Composite density maps for multivarate trajectories

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

Download date: 23. Aug. 2022

Composite Density Maps for Multivariate Trajectories

Roeland Scheepens, Niels Willems, Huub van de Wetering,
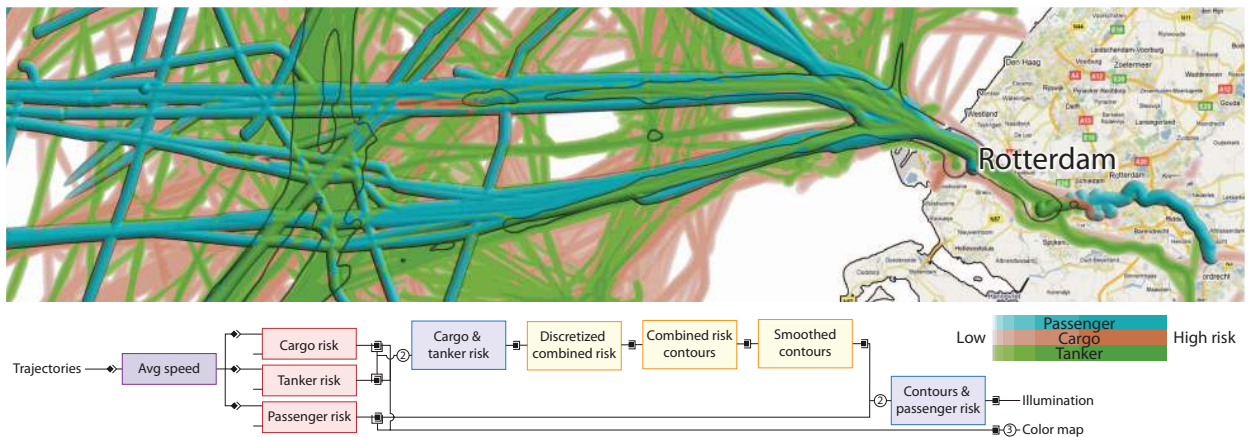Gennady Andrienko, Natalia Andrienko, and Jarke J. van Wijk

Fig. 1. An accident risk map of passenger vessels (turquoise), cargo vessels (orange), and tanker vessels (green) in front of Rotterdam harbor. The highest density at a location determines the color of the cell. The areas that need extra attention are highlighted in a separate illuminated height field containing the passenger risk map and contours as ridges. The contours are derived from a combination of tanker and cargo risks marking where they both have a high risk. The block diagram shows how the map is created.

Abstract—We consider moving objects as multivariate time-series. By visually analyzing the attributes, patterns may appear that explain why certain movements have occurred. Density maps as proposed by Scheepens *et al.* [25] are a way to reveal these patterns by means of aggregations of filtered subsets of trajectories. Since filtering is often not sufficient for analysts to express their domain knowledge, we propose to use expressions instead. We present a flexible architecture for density maps to enable custom, versatile exploration using multiple density fields. The flexibility comes from a script, depicted in this paper as a block diagram, which defines an advanced computation of a density field. We define six different types of blocks to create, compose, and enhance trajectories or density fields. Blocks are customized by means of expressions that allow the analyst to model domain knowledge. The versatility of our architecture is demonstrated with several maritime use cases developed with domain experts. Our approach is expected to be useful for the analysis of objects in other domains.

Index Terms—Trajectories, Kernel Density Estimation, Multivariate Data, Geographical Information Systems, and Raster Maps.

✦

1 INTRODUCTION

Object behavior is examined by analyzing their state as they move during a certain period. A state contains sampled sensor input. Additionally, data may be obtained from external data sources or by means of reasoning with existing knowledge [32]. The fused data sources result in multivariate trajectories: a time-series with multiple attributes per object capturing its movement. Visual inspection of attributes along these trajectories may reveal patterns that, for instance, may explain why movements have occurred.

Aggregation is a suitable technique for analyzing massive trajectory data sets [3]. A density map [25] is a visual aggregation method, based on kernel density estimation [26], that generalizes characteristics of multivariate trajectories. Many relevant analyses, such as anomaly detection and risk analysis, can be conducted with these density maps,

but their expressibility is limited. If the proposed fixed architecture is used, density maps only consist of predefined compositions of density fields generated from a subset of the trajectory data. For realistic analyses often more powerful expressions are needed.

We distinguish two user roles: analyst and operator. An analyst is a domain expert who defines, possibly complex, parameterized computations of density maps in order to find certain features. An operator, e.g., a coast guard surveillance operator, visually inspects a computed density map for occurrences of the features and interactively changes the parameters of its computation to facilitate the inspection.

To allow the analysts to express their interest in terms of density fields, the density contribution requires a more complex definition. First of all, attributes values could be taken into account, for instance as a weight for the contributions. Secondly, additional aggregations may be required, such as an average attribute value. Finally, local structures may need to be emphasized, for instance by highlighting areas where movements are vivid, such as meeting areas.

To increase the expressibility, we propose to extend the density field computation with a kernel with varying radius, in a similar fashion as Brunsdon [6], and a user-definable expression for the density contribution that may take other density fields into account. Furthermore, the analyst can combine multiple density field computations to compute a complex density field. With such density fields we can perform more analyses: First of all, we can compute non-spatial deviations from normal behavior, such as finding relatively slow objects moving

- R. Scheepens, N. Willems, H. v. d. Wetering, and J. J. v. Wijk are with Eindhoven University of Technology, The Netherlands, E-mail: {R.J.Scheepens, C.M.E.Willems, H.v.d.Wetering, J.J.v.Wijk}@tue.nl.
- G. Andrienko and N. Andrienko are with Fraunhofer Institute IAIS, Germany, E-mail: {Gennady, Natalia}.Andrienko@iais.fraunhofer.de.

against the main traffic flow. In fact, this is a contextual computation, where the environment is given by the data. Furthermore, we can show attribute variations in the density by relating attributes to the kernel radius. Finally, by defining the desired result in a composition of simple computations (e.g., bottom of Fig. 1), the analyst keeps an overview while building a complex density field. These improvements together enable the analysts and operators to investigate real-world situations.

In Section 2 we describe related approaches in literature. In Section 3 we define the basic models for trajectories and densities. Section 4 explains our approach to compute composite density maps, and in Section 5 we elaborate on the implementation. In Section 6 we apply our method to real-world use cases and in the final section we draw conclusions and suggest some future work.

2 RELATED WORK

Our contribution touches the literature in different ways. This includes visualizations of moving objects, and more specific, trajectory aggregation in the first two sections. User-defined expressions for building information visualizations are discussed in Section 2.3. In the last section we mention connections with geographical information systems.

2.1 Moving Object Visualizations

In the analysis of moving objects, we search for patterns that explain why a movement has occurred. Visualization is one of the methods that may reveal these patterns. Slingsby et al. [27] define a hierarchy based on attributes of trajectories and use tree maps to show the resulting distribution. We limit ourselves to spatial distributions given by the attributes. Trajectories of eye movements can be analyzed using a gaze plot, or a space-time cube as shown by Li et al. [22], who adapted the latter with the TimeWave principle, where a clock is traced along a time line. TimeWave enables to find and compare repetitive patterns, patterns that are hard to tackle with our approach. Eccles et al. [14] use a space-time cube to order and annotate events in time, allowing the user to tell the story encapsulated in the data. Another approach for event data is shown by Andrienko et al. [4] where events of public interest are found. Both works on events are too coarsely sampled in time to get useful interpolation between data points, as assumed in our method. The relative motion between objects is emphasized in the visualizations proposed by Crnovrsanin et al. [10] where abstract representations may solve possible confusions when movements are only displayed spatially. We try to solve this confusion by only showing the movement feature of interest as given by the user-defined expressions. Guo et al. [15] propose a visual analytics tool with multiple linked views to investigate a road crossing. Visual analytics tools as proposed by Andrienko et al. [2] and Liao et al. [23] allow visual exploration supported by machine learning techniques that are, unlike our method, limited in the use of domain knowledge. All these approaches have in common that they have a fixed architecture; for allowing the user to find new types of movement features the tool needs to be adapted, while we only define a new computation.

2.2 Trajectory Aggregation

Aggregation methods can be used to analyze large numbers of trajectories. A common approach to aggregate trajectories is with a grid structure. Andrienko et al. [1] show the aggregated cells with glyphs on top of a map. Density approaches show a smoothed view on the data, and despite the computational complexity, they can be generated fast using graphics hardware. Hurter et al. [17] splat smoothed points on a map, while Lampe et al. [21] convolve trajectory segments with a constant speed. A constant acceleration along a trajectory segment is assumed by Scheepens et al. [25], who also take other attributes into account in the visualization. Downs [13] proposes a density for trajectories by splatting a kernel adapted to the length of a segment and assumes a constant velocity. By convolving the path in a space-time cube and display the result with volume rendering techniques, temporal information can be revealed, as proposed by Demšar et al. [11]. Trajectories may also be aggregated by means of other structures than a grid. For instance, Andrienko et al. [5] extract a graph from trajectories and aggregate them along the graph to construct a flow map. In

our approach we also aggregate trajectories, but we use user-definable expressions that relate attributes in the data to the density contribution.

2.3 Expressions for Defining Visualizations

The user is allowed to influence the visualization of the density field with expressions; this is known as a software pattern for information visualization [16]. We found three types of expressions occurring in visualizations. Firstly, there are expressions for filtering, such as queries. These expressions occur as a filter expression in our method. Secondly, expressions occur to compute derived data. For instance, Dinkla et al. [12] propose to let the domain experts define expressions that tune the output of analysis algorithms as used in their visualization. In our case such expressions are used for smoothing, the kernel size, composition, and aggregation. Finally, expressions occur to define how complete renderings are related as they occur in computer graphics as Constructive Solid Geometry (CSG) or in visual spreadsheets as proposed by Chi et al. [9]. One could argue that our scripts are examples of such expressions.

2.4 Geographical Information Systems

A density field may be represented as a basic raster map in a Geographical Information System (GIS) [7]. To conduct spatial analysis, raster maps can be composed using mathematical expressions in map algebra as proposed by Tomlin et al. [28]. We can use similar expressions in the composition of density fields, with cell-wise expressions. The main difference with Tomlin's approach is that we can create new density fields based on other data than field data. Even though first steps have been done by Câmara et al. [8] to incorporate non-field data in map algebra, it has not been done for trajectories with multiple attributes. Nevertheless, the final result, a composite density map, can be shown in a GIS as a raster map as well.

To create maps for a GIS there is often a notion of architecture; the composition of processes that together create the end-result. There are various levels of integration within these architectures. For instance in a mashup as proposed by Wood et al. [34] we see that completely different tools perform each a part of the computation of the final map. Since GIS is known for a large number of compatible data formats it is relatively easy to incorporate new tools or tools made by others in the tool chain. From an information visualization perspective this is a suboptimal solution, since the user cannot easily change the parameters for the visualization. For this reason we chose to have a flexible architecture within our own tool that allows the analyst to define manipulable parameters resulting in an interactive visualization system for the operator.

3 DATA MODELS

This section describes data models for trajectories with multiple attributes, density field computations, and density maps.

3.1 Multivariate Trajectories and Density Fields

The state $\boldsymbol{\alpha}^o(t)$ of the movement of an object o is given by various attributes, such as a timestamp t, with $t \in [0, T]$, a position $\mathbf{p}(t)$, a velocity $v(t)$, and other attributes such as object type or orientation. The continuous trajectory $\boldsymbol{\alpha}^o$, or $\boldsymbol{\alpha}$ for a single object, is captured in a time-series $\boldsymbol{\alpha}_0 \ldots \boldsymbol{\alpha}_N$, where between consecutive tuples in time a trajectory can be reconstructed. For two states $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_1$, Willems et al. [30] model the reconstruction of a continuous path $\mathbf{p}$, for timestamps t_i, positions $\mathbf{p}_i$, velocities v_i, and assuming a constant acceleration. To obtain a continuous density field C_o for an object o, we convolve $\mathbf{p}$ in point $\mathbf{q}$ with a radially symmetric kernel k_r, i.e.,

$$C_o(\mathbf{q}) \quad = \quad \frac{1}{T} \int_0^T k_r(\mathbf{q} - \mathbf{p}(t)) dt, \tag{1}$$

where r is the kernel radius. This density field is normalized with the time to allow comparison of fields with different durations. A density field C of all trajectories is defined by summing the individual densities C_o per point. A discrete density field D, with a value $D(Q)$ in cell Q, is obtained by sampling C in the center of each cell.

3.2 Density Maps

Multiple density fields can be composed into a single image, called a density map [25]. After composing various density fields for subsets of trajectories chosen with a multivariate filter, relations between attributes may be revealed in the density map. Figure 2 shows a fixed architecture used to create a density map from trajectories [25]: First, the trajectories are divided into subsets using filters, where in the density model for each subset a density field is computed with a given kernel radius r_i and scaled with weight w_i. From here we can either compose the density fields with density aggregation, where the scalars are combined in a new density field, or the density fields are rendered with a color map c_i and their images are composed. The density map is a result of the rendering of the aggregated density or the composed images together with a density field displayed as an illuminated height field similar to enridged contour maps [**?**] by van Wijk and Telea. It turns out from a user study by Willems *et al.* [31] that density maps are as effective as animation and the space-time cube in a controlled environment.
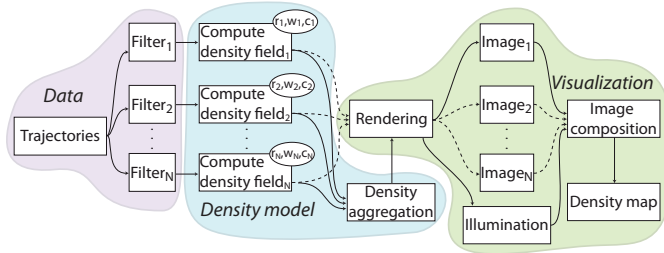

Fig. 2. The fixed architecture for density maps by Scheepens *et al.* [25].

4 COMPOSITE DENSITY MAPS

As an alternative for the above fixed architecture for density maps, we propose a flexible architecture for the density model part as shown in Figure 3. The new density model allows an analyst to define a computation for a set of density fields with a set of composable blocks. By including parameters the computation is usable for an operator.

Each individual block creates a new density field and may take density fields generated by other blocks as input. Some blocks also have trajectories as input, which may be filtered with a user-definable expression F_α, expressed in the attributes of the trajectory α. The filter F_α results in a filter attribute $f(t)$, which is 1 for states $\alpha(t)$ that satisfy

F_α and 0 otherwise. We define six types of building blocks: convolution, composition, counter, enhancement, enrichment, and iteration, which are discussed in the following sections.

In the visualization part of the architecture, density fields are either connected to the color map or the illumination component, both of which result in images to be composed into a density map. In the former component a color map is applied to one or more density fields, while in the latter a single density field is illuminated as a height field. In the color mapping, a single density field is drawn with a multi-hue color map, such as yellow-to-red. We use a logarithmic scale to display different orders of magnitude. With multiple density fields, each field gets a single hue color map with varying saturation, such as red-to-white. The hue values are given by sampling a perceptually balanced rainbow color map generated with PaletteView [29].

4.1 Convolution Block

A convolution block has density fields D_i and trajectories as input, and is parameterized by expressions for the filter F_α, for smoothing S_α, and for the kernel radius r_α. The density fields D_i are used as additional attributes $D_i(t)$ in state $\alpha(t)$. A density attribute $D_i(t)$ is obtained by bilinearly sampling the density field D_i at position $\mathbf{p}(t)$. With these new concepts, the contribution C_o to the density given in Equation (1) is generalized as follows

$$C_o(\mathbf{q}) = \frac{1}{T}\int_0^T f(t)\cdot S_\alpha(k_{r_\alpha(t)}(\mathbf{q}-\mathbf{p}(t)),t)dt, \quad (2)$$

with $f(t)$ the filter attribute and $\mathbf{p}(t)$ the position in state $\alpha(t)$. The user defines $S_\alpha(\rho,t)$ based on the density contribution ρ given by the kernel k and the attributes in $\alpha(t)$, including the density attributes. Similarly, the user defines an expression $r_\alpha(t)$ for the kernel radius.

We illustrate the convolution block with a time-based semantic depth of field [20], where recent movements of tankers are considered more of interest than older movements. We can compute such a field allowing only tankers in the filter F_α and vary both the smoothing S_α and the kernel radius r_α along the trajectory with these expressions:

$$S_\alpha(\rho,t) = \rho\ (t/T)^2, \quad (3)$$
$$r_\alpha(t) = 3-2.9t/T. \quad (4)$$

The non-linear decay of the density and the decreasing kernel radius r_α in time emphasize the recent movements with a sharp and small kernel as shown in Figure 4b. Compared with a normal density map in Figure 4a, we can now for instance see the order in which anchor zones are occupied based on the size of the dot.
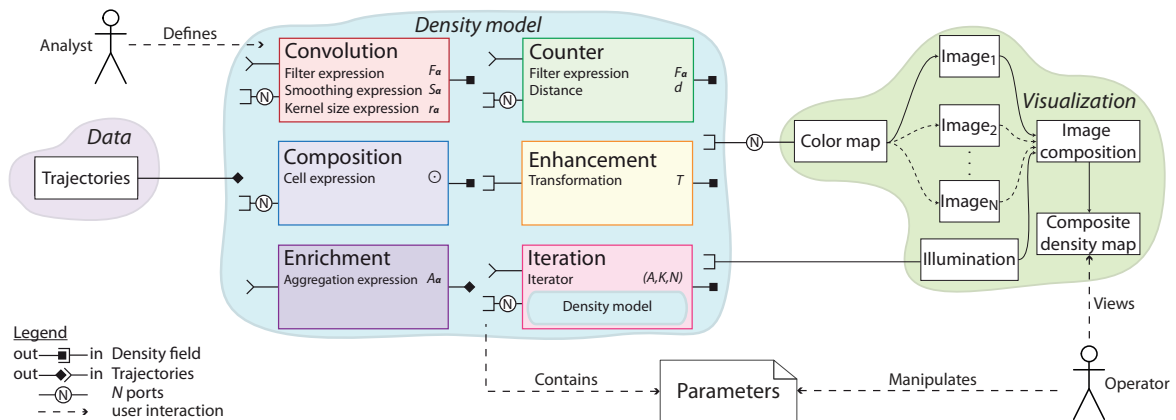
Fig. 3. The architecture for composite density maps consisting of three parts: data, density model, and visualization. The analyst composes the blocks to compute a set of density fields, while the operator uses the defined model by manipulating its parameters. Each block takes N density fields and, optionally, trajectories as input to compute a new density field, with the expressions enumerated in the corresponding block. These density fields are rendered as a colored image and an illuminated height field, which together give a composite density map.

Fig. 4. Composite density maps of vessels during a week, using (a) a normal density map, (b) a varying kernel radius emphasizing new movements with a small kernel using Equation (3), and (c) a varying kernel weight emphasizing sea lanes using Equation (5).
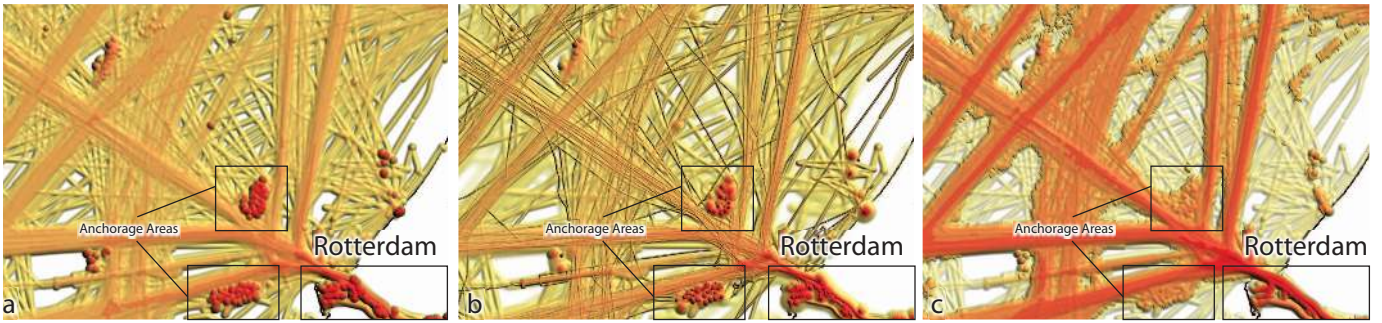
4.2 Composition Block

The composition block takes one or more density fields as input and produces a single density field as output. The output is generated based on a user-definable expression $\odot$ that composes the input density fields per cell. Examples of $\odot$ are the density aggregation operators of [25]: weighted addition and difference.

4.3 Counter Block

The counter block returns the number of occurrences of objects that have been active within a given distance from a cell. This building block is useful to quantitatively segment busy areas. For instance, with a count field D_c generated by a counter block, a density field enhancing traffic lanes can be constructed using a convolution block. If we define traffic lanes as areas with more than 8 moving ships, we can use a constant kernel radius r and a smoothing expression S_α defined by

$$S_\alpha(\rho,t) = \begin{cases} \rho \, v(t) & \text{if } D_c(t) > 8; \\ \rho & \text{otherwise,} \end{cases} \qquad (5)$$

where $D_c(t)$ and $v(t)$ are attribute values along trajectory α. Figure 4c shows enhanced traffic lanes while busy anchorage areas are reduced.

4.4 Enhancement Block

The enhancement block encapsulates a set of basic image processing operations, such as Gaussian blur, discretization, and edge detection as shown in Figure 5. For these operations, density fields are interpreted as a gray-scale image. Gaussian blur is used to remove distracting noise patterns in density fields. Discretization is used to segment the density field in areas with similar values using a step function. For example, by setting the number of steps to 2, a density field can be thresholded, and turns into a binary density field that can be used as a mask. Edge detection is implemented with a Sobel filter and is used to create contours in combination with the discretization.
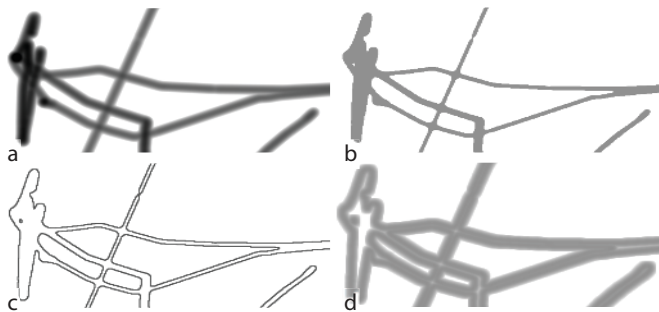

Fig. 5. Enhancements applied to a density field: (a) a normal density field, (b) the density field discretized using two steps, (c) contours generated from the discretization, and (d) the contours after Gaussian blur.

4.5 Enrichment Block

The enrichment block adds new attributes to trajectories. The state α is extended with attributes containing aggregations of existing attribute values, defined by an aggregation expression A_α, such as minimum, maximum, or average. The input of the block is a set of trajectories and the output is the same set of trajectories with the additional attribute.

4.6 Iteration Block

An iteration block is parameterized with an attribute A, an attribute value range $[0, K]$, and a number of steps N. The block computes a density model N times for consecutive intervals of length K/N. Figure 6 shows that the output density field is incrementally built by adding a density field for each interval. With an iteration block it is possible to force an order in which computations are performed. This order is required for instance to compute interactions (see Section 6.6).

Fig. 6. Intermediate density fields for 1, 2, and 3 consecutive iterations over (a) time intervals and (b) individual objects.

5 IMPLEMENTATION

We use a Graphics Processing Unit (GPU) to compute density fields fast. A density field is stored in a texture, where each texel represents a cell, and is computed with shader programs that run on the GPU.

The blocks and their relations are defined in a script language. To avoid long code listings, these scripts are depicted as block diagrams in this paper, such as at the bottom of Figure 8. For each block, shader code is generated on-the-fly by instantiating building block-type specific templates with expressions defined in the script. The CPU controls the flow of the computations and sends both the intermediate density fields as textures and the appropriate shader programs to the GPU. Unbound variables in the script are treated as parameters of the density model and can be easily adapted and animated with automatically generated user interface widgets (see Fig. 7) similar to the `Manipulate` function in Mathematica [33]. This allows the analyst to include useful parameters in the computation used by the operator.

Fig. 7. A parameter `minSpeed` can be varied with a widget.

5.1 Performance

The performance loss of a single density field computation due to the generalization is negligible compared to previous implementations. A single density field is computed in a split second by a machine equipped with an Intel Core i7 CPU at 2.8 GHz, 6 GB of RAM memory, and an NVidia GeForce GTX 285 with 1 GB of video memory, for given data size (400,000 points), resolution (1000x1000 cells), and kernel radius (2km). But, experienced users may create complex density fields that may require tens or more intermediate density fields. These computations in total may result in wall times of several minutes. For these complex computations, the memory size of the graphics hardware limits the usage of different intermediate density fields.

6 MARITIME USE CASES

We have constructed use cases together with two Dutch maritime partners, who analyze vessels captured with the Automatic Identification System (AIS) [18]. They investigate vessels with different perspectives: One wants to find anomalous behaving ships, while the other is interested in maritime safety. The presented use cases triggered the partners to experiment further with the tooling for their own analyses.

Reveiro *et al.* [24] describe the sense-making process for vessel tracking systems, rephrased with our definition of analysts and operators, as follows: The analyst first performs a foraging loop where the data is explored, filtered, and exploited by extracting information or noticing patterns. Using a visual representation and domain knowledge, the operator then gathers information and develops insight into the data set by manipulating the visual representation using interaction and by setting parameters. Finally, the operator may take some action based on newly developed insight or gathered information. We present use cases to show how our method can play a role within this process, whether it is in the process of defining a visualization of normal behavior, finding anomalies using a rule-based or a data-based approach, or providing interaction to drill down.

These use cases are demonstrated with a vessel trajectory data set in an area of approximately 400 by 400 kilometer of the North Sea off the Dutch coast. The data set contains 425,591 states covering 2,268 vessels over a period of one week. The trajectories have the following attributes in state $\boldsymbol{\alpha}^o(t)$: Vessel type s, mass m, course $c(t)$, actual orientation $h(t)$, velocity $v(t)$, and acceleration $a(t)$. The course and orientation of a vessel are angles relative to some reference direction.

6.1 Detailed Harbors

A density field may be an aggregation of a large amount of movement data, and as a result details might be lost, especially around the busier areas, such as harbors and anchorage areas as shown in Figure 4a. To emphasize these areas, we could decrease the kernel radius r, but this also adds more details to areas where we do not require extra detail. With previous density maps, density fields could only be computed with a fixed kernel radius or, at best, the trajectories could be divided into subsets with each assigned a different kernel radius. In this section we describe multi-resolution approaches focussing on details in harbors in a smoothed context.

A composite density map is generated with a convolution block (see top Fig. 8a) with an adaptive kernel radius that gives us more insight into what is happening in harbor areas. Since vessels generally sail slower in harbor areas, we accomplish this by using the velocity $v(t)$ to vary the kernel radius, i.e., $r_{\boldsymbol{\alpha}}(t) = \max(\omega v(t), 0.1)$, with ω a scaling factor. Figure 8a shows that a small kernel enhances movements in anchor and slow-moving areas while a large kernel blurs fast movers.

One of our partners has expressed interest in a variable kernel based on traffic density such that movements in traffic lanes and other busy areas also have more detail through a smaller kernel radius. This is expressed in a block diagram (see top Fig. 8b), by first generating a density field B with a constant kernel radius in a convolution block 'Area usage' to get an overview of the general area usage. The result is then fed to a second convolution block 'Area usage as kernel radius' with a varying kernel radius based on area usage defined as follows:

$$r_{\boldsymbol{\alpha}}(t) = \max(R - \omega \log(1 + B(t)), 0) \qquad (6)$$

where R is a maximum kernel radius, $B(t)$ the density attribute 'Area Usage', and ω some scaling factor. The second block is connected to both the color map and illumination component, as shown in Figure 8b. We are able to infer from the amount of detail where the busier areas are, but by separating the varying kernel in the illuminated height field and the area usage in the color map, we obtain a density map that is more smooth and therefore better legible. Figure 8c shows an example where individual vessels stopped in the anchorage areas and the harbor of Rotterdam. The detailed trajectories reveal the actual structure of these areas with slow movers.

6.2 Drifters

A dangerous situation may arise when a vessel has an engine failure, becomes uncontrollable, and starts drifting, which may result in a collision. A vessel is said to be drifting when it is moving slowly (i.e., $3 \le v(t) \le 5$ knot) and its course $c(t)$ and actual orientation $h(t)$ have a significant difference (i.e., $> 30°$). These specific values may be tuned by domain experts using the widgets in Figure 7.

To visualize these potentially dangerous situations, we create a block diagram as shown in Figure 9, where a single convolution block is created with $S_{\boldsymbol{\alpha}}(\rho, t) = \rho|c(t) - h(t)|$ and a filter $F_{\boldsymbol{\alpha}}$ to select only slow movers. The resulting density field is connected to the color component. For contextual information, we also create a density field for the entire set of trajectories and connect it to the illumination component. This reveals a number of potentially dangerous situations in the data set. In Figure 9b, we show an overview of the drifters we find with this block diagram and zoom in on several particular cases.

In Figure 9c we can see a vessel with drifting patterns around its turns. Closer investigation shows that this is research vessel which is expected to make sharp turns. In Figure 9d, we see potential drifters in the areas shown in red. By clicking in this area, we notice that these trajectories are caused by a single cargo vessel that originated from the North, loitered in the area for approximately two days and then proceeded into the harbor of Antwerp. It is possible that the vessel was too early and had to wait two days before it could head to its destination harbor. This vessel has most likely been drifting on several occasions while it was waiting. The vessel appears to alternate between drifting in a South-West direction and actively moving in a North-East direction. In Figure 9e cargo vessels move between ocean platforms, which are visualized as black squares. These vessels are most likely resupplying ocean platforms and making short stops and sharp turns. In Figure 9f we see two parallel drifter trajectories. Since these trajectories occur within a busy area, this may be a potentially dangerous situation. Closer investigation reveals that these drifter patterns are caused by a research vessel as well.

6.3 Sea Lanes

Traffic Separation Schemes (TSSes) are defined in busy areas to guide large vessels by means of sea lanes. In the North Sea, several TSSes are in effect, but they are not mandatory for all vessels. We separate the sea lanes and other frequently used routes from other movements to get an overview of normal shipping traffic as shown in Figure 10.

To do this, we partition the data into N ($N \ge 8$ to avoid side-effects in crossings) sectors η, each representing a course range. For each sector, we compute a density field and sum the resulting fields into the final field shown in 10b. The partial lanes are masked with a thresholded smoothed count field to obtain only busy lanes. We filter out all movements with a course that differs from the average course by at least some constant using an average course field. The 'Average course' field is computed using the cell-wise division of two density fields: a density field 'Convolved course' using the smoothing function $S_{\boldsymbol{\alpha}}(\rho, t) = \rho c(t)$ and a density field called 'Convolved duration' with $S_{\boldsymbol{\alpha}}(\rho, t) = \max(\rho, \varepsilon)$, for a small ε to avoid divisions by zero.

The sea lanes appear to disintegrate in the marked areas in Figure 10b. This happens where vessel directions start diverging or where there simply is not enough data to extract a reliable route. The latter is the case within the encircled area for a little used sea lane.
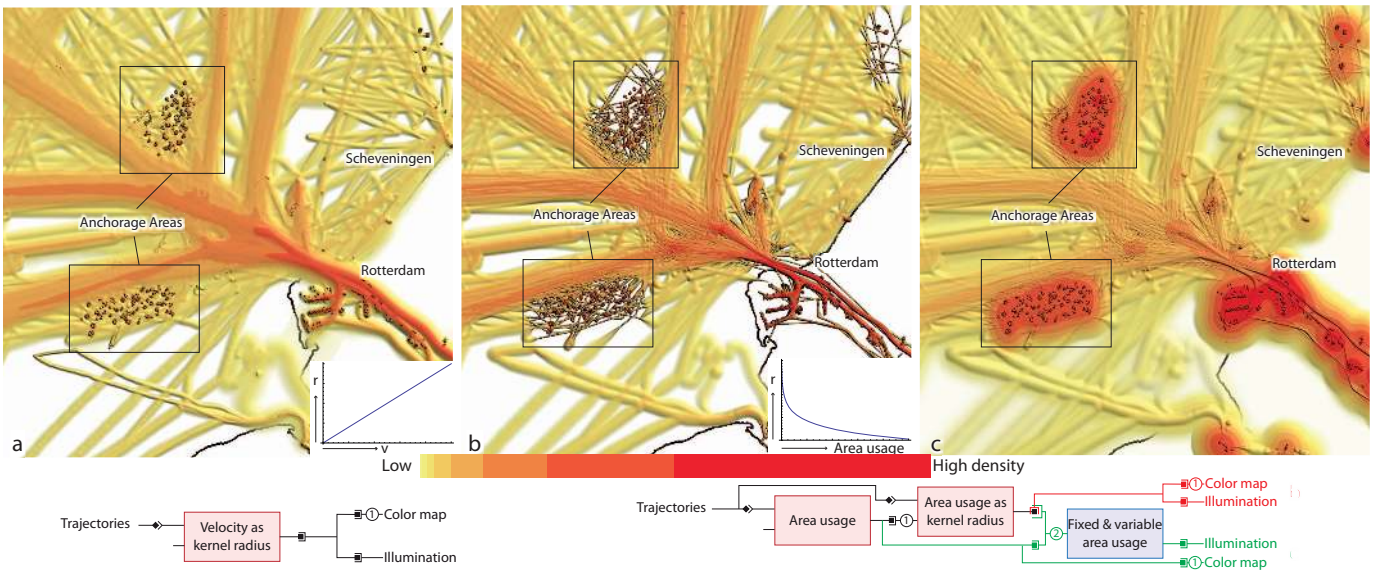
Fig. 8. Density maps of the area around the harbor of Rotterdam generated using a variable kernel based on (a) the velocity and (b) on traffic density. In (c) we use the density field of (b) as a illuminated height field and connect the area usage to the color component. In (a) more detail is visible in areas with slow moving or stationary vessels, such as inside the harbor and in anchorage areas. In (b) more detail is visible in dense areas such as the harbor mouth, the anchorage areas outside the harbor as well as the sea lanes. The block diagram for (a) is shown on the left, while the block diagram for (b) and (c) is shown on the right, with (b) with black and red links and (c) with black and green links.

Fig. 9. (a) The course $c(t)$ and actual orientation $h(t)$ of a vessel. (b) An overview of all drifter movements shown in color with all data shown in the illuminated height field. In areas (c), (d), (e), and (f) we see several closeups of a number of unusual drifter patterns.
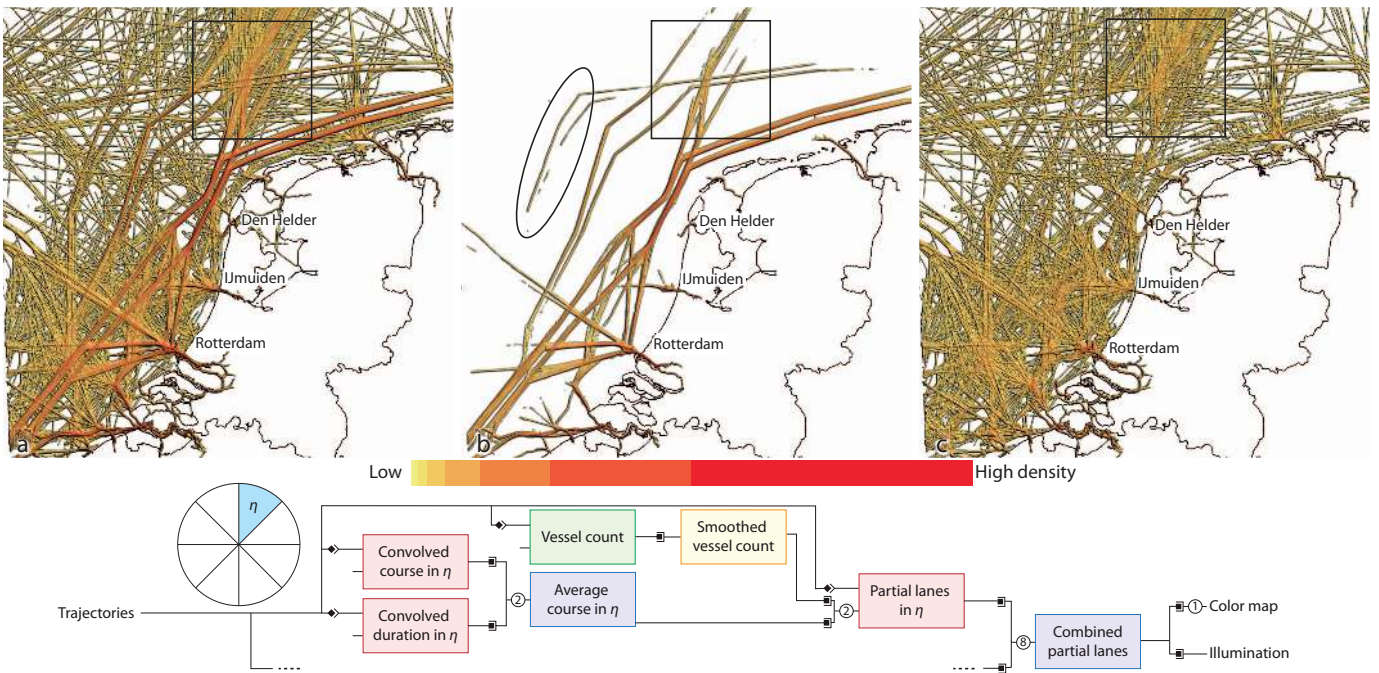
Fig. 10. A density map (a) containing all movements, (b) containing only movements over sea lanes with a minimum of 10 vessels and a minimum course deviation of $10°$. (c) The complement of (b), showing traffic outside of the lanes. The block diagram is shown at the bottom.
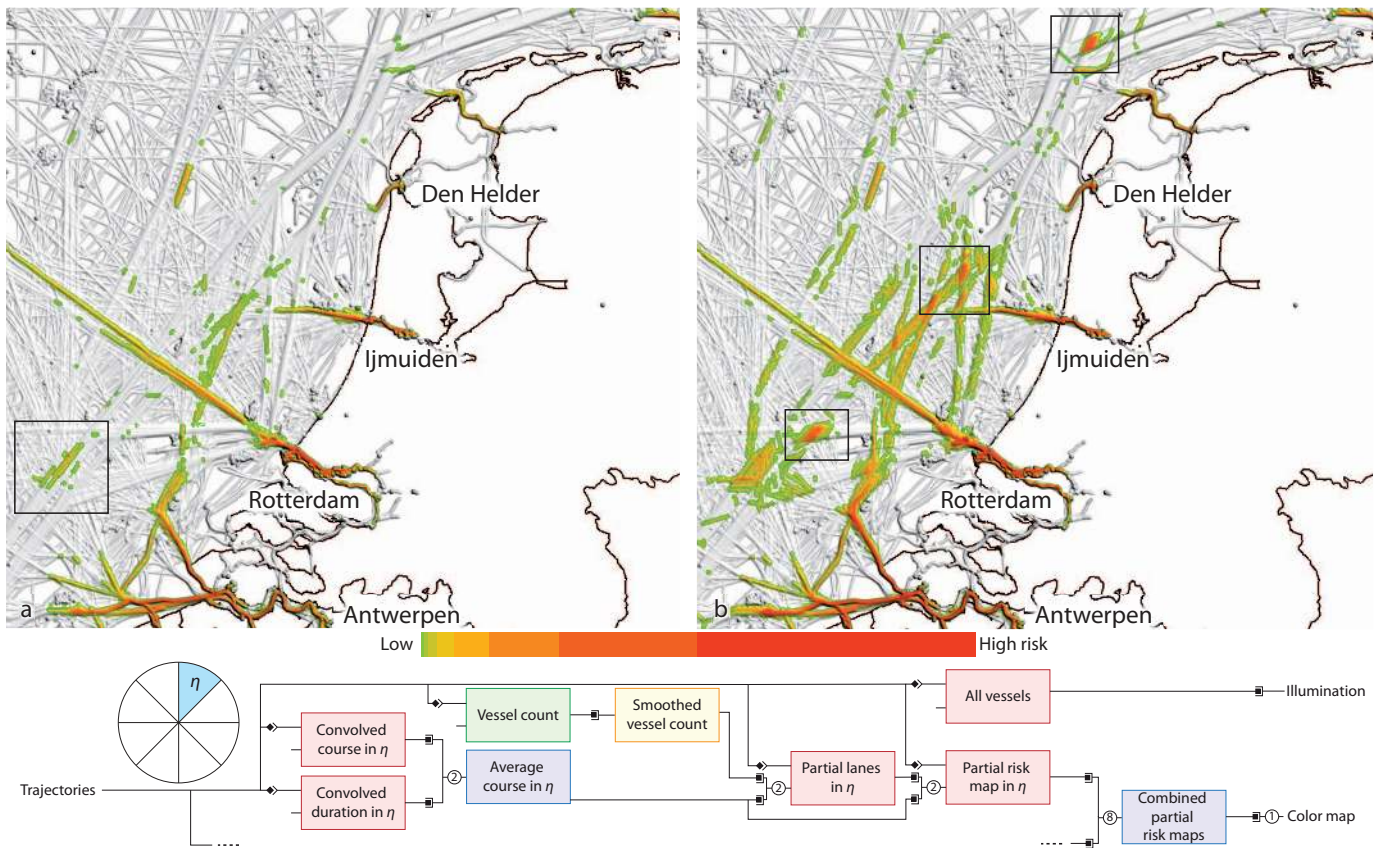

Fig. 11. Collision risk maps shown with a green to red color map with all data shown in the illuminated height field as context. The maps use a minimum deviation of (a) $175°$ and (b) $125°$. The block diagram is shown at the bottom. In the marked area in (a) we see a single research vessel moving in the opposite direction of a sea lane. In the marked areas in (b) we see sea lanes intersect with an angle larger than $125°$.

6.4 Traffic Direction

The TSSes separate vessel traffic heading in opposite directions, but vessels may ignore these schemes and move in a potentially dangerous direction. We use the average course maps generated in the previous use case to see whether there are vessels moving against the main traffic flow or whether there are areas containing major traffic flows in both directions. To do this, another convolution block is added to each sector; it filters out all movements that have a course difference less than some threshold compared to the average course. The results of these blocks are then summed into a single density field that shows the risk of collision in each area given some threshold. In Figure 11 we show an example with a minimum deviation of $175°$ (Fig. 11a) and an example of a minimum deviation of $125°$ (Fig. 11b). The resulting density map can be interpreted as a risk map that shows the distribution of possible collisions. From the average course maps of the previous use case we have only used those with many vessels, since the ones with a low number show low to no risk of collision.

In the example with a minimum deviation of $175°$, we immediately see hotspots in and near harbors. Since harbors typically have narrow access corridors, vessels are likely to move in opposite directions close to each other, which in turn increases the likelihood of collisions. Note that the kernel radius defines this 'closeness' relation. We can also see a small green pattern in a busy sea lane to the West, which is not visible in a standard density map at all. This means a vessel has been moving in the opposite direction within a sea lane, which is potentially dangerous. Closer investigation reveals that this is a vessel moving back and forth over a relatively small area to collect hydrographic data. In the example with a minimum deviation of $125°$, we can see a high risk of collision in and near harbors and an increased risk in intersections between sea lanes where the intersecting angle is at least $125°$.

6.5 Risk Analysis

A common task in movement analysis is risk assessment. Each vessel is assigned a risk, which is defined as the probability of an accident multiplied by the consequences of the accident, as described by Koldenhof et al. [19]. Both the accident probability and consequences are derived from vessel attributes such as size, age, type, cargo, and environmental attributes, e.g., depth or distance to the shore. The risk of each individual ship is then aggregated into a risk map.

This risk model can be implemented as a block diagram (see bottom Fig. 1). First we enrich the trajectories with the average speed v_{avg} per vessel, which is used in the filter $F_{\boldsymbol{\alpha}}$ to exclude slow movers from the analysis. We model the risk κ for a single vessel o with the attribute values in state $\boldsymbol{\alpha}^o(t)$ as $\kappa(t) = \omega m v^2(t)$, where ω is a risk factor based on the vessel type s and the remainder of the risk is the kinetic energy function $E_k = \frac{1}{2}mv^2$, with mass m and velocity v, as a simplified model for the result of a collision. The risk factor ω is high for passenger ships due to potential loss of life due to an accident, ω has a medium value for tanker and cargo vessels carrying potentially hazardous materials, and ω is low for normal tanker and cargo vessels.

Using this model, we create a density field with $S_{\boldsymbol{\alpha}}(\rho, t) = \rho \kappa(t)$ for each vessel type (passenger, cargo, and tanker) and visualize these using three distinct color maps (see Fig. 1). The colored density fields are combined using the maximum operator of [25], showing the color of the highest density field. The areas with a high risk for both tankers and cargo vessels are invisible, therefore we create contours around them. These contours combined with the passenger ship risk map are connected to the illumination component, which indicates the areas that need attention due to high risks. They are created by discretizing the combined cargo and tanker risk map into a binary field using an enhancement block. The binary fields is then fed into another enhancement block with a Sobel filter, resulting in crisp edges. We smooth the contours with a Gaussian using another enhancement block.

Looking at Figure 1, we notice that for instance cargo vessels and tankers sail in less well-defined routes than passenger ships. Despite this high spread, the combined high risk is mainly in the same routes as passenger ships in the sea lanes between West and East, and, in a sea lane between North and South where they move towards the mandatory sea lanes for merchant vessels. By comparing the contours

and the illuminated height field of the passenger ships, this density map allows us to distinguish between areas with a high risk caused by loss of life or caused by financial or environmental damage.

6.6 Vessel Interactions

Due to the high amount of traffic and the limited manoeuvrability of vessels, the collision risk is relatively high, not only in the harbor areas, but also on the open sea. We define that a vessel o' interacts with another vessel o at time t if o' is within interaction distance $d_o(t)$ of vessel o as shown in Figure 12a, where these interactions are then convolved. We propose a method, using our architecture, to visualize such interactions with a flexible definition for $d_o(t)$.

Since both velocity and mass reduce the manoeuvrability of a vessel, we express the interaction distance $d_o(t)$ in a similar fashion as in Section 6.5, i.e., $d_o(t) = \omega m v^2(t)$, with ω a scaling factor. For kernel radius $r_{\boldsymbol{\alpha}}(t) = d_o(t)$, we obtain a density map as shown in Figure 13a. We see that the larger and faster a vessel is, the larger its radius is.

Since continuously sampling each point in time is infeasible, we perform our computations in discrete steps. Therefore, we use the iteration block to find vessels that are within interaction distance of another vessel for separate time intervals. In each iteration step, we use a count block to find overlapping movements. This result is used as an extra attribute for a convolution block in which only movements for which the result of the count block is two or more are convolved as displayed in Figure 12cd. This method, using a two minute time interval, results in the density map as shown in Figure 13b.

With this method we can see stretched out patterns marked with arrows in sea lanes where vessels overtake each other. Although these vessels pass each other close by, this is not considered dangerous. Also, tugs and pilot vessels cause large patterns in and near the harbor mouths when they are tugging or guiding other vessels. In both cases, the interacting vessels move in a similar direction. Dangerous interactions are found by subtracting the density fields for these parallel movements from a density field containing all movements.

The data is partitioned into N sectors η, each representing a range of $\gamma°$ courses (see Fig. 13). For each sector η, we count the occurrences of parallel movements within a sector η widened with β on both sides, called η^+. If an object moves with a direction in sector η, then a parallel movement occurs if another object moves with a direction within η^+, and hence deviates less than an angle β with sector η. If the number of parallel movements is two or more then movements within η are convolved.

This results in the density map shown in Figure 13c. All areas shown in color now indicate what we have defined as interactions without movement in a similar direction. We take a closer look at a few of them using interactive inspection in our tool. In Figure 13d we can see a small dot caused by two vessels passing each other at close range, where the vessel heading in South-Western direction has to slightly change course to avoid the other vessel. In Figure 13e we see another small dot caused by a vessel heading in a North-Western direction having to stop for a vessel heading in a North-Eastern direction. All visible dots in the density map in Figure 13c are caused by similar situations, either by vessels passing each other close by without action taken, or by vessels having to perform special manoeuvres to avoid colliding with other vessels.

In Figure 13f we see a number of small dots near the harbor mouth of Rotterdam, which is an area with high pilot vessel activity. Pilot vessels carry trained professionals to guide larger vessels through busy and/or dangerous areas around and in harbors. The pilot vessel latches onto a large vessel and the pilot boards the larger vessel to guide it safely into the harbor. The dots in this area are possibly caused by these pilot vessels latching onto larger vessels. To investigate whether this is true, we slightly adapt the method described above. We do not only filter out movements in areas with less than two overlapping vessel movements, but also filter out interactions where there are no pilot vessels involved. We do this by adding another count block that counts the number of pilot vessels within a time interval and use it as an attribute for our convolution block. We now get a density map where the only remaining interactions occur at the harbor mouth.
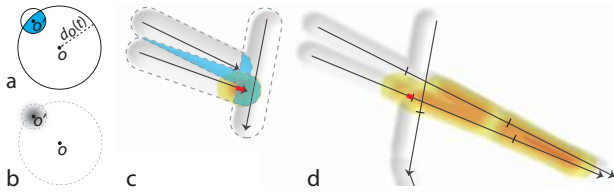
Fig. 12. (a) Interaction distance $d_o(t)$ of a vessel o at time t. (b) Vessel o' is within $d_o(t)$ of vessel o and is therefore convolved. The vessel interaction procedure applied to (c) a single and (d) three time intervals. Areas where multiple trajectories overlap are shown in blue and the movements within these areas are shown in red and convolved.

7 CONCLUSION AND FUTURE WORK

We have introduced a flexible architecture for composite density maps, where a density field is computed by means of a composition of simple density and trajectory computations. The density field computation itself is extended with a variable kernel radius along a trajectory and user-definable expressions for filtering, smoothing, and composing density fields. These expressions enable to apply domain knowledge as demonstrated with maritime use cases showing patterns for collisions risks, drifters, common routes, and interactions.

To extend this work, we are looking in the following directions. First of all, we would like to include environmental context given by raster maps, such as weather or elevation, and use them as input in the density field computations. To increase the performance of the density field computation we may use the one as proposed by Lampe *et al.* [21], where the trajectories are convolved with a constant velocity instead of a constant acceleration. However, it is unclear what the speedup will be when using our multivariate approach. Furthermore, we would like to extend the set of building blocks and experiment with encoding attributes in a non-symmetrical kernel. User interaction needs to be developed further such that, for example, objects participating in an interesting pattern or anomaly can be easily identified and traced. Also, cyclical movement patterns and density fields of other attributes than spatial position are of interest. Lastly, we would like to collaborate with domain experts from other fields, to check whether trajectories of other objects can be analyzed using our method as well.
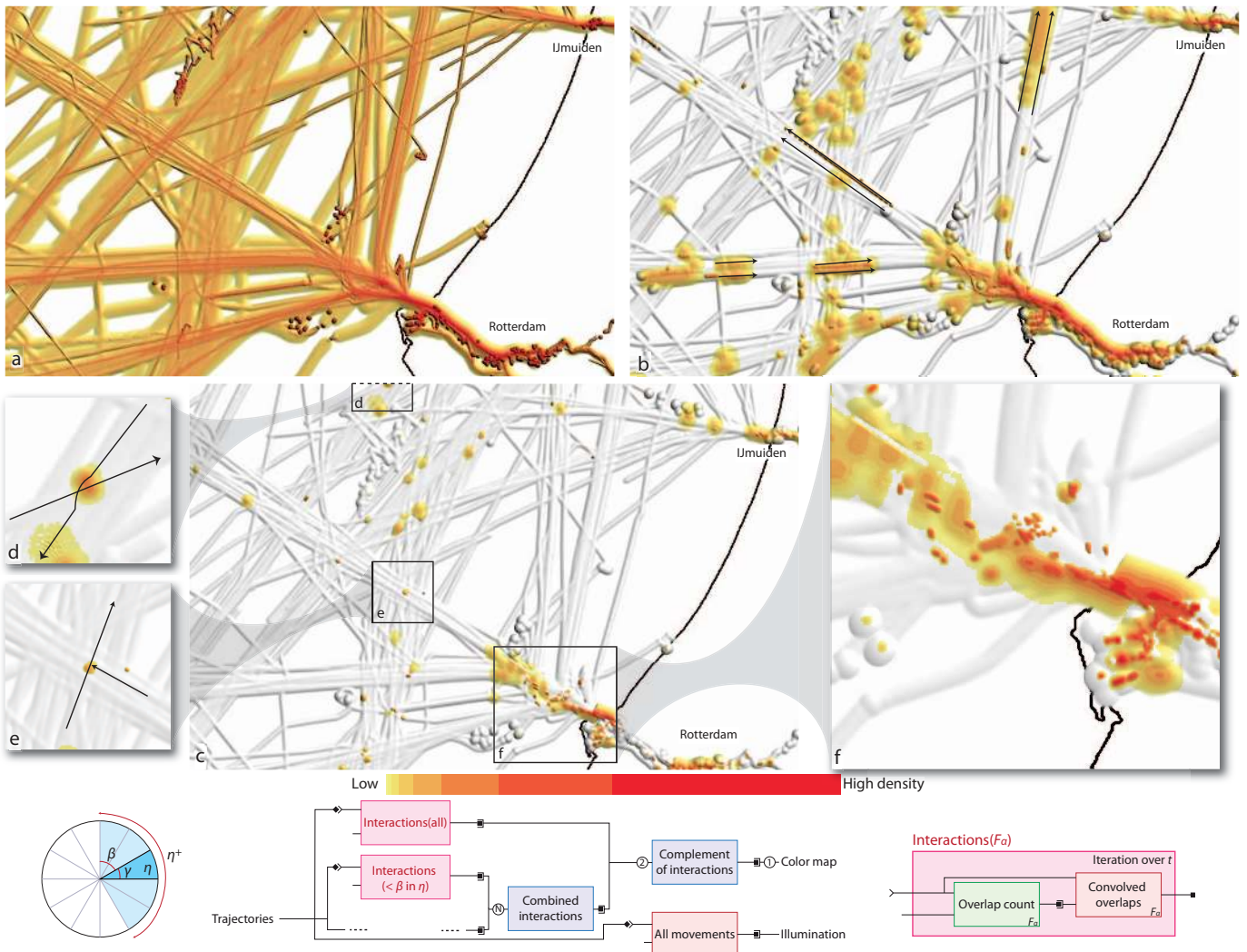
Fig. 13. Interactions around Rotterdam harbor with $\beta = 60°$, $\gamma = 30°$, $N = 12$, and the arrows show vessels overtaking each other. (a) An overview of all vessels using $r_{\boldsymbol{\alpha}}(t) = d_o(t)$, (b) all interactions within a time interval of two minutes in color and an overview of all movements used as illuminated height field, and (c) all vessel interactions with a course difference of at least β. The block diagram for density map c is shown at the bottom.

REFERENCES

[1] G. Andrienko and N. Andrienko. Spatio-temporal aggregation for visual analysis of movements. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 51–58, 2008.

[2] G. Andrienko, N. Andrienko, S. Bremm, T. Schreck, T. Von Landesberger, P. Bak, and D. Keim. Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. *Computer Graphics Forum*, 29(3):913–922, 2010.

[3] G. Andrienko, N. Andrienko, U. Demšar, D. Dransch, J. Dykes, S. I. Fabrikant, M. Jern, M.-J. Kraak, H. Schumann, and C. Tominski. Space, time and visual analytics. *International Journal of Geographical Information Science*, 24(10):1577–1600, 2010.

[4] G. Andrienko, N. Andrienko, M. Mladenov, M. Mock, and C. Pölitz. Discovering bits of place histories from people's activity traces. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 59–66, 2010.

[5] N. Andrienko and G. Andrienko. Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, 17(2):205–219, 2011.

[6] C. Brunsdon. Estimating probability surfaces for geographical point data: An adaptive kernel algorithm. *Computers and Geosciences*, 21(7):877–894, 1995.

[7] P. A. Burrough and R. A. McDonnell. *Spatial analysis using continuous fields*, chapter 8, pages 200–236. Oxford University Press, 1998.

[8] G. Câmara, D. Palomo, R. Cartaxo, M. Souza, and R. F. D. Oliveira. Towards a generalized map algebra: principles and data types. In *Proceedings of the Workshop Brasileiro de Geoinformática*, 2005.

[9] E. Chi, J. Riedl, P. Barry, and J. Konstan. Principles for information visualization spreadsheets. *IEEE Computer Graphics and Applications*, 18(4):30–38, 1998.

[10] T. Crnovrsanin, C. Muelder, C. Correa, and K.-L. Ma. Proximity-based visualization of movement trace data. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 11–18, 2009.

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

[12] K. Dinkla, M. A. Westenberg, H. M. Timmerman, S. van Hijum, and J. J. van Wijk. Comparison of multiple weighted hierarchies: Visual analytics for microbe community profiling. *Computer Graphics Forum*, 30(3):1141–1150, 2011.

[13] J. Downs. Time-geographic density estimation for moving point objects. In *Geographic Information Science*, volume 6292 of *Lecture Notes in Computer Science*, pages 16–26. 2010.

[14] R. Eccles, T. Kapler, R. Harper, and W. Wright. Stories in geotime. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 19–26, 2007.

[15] H. Guo, Z. Wang, B. Yu, H. Zhao, and X. Yuan. Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 163–170, 2011.

[16] J. Heer and M. Agrawala. Software design patterns for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1077–2626, 2006.

[17] C. Hurter, B. Tissoires, and S. Conversy. Accumulation as a tool for efficient visualization of geographical and temporal data. In *Proceedings of the AGILE Workshop Geospatial Visual Analytics: Focus on Time*, 2010.

[18] ITU. Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile band. *Recommendation ITU-R M.1371-1*, 2001.

[19] Y. Koldenhof, C. van der Tak, and C. Glansdorp. Risk awareness; a model to calculate the risk of a ship dynamically. *International Scientific and Technical Conference on Maritime Traffic Engineering*, 2009.

[20] R. Kosara, S. Miksch, and H. Hauser. Focus+context taken literally. *IEEE Computer Graphics and Applications*, 22(1):22–29, 2002.

[21] O. D. Lampe and H. Hauser. Interactive visualization of streaming data with kernel density estimation. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 171–178, 2011.

[22] X. Li, A. Çöltekin, and M.-J. Kraak. Visual exploration of eye movement data using the space-time-cube. In *Geographic Information Science*, volume 6292 of *Lecture Notes in Computer Science*, pages 295–309. 2010.

[23] Z. Liao, Y. Yu, and B. Chen. Anomaly detection in GPS data based on visual analytics. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 51–58, 2010.

[24] M. Riveiro and G. Falkman. Supporting the analytical reasoning process in maritime anomaly detection: Evaluation and experimental design. In *Proceedings of Information Visualization (IV)*, pages 170–178, 2010.

[25] R. Scheepens, N. Willems, H. van de Wetering, and J. J. van Wijk. Interactive visualization of multivariate trajectory data with density maps. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 147–154, 2011.

[26] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Nr. 26 in Mono. on Stat and Appl. Probability. Chapman & Hall, 1992.

[27] A. Slingsby, J. Dykes, and J. Wood. Using treemaps for variable selection in spatio-temporal visualization. *Information Visualization*, 7(3–4):210–224, 2008.

[28] C. D. Tomlin and J. K. Berry. A mathematical structure for cartographic modeling in environmental analysis. In *Proceedings of the Symposium of the American Congress on Surveying and Mapping*, pages 269–283, 1979.

[29] M. Wijffelaars, R. Vliegen, J. J. van Wijk, and E.-J. van der Linden. Generating color palettes using intuitive parameters. *Computer Graphics Forum*, 27(8):743–750, 2008.

[30] N. Willems, H. van de Wetering, and J. J. van Wijk. Visualization of vessel movements. *Computer Graphics Forum*, 28(3):959–966, 2009.

[31] N. Willems, H. van de Wetering, and J. J. van Wijk. Evaluation of the visibility of vessel movement features in trajectory visualizations. *Computer Graphics Forum*, 30(3):801–810, 2011.

[32] N. Willems, W. R. van Hage, G. de Vries, J. H. Janssens, and V. Malaisé. An integrated approach for visual analysis of a multisource moving objects knowledge base. *International Journal of Geographical Information Science*, 24(10):1543–1558, 2010.

[33] Wolfram research Inc. Manipulate. Accessed: 3 February 2011, http://reference.wolfram.com/mathematica/ref/Manipulate.html.

[34] J. Wood, J. Dykes, A. Slingsby, and K. Clarke. Interactive visual exploration of a large spatio-temporal dataset: Reflections on a geovisualization mashup. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1176–1183, 2007.