Density Based, Visual Anomaly Detection
Roeland Scheepens, Niels Willems, Huub van de Wetering, Jarke J. van Wijk
{R.J.Scheepens, C.M.E.Willems, H.v.d.Wetering, J.J.v.Wijk} @tue.nl
Eindhoven University of Technology, The Netherlands

Figure 1: A density field is created for both live (L) and historical (H) data, serving as input for anomaly operator ⊖ and resulting in a new density field L⊖H. This field is visualized using a multi-hue color map and composed, with the historical density field, into an illuminated height field.

1 Introduction

Anomaly detection in a visual setting allows a user to more easily recognize and interactively investigate anomalies. We show how density maps by Scheepens et al. [1] can support the analysis. Density maps are created in a four-way procedure–see Figure 1. First, the user divides the data into one or more subsets using interactive controls. Second, for each subset, smoothed trajectories are aggregated in a density field. Third, the density fields are combined using an aggregation operator and, finally, the resulting density field is rendered and composed into a density map. With modern graphics hardware, this process can be performed at interactive speeds.

1.1 Density Model

The trajectory \( \alpha(t) \) of an object, defined over a continuous range of time \([t_0, t_1]\), is given at time \( t \) as a tuple \( \alpha(t) = (p(t), v(t)) \) containing position \( p(t) \) and velocity \( v(t) \). A density field is generated by moving a smoothing kernel along the trajectories. A user can create multiple density fields, one for each subset of data defined by an attribute filter \( F_\alpha \). Space is subdivided into a regular grid of cells with equal area, for which a density field is computed per cell \( Q \). The density fields are then combined using a per-cell density aggregation and visualized using a multi-hue color map. The combined density field is rendered as an illuminated height field and composed with the color mapped density field into a density map.

Large vessels usually follow predefined shipping lanes. If such a vessel moves outside these shipping lanes, this constitutes a possible anomaly. We can find and investigate such anomalies by comparing a density field \( H \) of a sufficiently large set of historical movements with a density field \( L \) of live data containing recent movements. Where the historical data represents normal behavior, we define anomalous movements as current movements that differ significantly from historical movement patterns. We introduce an anomaly operator \( \ominus \) as a density aggregation and define the density field \( (L \ominus H)(Q) = \max(0, \omega L(Q) - H(Q)) \) to reveal anomalous areas, where \( \omega \) is some weight factor.
Figure 2: Vessel traffic in front of the Dutch coast. (A) Live traffic with a half hour tail. The colored trajectory is detected as an anomaly in the context of historical data, which turns out to be a single cargo vessel after inspection. (B) Anomalies among vessels carrying potentially hazardous material in color and the context of the entire historical set in gray.

2 Use Case: Anomaly Detection

We investigate a small area North of Rotterdam. For our live data set we use half an hour of trajectories, for our historical data set we use six days of trajectories and we set the weight $\omega$ to 1. The smoothing kernel varies over time such that older movements are displayed smaller, indicating the direction. Anomalies are visualized using a green-to-red color map and a combination of the historical density field $H$ and live density field $L$ is shown as context using illumination—see Figure 2A. We now see an anomalous area worth investigating by looking at the colors. We can click on this area to retrieve a list of vessels contributing to this anomaly. This tells us that the anomaly is caused by a cargo vessel. This vessel is likely outside of the designated shipping lanes to resupply ocean platforms in the area.

In Figure 2B we assign a filter $F_\alpha$ such that only vessels carrying potentially hazardous material are considered. A density field of the entire historical data set is shown as context. We see several vessels moving in areas, mostly outside the shipping lanes, where there are normally no or only few vessels carrying potentially hazardous material.

3 Future Work

We have concentrated on spatial anomalies, however, in future research we hope to define more operators for other types of anomalies such as drifters, speeders, or vessels moving against the main traffic flow. Furthermore, more complicated anomalies such as behavioral patterns within single trajectories or interactions between vessels may be recognized and visualized as anomalies.

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.