Eurographics / IEEE Symposium on Visualization 2011 (EuroVis 2011)
H. Hauser, H. Pfister, and J. J. van Wijk
(Guest Editors)

Volume 30 (2011), Number 3

Evaluation of the Visibility of Vessel Movement Features in Trajectory Visualizations

Niels Willems[1], Huub van de Wetering[1], and Jarke J. van Wijk[1]

[1]Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands

Abstract

There are many visualizations that show the trajectory of a moving object to obtain insights in its behavior. In this user study, we test the performance of three of these visualizations with respect to three movement features that occur in vessel behavior. Our goal is to compare the recently presented vessel density by Willems et al. [WvdWvW09] with well-known trajectory visualizations such as an animation of moving dots and the space-time cube. We test these visualizations with common maritime analysis tasks by investigating the ability of users to find stopping objects, fast moving objects, and estimate the busiest routes in vessel trajectories. We test the robustness of the visualizations towards scalability and the influence of complex trajectories using small-scale synthetic data sets. The performance is measured in terms of correctness and response time. The user test shows that each visualization type excels for correctness for a specific movement feature. Vessel density performs best for finding stopping objects, but does not perform significantly less than the remaining visualizations for the other features. Therefore, vessel density is a nice extension in the toolkit for analyzing trajectories of moving objects, in particular for vessel movements, since stops can be visualized better, and the performance for comparing lanes and finding fast movers is at a similar level as established trajectory visualizations.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/Methodology

1. Introduction

The movement of objects has been investigated by analysts for decades. They search for movement features, such as meetings, flocks, congestion, or hotspots. For a taxonomy see Dodge *et al.* [DWL08]. Many approaches exist to find these features, such as geometric algorithms [AGLW08], reasoning [WvHdV*10], clustering [AAR*09], and visualization [ZFH08]. We focus on visualization as a means to extract vessel movement features in trajectories, and measure the performance among different visualizations.

Many types of visualizations can be used to display moving objects, and most of them have in common that they are an aggregated view on the (x,y,t)-space as defined by Andrienko and Andrienko [AA10]. We evaluate the performance of a dot animation, the space-time cube, and vessel density [WvdWvW09]. In terms of aggregated views, an animation of dots moving along the position of the objects is an aggregated view of all current locations (x,y) for varying times t. The other two visualizations are an aggregated view showing the complete trajectories, either in a 3D representation with separation of time for a space-time cube, or in a single plane aggregating time for vessel density. Each of these views allow a user to determine whether or not movement features occur in the trajectories, but some views are expected to perform better than others for certain features.

For common maritime tasks in the analysis of vessel trajectories, we are interested in the performance of vessel density with respect to other well-known trajectory visualizations. The response time and correctness is investigated for solving tasks matching with the following vessel movement features: *co-location in space*, *stops*, and *fast movement*. In vessel traffic monitoring, co-located ships appear in sea lanes and stops occur in front of harbors. Stops should be noticed in sea lanes to avoid dangerous situations. Fast movement at sea may indicate a potential threat of intruders. Furthermore, real-world data consist of many trajectories that are tracked with slightly inaccurate measuring devices. Therefore, we also test for the influence of complex trajectories and scalability of these visualizations. In our understanding these properties have not been tested before.

The paper is organized as follows: First, we elaborate on existing work in Section 2. In Section 3, we describe how a single trial in the experiment consisting of data and visualizations is created. In Section 4, we explain how trials are composed into a user test and explain how this test is executed. The test results are analyzed in Section 5. Finally, in Section 6, we draw conclusions and suggest future work.

2. Related Work

In this section we elaborate on existing work regarding moving object data visualizations and on user studies conducted to the two alternative visualizations, space-time cube and animation, which we investigate next to vessel density.

2.1. Moving Object Visualizations

Apart from the visualization methods as studied in this paper there are various other ways to display the trajectories of moving objects. Demšar *et al.* [DV10] combine a density and space-time cube approach in a single volume rendering to reduce ambiguity in the temporal aggregation of trajectories. World-wide trajectories and their aggregates are displayed by Grundy *et al.* [GJL*09] in a special version of the space-time cube where the cube is folded around the globe. Romero *et al.* [RSSA08] use the space-time cube to analyze activities in a video stream, while Eccles *et al.* [EKHW07] use it for story telling based on events. Hurter *et al.* [HTC09] introduce a GPU-tool to visualize various attributes along airplane trajectories using interactive masking and rolling the dice paradigm. Multiple attributes along trajectories of vessel trajectories are aggregated by Scheepens *et al.* [SWvdWvW11] in density maps. Bak *et al.* [BMJK09] show multivariate summaries of events in mouse trajectories with growth ring maps. Zhao *et al.* [ZFH08] find repeating patterns by displaying trajectories along a circular diagram.

2.2. Evaluation of the Space-time Cube

Demissie [Dem10] evaluates the space-time cube against static maps by means of a single map with polylines or small multiples with polylines, and an animation of dots moving along a polyline. Upto four trajectories from pedestrians walking in a city are analyzed. Subjects count the following movement features: stops, returns along a same path, returns to a same location, and speed changes. Based on 16 participants they conclude that animation excels for most features, since it captures the temporal dynamics best.

Patterns of pedestrian movements are evaluated by Kristensson *et al.* [KDA*09] in a space-time cube and a map with polylines annotated with labels containing timestamps. Subjects solve spatio-temporal queries with various levels of complexity for upto four trajectories. Based on the data captured from 30 participants, they conclude that the space-time cube is more prone to error for simple queries with equal response time, while for complex queries the response time for the space-time cube is better. In future work they suggest to test the space-time cube for data with higher density.

In virtual environments, the space-time cube can be rendered in either strong or weak 3D. Kjellin *et al.* [KPSL10] conduct a user study to see the effect between these rendering techniques for finding the largest spread between various disease propagation. The test is executed by 32 participants, from which they conclude that there is no significant difference between the two rendering techniques.

The space-time cube paradigm can also be used for displaying sequences of events by using abstract *xy*-axes, such as participants and activities. In a user study conducted by Vrotsou *et al.* [VFC10] the effect of 2D versus 3D representations is studied for different camera positions of the space-time cube: front view (2D), side view (2D), and rotation along a single axis as interaction (3D). The task is to find simultaneously occurring events in different sequences. Based on the data of 20 participants, they argue that 3D works significantly better than 2D, not only for simple tasks containing small number of sequences, but also for more complex tasks containing more sequences. Moreover, they expect that 2D results in a memory overload for complex tasks.

Our test differs from these studies, since we investigate properties of real-world data, such as a high density and visual clutter, and we investigate other movement features.

2.3. Evaluation of Animation

Animated graphics can be used to show, among others, state transitions, view transitions, explanations of how something works, and trends in data. By animating the movement of objects, we may understand the situation and judge whether or not the movements were as expected, or were normal. Many user studies comparing the effectiveness of animations and static representations have been conducted. Tversky *et al.* [TMB02] studied several of them that concern animations used in teaching complex systems, and concluded that animation has not be proven to be more effective. They suggest the *congruence principle* and the *apprehension principle* to come to at least successful animated graphics. Gonzalez [Gon96] evaluates the effectiveness of animation in decision making and suggests to use realistic graphics, smooth transitions, and an appropriate interaction style. In their study of effectiveness of animation in trend visualization Robertson *et al.* [RFF*08] compare animated bubble charts to two static visualizations and find that animations are both more time consuming and less accurate. On the other hand they find that the users preferred working with animations.

3. Scene Design

We define a scene to be a combination of a data set and a visualization. In Section 3.1 we explain how we can generate variations of trajectory data and in Section 3.2 we elaborate on various well-known trajectory visualizations.

3.1. Data Generation

For our experiment, we prefer synthetic data over real-world data to control the properties to be tested. We aim for realistic data, which leads to the following requirements, based on observations of vessel traffic:

1. vessels typically do not travel along the same path;
2. vessels hardly occur at the same time and position;
3. vessels often move in straight one-way sea lanes;
4. a cross-section of a sea lane has a normal distribution,
5. vessels move with a nearly constant speed.

We generate data with these requirements and vary the number of trajectories $\mathcal{C}$, the visual clutter $\mathcal{N}$, and different movement features $\mathcal{M}$. In Appendix A a detailed model is described for generating this synthetic data.

The number of trajectories $\mathcal{C}$ is varied to test for busy lanes. We have defined presets $\mathcal{C}_{sparse}$ and $\mathcal{C}_{dense}$, with 4 and 15 trajectories, where $\mathcal{C}_{sparse}$ is chosen to trigger preattentive vision, while $\mathcal{C}_{dense}$ forces a possible cognitive overload as a result of a high density of trajectories at one location. $\mathcal{C}_{dense}$ has a similar local density as real-world data as we encounter in the maritime domain.

For visual clutter $\mathcal{N}$, we distinguish simple $\mathcal{N}_{simple}$ and complex $\mathcal{N}_{complex}$ trajectories. Simple trajectories are associated with inert large tankers that tend to move along straight lines. Complex trajectories appear in ship traffic during stormy weather, but also in animal and pedestrian movements. Combined complex trajectories result in entangled trajectories with more visual clutter. Figure 1 shows how our data generation compares to real-world data for visual clutter with vessels ($\mathcal{N}_{simple}$) and pedestrians ($\mathcal{N}_{complex}$).

Trajectories should include a certain movement feature $\mathcal{M}$ to create sensible tasks. The movement features are limited to those occurring in vessel traffic. We aim to evaluate the claims about strong properties of vessel density [WvdWvW09]: stoppers $\mathcal{M}_{stop}$, fast movers $\mathcal{M}_{fast}$, and lanes $\mathcal{M}_{lanes}$. Since vessel density does not support showing direction, we have excluded related movement features.

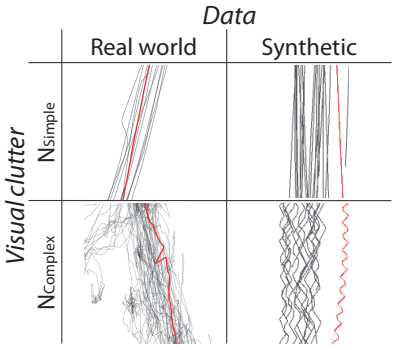

Figure 1: *Real-world versus synthetic movement data, with visual clutter resulting from simple (ships in a sea lane) or complex (pedestrians in a street [vdSvS08]) trajectories.*

Between scenes $\mathcal{C}$, $\mathcal{N}$, and $\mathcal{M}$ are varied. A scene is built with four lanes where all objects move in the same direction. Between lanes the number of stoppers for $\mathcal{M}_{stop}$, the speed in $\mathcal{M}_{fast}$, and the number of trajectories in $\mathcal{M}_{lanes}$ are varied. Vessel traffic is composed of many lanes. We do not address this complexity in our study, but focus on single lanes, to keep the experiment manageable. However, we argue that understandability of vessel traffic in single lanes is a prerequisite for understanding more complex scenarios.

3.2. Visualization

This section describes visualizations $\mathcal{V}$ that we investigate in our study: An animation of moving dots $\mathcal{V}_{anim}$, the space-time cube $\mathcal{V}_{cube}$, and vessel density $\mathcal{V}_{dens}$. We have chosen animation, since it appears often in vessel monitoring systems. The space-time cube is chosen because it is a visualization showing many movement features, can be used as a geographical visualization [Kra03], and is widely evaluated (see Section 2.2), which sets a baseline for our comparison. To avoid that extra features result in cross-effects, only the core elements of the visualizations are implemented.

3.2.1. Animation

Trajectories are animated by displaying a sequence of frames $[f_0,\ldots,f_T]$, where each frame f_t contains the position at a certain time t for all trajectories. The objects are rendered as circular dots with a diameter of 8 pixels. The whole animation takes about ten seconds. We do not allow any interaction, i.e., no pause and no rewind, since for the movement features $\mathcal{M}$ chosen there is no extra value: the animation is long enough to make a decision, hence a rewind action is not needed. For $\mathcal{M}_{stop}$ and $\mathcal{M}_{fast}$ a pause action is useless, since these features are not distinguishable anymore in a single frame. For $\mathcal{M}_{lanes}$ a pause action could be beneficial to count the dots, though the differences between lanes is large enough to see the difference without counting. The speed is tuned to be acceptable for an average person.

3.2.2. Space-Time Cube

In the seventies, Hägerstrand [Häg70] proposed to show trajectories in a 3D cube with two axes for space and one for time: the space-time cube . Trajectories are displayed as a 3D curve $(x(t),y(t),t)$, called the space-time path. The footprint, the projection of the space-time path on the (x,y)-plane, is omitted for possible cross-effects with vessel density, which is a static 2D map as well. In Figure 2 the space-time cube reveals a busy lane as a lane containing many space-time paths, a stop as a vertical segment in a space-time path, and a fast object as a space-time path with a flat slope. When using the space-time cube, occlusion is a problem that often occurs. For this reason we force occlusion by testing with four

lanes instead of two or three lanes. To overcome the resulting overdrawing we have used alpha blending with an opacity of 0.85, which shows a subtle highlight for dense areas. Furthermore, to keep the lanes distinguishable for comparison, each lane is rendered with perceptually balanced colors generated with PaletView by Wijffelaars *et al.* [WVvWvdL08]. Interaction is minimalized to overcome occlusion effects; the subject may rotate the cube with the mouse around the time axis and change the angle to look into the cube.

3.2.3. Vessel Density

The third visualization method is vessel density [WvdWvW09], which is a Kernel Density Estimation (KDE) [Sil92] where trajectories are convolved in the (x,y)-plane with a kernel moving along the trajectory with the same speed as the object. The density represents the sum of the durations that objects are in the neighbourhood. So, fast objects have a relatively low contribution and slow objects have a relatively high contribution. Just density is ambiguous, since many objects passing by also give a high contribution. To distinguish between a slow vessel and many fast vessels, two KDEs with two kernel sizes (large and small) are used simultaneously. The rendering then consists of a color mapped large kernel KDE and an illuminated height map of the sum of both a small and a large kernel KDE. By means of the colors we can see the global route structure, while in the shading detailed trajectories pop up, in particular stopping vessels, which are shown as dark dots in the shading (see inset in Figure 2). The ambiguity of density is solved by comparing the color mapped density with the number of trajectories in the shading.

4. Evaluation Setup

We have everything in place to generate trials. Table 1 shows an overview of the independent variables and their presets in our user test, as given in the previous section. In this section we generate trials and compose them in a test.

Name	Symbol	Presets
Movement feature	$\mathcal{M}$	$\mathcal{M}_{stop}$, $\mathcal{M}_{lanes}$, $\mathcal{M}_{fast}$
Visual clutter	$\mathcal{N}$	$\mathcal{N}_{simple}$, $\mathcal{N}_{complex}$
Trajectory count	$\mathcal{C}$	$\mathcal{C}_{sparse}$, $\mathcal{C}_{dense}$
Visualization	$\mathcal{V}$	$\mathcal{V}_{anim}$, $\mathcal{V}_{cube}$, $\mathcal{V}_{dens}$

Table 1: *An overview of the experiment parameters.*

4.1. Trial Design

The task in each trial is a multiple choice question that differs per movement feature (see Table 2). First the question is shown on a blank screen, and when the user presses the space bar, a scene with four numbered lanes is shown, and the response time measuring starts. For clarity the question is repeated at the bottom of the screen. To avoid reading mistakes the three questions are displayed in distinct colors. The subject chooses a lane by entering its number on the keyboard after which the response time is measured and the correctness is evaluated (correct scores 100%, wrong 0%).

Preset	Question
$\mathcal{M}_{fast}$	In which lane do objects have the highest speed?
$\mathcal{M}_{stop}$	Which lane contains the most stopping objects?
$\mathcal{M}_{lanes}$	Which lane contains the most objects?

Table 2: *Questions per movement feature.*

4.2. Test Design

The user test is a within-subject evaluation, is executed according to the schedule given in Table 3, and takes about one hour per subject. First, the subjects are instructed using a poster of Figure 2 containing screenshots of all the visualizations with all different movement features to explain the significant visual features. Then the subjects get familiar with the test environment using test trials with twice all combinations of visualizations and movement features for a data set with a small number of trajectories. In the first run the subjects are shown only two lanes to get familiar with the visualizations and in the second run four lanes to get familiar with the size of the scene in the actual test.

	Description	Duration
1.	Explain test with poster	10 min
2.	Test trials with feedback	10 min
3.	Test (180 trials à 10 sec)	30 min
4.	Fill out questionaire	10 min
	Total	60 min

Table 3: *The user test procedure.*

The test consists of a sequence of trials generated by all combinations of $\mathcal{V} \times \mathcal{M} \times \mathcal{N} \times \mathcal{C}$. These sequences are grouped in nine pairs of visualizations in $\mathcal{V}$ and movement features in $\mathcal{M}$. Each subject gets the sequence in a different order to compensate for learning effects and fatigue effects. For each pair, five groups of trials consisting of all four randomly ordered combinations of visual clutter $\mathcal{N}$ and trajectory count $\mathcal{C}$ are generated. This results in $3 \times 3 \times 2 \times 2 \times 5 = 180$ trials per subject. After completion of the test, the subject is interviewed by means of a questionaire about preferences for visualizations and possible reasons why the subject could not optimally perform.

4.3. Test Apparatus

The tests are conducted with a 20.1" LCD screen with 1680x1050 resolution and 60*cm* view distance. The framework is implemented using Java and JOGL, with shader programs for generating vessel density on-the-fly. The trials are shown in full-screen mode to avoid distractions.

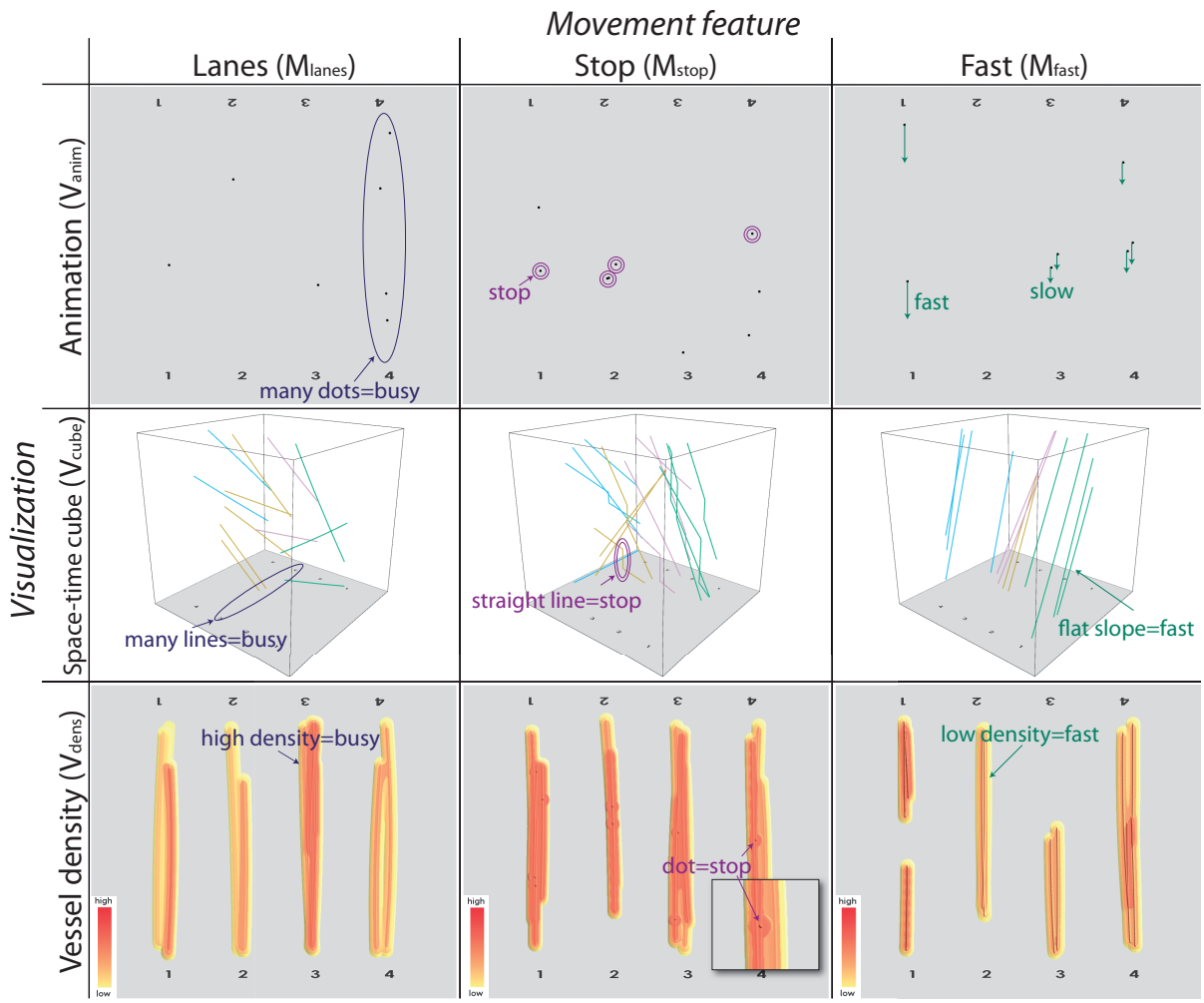

Figure 2: *A table showing all movement features with all visualizations. Each data set with a certain movement feature contains the same lanes with trajectories, but the order is changing in each trial and hence in each visualization shown.*

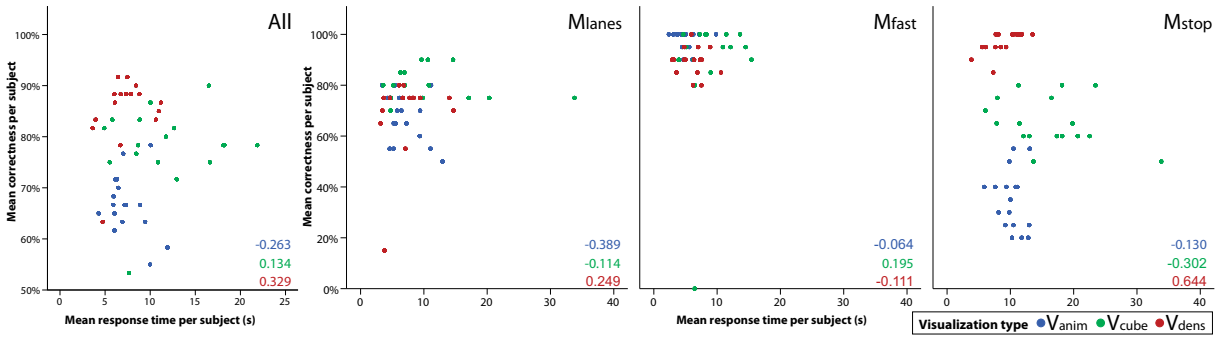

Figure 3: *Scatter plots with the mean correctness and mean response time per subject. The Pearson's correlation coefficients are shown in the bottom right corner. From left to right: All data and data grouped by $\mathcal{M}_{lanes}$, $\mathcal{M}_{fast}$, and $\mathcal{M}_{stop}$.*

4.4. Participants

The test is executed by students and staff of the mathematics and computer science department of our university. The age of the 17 participants ranges between 23 and 49 years old, all are right-handed, none of them is color-blind, and all are familiar with information visualization concepts. None of them is experienced with the space-time cube and two are experts in density visualizations.

5. Evaluation Results

In this section we evaluate the results from the user test described in the previous section.

5.1. Hypotheses

Our expectation of the performance of the visualizations is captured in the following hypotheses:

H1: There is a high positive correlation between response time and correctness.
We expect that if a visualization takes a longer exploration time, the answers are more likely to be correct.

H2: No visualization excels for all movement features.
A priori, we can see that some visualizations perform better than others: $\mathcal{V}_{anim}$ may perform best for detecting fast objects, and $\mathcal{V}_{dens}$ is likely to perform best for lanes, but has to compete with $\mathcal{V}_{cube}$ for stops.

H3: $\mathcal{V}_{cube}$ is (a) least robust to clutter, (b) least scalable.
The complexity introduced in the trajectory data as well as showing more trajectories will clutter the view on the space-time paths, resulting in less readability.

H4: For $\mathcal{M}_{stop}$, (a) $\mathcal{V}_{anim}$ performs worst, independent of the number of trajectories, (b) $\mathcal{V}_{cube}$ performs best for a small number of trajectories, while (c) $\mathcal{V}_{dens}$ performs best for a large number of trajectories, all independent of clutter.
We expect that a static representation will perform better than $\mathcal{V}_{anim}$, since animation lacks giving an overview. $\mathcal{V}_{cube}$ is expected to perform best for a small number of trajectories, since the vertical lines in the space-time path tend to be more visible than the small dots in $\mathcal{V}_{dens}$, though for many trajectories $\mathcal{V}_{dens}$ performs best, since then in $\mathcal{V}_{cube}$ the vertical lines are not visible due to occlusion.

H5: For $\mathcal{M}_{lanes}$, (a) $\mathcal{V}_{dens}$ performs best, followed by respectively (b) $\mathcal{V}_{cube}$ and (c) $\mathcal{V}_{anim}$, independent of the number of trajectories or clutter.
$\mathcal{V}_{dens}$ is expected to perform best, since it has a clean representation for lanes and counting happens instantly. $\mathcal{V}_{anim}$ is expected to perform least, since no overview is given.

H6: For $\mathcal{M}_{fast}$, (a) $\mathcal{V}_{dens}$ performs worst, independent of the number of trajectories, (b) $\mathcal{V}_{cube}$ performs best for a small number of trajectories, while (c) $\mathcal{V}_{anim}$ performs best for a large number of trajectories, all independent of clutter.
It is expected to be hard to see fast moving objects in $\mathcal{V}_{dens}$, since the smoothing differences in density are smaller than the differences in speed. $\mathcal{V}_{anim}$ is expected to perform best for a large number of trajectories, since there is a flow of dots appearing that catches the eye. This flow is missing for a small number of trajectories, therefore we expect the opposite order for $\mathcal{V}_{anim}$ and $\mathcal{V}_{cube}$.

5.2. Analysis

The performance of the test is measured in terms of correctness (overal score: $\mu = 77\%, \sigma = 42.3\%$) and response time in seconds (overal score: $\mu = 8.8s, \sigma = 8.3s$). In Figure 3 the mean correctness and mean response time per subject is displayed in scatter plots with different visualizations for all data and for each movement feature. In the scatter plot with all data we see clear clusters for $\mathcal{V}_{dens}$ performing good (high correctness in short response time) and $\mathcal{V}_{anim}$ performing less good (lower correctness than $\mathcal{V}_{dens}$ with low response time), while the spread for $\mathcal{V}_{cube}$ is high in both correctness and response time. By differentiating between movement features, we can see that the spread of $\mathcal{V}_{cube}$ is mainly because of a high spread for stops. Since for all data there is no high positive correlation per visualization, we reject H1.

The response time is not an appropriate measure in the performance analysis, since each visualization is explored differently due to interaction in $\mathcal{V}_{cube}$ and the obligation to watch $\mathcal{V}_{anim}$ for a while to be able to answer the question. Therefore, we only consider correctness as a measure in the performance analysis for the remaining hypotheses.

In Figure 4 the mean correctness is given for each visualization and grouped by movement feature, with a 95% confidence interval. The difference in correctness between the visualizations for stops is noticeably high, while for the other movement features the correctness seems to be in the same order of magnitude. Based on the mean correctness we can see that each visualization excels for one movement feature. In order to find out if the visualization with the highest correctness is significantly better than the others, we analyze the data with a one-way ANOVA test and a Tukey's HSD post hoc test for pairwise comparison. This test can be done for each movement feature, in fact for any subset of the data. As an example, we show a single complete ANOVA test for $\mathcal{M}_{lanes}$, which results in homogeneous subsets (HSs) of visualizations, i.e., visualizations without significant difference are contained in one homogeneous subset, while between these subsets a significant difference exists.

The ANOVA analysis with the subset of the data containing $\mathcal{M}_{lanes}$ shows that there is a significant effect for the mean correctness between the various visualizations, since $F(2,1017) = 7.785$, for $p < 0.001$. From Figure 4 for $\mathcal{M}_{lanes}$, we can see an ordering in the mean correctness for the three visualizations. To reveal whether the pairwise differences are significant, we conduct a Tukey's HSD test. The

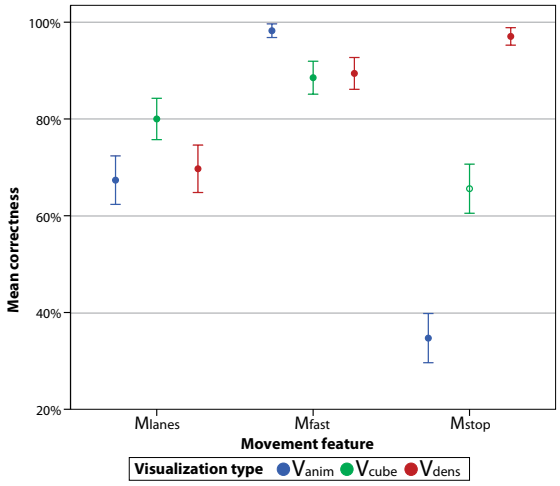

Figure 4: *The mean correctness with error bars (95% confidence interval) for visualizations per movement feature.*

results are shown in Table 4 where significant differences are printed bold. From this table we can derive the following HSs: $[\mathcal{V}_{cube},\{\mathcal{V}_{anim},\mathcal{V}_{dens}\}]$, which should be read as: $\mathcal{V}_{cube}$ has a significantly higher correctness than the others, while $\mathcal{V}_{anim}$ and $\mathcal{V}_{dens}$ do not have a significant difference.

	$\mathcal{V}_{anim}$	$\mathcal{V}_{cube}$	$\mathcal{V}_{dens}$
$\mathcal{V}_{anim}$	X	**5.39**	1.24
$\mathcal{V}_{cube}$	**5.39**	X	**4.14**
$\mathcal{V}_{dens}$	1.24	**4.14**	X

Table 4: *Tukey's HSD post hoc test results. For* $\alpha = 0.05$*, 3 treatments (visualizations), and a within-groups df>1000 (for our data df=1017), all values larger than 3.31 are shown in bold and have a significant difference.*

Similarly as in the paragraph above, we have computed HSs for all the combinations of the independent variables $\mathcal{M}$, $\mathcal{N}$, and $\mathcal{C}$, which are shown in the graph in Figure 5, with the result of the example displayed in the green node $\mathcal{M}_{lanes}$. The root *All* is shown in dark purple, representing all data, from which we branch in all movement features $\mathcal{M}$ colored in various colors, and we branch further in all the combinations of visual clutter $\mathcal{N}$ and trajectory count $\mathcal{C}$. Two nodes are connected if just one new variable is added to the list of variables in the direction from the root to a leaf. With this tree we can analyze the remaining hypotheses.

We cannot reject H2 for the correctness, since from nodes $\mathcal{M}_{stop}$, $\mathcal{M}_{lanes}$, and $\mathcal{M}_{fast}$ follows that each movement feature has a visualization that performs significantly better.

For the robustness to clutter in H3a, we have to look at the purple node *All*-$\mathcal{N}_{complex}$ with $[\mathcal{V}_{dens}, \mathcal{V}_{cube}, \mathcal{V}_{anim}]$, from which we conclude that $\mathcal{V}_{anim}$ is the least robust to clutter and therefore we reject H3a. For scalability in H3b, we inspect node *All*-$\mathcal{C}_{dense}$ with $[\mathcal{V}_{dens}, \{\mathcal{V}_{anim}, \mathcal{V}_{cube}\}]$ and conclude that indeed $\mathcal{V}_{cube}$ is in the least HS, though together with $\mathcal{V}_{anim}$, therefore we cannot reject H3b.

For stoppers in all red nodes, $\mathcal{V}_{anim}$ is always in the least performing HS, therefore we cannot reject H4a. For $\mathcal{M}_{stop}$-*-$\mathcal{C}_{sparse}$ nodes, $\mathcal{V}_{cube}$ is always in the best performing HS, though always together with $\mathcal{V}_{dens}$, therefore we cannot reject H4b. For $\mathcal{M}_{stop}$-*-$\mathcal{C}_{dense}$ nodes, $\mathcal{V}_{dens}$ is always the best performing visualization, therefore we can also not reject H4c. Combined with the observations in H4b, we can say that $\mathcal{V}_{dens}$ is the best performing visualization independent of visual clutter and the number of trajectories.

Regarding detecting lanes, $\mathcal{V}_{dens}$ is only in the best performing HS for $\mathcal{M}_{lanes}$-$\mathcal{N}_{complex}$-$\mathcal{C}_{dense}$, $\mathcal{M}_{lanes}$-$\mathcal{N}_{simple}$, and $\mathcal{M}_{lanes}$-$\mathcal{N}_{simple}$-$\mathcal{C}_{sparse}$, therefore we reject H5a. $\mathcal{V}_{cube}$ is for all green nodes for lanes in the best performing HS and $\mathcal{V}_{anim}$ is always in the least performing HS, therefore we reject H5b and we cannot reject H5c.

For finding fast movers, $\mathcal{V}_{dens}$ is only not performing worst for $\mathcal{M}_{fast}$-$\mathcal{C}_{sparse}$, and $\mathcal{M}_{fast}$-$\mathcal{N}_{complex}$-$\mathcal{C}_{sparse}$. Due to the dependence on the number of trajectories we should reject H6a. For all $\mathcal{M}_{fast}$-*-$\mathcal{C}_{sparse}$ nodes $\mathcal{V}_{cube}$ performs the worst, hence we reject H6b, while for all $\mathcal{M}_{fast}$-*-$\mathcal{C}_{dense}$ nodes $\mathcal{V}_{anim}$ performs best, therefore we cannot reject H6c.

Next to the analysis for the hypotheses we observe another noticeable significant difference. When not considering the differentiation between movement features, that are the nodes *All*-*, we can see that $\mathcal{V}_{dens}$ performs best overall, independent of the number of trajectories and independent of the complexity of the trajectories. From Figure 4 we should conclude that this high overall performance is due to a high mean performance of $\mathcal{V}_{dens}$ for stops, which cannot be compensated by the other movement features.

5.3. Discussion

From the analysis in the previous section we can conclude that $\mathcal{V}_{dens}$ performs best for $\mathcal{M}_{stop}$, $\mathcal{V}_{anim}$ for $\mathcal{M}_{fast}$, and $\mathcal{V}_{cube}$ for $\mathcal{M}_{lanes}$. The first combination seems to be logical, since stops are shown in a clean overview using density, while for the others stops are not always visible. That animation performs best for fast movements is logical as well, since it is easier to compare motion than comparing the other visual features. We are surprised that the space-time cube does not perform best, since comparing slopes tends to be easy as well, however in this particular visualization the occlusion and viewpoint might obscure the slope, as also mentioned by the subjects in the questionnaire. It is remarkable that the space-time cube works best for the lanes task, since the density is expected to perform best here, due to the nice overview of sea lanes in [WvdWvW09]. Reasons why density does not excel for comparing lanes could be that: (1) due

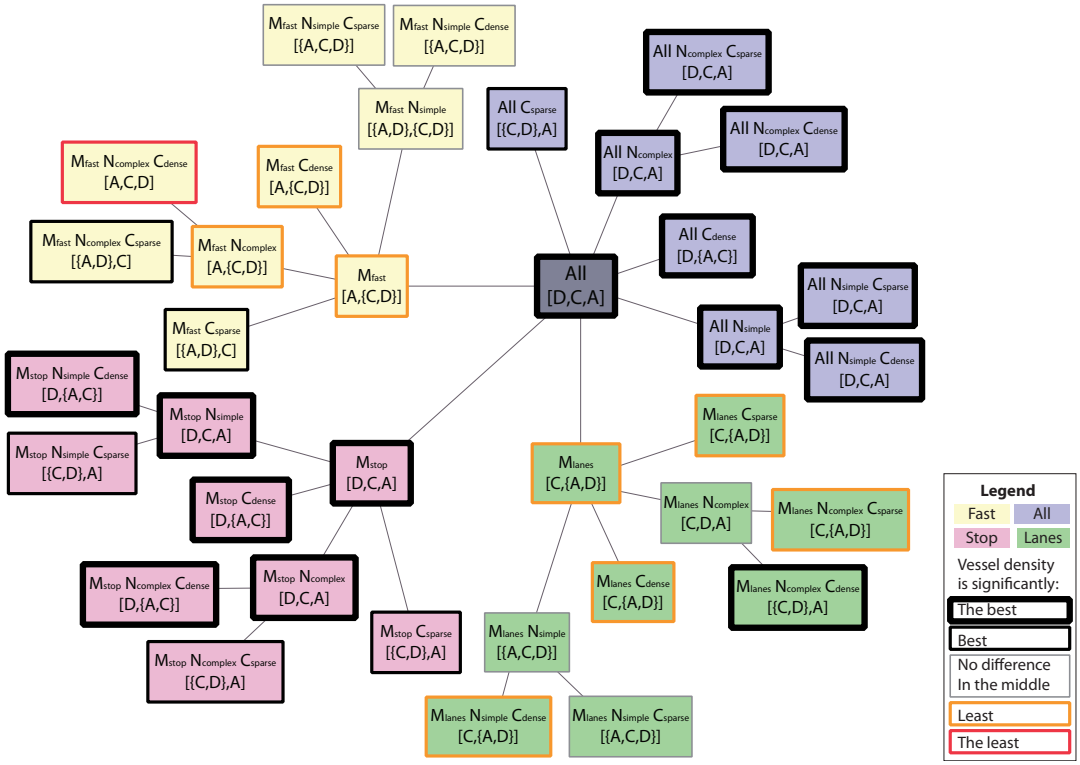

Figure 5: *A tree showing in each node subsets of the test data given by the filter criteria. For each node the homogeneous subsets of visualizations (A=$\mathcal{V}_{anim}$, C=$\mathcal{V}_{cube}$, D=$\mathcal{V}_{dens}$) are given, which are obtained by conducting an ANOVA test with correctness as a dependent variable and visualization as an independent variable with Tukey's HSD post hoc tests for* $\alpha = 0.05$.

to the smoothing, variations between the number of trajectories diminish; (2) due to the spatial variation, the lane with the highest number of trajectories may not have the highest density if the lane is slightly wider than the other lanes; (3) for a small number of trajectories, which may not cover the whole lane, holes may occur as in Figure 2 for $\mathcal{M}_{fast}$.

To come to a final judgement about the performance of vessel density in comparison to the other visualizations, we have annotated the nodes in Figure 5 with borders for vesseldensity being the best (fat black), best (thick black), least (orange), and the least (red). Based on these annotations we can conclude that density performs best for finding stoppers. For the other movement features density is competitive with the other visualizations, since only in a single case ($\mathcal{M}_{fast}$-$\mathcal{N}_{complex}$-$\mathcal{C}_{dense}$) density performs the least, whereas in all other cases it performs as good as another visualization. Furthermore, from feedback of the participants, density is often chosen to be the favorite visualization: overall 88%, for stoppers 100%, for lanes 82%, and for many, for simple, and for complex trajectories each 71%. Overview, no interaction, clear features, no constant attention while watching, and aesthetics are the reasons why participants preferred density.

This study is focussed on vessel traffic, but the chosen movement features also occur for other objects, such animals stopping at drinking places. However, vessel density excels for route-bound traffic, where straight lanes often occur, resulting in a strong visual cue. The results are likely to generalize to real-world data with the tested movement features occurring within a single lane. For not tested movement features it is unknown how vessel density will perform, but we know that features related to direction are invisible in vessel density. Since the tested movement features are not heavily dependent on direction, we expect the results to be similar for two-way lanes. For movements with more vivid speed variations than the assumed constant speed, the results are expected to differ, since for stops and fast movers it will be harder to determine what a stop or fast mover is. For busy lanes the results may be similar, since the temporal component is not much of influence. The settings for the number of trajectories $\mathcal{C}$ and visual clutter $\mathcal{N}$ are optimized for vessel traffic, and need recalibration for non-route bound traffic.

6. Conclusion and Future Work

We have presented a controlled user experiment to test how vessel density performs with respect to animation and the space-time cube for sets of simulated vessel trajectories with

properties of real-world data. Therefore, we have tested three different maritime movement features (stoppers, busy lanes, fast movers), for robustness with respect to complexity and scalability. For stops density performs significantly best, is robust to visual clutter and scales well. For the other features, density performs at least as good as the other visualizations.

For future work we would like to test vessel density with real operators to find out how useful it is in the real-world. Furthermore, in our test the settings for visual clutter and scalability are limited to two cases and we would like to find the critical values by increasing these number of cases. This was not yet possible due to the combinatorial explosion of four independent variables. A follow-up test with combinations of the visualizations, such as an animation with vessel density is of interest. Other movement features, such as congestions, should be tested to find out how vessel density performs in the whole spectrum of tasks. Finally, we are curious if other variables than visual clutter and scalability that play a role in finding movement features in real-world data.

Acknowledgements

We thank the participants for their time. This work has been carried out as a part of the Poseidon project at Thales Nederland under the responsibilities of the Embedded Systems Institute (ESI). This project is partially supported by the Dutch Ministry of Economic Affairs under the BSIK program.

Appendix A: Data Generation Model

Trajectory Model

A trajectory τ models the movement of an object and maps a finite time interval to a position in two-dimensional space. We model a trajectory with several parameters that are varied later on to create lanes, sets of trajectories. First we model the geometry of a 'horizontal' trajectory as a parametric function in $u \in [0,1]$ representing half a period of a sine: $(u, \sin(\pi\ u))$. We modulate this trajectory along its normal with an offset function $o(u) = A\sin(\omega u + \varphi)$ to gain more visual clutter with parameters amplitude A, frequency ω, and phase φ. To get a trajectory geometry with given start point and end point, an affine transformation is applied.

In a second step, we add timing information to the geometry. Assuming that we have a natural parametrization $\mathbf{p}(s)$ of the geometry and given a fixed velocity parameter v and a start time t_0 the timing of the curve is given by $\mathbf{p}((t-t_0)v)$. This can easily be extended for the additional timing parameters we require in this model: a starting position given by a time t_1, a stop time t_2, and a stop duration Δt. Note that a stop occurs just once in a trajectory to avoid ambiguity later on in the task of counting stopping objects.

For generating either simple trajectories or complex trajectories with visual clutter we have defined two presets $\mathcal{N}$: for simple data $\mathcal{N}_{simple} \equiv \{B = 0.02, A = 0\}$ and for complex data $\mathcal{N}_{complex} \equiv \{B = 0.05, A = 0.04, \omega = 2\}$.

Lane Model

A lane is a set of trajectories and takes the following parameters: the number of trajectories N, the number of stopping objects S, the average speed V of all objects, a start point $\mathbf{p}_0$, an end point $\mathbf{p}_1$, and a visual clutter preset for $\mathcal{N}$.

A lane is created by generating N trajectories according to preset $\mathcal{N}$ and with normally distributed start and end points with $\mu = \mathbf{p}_0$ and $\mu = \mathbf{p}_1$ respectively, and a fixed variance $\sigma_{\mathbf{p}} = (0.01, 0.01)$ ignoring points at a distance larger than 2σ. The begin time t_0 of a trajectory is normal distributed with $\mu = 0$ and with variance σ_t varying between 3 and 10, depending on the movement feature, which is described in the next paragraph. A uniformly distributed time t_1 is chosen for some of the trajectories that will not start in $\mathbf{p}_0$. The speed is chosen with a normal distribution with $\mu = V$ and small variance σ_v to obtain a more realistic movement profile within a lane. For S of these trajectories stopping time t_2 is chosen with a uniform distribution in the center 80% of the trajectory. The stop duration is chosen uniformly distributed between 4% and 15% of the total duration of the trajectory. The stopping parameters are chosen according to a pilot study to tune for clear visibility of the stop features, in particular in a space-time cube containing many trajectories.

In the user study we use three types of lanes containing different movement features defined by presets for $\mathcal{M}$: for lanes containing stopping objects $\mathcal{M}_{stop} \equiv \{V = v\}$, for lanes containing fast objects $\mathcal{M}_{fast} \equiv \{S = 0\}$, and for other lanes $\mathcal{M}_{lanes} \equiv \{S = 0, V = v\}$, where v is one-tenth of the distance between the endpoints of the lane per second, resulting in trajectories lasting approximately ten seconds.

Lane Composition

A scene is composed of four lanes, all with the same movement preset for $\mathcal{M}$, the same clutter preset for $\mathcal{N}$, and the same approximate number $\#\tau$ of trajectories given by a preset for $\mathcal{C}$. $\mathcal{C}$ has two presets: a sparse lane $\mathcal{C}_{sparse} \equiv \{\#\tau = 4\}$ and a dense lane $\mathcal{C}_{dense} \equiv \{\#\tau = 15\}$.

The begin points and end points of the lanes are given by $\mathbf{p}_0 = (0,0)$ and $\mathbf{p}_1 = (0,1)$, resulting in trajectories in a vertical configuration. For the movement features some of the lane parameters still need to be specified. Where these parameters differ between the four lanes, they are chosen such that two equal a so-called base level, while the other two differ by a factor γ or γ^{-1}. The order of the lanes is randomized. We have chosen $\gamma = 1.5$ to obtain subtle, but noticeable, differences between lanes. For $\mathcal{M}_{stop}$ the number of trajectories $N = \#\tau$ and the number of stops S has base level $\#\tau/\gamma$. For $\mathcal{M}_{fast}$ the number of trajectories $N = \#\tau$ and the speed V has base level v. Finally, for $\mathcal{M}_{lanes}$ the number of trajectories N has base level $\#\tau$.

References

[AA10] ANDRIENKO N., ANDRIENKO G.: A general framework for using aggregation in visual exploration of movement data. *The Cartographic Journal 47*, 1 (2010), 22–40.

[AAR*09] ANDRIENKO G., ANDRIENKO N., RINZIVILLO S., NANNI M., PEDRESCHI D., GIANNOTTI F.: Interactive visual clustering of large collections of trajectories. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology* (2009), pp. 3–10.

[AGLW08] ANDERSSON M., GUDMUNDSSON J., LAUBE P., WOLLE T.: Reporting leaders and followers among trajectories of moving point objects. *GeoInformatica 12* (2008), 497–528.

[BMJK09] BAK P., MANSMANN F., JANETZKO H., KEIM D.: Spatiotemporal analysis of sensor logs using growth ring maps. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 913–920.

[Dem10] DEMISSIE B.: *Geo-Visualization of Movements: Moving Objects in Static Maps, Animation and The Space-Time Cube*. VDM, 2010.

[DV10] DEMŠAR U., VIRRANTAUS K.: Space-time density of trajectories: Exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science 24*, 10 (2010), 1527–1542.

[DWL08] DODGE S., WEIBEL R., LAUTENSCHÜTZ A.-K.: Towards a taxonomy of movement patterns. *Information Visualization 7*, 3–4 (2008), 240–252.

[EKHW07] ECCLES R., KAPLER T., HARPER R., WRIGHT W.: Stories in geotime. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology* (2007), pp. 19–26.

[GJL*09] GRUNDY E., JONES M. W., LARAMEE R. S., WILSON R. P., SHEPARD E. L.: Visualisation of sensor data from animal movement. *Computer Graphics Forum 28*, 3 (2009), 815–822.

[Gon96] GONZALEZ C.: Does animation in user interfaces improve decision making? In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground* (1996), pp. 27–34.

[Häg70] HÄGERSTRAND T.: What about people in regional science? *Papers in Regional Science 24* (1970), 6–21.

[HTC09] HURTER C., TISSOIRES B., CONVERSY S.: FromDaDy: Spreading aircraft trajectories across views to support iterative queries. *Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1017–1024.

[KDA*09] KRISTENSSON P., DAHLBACK N., ANUNDI D., BJORNSTAD M., GILLBERG H., HARALDSSON J., MARTENSSON I., NORDVALL M., STAHL J.: An evaluation of space time cube representation of spatiotemporal patterns. *IEEE Transactions on Visualization and Computer Graphics 15*, 4 (2009), 696–702.

[KPSL10] KJELLIN A., PETTERSSON L. W., SEIPEL S., LIND M.: Different levels of 3D: An evaluation of visualized discrete spatiotemporal data in space-time cubes. *Information Visualization 9*, 2 (2010), 152–164.

[Kra03] KRAAK M.-J.: The space-time cube revisited from a geovisualization perspective. In *Proceedings of the International Cartographic Conference (ICC)* (2003), pp. 1988–1995.

[RFF*08] ROBERTSON G., FERNANDEZ R., FISHER D., LEE B., STASKO J.: Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1325–1332.

[RSSA08] ROMERO M., SUMMET J., STASKO J., ABOWD G.: Viz-a-vis: Toward visualizing video through computer vision. *IEEE Transactions on Visualization and Computer Graphics 14* (2008), 1261–1268.

[Sil92] SILVERMAN B. W.: *Density Estimation for Statistics and Data Analysis*. No. 26 in Monographs on Statistics and Applied Probability. Chapman & Hall, 1992.

[SWvdWvW11] SCHEEPENS R., WILLEMS N., VAN DE WETERING H., VAN WIJK J. J.: Interactive visualization of multivariate trajectory data with density maps. In *Proceedings of IEEE Pacific Visualization Symposium* (2011), pp. 147–154.

[TMB02] TVERSKY B., MORRISON J. B., BETRANCOURT M.: Animation:can it facilitate? *International Journal of Human-Computer Studies 57* (2002), 247–262.

[vdSvS08] VAN DER SPEK S., VAN SCHAICK J.: *Urbanism on Track - Application of Tracking Technologies in Urbanism*, vol. 1 of *Research in Urbanism Series*. IOS Press BV, 2008.

[VFC10] VROTSOU K., FORSELL C., COOPER M.: 2D and 3D representations for feature recognition in time geographical diary data. *Information Visualization 9*, 4 (2010), 263–276.

[WvdWvW09] WILLEMS N., VAN DE WETERING H., VAN WIJK J. J.: Visualization of vessel movements. *Computer Graphics Forum 28*, 3 (2009), 959–966.

[WvHdV*10] WILLEMS N., VAN HAGE W. R., DE VRIES G., JANSSENS J. H., MALAISÉ V.: An integrated approach for visual analysis of a multisource moving objects knowledge base. *International Journal of Geographical Information Science 24*, 10 (2010), 1543–1558.

[WVvWvdL08] WIJFFELAARS M., VLIEGEN R., VAN WIJK J. J., VAN DER LINDEN E.-J.: Generating color palettes using intuitive parameters. *Computer Graphics Forum 27*, 8 (2008), 743–750.

[ZFH08] ZHAO J., FORER P., HARVEY A. S.: Activities, ringmaps and geovisualization oflarge human movement fields. *Information Visualization 7*, 3-4 (2008), 198–209.