Computer Graphics International (2009)
Brian Wyvill, Nadia Magnenat-Thalmann (Editors)

# Visual Analytics of Time Dependent 2D Point Clouds

Tatiana von Landesberger and Sebastian Bremm and Peyman Rezaei and Tobias Schreck

Interactive Graphics Systems Group
Technische Universität Darmstadt, Germany

**Abstract**

*Two dimensional point data can be considered one of the most basic, yet one of the most ubiquitous data types arising in a wide variety of applications. The basic scatter plot approach is widely used and highly effective for data sets of small to moderate size. However, it shows scalability problems for data sets of increasing size, of multiple classes and of time-dependency. In this short paper, we therefore present an improved visual analysis of such point clouds. The basic idea is to monitor certain statistical properties of the data for each point and for each class as a function of time. The output of the statistic analysis is used for identification of interesting data views decreasing information overload. The data is interactively visualized using various techniques. In this paper, we specify the problem, detail our approach, and present application results based on a real world data set.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image GenerationDisplay Algorithms H.5.0 [Information Systems]: Information Interfaces and PresentationGeneral

## 1. Introduction

In many important applications, increasing amounts of data are acquired which need to be effectively analyzed. For example, in environmental monitoring, financial and economic analysis, traffic surveillance, etc., large quantitative data sets arise which need to be examined by experts and communicated to others. These data sets are often time-dependent and can be categorized into several entity classes. The multivariate character of the data is often simplified into low dimensional space by an appropriate projection technique (e.g., PCA, MDS, etc.). The classic *scatter plot* is one of the most widely used techniques to visualize such low dimensional quantitative data showing data point distribution.

However, *scalability limitations* exist regarding the amount of information that can be effectively displayed with the scatter plot technique in its basic form. If the number of data points grows, standard scatter plots get increasingly overplotted. Additionally, perceptional limitations regarding the number of different point classes that can be distinguished using color and shape exist [War04]. For time-dependent data, the problem is even more aggravated, as inter temporal characteristics of the point cloud data need to be considered.

In this short paper, *we introduce a system for effective visual analysis of time-dependent 2D point clouds*. We extend the approach of Wilkinson et al. [WAG05] for multi-class time-dependent data and combine it with effective data visualization. Statistic characteristics of point clouds over time (changes in point distributions, characteristics of point cloud movements, etc.) are used for determination of interesting data views. Various data visualizations using hulls and path traces are offered for the exploration of the data set. The system is applied on real world data, illustrating its effectiveness.

The remainder of this paper is structured as follows. Section 2 discusses related work. Section 3 introduces our approach. In Section 4, we apply our system on a real world data set. Finally, Section 5 summarizes and outlines future work in the area.

## 2. Related Work

In many application areas, point cloud visualization can be an effective tool for data analysis. In case of multi-dimensional data, suitable projection techniques are used in oder to obtain low dimensional point data. The prominent projection techniques include principal components analysis (PCA) [Jol02] or multi-dimensional scaling (MDS) [CC01]. The visualization of point cloud data has been used in finance [DG04], geography [PSKN06, HK98], microarray data [CKG05] or database exploration [PKJ*07].

*Scatter plots* are usually used for visualization of low-dimensional point clouds. Multivariate data can be visualized by so-called scatter plot matrices [WAG05]. Effective navigation through the space of axis parallel projections in high-dimensional point cloud data was introduced in [EDF08]. Point clouds may be represented by solid shapes, using various geometric constructs or using distance fields. In [SP07], the comparison of various *hulls* was shown. These include minimum bounding discs, boxes, and convex hulls. In [SSZW08], an algorithm for construction of compact, enclosing shapes was presented. *Distance fields* allow representation of point sets by smooth formation of visual areas by using appropriate transfer functions [KTSZ08]. On the efficiency side, the visualization of massive point cloud data sets may be accelerated by appropriate data structures as presented in [HE03].

The visualization of *time-dependent point clouds* employs animation [CKG05], trajectories [NFA01] or a combination of both [gap, TK07]. However these techniques concentrate on display of individual points without any statistics of motion or point cloud properties. An evaluation of the techniques was presented in [RFF*08]. The results show, that animation is more suitable for presentation whereas trajectories are more suitable for analysis. However, the display quickly gets crowded increasing the number and length of simultaneously shown trajectories. Therefore methods for automatic trajectory analysis have been proposed [STKF07].

The visualization of static one-class point data extended with statistical analysis of the data was explored in [WAG05] and [WAG06]. The statistic indicators of point cloud shape and point distribution is used for a proposal of interesting dimensions of the data for further visual inspection.

The above mentioned approaches, however, are not able to efficiently analyze and visualize dynamics of multi-class point clouds. In order to overcome this drawback, we combine effective visualization of data with statistic analysis over time.

## 3. Visual Analytics of Time-Dependent 2D Point Clouds

In this section, we describe our approach for interactive visualization of large time-dependent 2D point clouds. We present a systematization of possible aspects to be addressed when analyzing and visualizing such data types. We present important analytic measures for data analysis and show several techniques for visual abstraction of point cloud sets.

### 3.1. Aspects of Time-Dependent Point Clouds

When analyzing time-dependent 2D point clouds (or multi-class points/ point classes), several aspects need to be taken into consideration. We often need to discriminate between static and time-dependent indicators concerning single points, indicators of distribution of points within a class, and indicators for relative position of point classes.

In case of *single point in a point cloud*, its absolute and relative position in a point cloud is relevant. For example, it is relevant whether the point is an outlier in the class, lies in the center or at the border of the point cloud. For the point dynamics, the change of these indicators, of absolute positions (the trajectory), of movement directions, and speed is very interesting for the analysis.

For *a point cloud*, the distribution of points in a cloud, the cloud size and the cloud shape are of main interest. For time-dependent point clouds, the changes in these indicators as well as changes in the absolute position of the cloud are relevant. The size and shape of the cloud can be represented by enclosing hulls (alpha, convex, butterfly, circle, etc.) which can be visualized and for which we can calculate various descriptors (area, convexity, direction, etc.). Mid-points of point classes and their movement can serve as an indicator of cloud position and its change over time.

*Multiple point clouds* can be characterized by their relative position to each other (e.g., number of overlapping points, the relative overlapping area, number of overlapping point clouds) and by their dynamics (co-movements).

### 3.2. Visual Analytics of Time-Dependent Points

The absolute and relative position of points in a point cloud can be visualized by traditional scatter plots where the outlier points are highlighted (see Figure 1 (left)). The time development of points can be shown using trajectories (see Figure 1 (center)). However, trajectory visualization for long time periods and/or large number of points leads to overcrowded displays (see Figure 1 (right)). This can be overcome by (i) showing trajectories only for a small number of recent points, (ii) showing trajectories only for interesting points or (iii) aggregation of movements for the whole point clouds. The first option already increases the readability of the display by reducing the previewed time period. However, it stays problematic for large point clouds. The second option needs identification of interesting points. For this reason, an examination of movement features for points is needed. Point speed, local and global movement direction are monitored and turning points are indicated (see Figure 2). This supports selection of interesting views. The third option is discussed in the next section.

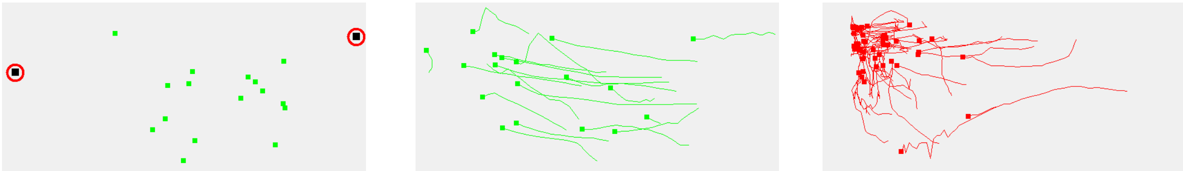

**Figure 1:** *Visualization of time-dependent points. Left: Point cloud with indicated outliers. Center: Point development visualization using trajectories. Right: Overcrowded display when using trajectories.*

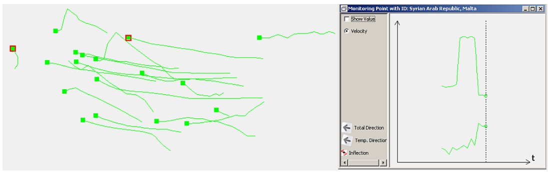

**Figure 2:** *Visual analysis of time-dependent points. Left: Point trajectories. Right: Point development analysis for selected points using speed and direction indicators.*

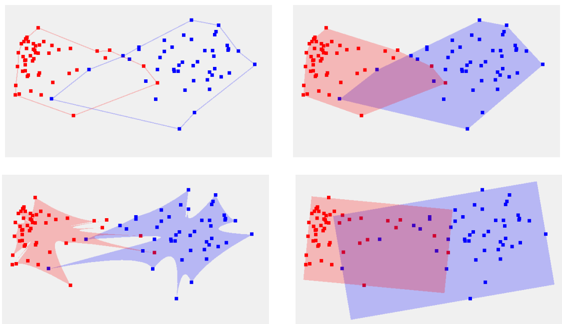

**Figure 3:** *A variety of hulls used for point cloud abstraction. From left to right: alpha hull, convex hull, butterfly hull and PCA-aligned minimum bounding box.*

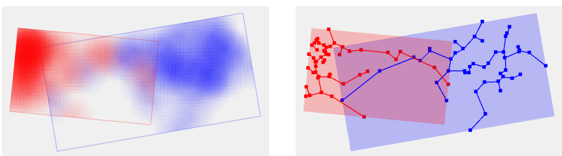

**Figure 4:** *Visualization of point cloud density. Left: density visualization, right: minimum-spanning tree showing point density and outliers.*

### 3.3. Visual Analytics of a Time-Dependent Point Cloud

In a static case, each point cloud can be characterized by the area of its hull, its convexity, density and number of points at the hull border. These indicators are based on [WAG05] [WAG06]. For modeling the cloud dynamics, the indicators can be monitored over time. This helps to identify interesting changes in point clouds (changes in shape, density etc.) (see Figure 5). In addition, the changes in point cloud positions can be visualized by their path traces using transparency (see Figure 5 (left)). The traces can be shown for the whole time period or only for a number of most recent observations. Hull traces show the general movement of the whole cloud, however they disregard movements of individual points within them. Another possible abstraction of the point cloud movement is visualization of trajectories of the cloud mid-point (see Figure 5 (left)). It shows the main direction of the movement of the majority of the points, however does not show differences in hull sizes and point distributions within the clouds.

### 3.4. Visual Analytics of Multiple Time-Dependent 2D Point Clouds

The analysis of multiple point clouds focuses on the comparison of shape and distribution properties and relative position of the clouds. The visualization of the point clouds reveals their position and the size of the overlap (see Figure 5). Statistic properties of each hull can be compared and used for selection of the most interesting hulls (e.g., with the highest density, the largest area, the lowest convexity etc.). These statistic properties can be automatically traced over time and used for identification of interesting changes in point clouds (see Figure 5 (right)). Correlations of these indicators can point to hulls behaving similarly. Figure 5 (right up) shows the overlap of the point clouds (overlapping area and relative number of points in the overlapping area) over time. It clearly identifies the time periods when the two clouds intersect (see Figure 5 (left)). The chart 5 (right down) shows the comparison of point cloud size development. It shows high correlation of the two indicators and the common increase and decrease of the cloud sizes through time. These monitoring tools can be used for identification of interesting views on the data especially in case of an overcrowded display for many point clouds over longer time periods.

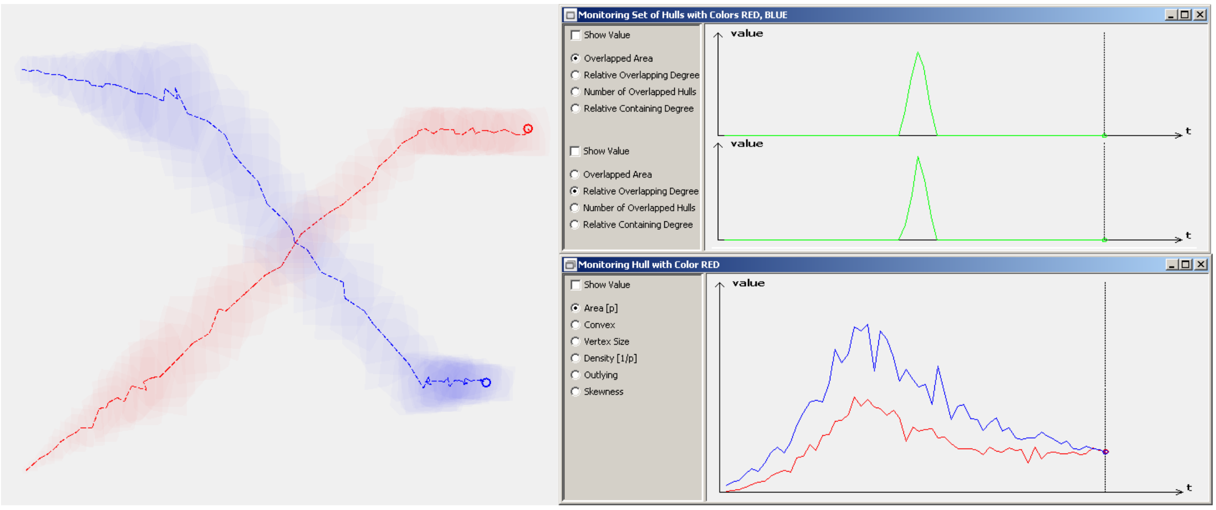


**Figure 5:** *Monitoring of point cloud development. Left: Point cloud dynamics visualization using hull traces and mid point trajectories. Right: Monitoring of point cloud statistics by choosing from various indicators. Right up: Monitoring of point cloud overlap. Right down: Monitoring of point cloud size.*

## 4. Application

In this section we use a real world data set on socio-economic development in world countries provided by OECD (http://www.oecd.org/statsportal/). The same data set has been previously used by [gap, RFF*08] for illustrating visualization techniques for time-dependent 2D points. The selected data show the relationship between average number of children per woman and woman work participation on annual basis in the period of 1980 - 2005 in world countries grouped by geographic region.

The Figure 6 (left) shows the system for visual analysis of dynamic point clouds. The socio-economic indicators for each country in the data set in in the year 2000 are shown as point clouds colored by the region (red = Europe, blue = Africa, green = middle east, light blue = south Asia). This shows the diversity of the values for countries with an overlap of several countries. For better analysis of the indicators by region, hull visualization with density is used (see Figure 6 (center)). It shows that the regions mostly differ with a smaller less dense overlap in the middle. The size of the hull is an indicator for the diversity within the regions. The development of the data are visualized and analyzed in the Figure 6 (right). It shows movement of each region towards upper left corner indicating a trend toward higher work participation rate with lower child birth rate, especially in south Asia. The statistical analysis shows the declines in the point cloud area for all regions, except for Africa. When looking at dynamics of selected countries (see Figure 2), we can see that the speed of progress in work-life quality changes over time with a temporal increase in the last years.

## 5. Conclusions and Future Work

We described a system for interactive visualization and analysis of large time-dependent point clouds. A rich set of visual (i.e., hull formation) and analytic (i.e., point statistics) abstraction tools was combined to form a system for visual analytics of point cloud data. The system offers various selections of interesting views. Its effectiveness was demonstrated by application on a real world data set. Our approach is suitable for dynamic point clouds of multiple classes. For very large datasets there can still be constraints to its efficiency.

Interesting future work includes on further broadening the system with more analytic functions and improved visual representations for the different point cloud features. It would be interesting to use time series analysis techniques for automatic indication of interesting time intervals in the data. The visual representation could be extended by combining statistic data properties with visual abstractions. We would like to further extend the system for visual analysis of weighted and 3D point clouds. It would be very interesting to conduct usability studies for verifying our approach.

## Acknowledgments

This work was partially funded by the German Research Foundation DFG within Priority Programme 1335: Scalable Visual Analytics under grant SCHR 1229/1-1.

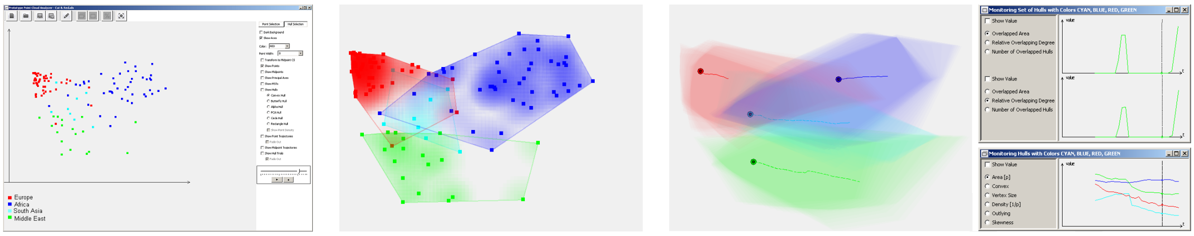

**Figure 6:** *Visualization of socio-economic data showing average number of children per woman (X-axis) versus woman work participation rate(Y-axis) for world countries grouped by geographic region. Left: Point cloud visualization. Center: Hull visualization. Right: Visual analysis of point cloud dynamics.*

## References


[CC01] COX M., COX M.: *Multidimensional Scaling*. Chapman and Hall, 2001.

[CKG05] CRAIG P., KENNEDY J., GUMMING A.: Coordinated parallel views for the exploratory analysis of microarray time-course data. *Coordinated and Multiple Views in Exploratory Visualization, 2005. (CMV 2005). Proceedings. Third International Conference on* (July 2005), 3–14.

[DG04] DWYER T., GALLAGHER D.: Visualising changes in fund manager holdings in two and a half-dimensions. *Information Visualization 3*, 4 (2004), 227–244.

[EDF08] ELMQVIST N., DRAGICEVIC P., FEKETE J.-D.: Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2008) 14* (2008), 1141–1148.

[gap] GAPMINDER. http://www.gapminder.org/.

[HE03] HOPF M., ERTL T.: Hierarchical splatting of scattered data. In *Proc. IEEE Visualization* (2003).

[HK98] HERRMANN A., KEIM D.: The gridfit algorithm: An efficient and effective approach to visualizing large amounts of spatial data. In *Proc. IEEE Visualization* (1998).

[Jol02] JOLLIFFE I.: *Principal Components Analysis*, 3rd ed. Springer, 2002.

[KTSZ08] KALBE T., TEKUŠOVÁ T., SCHRECK T., ZEILFELDER F.: GPU-accelerated 2D point cloud visualization using smooth splines for visual analytics applications. In *Spring Conference on Computer Graphics, SCCG 2008* (2008), pp. 111–125.

[NFA01] NORTH C., FAROOQ U., AKHTER D.: Datawear: Revealing trends of dynamic data in visualizations. In *Proc. IEEE Symposium on InfoVis 2001* (2001), pp. 8–11.

[PKJ*07] PU J., KALYANARAMAN Y., JAYANTI S., RAMANI K., PIZLO Z.: Navigation and discovery in 3d cad repositories. *IEEE Computer Graphics and Applications 27*, 4 (2007), 38–47.

[PSKN06] PANSE C., SIPS M., KEIM D., NORTH S.: Visualization of geo-spatial point sets via global shape transformation and local pixel placement. *IEEE Transactions on Visualization and Computer Graphics 12* (September-October 2006).

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

[SP07] SCHRECK T., PANSE C.: A new metaphor for projection-based visual analysis and data exploration. In *IS&T/SPIE Conference on Visualization and Data Analysis* (2007).

[SSZW08] SCHRECK T., SCHUESSLER M., ZEILFELDER F., WORM K.: Butterfly plots for visual analysis of large point cloud data. In *Proc. 16. Int. Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG'2008)* (2008).

[STKF07] SCHRECK T., TEKUŠOVÁ T., KOHLHAMMER J., FELLNER D.: Trajectory-based visual analysis of large financial time series data. *SIGKDD Explor. Newsl. 9*, 2 (2007), 30–37.

[TK07] TEKUŠOVÁ T., KOHLHAMMER J.: Applying animation to the visual analysis of financial time-dependent data. In *IV '07: Proceedings of the 11th International Conference Information Visualization* (2007), pp. 101–108.

[WAG05] WILKINSON L., ANAND A., GROSSMAN R.: Graph-theoretic scagnostics. In *Proc. IEEE Symposium on Information Visualization* (2005).

[WAG06] WILKINSON L., ANAND A., GROSSMAN R.: High-dimensional visual analytics: Interactive exploration guided by pairwise views of point distributions. *Visualization and Computer Graphics, IEEE Transactions on 12*, 6 (Nov.-Dec. 2006), 1363–1372.

[War04] WARE C.: *Information visualization: Perception for Design*. Morgan Kaufmann, 2004. 2nd edition.