

# EXPLORATORY GRAPHICS, MEMORY AND IMAGE FEATURES: TESTING THEIR RELATIONS

Ulanbek D. Turdukulov,\* Menno-Jan Kraak

ITC - International Institute of Geo-Information Science and Earth Observation  
PO Box 6, 7500 AA, Enschede, the Netherlands  
turdukulov, kraak @itc.nl

**KEY WORDS:** Time-series, Satellite images, Data exploration , Feature tracking , Visualization, Animation, Evaluation, Interaction

## ABSTRACT:

Visualization assist the researcher in observing and exploring sensor data. Exploration is a process where attention, memory and graphics play an important role. This paper discusses the relations among them. Current visualizations such as animations often do not offer sufficient support for the representation of image features and their evolution. This results in poor memorization and as a consequence, in poor understanding of dynamic features present in the images. In this paper, we test two exploratory tools: an animation and the prototype that tracks the features in time-series of sensor data and visualizes their evolution. Both were used for same set of exploratory tasks and with the same participants. The test persons comprised a novice and expert group who were offered the tools in varying sequence. Test results indicate that the novice participants could answer questions faster with the prototype while there was no such difference for the experts. Both groups had a higher error rate while using the animation and both groups reported a higher satisfaction level with the prototype. Overall, the test supports the notion that the representation of image feature and their evolution should be maintained in the visualizations intended for exploration of sensor data.

## 1 INTRODUCTION

Digital sensors are widely used pieces of equipment in research and management, particularly in spatially related disciplines. Examples include use of cameras, satellite and airborne sensors, pods in a sensor web, and so on. Visualization assist the researcher to observe and explore sensor data. Exploration is a learning process where attention, memory and graphics play an important role. This paper discusses the relations among them.

The main role of graphics in data exploration is to facilitate memorization and to guide visual search in the data (Bertin, 1983),(Ware, 2004). Bertin (Bertin, 1983) defined three main functions of graphic: inventory – when graphic is a storage that *replaces* human memory – corresponding to elementary graphic reading level; graphic processing – when information is still comprehensive but graphics becomes ‘processed’ for easier memorization and comparison – intermediate graphic reading level; and message – when graphic becomes a form that *inscribed* in human memory – overall level. To understand the graphical information is to visually memorize one or several images (Bertin, 1983). Visual memorization is also linked to a number of components that data contain. Thus following Bertin (Bertin, 1983), an efficient graphic can be defined as the one that provide access to all reading (memorization) levels of all data components that needs to be explored.

Sensor data are presently explored mainly by animating 2D or 3D image sequences. From graphical point of view, such representation contains two visual variables indicating location on the image plane (two spatial components), and one variable describing the image intensity (thematic component). Temporal component is the outcome of the process of taking a series of individual static images and stringing them together in a timed sequence to give the impression of continuous movement. Therefore, no graphical variables are utilized for the temporal component per se, instead they rely on memory, ‘temporal deduction’ (Koussoullakou and Kraak, 1992). Interaction with temporal component is usually facilitated by the user controlled video-type interactions (play

forward/ backward, pause/stop, etc.). However, evaluation studies show that despite having these interactions, animations can still lead to information overload (Morrison et al., 2000), (Lowe, 1999), (Lloyd, 2005), (Fabrikant and Goldsberry, 2005) (Blok, 2005).

Here it is argued that reasons for the current graphics shortcomings in supporting the exploration process are related a) to the data complexity and animations design, containing images that are mimicking reality (Turdukulov and Kraak, 2005) and b) to insufficient representation of *image features*. Temporal dimension shifts the user’s focus from the data components – thematic, spatial and temporal data properties, towards a knowledge component – image features, their lifetime and their evolution (Menis and Peuquet, 2003),(Peuquet, 1999),(Rana and Dykes, 2003). When multiple highly dynamic features are considered, they often exhibit ‘relationship’ and form ‘behaviors’ and therefore, such feature component needs an explicit representation. If represented, the component’s reading levels can assist users in searching for feature of interest, in comparing several image features and in perceiving the overall trend of evolving features in the time-series.

## 2 TEST PROCEDURE

To test the above argument both animation and research prototype are used and tested against each other with similar tasks and the same test persons. The research prototype, called ‘*DynoMap*’ (the prototype from now on) was build in our earlier work (Turdukulov et al., 2007), (Turdukulov and Kraak, 2005). It tracks image features in time-series, visualizes their evolution and provides several interactive option for controlling the image content and exploring the evolution of image features. Figure 1 displays snapshot of the prototype in application to exploration of convective clouds.

Typical animation consisted of video-type controls (play forward/backward, pause/ stop, change animation rate) and a few navigation options: zoom in/out, translate, reset. Figure 2 shows a snapshot of the animation.

\*Corresponding author

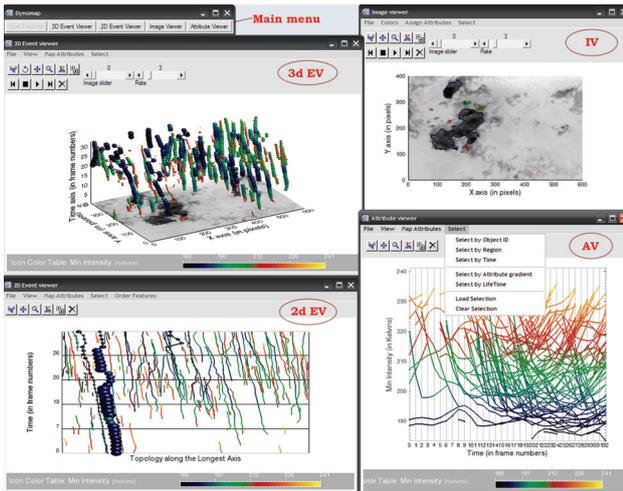


Figure 1: The prototype displaying the evolution of precipitating clouds. The prototype consists of four linked views: 3d Event Viewer (3dEV), 2d Event Viewer (2dEV), Image Viewer (IV) and Attribute Viewer(AV)

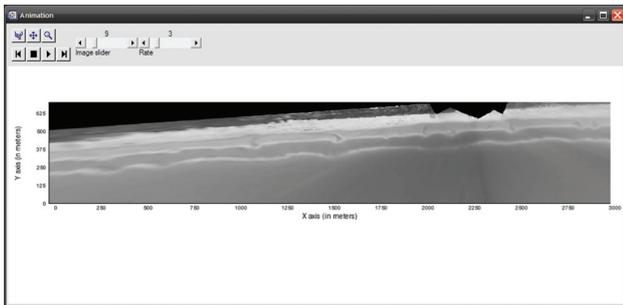


Figure 2: ‘Typical’ animation displaying the time-series of coastal morphology images used for the test.

Note that the objective was not to test the usability of the tools but rather to test the data learning and memorization process with the tools. However, since usability evaluation addresses the relationship between tools and their users, the same principle can be applied to measure geoscientists’ ability to explore, memorize and perceive trends in the evolution of image features with the tools. There are various ways where usability can be measured as described by different authors. International Organization for Standardization (ISO9241-11) identified three usability measures, which include effectiveness, efficiency and satisfaction. Nielson (Nielsen, 1993) identified five usability measures: learnability, efficiency, memorability, errors, and satisfaction. Hornabek (Hornbaek, 2006) reviewed 180 current Human Computer Interaction studies and indicated the following most frequent measures of usability that are also adopted for this paper:

- Measures of effectiveness: task completeness and accuracy of answers.
- Measures of efficiency: time and tool usage patterns.
- Measure of satisfaction: standard questionnaire and user preferences.

## 2.1 Task description

The main goal of monitoring the spatial dynamics through visual exploration is to gain insight of the processes and conditions that

lead to the occurrence of the phenomena. Here, geoscientist are mainly interested on learning more about the appearance of phenomena, relationship in space and time between image features, causes and trends or behavior of these features over time. Two case studies have been used. In the descriptions below a reference to the test questions is given:

- **Meteorology:** main task is monitoring, nowcasting convective clouds and studying the conditions in which the convections are formed. Nowcasting is short term forecasting about potentially dangerous weather phenomena that is usually ‘not caught’ by the long-term forecasting models. Convective clouds are highly dynamic phenomena that lasts for about 30 minutes (corresponding to 2 images of the MSG-2 satellite) and can cause severe weather conditions. Thus, the test tasks in exploring the convections must be related to the locating of convections: finding the temporal intervals when convections are formed (question 1), selecting clouds with the steepest and longest-lasting convections in each cluster (question 2) and observing the migration rates of those convections (question 3).
- **Coastal morphodynamics:** main task is to explore the evolution of rip channels in highly dynamic coastal zones. Rip channels are dangerous swimming sites and therefore, the exploration helps the users not only in warning about them, but also in understanding and associating rip channels with other relatively well known coastal phenomena: waves and sediment transport. Rip channels exist between storms because high energy waves during the storms flatten out any previous sediment accumulations. Rip channels also tend to be generated all at the same time (usually within 1 - 2 days), but disappear at different rates. Therefore the appearance boundary of each temporal cluster is generally sharper than the ragged disappearance boundary. Thus, the test tasks in rip channel exploration must be related to: finding temporal clusters (intervals) when the rip channels existed (question 1) and selecting rip channels with the longest lifetime in that clusters (question 2) and observing the migration rate and size variation of those samples (question 3).

The sequence of questions outlines the data exploratory levels: overview, zoom, details and were mainly oriented at temporal aspects of sensor data. Tasks were refined and agreed with the representative domain experts before the test execution.

## 2.2 Datasets

To reduce the learning effects, datasets explored were only used once. Therefore, for each case study, three datasets of different dates but with the same spatial coverage and same temporal duration were used. One datasets aimed for the tool familiarization stage to allow test participants to explore the user interfaces and menu options of the animation and the prototype. The next two were used for the actual test with varying succession of the animation and the prototype. Datasets had the following dimensions:

- three sets of coastal morphology images with resolution: 1201 pixels \* 281 pixels \* 50 days. These were extracted from the larger set of 280 images: days 170–220 were used for the for familiarization purpose; days 1–50 for the first tool and days 51-100 for the second tool;
- three sets of MSG satellite images for convective cloud test. Datasets had the following resolution: 500 pixels \* 600 pixels \* 33 frames and were acquired between 10-00 and 18-00

GMT on 27<sup>th</sup> July 2006 (for familiarization purpose), 23<sup>rd</sup> April 2007 (for the first tool) and 24<sup>th</sup> April 2007 (for the second tool) respectively.

### 2.3 Test persons

In order to evaluate the effectiveness of the selected visualization methods, a user group of both domain experts and novice users was chosen. The novice user group (11 persons) consisted mainly of PhD students working with time-series of remote sensing data; however none of them were exposed to coastal morphology data before. The expert group consisted of three PhD students involved in monitoring the rip channels on the daily basis. To get inside into the second case study, two meteorology experts were available for the test. Among all the tested persons, one expert and two novice users have used neither animation nor time-series of sensor data before.

### 2.4 Test sessions

The test sessions took place at different locations and times. On average, the test lasted for two hours and consisted of three sessions: introduction, tool familiarization and an actual test. During the introduction, which took approximately 20 minutes, each participant was told about the case study, the prototype, the objective of the test, and what participants were supposed to do during the test. The time scheduled for familiarization with the prototype was 25 minutes and about 5 minutes for ‘playing’ with the animation. Here participants were asked to explore on their own every part of the software, user interface and all menu options. After the initial familiarization stage, participants were exposed for 10 minutes to the sample tasks that represented the actual test questions. Thus, we made sure ‘that users understand the purpose of the tools and can operate them’ (Andrienko et al., 2002). After the test persons were acquainted with the tools, they were left to execute the test tasks independently. To reduce the tool learning effect and to discourage the cooperation between adjacent participants, the tool sequence was varied among the test persons. At the end, participants were asked to complete the questionnaire as their final session of the task.

## 3 TEST RESULTS

At the end, 14 full test results were obtained. Two out of 16 test persons gave up the test when they were using the animation. Results of one of the above drop out test persons (N11) were discarded in the following analysis. The rest 14 participants performed the test exhaustively and enthusiastically.

### 3.1 Effectiveness

Effectiveness is measured in terms of completeness of answers and the accuracy of the responses. In coastal morphology case, the tasks and answers were agreed with the domain expert before the test. Therefore, the effectiveness indicators were simply derived by comparing test answers to the expert’s knowledge. Figure 3 and 4 shows that the answers that were given by the test participants agreed well with the expert’s conclusions.

**3.1.1 Completeness** One major difficulty in measuring the task completion rate was related to the exploratory nature of questions and their *overview, zoom and details* hierarchical structure. Thus, the outcome of the first question (number of temporal clusters) would influence the completeness of the following questions (find the longest existed rip channel in each cluster). Therefore, to take into account the nature of question, the following weighted

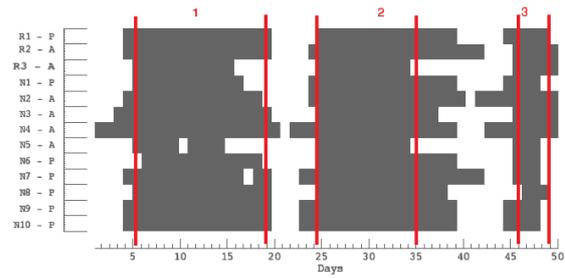


Figure 3: Temporal boundaries of rip clusters (in days) for first task in a sequence. Vertical lines indicate boundaries that were pre-defined by the expert. The Y axis refers to the test persons and tool they used in Task 1. Among test persons: ‘C’ refers to cloud expert, ‘R’ – to rip channel expert, ‘N’ – to novice test person respectively. ‘A’ denote the animation, ‘P’ – the prototype.

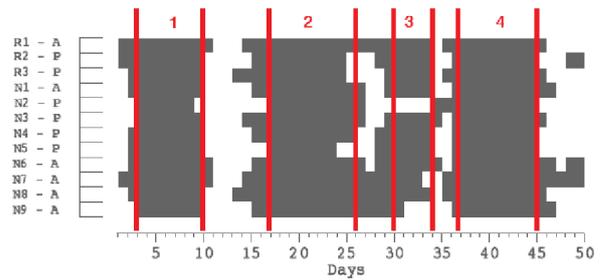


Figure 4: Temporal boundaries of rip clusters (in days) for the second task in a sequence. Vertical lines indicate boundaries that were pre-defined by the expert. The Y axis refers to the test persons and tool they used in Task 1. Among test persons: ‘C’ refers to cloud expert, ‘R’ – to rip channel expert, ‘N’ – to novice test person respectively. ‘A’ denote the animation, ‘P’ – the prototype.

scheme was adopted: 1 (100% rate) shows fully answered question; rates greater than 1 (greater than 100%) indicate that a test person provided several alternative options or additional temporal cluster(s) (in % equal to a number of alternatives provided, Table 1).

	TP	AQ1	AQ2	AQ3	PQ1	PQ2	PQ3
1	C1	0.5	0.25	0.25	1	0.3	0.3
2	C2	0.3	0.3	0.3	1	1	1
3	R1	0.75	0.75	0.75	1	1	1
4	R2	1	1	1	1.2	1.2	1.2
5	R3	1	1	1	1	1	1
6	N1	1	1	1	1	1	1
7	N2	1	1	0.67	0.75	0.75	0.62
8	N3	1	1	0.67	1	1	0.67
9	N4	1	1	1	1	1.5	1
10	N5	1.2	0.5	0.25	1	0.5	0.25
11	N6	1.2	1.2	1.2	1	1	1
12	N7	0.75	0.75	0.75	1.2	1.2	1.2
13	N8	0.75	0.75	0.75	1	1	0.84
14	N9	0.75	0.75	0.75	1	1	1
16	N10	0	0	0	1	1	1
	<b>Average</b>	<b>0.81</b>	<b>0.75</b>	<b>0.69</b>	<b>1.01</b>	<b>0.97</b>	<b>0.87</b>

Table 1: Tasks completion rate (see text for explanations). ‘PQ’ implies ‘Prototype Question’, whereas ‘AQ’ means ‘Animation Question’. Test groups denote as: ‘C’ – cloud expert, ‘R’ – rip channel expert, ‘N’ – novice test person respectively.

Table 1 shows that excluding those who quit the test, 6 out of 14 participants (including 3 experts) indicated less temporal clus-

ters while they used the animation as their exploratory tool. Only 1 such test person provided less temporal cluster when he used the prototype (N2). Also, 4 out of 14 participants (including 2 experts) gave extensive answers by providing several alternatives (N4, R3) or by providing an extra temporal cluster (R2, N7) when they used the prototype. Similarly, 2 out of 14 tested people presented an extra temporal cluster while utilizing the animation (N5 and N6). Consequently, the average task completion rate was higher when participants were using the prototype (Table 1).

**3.1.2 Accuracy of answers** Figure 5 shows errors of detecting temporal clusters. Half of the participants had higher error rate in the animation than in the prototype.

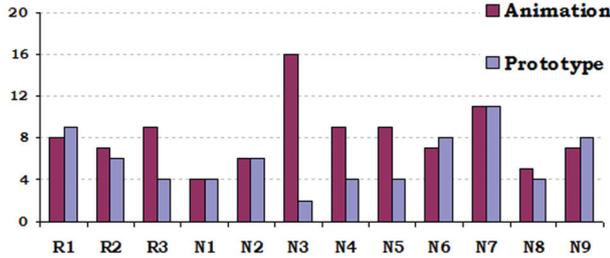


Figure 5: Cumulative errors (in days) in defining the appearance and disappearance borders of rip channels.

The same trend is observed in the mean error values per task sequence and per tool (Table 2). The absolute values of the mean errors have decreased with the second task showing that the participants have learnt the phenomena in their first task and were more accurate in the following task. On contrary, the relative error difference between the tools in the second task have increased. Although it was influenced by a few outliers (i.e., N7 and N3), the trend remained constant even when those results were excluded.

Task 1		Task 2	
Animation	Prototype	Animation	Prototype
R2	9	R1	7
R3	7	N1	9
N2	6	N6	8
N3	16	N7	11
N4	9	N8	4
N5	9	N9	8
	N10		9
<b>Mean</b>	<b>9.3</b>	<b>Mean</b>	<b>6.8</b>
	<b>8</b>		<b>4.3</b>

Table 2: Cumulative errors (in days) per task and per tool in defining the appearance and disappearance borders of rip channels.

Similar error tendency is observed when the participants were locating the rip channel with the longest lifetime (Figure 6 and Table 3).

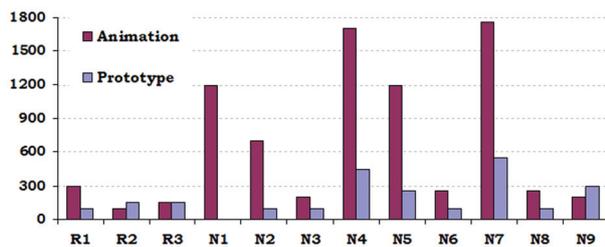


Figure 6: Cumulative errors (in meters) of locating the longest existed rip channel in each temporal cluster.

Task 1		Task 2	
Animation	Prototype	Animation	Prototype
R2	150	R1	100
R3	100	N1	0
N2	700	N6	100
N3	200	N7	550
N4	1700	N8	100
N5	1200	N9	300
	N10		200
<b>Mean</b>	<b>675</b>	<b>Mean</b>	<b>192.9</b>
			<b>633.8</b>
			<b>200</b>

Table 3: Cumulative errors (in meters) per task and per tool of locating the longest existed rip channel in each temporal cluster.

Results of the test that involved identifying appearance/disappearance intervals of convective clouds confirm the trend of the coastal data exploration. Both cloud experts delineated less temporal clusters (Figure 7) and demonstrated difficulty in locating the clouds with the steepest and longest-lasting convections while using the animation.

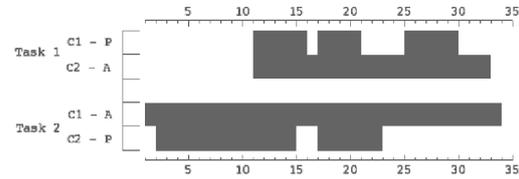


Figure 7: Experts' answers on appearance/disappearance clusters of convective clouds.

Accuracy analysis of the third question concerning the migration rates of the selected image features (and their size variation in rip channels case) had insignificant differences between the tools. This is largely due to the fact that these spatial properties were easy to observe in the animation (which is also included in the prototype in the form of Image Viewer and 3d Event Viewer, see Figure 1) and that the prototype had also additional option of displaying the coordinates of the queried image features.

## 3.2 Efficiency

**3.2.1 Timing** In the test, the actual answer time and the completeness of answers are related issues: incomplete answer would require less time. In order to make the timing comparable, the actual answer times must be compensated by the completeness factors presented on Table 1. Average actual answer time per tool session was 23.5 minutes with periods ranging from 8 minutes to maximum of 40.8 minutes. A method to compare the answer times and to avoid the individual differences is to take ratio of the first session period to the second. Table 4 summarizes the proportions for both actual (row) and compensated answer times. It shows that the novice test persons ( $n=9$ ) generally had longer duration with the first tool than they had with the second, disregard of the tool – average ratio of first tool usage time to the second is 1.72 and 1.53 for raw and compensated intervals respectively, see Table 4. This seems to be obvious as they were familiarizing with the phenomena in their first session. However, when the sequence was ‘animation–prototype’, novice participants gave faster answers with the prototype than the average ratio (Mean A/P = 2.05,  $n=4$ , Table 4). On contrary, when the sequence was ‘prototype–animation’, novice tester responded slower with the animation than the average ratio (Mean P/A = 1.45,  $n=5$ ). This trend also preserved when the answer times were weighted by the completeness factors (Table 4).

	Raw time		Compensated time	
	Novice	Expert	Novice	Expert
Total Mean	1.72	1.0	1.53	0.96
Mean of A/P	2.05 (n=4)	0.84 (n=3)	1.83	1.33
Mean of P/A	1.45 (n=5)	1.24 (n=2)	1.29	0.95

Table 4: Summary of answer times: Mean ratio of first tool time to the second. ‘A’ denotes the animation, ‘P’ indicates the prototype.

In raw times, the expert participants demonstrated the opposite tendency – they were slower responding while using the prototype than the average value. This implies that the experts are generally accustomed to the exploring the phenomena using animations. However, the difference between two sequences in experts’ group disappeared after the completeness weights have been applied: average ratio became around 0.96, mean of A/P was 1.33 – but largely due to an expert participant that had no previous experience in using the animation (C2). Excluding that person, all the compensated proportions on the right side of Figure 8 are around 1, indicating that the experts could learn the prototype in 25 minutes time and were able to utilize the prototype as fast as they did the animation.

The ratio of the first tool time usage to the second also varies within the test questions. Figure 8 shows that the first question had a highest differences in time between two tool sequences (n=14) and this difference was gradually diminishing towards the third question. This indicates the trend that the first question took longer time to answer with the animation than with the prototype (Mean A/P is large than the overall mean); towards the third question participants were able to get answers as fast, or even faster in raw times, with the animation than with the prototype.

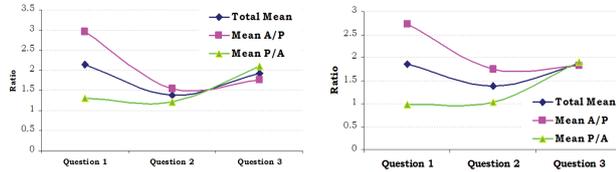


Figure 8: Summary of answer times per question for all participants (n=14): mean ratio of first tool usage time to the second. Raw times are displayed on the left graph, compensated times are on the right.

**3.2.2 Tool usage patterns** Tool usage patterns of the prototype were derived from the mouse movement observations. Table 5 summarize the observations and show that 2d Event Viewer (2dEV) was used most frequently (96.2 minutes), followed by the Attribute Viewer (AV–90 minutes). Their combination was also the preferred strategy in answering the test questions. As expected, the 3dEV was used more frequently in case of convective clouds (Table 5). Mouse movement data revealed that cloud experts utilized the 3dEV mainly in the beginning of the session to check the cloud tracking results. For that purpose, coastal experts mainly relied on the IV. Once the tracking was confirmed, experts diverted their attention to the interactive options of 2dEV and AV graphs.

Only 2 (novice) participants (N3 and N2) out of all 16 tested persons used no interactive menu options at all. Highlighting objects in several views (without filtering other objects) was the most frequently exploited option (109.6 minutes). ‘Behavioral’ brushes – selecting image objects by their lifetime or by gradients in their history – was the second preferred option. Brushing objects by their IDs was mainly utilized when participants were

	3dEV	2dEV	AV	IV
C1	2.9	8.8	6.9	1.7
C2	2.3	7.7	19.7	2.6
R1		6.9	6.5	3.3
R2		5.8	5.3	
R3		4.9	10.2	3.8
N1	10.8	3.5	6.4	0.5
N2	1.9	3.5		
N3		9.4		
N4	1.5	6.3	7.9	0.9
N5		2.8	3.0	
N6	12.3	8.2	3.0	2.9
N7	6.6	11.8	8.6	8.6
N8	0.4	13.3	2.1	2.0
N9	2.3	3.4	10.6	
<b>Total</b>	<b>40.8</b>	<b>96.2</b>	<b>90</b>	<b>26.1</b>

Table 5: Time (in minutes) spent by the test persons using the views of the prototype.

answering the third question. The less attended interactive options were ‘Temporal brushing’, ‘Attribute mapping’ and ‘Animate’ (both manually sliding images and rate-controlled animation). Except of one novice participant, the ‘Animate’ option was mainly used by the expert group to examine the tracking results or to confirm their answers. Mouse movement data indicate, that the novice participant (N7) that used the animation functionality of the 3dEV applied it in manual mode to measure the appearance/disappearance boundary of the selected objects. The rest of the novice participants completely ignored the dynamic aspect of animation in 3dEV and in IV views.

### 3.3 Satisfaction and Preferences

All participants recorded high overall level of satisfaction with the prototype. Average satisfaction score was 4.3 on scale from 1 – totally unsatisfactory, 2 – unsatisfactory, 3 – neutral, 4 – moderately satisfactory, to 5 – satisfactory. Satisfaction with the prototype compared to the animation varied for each test question. Participants were satisfied with the prototype while answering the overview question (average score is 4.3), equally or slightly more enjoyed when comparing several image features (4.4) and had the lowest satisfaction scores (i.e., moderately satisfying – 4.0) for the question associated with focusing on the selected object and observing properties and movements.

As it was mentioned earlier, two out of 16 test persons gave up the test when they were using the animation. Interestingly, it happened only when they had the animation as their second tool. The causes for the drops out were: ‘Difficult to accomplish this task using the animation’ or ‘Became too confused’.

Regarding the comments, rip channel experts were mainly concerned with the connection to the extraction stage (‘Interactive link to the extraction algorithm’) and choice of parameters associated with the rip extraction stage. None of the two cloud experts commented about the extraction of precipitating cloud. Interestingly, none of the novice and experts participants commented on the tracking algorithm and accuracies associated with the tracking.

Various comments regarding the user interface, interactive options and better integration between the views of the prototype were received.

#### 4 DISCUSSIONS AND CONCLUSIONS

Overall, the test supports the notion that the image feature representation should be maintained in the exploratory visualizations. In contrast to the animation (which only represent elementary feature reading level), prototype's feature representation gives access to feature's intermediate and overall reading levels. These levels assist the user in searching for features of interest, in comparing several image features and in perceiving the the overall trend of evolving features in the images.

The first two questions of the test were specifically targeting these levels. 6 out of 14 participants supplied incomplete answers by indicating a lower number of temporal intervals while using the animation. This trend is also shared by the expert group (3 out of those 6 were experts). Average time responses were faster using the prototype for the first two questions. Average error rate in detecting the boundaries of those clusters was higher in the animation. In the question that involved the comparison of several image features, participants of both groups who used the animation had the higher error rates in locating the longest existed rip channel. The same trend can be confirmed in experts that were detecting convective clouds in the animation: less temporal intervals and difficulty in comparing and locating the longest–lasted and the steepest convections.

However, on the elementary (spatial) reading level – i.e., focusing on particular image object and following its migration rate, there was no apparent tendency. Participants responded equally fast and equally accurate with both tools. Also, according to the satisfaction levels, this task has received the lowest satisfaction score in the prototype (moderately satisfying compare to the animation).

The participants had a short exposure to many interactive options and multiple views of the prototype. Nevertheless, 14 out of 16 participants actively exploited menu options and combinations of various visual queries. With respect to the comments received, the rip channel experts were mainly concerned with the prototype's connection to the extraction stage and with the choice of the parameters associated with the rip channel classification. This shows the direction to further research: that a fully integrated visual exploratory environment should allow users access to the extraction stage, change parameters or extraction methods (such as for instance in (Lucieer and Kraak, 2004)), and see the effects on tracking results in real time.

To summarize: the test participants provided faster and more complete and accurate answers for the first two questions using the prototype than the animation. They were also more satisfied with the prototype than with the animation when answering these questions. The results support the initial hypothesis – that representing image features and their evolution assist users in comparing several image features, in perceiving the overall trend of evolving features and in locating the features of interest in the time-series. However, more extensive testing is needed, especially among the expert users, to gain insight into the experts' usage of the prototype and their problem–solving strategies.

#### REFERENCES

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J. and Kretchmer, U., 2002. Testing the usability of interactive maps in commongis. *Cartography and Geographic Information Science* 29(4), pp. 325–342.

Bertin, J., 1983. *Semiology of graphics : diagrams, networks, maps*. The University of Wisconsin Press, Madison.

Blok, C., 2005. *Dynamic visualization variables in animation to support monitoring of spatial phenomena*. PhD, Universiteit Utrecht/International Institute for Geo-Information Science and Earth Observation ITC.

Fabrikant, S. I. and Goldsberry, K., 2005. Thematic relevance and perceptual salience of dynamic geovisualization displays. In: *Proceedings of 22nd ICA international cartographic conference: mapping approaches into a changing world*, A Coruna, Spain.

Hornbaek, K., 2006. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies* 64(2), pp. 79–102.

Koussoullakou, A. and Kraak, M.-J., 1992. Spatio-temporal maps and cartographic communication. *The cartographic Journal* 29, pp. 101–108.

Lloyd, R. E., 2005. Attention on maps. *Cartographic Perspectives* (52), pp. 27–57.

Lowe, R. K., 1999. Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education* 14, pp. 225–244.

Lucieer, A. and Kraak, M.-J., 2004. Interactive and visual fuzzy classification of remotely sensed imagery for exploration of uncertainty. *International Journal of Geographical Information Science* 18(5), pp. 491–512.

Mennis, J. and Peuquet, D., 2003. The role of knowledge representation in geographic knowledge discovery: A case study. *Transactions in GIS* 7(3), pp. 371–391.

Morrison, J. B., Tversky, B. and Betrancourt, M., 2000. Animation: Does it facilitate learning? In: *AAAI Spring Symposium on Smart Graphics*, Stanford, California, pp. 53–59.

Nielsen, J., 1993. Iterative user-interface design. *IEEE Computer* 26(11), pp. 32–41.

Peuquet, D., 1999. Chapter 8: Time in GIS and geographical databases. In: P. Lonely, M. Goodchild, D. Maguire and R. D.W. (eds), *Geographical information Systems, Volume 1. Principles and Technical Issues*, 2 edn, Vol. 1, John Wiley and sons, inc., New-York.

Rana, S. and Dykes, J., 2003. A framework for augementing the visualization of dynamic raster surfaces. *Information Visualization* (2), pp. 126–139.

Turdukulov, U. D. and Kraak, M. J., 2005. Visualization of events in time-series of remote sensing data. In: *Proceedings of 22nd ICA international cartographic conference: mapping approaches into a changing world*, A Coruna, Spain.

Turdukulov, U., Kraak, M.-J. and Blok, C. A., 2007. Designing a visual environment for exploration of time series of remote sensing data: in search for convective clouds. *Computers & Graphics*.

Ware, C., 2004. *Information Visualization. Perception for Design*. 2 edn, Morgan Kaufmann/Elsevier, Amsterdam.