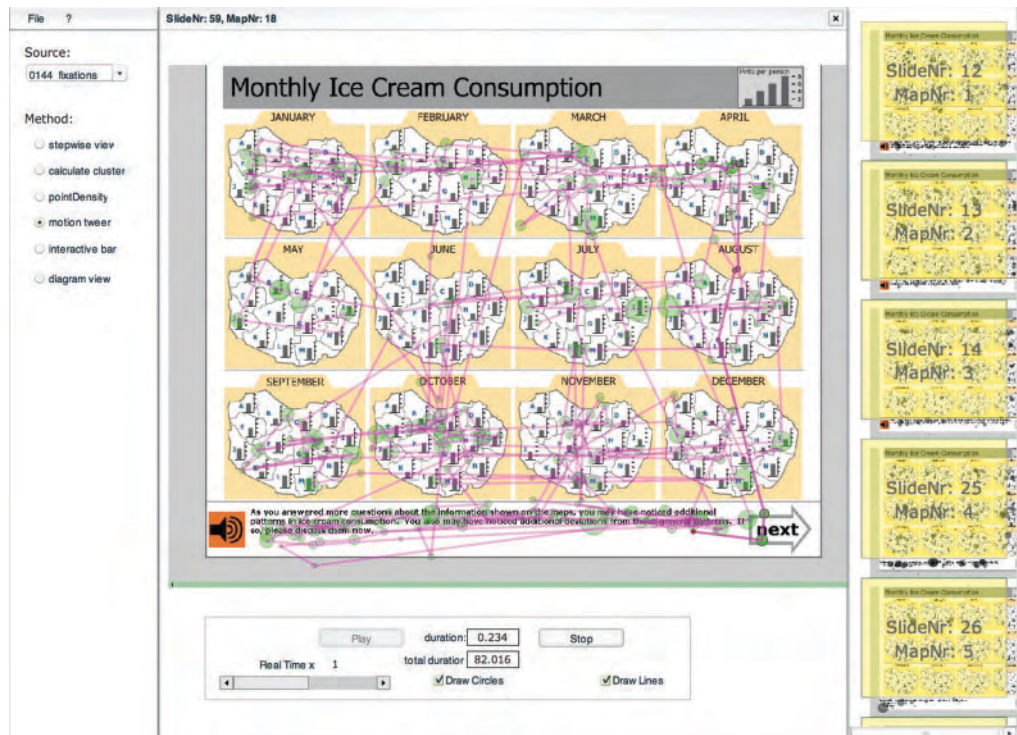

Figure 6. Visual analytics interface to depict inference-making behaviour through eye movements

for our work. For one, in exploratory work, the process is often unknown, thus empirical data cannot easily be compared with an idealised (theoretical) model sequence. Second, Markov models assume that the likelihood of an event occurring is conditional only on the immediate predecessor event, which is too limiting for modelling inference behaviour based on eye movements. In our work, we do not know what the process is at the outset. We need first to identify patterns hidden in the large eye-movement data collections by summarising and comparing various inference-making histories as a whole. We are also interested in identifying similarities across people, tasks, and modalities that might tell us something about the underlying process being affected by varying inference affordances.

Sequence alignment methods discussed in the next section seem particularly promising for our purpose, because they are good at identifying prototypical inference patterns by means of summarising and categorising eye-movement sequences (i.e., chains of attention events) across people and tasks.

SEQUENCE ALIGNMENT ANALYSIS (SAA)

Sequence alignment analysis (SAA), another technique of relevance to us, has been indispensable in bio-medical research for uncovering patterns and similarities in vast DNA and protein databases. Sequence alignment algorithms were developed in biology and computer science in the 1980s (Sankoff and Kruskal, 1983), and respective software packages became available soon thereafter (e.g., ClustalW). On a most general level, SAA algorithms

identify similarities between character sequences, based on the frequency and positions of characters representing objects or events, and on character transitions that are necessary for similarity assessment (Wilson, 2006). SAA has also become popular in the social sciences (Abbott, 1995), including geography (Joh *et al.*, 2002; Shoval and Isaacson, 2007), but has hardly been looked at by the cognitive community working with eye-movement data (West *et al.*, 2006).

SEQUENCE ALIGNMENT ANALYSIS OF EYE-MOVEMENT RECORDINGS

We employed the ClustalG software (Wilson *et al.*, 1999) to systematically compare and summarise individual inference-making histories collected through eye-movement data analysis. ClustalG is a generalisation of the various Clustal software packages widely used in the life sciences to analyze gene sequences in DNA and proteins (represented by characters with a limited alphabet). ClustalG has been developed specifically to deal with social-science data that require more complex coding schemes (i.e., an extended alphabet) for describing more complex event histories and social processes (Wilson *et al.*, 1999). The proposed SAA on collected eye-movement data includes a two-step approach: (1) data reduction of overt inference behaviour by summarisation of collected eye-movement sequences (across people and inference tasks), and (2) categorisation of found behavioural patterns by aggregating similar sequences into groups through cluster analysis. The steps can be applied in any order. In the discussion below, we inverted the analysis step sequence, exemplified for one
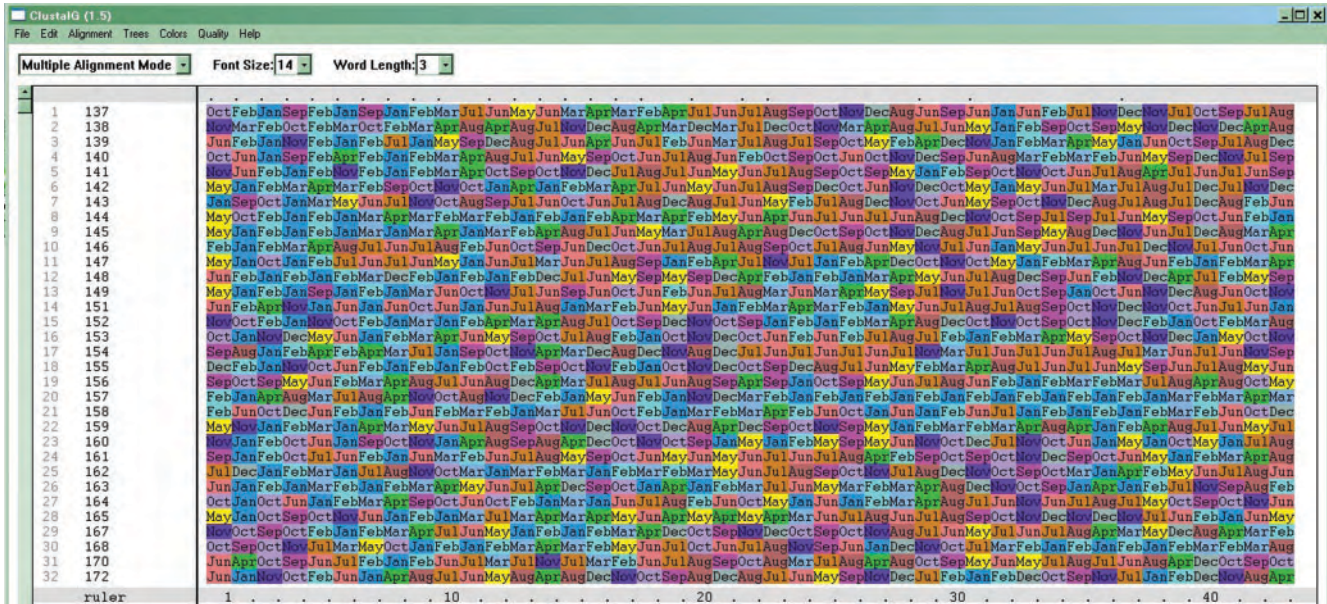
Figure 7. Participants' eye-movement sequences loaded into ClustalG

inference task with the SMMD (sample data shown in Figure 1).

CATEGORISATION OF EYE-MOVEMENT BEHAVIOUR

As mentioned earlier, aside from raw X-, Y-coordinates, we also collected fixation sequences based on pre-defined areas of interest (AOI), one area for each map in the SMMD. We post-processed the AOI data for each test participant and stored categorical character sequences into one ASCII text file (for one exploratory inference task, see Figure 2). Sequences vary considerably in length, from about 300 words to over 1100 words, where a word includes 3-character abbreviations for the months in the depicted SMMD time series (i.e., 'Jan', 'Feb', etc.).

The loaded sequences are colour-coded based on the months of the year. One row represents a viewing sequence for one participant. The viewing sequence begins on the left hand side of Figure 7, at starting position '1' found on the bottom row (x-axis) labelled 'ruler'. One can immediately see the winter months cluster at the beginning in cold colours (blue to purple) followed by the summer months in warm colours (yellow to brown). Next, a multiple alignment process is carried out, based on recommended input values by the ClustalG developers (Wilson et al., 1999). The first alignment phase includes a global pairwise-alignment procedure to identify similarities between whole sequences. The result is a resemblance matrix that is input to an unrooted phylogenetic-tree model (Saitou and Nei, 1987). This tree model (not depicted) represents branch lengths proportional to the estimated sequence uniqueness along each branch and is subsequently applied to guide the multiple alignment phase. Phase two, multiple alignment, is in essence a series of pairwise alignments following the branching order of the previously computed tree model.

Figure 8 portrays an extract of aligned sequences. One can see that the Jan–Feb pattern (in blue) is well aligned,

followed by gaps where sequences do not align (indicated in Figure 8 with dashes) and aligned portions of a Nov–Dec pattern. This pattern suggests that a significant group of people may have treated the temporally adjacent winter months as an inference unit but not at the same moment during the exploration. Perhaps this is due to Jan–Feb and Nov–Dec months being spatially far away from each other on the SMMD, and people seem to have employed varying viewing strategies and orders to compare them.

The uniqueness information contained in the clustering tree can be further analyzed to categorise aligned sequences. Based on the dendrogram, we identified three clusters. One cluster (containing three participants) can be characterised by viewing behaviour with considerable noise due to significant eye-tracking signal loss, as shown in Figure 9 (most and longest fixations outside the viewing area in the upper left corner).

The other two clusters are more difficult to analyze by simply playing back the viewing behaviour or by visually comparing the groups of gaze plots. For this reason, we decided to employ a powerful geovisual analytics toolkit specifically targeted for the analysis of movement data (Andrienko et al., 2007). Details of the software and provided analysis routines can be found in Andrienko et al. (2007).

SUMMARISATION OF EYE-MOVEMENT BEHAVIOUR

Trying to make sense of gaze data for one single test participant on one inference task is already difficult enough, due to extensive overplotting (as shown in the figures above). Trajectory data from Figure 1 shown earlier has been processed with a summarisation method from Andrienko et al. (2007), and the aggregated eye-movement path for that same participant is visualised in Figure 10.

The summarisation analysis depicted in Figure 10b includes directional information for the trajectories in the
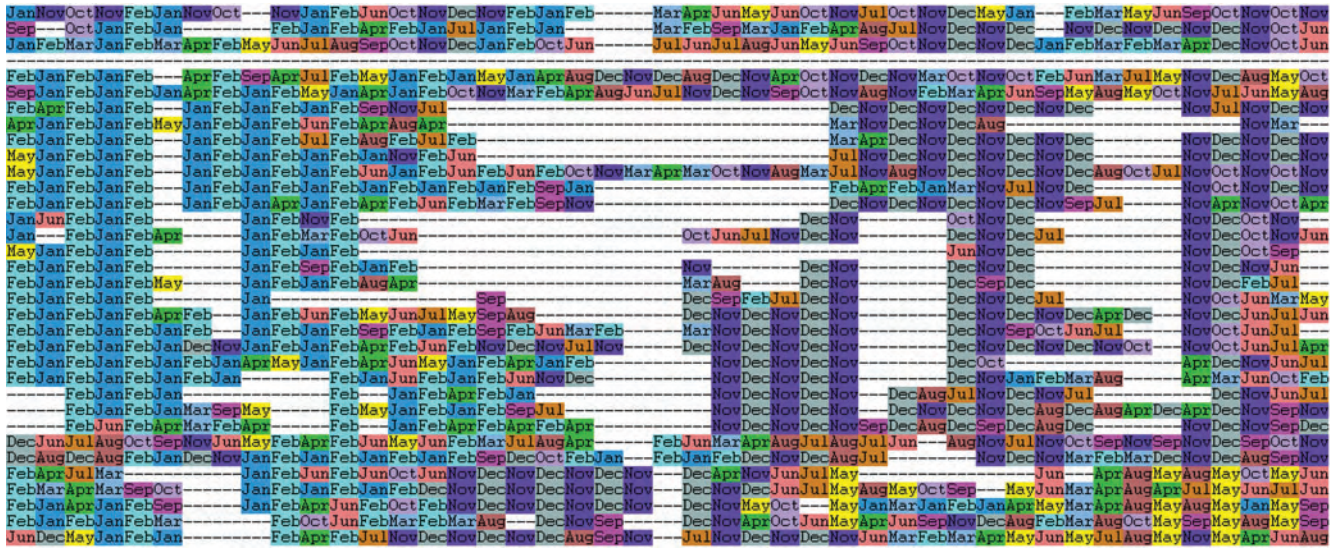
Figure 8.   Subset of aligned sequences

gaze plot (blue lines with arrows). Thicker lines indicate more movements. The depicted pattern suggests that this participant did not divide his/her attention equally over all maps. The first row was investigated more frequently, in both directions and in various spatial intervals (e.g., one/two steps forward, one/two steps backwards, etc.). Short vertical lines between rows suggests that the participant also chose a spatial viewing strategy, that is, viewing nearby displays irrespective of the suggested temporal sequence. Longer trajectories (missing arrowheads)
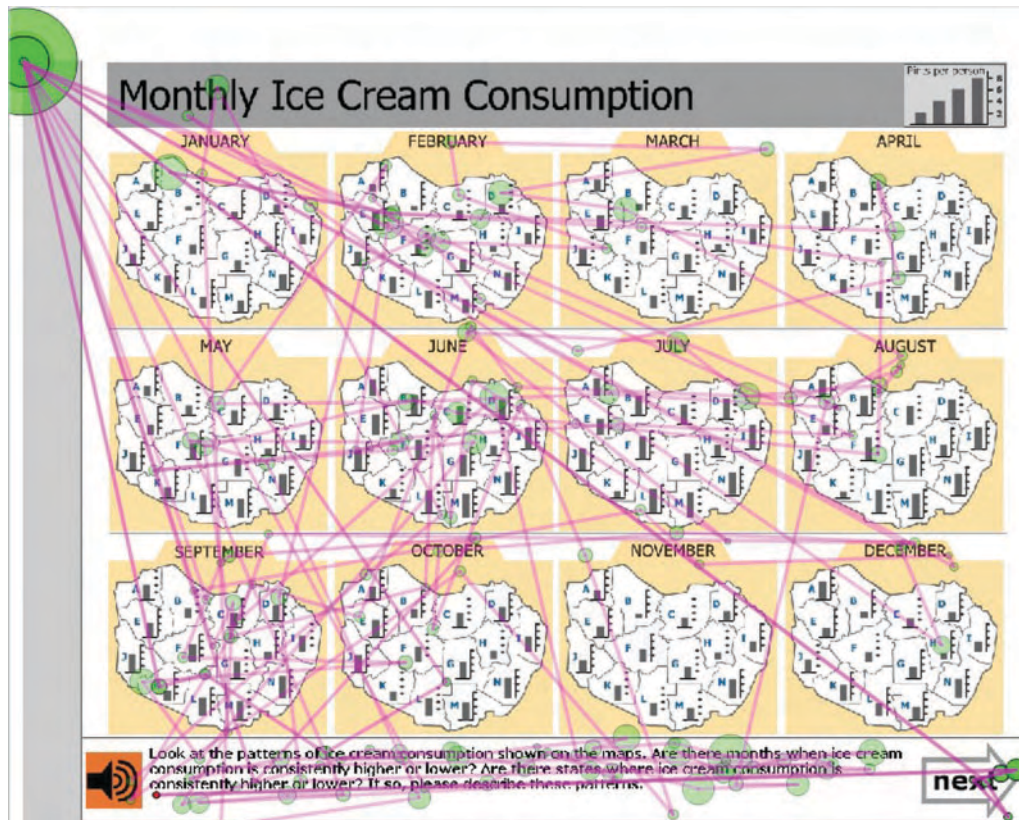
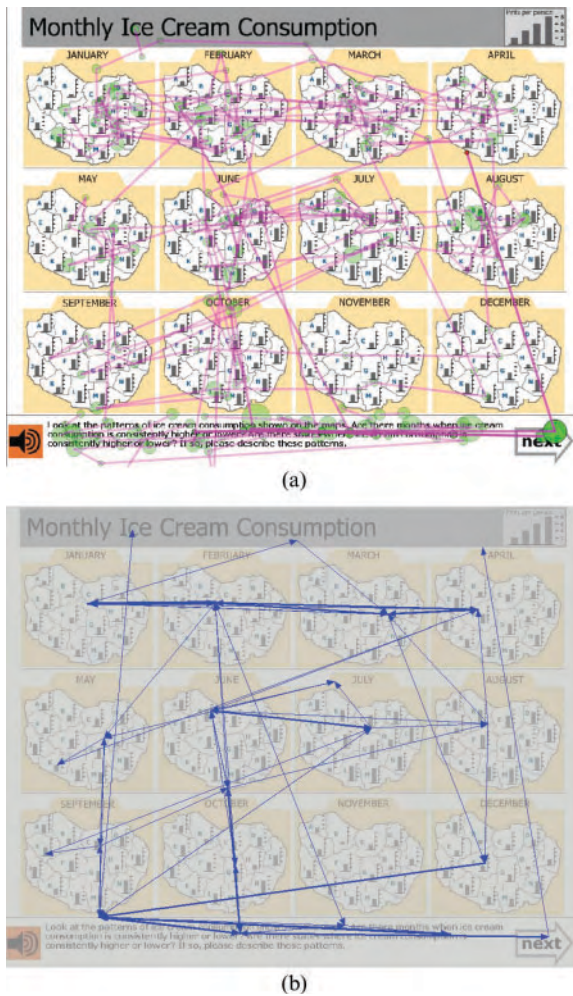Figure 9.   Outlier eye movement sequence due to eye tracking recording problems

Figure 10. Effect of data reduction, (a) original and (b) summarised eye movements

mean that information below the line was looked at 'in passing', if at all. For example, the last row, including October, November, and December, has comparatively few fixation locations (see next Figure 11) and were looked at in reverse order from the suggested viewing sequence. To validate the summarisation procedure, it also helps just to look at fixation patterns, as visualised in Figure 11.

The overplotting problem gets exacerbated when trying to inspect trajectories across all test subjects, as shown in Figure 12 below.

As Figure 12 shows, severe overplotting does not allow one visually to discover anything! To identify potential viewing strategies on a single inference task, we summarised all participant data based on cluster membership discussed earlier, identified during phase two of the sequence alignment procedure. As mentioned earlier, participants are clustered based on similarities in viewing behaviour (i.e., viewing sequences). The results of the three summarisations by participant clusters are displayed in Figure 13.

In other words, the following discussion of results and conclusions are based on summarisations across all participants. Generally, the spatial trajectory patterns can be described in terms of completed distances (i.e., long or short moves) and/or movement headings (i.e., vertical,

horizontal and diagonal moves). The horizontal trajectories at the bottom of each panel in Figure 13 are generally related to reading the test question, even if the lines are not displayed exactly over the respective text portion in the above displays. This visual mismatch is dependent on the aggregation algorithm used. Horizontal trajectories within a row of maps suggest that participants are moving their eyes in the suggested temporal sequence. Sequential viewing behaviour is also indicated when horizontal trajectories are connected by diagonals from the end of one row of maps to the beginning of the next row below. When playing back eye movement behaviours one can see that diagonal moves are always performed in the forward direction, while horizontal moves can be both performed forwards and backwards. Vertical moves across map rows suggest two things. Firstly, longer vertical moves (starting or ending from the question) are performed when participants initially read the test question and then start inspecting the maps, or when eyes are returning to the test question during the map exploration task. Second, shorter vertical moves within and across map rows indicate spatial exploration behaviours, for example when nearby maps are inspected instead of following the suggested temporal arrangement.

Visual pattern inspection suggests a couple of distinguishing features across behavioural clusters. 'Spatial search' behaviour is depicted noticeably in the star-like trajectory pattern shown in Cluster 1 in Figure 13a (representing 30% of the participants). The centre of the star is the second map from the left, in the centre row. A similar star pattern is visible in Cluster 3 (8% of the participants), and its centre at the same location (i.e., the June map) as in Cluster 1. Cluster 2 shown in Figure 13b includes the largest proportion of participants (62%), and features dominantly horizontal trajectories. By animating the eye movement behaviours for this cluster one can detect that the horizontal trajectories include forward moves and backtracking within map rows. A participant's summarised trajectory exhibiting this kind behaviour is shown in Figure 1. Interestingly, the horizontal moves within the rows are not only connected with diagonals in Cluster 2, but also with vertical lines at respective row ends. When inspecting these eye movements again by animation one can see that people combine temporal and spatial search strategies. The map sequences are looked at in reverse temporal order in the middle row, perhaps to increase spatial search efficiency.

These empirical findings on static small multiple displays suggest the following design principles for providing computationally equivalent animations: Animations should not only provide a play 'forward' button, and/or 'forward' sequencing interactivity, but also include backwards animation and reverse sequencing options, to provide at least equally efficient inference affordances compared with small multiples. Making SMMDs interactive, so that users can rearrange the map sequence according to the spatial, temporal, or spatio-temporal inference making tasks, and respective knowledge extraction goals can alleviate layout problems in static SMMDs.

In terms of methodology this research proposes a combined geovisualisation and visual geoanalytics
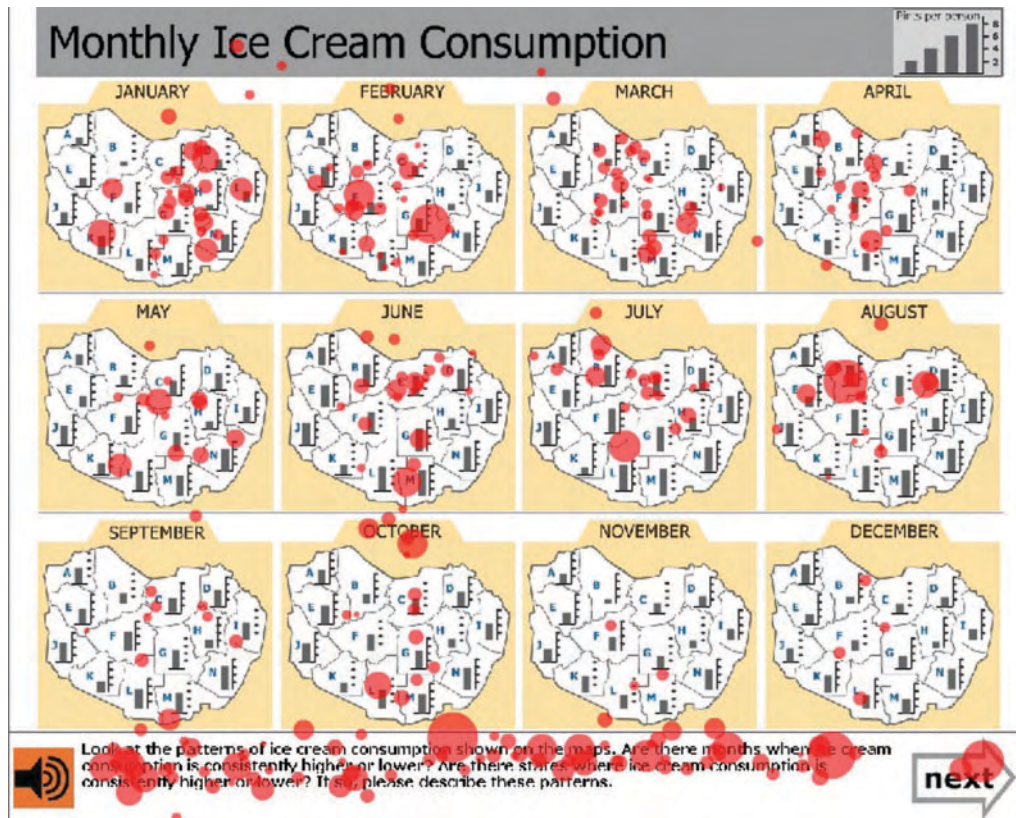
Figure 11.   Fixation pattern of same participant as in Figure 10

approach, to better quantify people's inference making processes from and with visuo-spatial displays. Considering that eye-movement recordings are location-based, they can be easily imported into an off-the-shelf GIS or, as in our case, a specifically developed visual geoanalytics tool. Eye movements can be displayed and analyzed in more detail with powerful spatial analytical tools in a similar fashion to the display and analysis geographic movement data. Geovisualisation methods are helpful for getting first insights on inference behaviours of individuals, for example, by simply being able to display gaze plots, and/or play back

Figure 12.   Gaze plots for several test participants

people's gaze trails over the explored graphic stimuli. Highly interactive visual geoanalytics toolkits such as proposed by Andrienko *et al.* (2007) provide an additional excellent framework to more efficiently handling massive fine grained spatio-temporal movement data, by summarising and categorising groups of behaviours. Empirical results based on the methods described earlier can be additionally linked to the more traditional success measures such as task completion time and accuracy of response. For example, in future work we will be exploring the potential relationship between viewing strategies based on identified cluster membership with the quality and speed of response.

CONCLUSIONS

A new concept, coined *inference affordance*, is proposed to overcome drawbacks of traditional empirical 'success' measures when evaluating static visual analytics displays and interactive tools. In doing so, we hope to respond to the ICA Commission on Geovisualisation's third research challenge on cognitive issues and usability in geovisualisation, namely, to develop a theoretical framework based on cognitive principles to support and assess usability methods of geovisualisation that take advantage of advances in dynamic (animated and highly interactive) displays (MacEachren and Kraak, 2001). Furthermore, a novel research methodology is outlined to quantify inference affordance, integrating visual geoanalytics approaches with sequence alignment analyses techniques borrowed from bioinformatics. The presented visual analytics approach
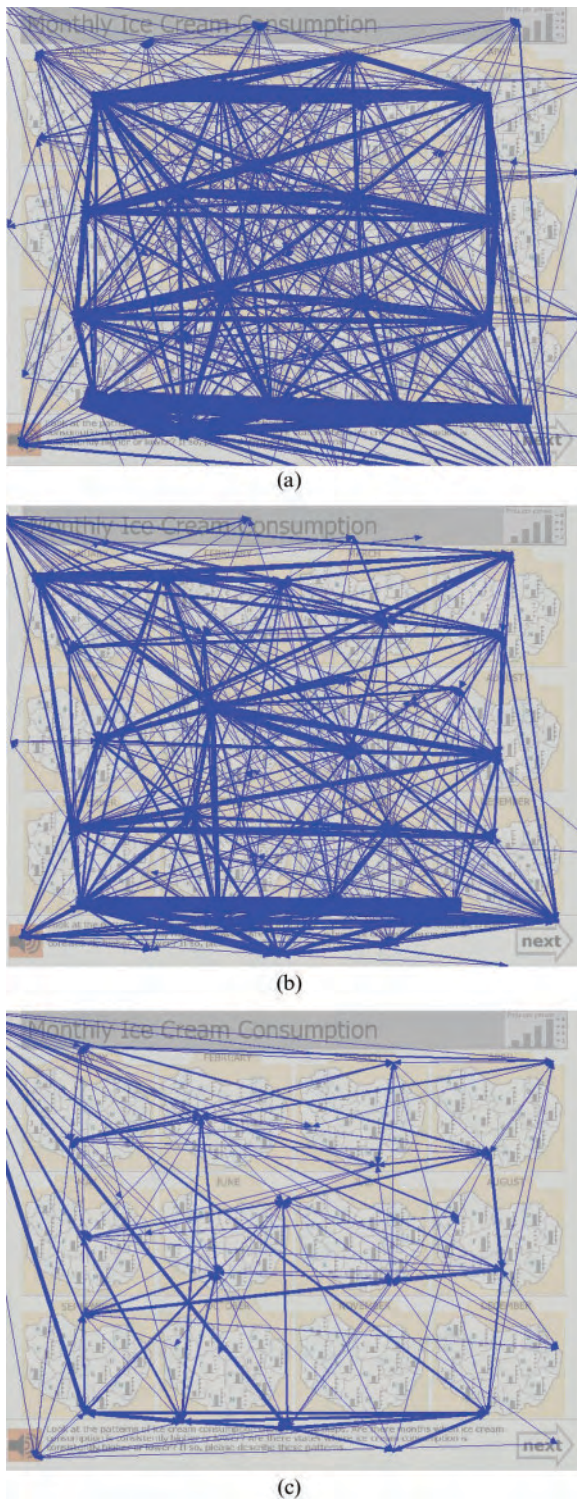
Figure 13. Summarised eye movements across participant clusters, based on viewing behaviour: (a) movement cluster 1, (b) movement cluster 2, (c) movement cluster 3

focuses on information reduction of large amounts of fine-grained eye-movement sequence data, including sequence categorisation and summarisation.

Presented inference-making behaviours extracted from eye movement records provide first support to the contention that small-multiple displays cannot generally

be computationally or informationally equivalent to non-interactive animations (in contrast to claims by cognitive scientists cited above); the computational and informational equivalence of displays do depend on the task, the information extraction goal, and the decision-making context.

By applying the outlined framework to collected empirical evidence on static small multiple displays, we hope to provide a better understanding of how people use static small-multiple displays to explore dynamic geographic phenomena, and how people make inferences from static visualisations of dynamic processes for knowledge construction in a geographical context.

BIOGRAPHICAL NOTES

Sara Irina Fabrikant is an associate professor of geography and head of the Geographic Visualisation and Analysis Unit in the Department of Geography at the University of Zürich, Switzerland. Her research interests are in geographic information visualisation, GIScience and cognition, graphical user interface design, and dynamic cartography. She earned a PhD in geography from the University of Colorado-Boulder (USA), and an MS in geography from the University of Zurich (Switzerland).

ACKNOWLEDGMENTS

REFERENCES

Abbott, A. (1990). A Primer on Sequence Methods. **Organisation Science** 1(4): 375–392.
Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. **Annual Review of Sociology** 21: 93–113.

Andrienko, G., Andrienko, N., and Wrobel, S. (2007). Visual Analytics Tools for Analysis of Movement Data. **ACM SIGKDD Explorations** 9(2): 38–46.

Andrienko, N. and Andrienko, G. (2007). Designing Visual Analytics Methods for Massive Collections of Movement Data. **Cartgraphica** 42(2): 117–138.

Bertin, J. (1967). **Sémiologie Graphique: Les Diagrammes – les Réseaux** – les Cartes, Mouton, Paris.

Bétrancourt, M. and Tversky, B. (2000). Effect of Computer Animation on Users' Performance: A Review. **Le travail Humain** 63(4): 311–330.

Bétrancourt, M., Morrison Bauer, J., and Tversky, B. (2000). Les Animations Sont-Elles Vraiment Plus Efficaces? **Revue D'Intelligence Artificielle** 14: 149–166.

Brodersen, L., Andersen, H. H. K., and Weber, S. (2002). **Applying Eye-Movement Tracking for the Study of Map Perception and Map Design**, Kort and Matrikelstyrelsen. National Survey and Cadastre Denmark, Copenhangen, Denmark.

Cutler, M. E. (1998). **The Effects of Prior knowledge on Children's Ability to Read Static and Animated Maps**. Unpublished MS thesis. Department of Geography, University of South Carolina, Columbia, SC.

Duchowski (2007). **Eye Tracking Methodology**, Springer, Berlin, Germany.

Encyclopædia Britannica (2008). **Muybridge, Eadweard**. (http://www.britannica.com/eb/article-9054508/Eadweard-Muybridge, Jan. 8, 2008).

Fabrikant, S. I. (2005). Towards an Understanding of Geovisualisation With Dynamic Displays: Issues and Prospects. **Proceedings**, American Association for Artificial Intelligence (AAAI) 2005 Spring Symposium Series: Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance, Stanford University, Stanford, CA, Mar. 21–23, 2005: 6–11.

Fabrikant, S. I. and Goldsberry, K. (2005). Thematic Relevance and Perceptual Salience of Dynamic Geovisualisation Displays. **Proceedings**, 22th ICA/ACI International Cartographic Conference, A Coruña, Spain, Jul. 9–16: (CDROM).

Griffin, A. L., MacEachren, A. M., Hardisty, F., Steiner, E., and Li, B. (2004). A Comparison of Animated Maps with Static Small-Multiple Maps for Visually Identifying Space-Time Clusters. **Annals of the Association of American Geographers** 96(4): 740–753.

Grossmann, T. (2007). **Ansatz zur Untersuchung der Wahrnehmung bei geographischen Darstellungen. Ein Werkzeug zur visuellen Exploration von Blickregistrierungsdaten**. Unpublished Master Thesis. UNIGIS Program, Salzburg.

Hacisalihzade, S. S., Stark, L. W., and Allen, J. S. (1992). Visual Perception and Sequences of Eye Movement Fixations: A Atochastic Modeling Approach. **IEEE Transactions on Systems, Man and Cybernetics** 22(3): 474–481.

Harrower, M. (2003). Designing Effective Animated Maps. **Cartographic Perspectives** 44: 63–65.

Harrower, M. (2007). The Cognitive Limits of Animated Maps. **Cartographica** 42(4): 349–357.

Harrower, M. and Fabrikant, S. I. (in press). The Role of Map Animation in Geographic Visualisation. In: Dodge, M., Turner M., and McDerby, M., (eds.), **Geographic Visualisation: Concepts, Tools and Applications**, Wiley, Chichester, UK, pp. 49–65.

Hegarty, M. (1992). Mental Animation: Inferring Motion from Static Displays of Mechanical Systems. **Journal of Experimental Psychology: Learning, Memory, and Cognition** 18(5): 1084–1102.

Hegarty, M. and Sims, V. K. (1994). Individual Differences in Mental Animation During Mechanical Reasoning. **Memory and Cognition** 22: 411–430.

Henderson, J. M. (2007). Regarding Scenes. **Current Directions in Psychological Science** 16: 219–222.

Henderson, J. M. and Hollingworth, A. (1998). Eye Movements During Scene Viewing: An Overview. In: Underwood, G., (ed.), **Eye Guidance in Reading and Scene Perception: Eye Guidance while Reading and While Watching Dynamic Scenes**, Elsevier, Oxford, UK: 269–293.

Irwin, E. (2004). Fixation Location and Fixation Duration as Indices of Cognitive Processing. In: Henderson, J. M., and Ferreira, F., (eds.), **The Integration of Language, Vision, and Action: Eye Movements and the Visual World**, Psychology Press, New York, NY: 105–134.

Joh, C.-H., Arentze, T., Hofman, F., and Timmermans, H. (2002). Activity Pattern Similarity: A Multidimensional Sequence Alignment Method. **Transportation Research Part B** 36: 385–403.

Koussoulakou, A. and Kraak, M. J. (1992). Spatio-temporal Maps and Cartographic Communication. *The* **Cartographic Journal** 29: 101–108.

Kriz, S. and Hegarty, M. (2007). Top-down and Bottom-up Influences on Learning from Animations. **International Journal of Human-Computer Studies** 65: 911–930.

Krygier, J. B., Reeves, C., DiBiase, D., and J. Cupp, J. (1997). Multimedia in Geographic Education: Design, Implementation, and Evaluation. **Journal of Geography in Higher Education** 21(1): 17–39.

Laube, P. and Purves, R. (2006). An Approach to Evaluating Motion Pattern Detection Techniques in Spatio-Temporal Data. **Computers, Environment and Urban Systems** 30(3): 347–374.

Laube, P., Dennis, T., Forer, P., and Walker, M. (2007). Movement Beyond the Snapshot – Dynamic Analysis of Geospatial Lifelines. **Computers, Environment and Urban Systems** 31(5): 481–501.

Lowe, R. K. (1999). Extracting Information from an Animation during Complex Visual Learning. **European Journal of Psychology of Education** 14(2): 225–244.

MacEachren, A. M. and Kraak, M.-J. (2001). Research Challenges in Geovisualisation. **Cartography and Geographic Information Science** 28(1), 13–28.

MacEachren, A. M., Dai, X., Hardisty, F., Guo, D., and D., L. (2003). Exploring High-D Spaces with Multiform Matrices and Small Multiples. **Proceedings**, IEEE Symposium on Information Visualisation, Seattle, WA, Oct. 19–24, 2005: (CDROM).

Montello, D. R. (2002). Cognitive Map-Design Research in the 20th Century: Theoretical and Empirical Approaches. **Cartography and Geographic Information Science. Special Issue on The History of Cartography in the 20th Century** 29(3): 283–304.

Morrison, J. B. and Tversky, B. (2001). The (in)effectiveness of Animation in Instruction. *Proceedings,* Jacko J., and Sears, A., (eds.), Extended Abstracts of the ACM Conference on Human Factors in Computing Systems, Seattle, WA: 377–378.

Morrison, J. B., Bétrancourt, M., and Tverksy, B. (2000). Animation: Does it Facilitate Learning? **Proceedings**, Papers from the 2000 AAAI Spring Symposium, Smart Graphics: 53–60.

Rayner, K., (ed.). (1992). **Eye Movements and Visual Cognition: Scene Perception and Reading**, Springer Verlag, New York, N.Y.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. **Psychological Bulletin** 124(3): 372–422.

Rensink, R. A., O'Regan, J. K., and Clark, J. J. (1997). To See or Not to See: The Need for Attention to Perceive Changes in Scenes. **Psychological Science** 8: 368–373.

Saitou, N. and Nei, M. (1987). The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. **Molecular Biology and Evolution** 4: 406–425.

Sankoff, D. and Kruskal, J. (1983). **Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparision**, Addison-Wesley, Reading, MA.

Scaife, M. and Rogers, Y. (1996). External Cognition: How Do Graphical Representations Work? **International Journal of Human-Computer Studies** 45: 185–213.

Shoval, N. and Isaacson, M. (2007). Sequence Alignment as a Method for Human Activity Analysis in Space and Time. **Annals of the Association of American Geographers** 92(2): 282–297.

Simon, H. A. and Larkin, J. H. (1987). Why a diagram is (sometimes) worth ten thousand words. **Cognitive Science** 11: 65–100.

Slocum, T. A., Sluter, R. S., Kessler, F. C., and Yoder, S. C. (2004). A Qualitative Evaluation of MapTime, A Program for Exploring Spatiotemporal Point Data. **Cartographica** 39(3): 43–68.

Steinke, T. R. (1987). Eye Movement Studies in Cartography and Related Fields. **Cartographica** 24(2): 40–73.

Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty, and Instructional Design. **Learning and Instruction** 4: 295–312.

Thomas, J. J. and Cook, K. A. (2005). Illuminating the Path: Research and Development Agenda for Visual Analytics, IEEE Press, Richland, WA.

Tufte, E. (1983). **The Visual Display of Quantitative Information**, Graphics Press, Cheshire, Connecticut.

Tversky, B., Bauer Morrison, J., and Bétrancourt, M. (2002). Animation: Can it Facilitate? **International Journal of Human-Computer Studies** 57: 247–262.

Wade, N. and Tatler, B. (2005). **The Moving Tablet of the Eye: The origins of modern eye movement research**, Oxford University Press, Oxford, UK.

West, J., Haake, A. R., Rozanski, E. P., and Karn, K. S. (2006). eyePatterns: Software for Identifying Patterns and Similarities Across Fixation Sequences. **Proceedings**, 2006 Symposium on Eye tracking Research & Applications, San Diego, CA, Mar. 27–29, 2006: 149–154.

Wilson, C. (2006). Reliability of Sequence Alignment Analysis of Social Processes: Monte Carlo tests of ClustalG software. **Environment and Planning A** 38: 187–204.

Wilson, C., Harvey, A., and Thompson, J. (1999). ClustalG: Software for Analysis of Activities and Sequential Events. **Proceedings**, Longitudinal Research in Social Sciences: A Canadian Focus, Windermere Manor, London, Ontario, Canada, Oct. 25–27, 1999.