Exploratory visualization based on multidimensional transfer functions and star coordinates

ALEX BORDIGNON, RENER CASTRO, HÉLIO LOPES, THOMAS LEWINER AND GEOVAN TAVARES

Department of Mathematics — Pontifícia Universidade Católica — Rio de Janeiro — Brazil
{alexlaier, rener, lopes, tomlew, tavares}@mat.puc--rio.br.

Abstract. Exploration and analysis of multivariate data play an important role in different domains. This work proposes a simple interface prototype that allows a human user to visually explore multivariate spatial objects, such as images, sequence of images or volume. It uses star coordinates as a widget to display the multivariate data on the computer 2D screen. The user then identifies a feature on this powerful coordinate system by mapping a selected feature region on that widget to a color and opacity. As a visual result, the feature is rendered on the object's space composing the use of this maps and the star coordinates projection. Some examples illustrate the potential of this interface.
Keywords: ***Visual data mining. Interfaces. Stellar coordinates. Transfer function. Visualization.***

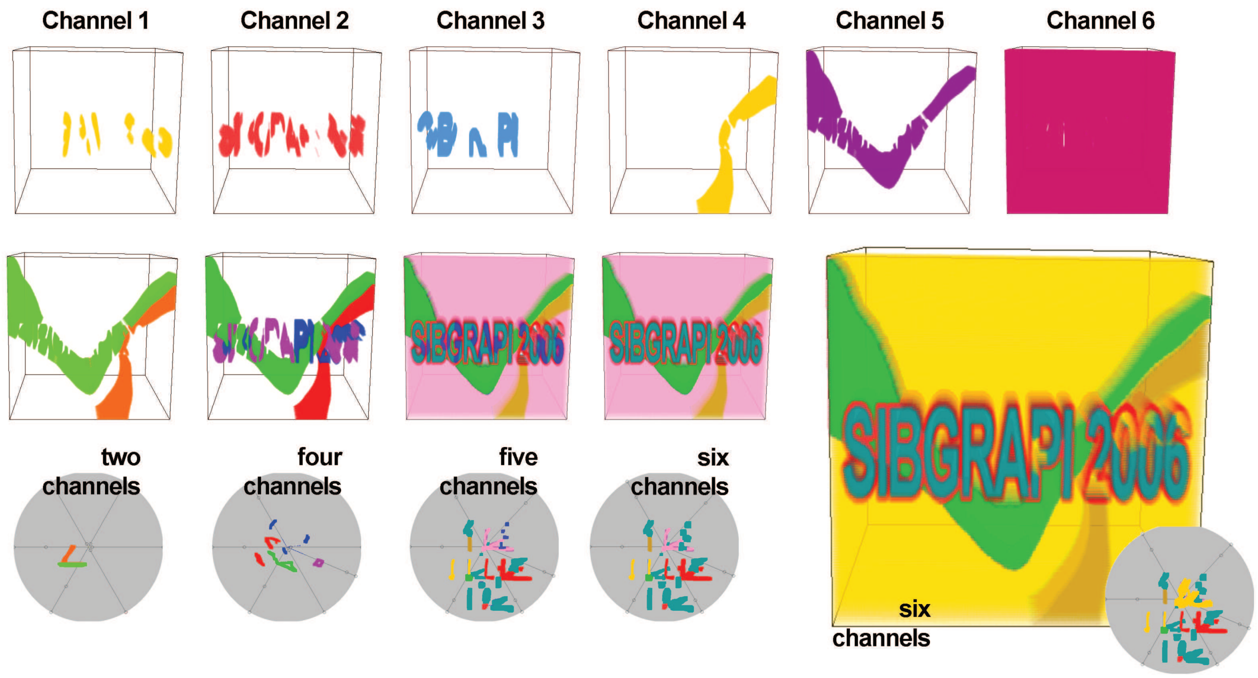

Figure 1: *Exploration of an image decomposed into 6 color channels.*

1 Introduction

Motivation. The exploration and analysis of multivariate data are important issues for a wide class of applications [8], like medical imaging, geology, geophysics and computational fluid flow simulations, just to cite a few. Researchers in the field of *Information Visualization* have been developing several techniques to explore and to analyze interactito explore and to analyze interactively multivariate data with a high dimensionality [10]. sk to find valuable information structures in today's data sets, which are usually huge and complex. An effective solution is to include the human's knowledge and interaction in the data exploration process. How to do this work is the main object of study in the *Visual Data Mining* area of research (see [9], for a survey).

On the one hand, an important step in visual data exploration process is to present visually multivariate data. *Par-*

Preprint MAT. 10/06, communicated on May 14^{th}, 2006 to the Department of Mathematics, Pontifícia Universidade Católica — Rio de Janeiro, Brazil. The corresponding work was published in the proceedings of the Sibgrapi 2006, pp. 273–280. IEEE Press, 2006.

allel Coordinates [1], *Star Coordinates* [6], *RadViz* [5] and *Viz3D* [10] are examples of multidimensional visualization techniques that display on the computer 2D screen a graphical object for each element on the data set.

On the other hand, transfer functions have been recognized as an important tool to *Volume Rendering* [8]. It assigns a color and opacity to each voxel according to its attributes in order to make a feature on the data visible. The problem of identifying a good transfer function is again a difficult task. However, there are direct manipulation schemes that make transfer functions an efficient tool for spatial data exploration [4].

Objective. This work deals with discrete domains such as images, sequences of images or volumes. On these objects, we suppose that each element (pixel or voxel) has several attributes. From now on, such data sets are called *multivariate spatial objects*, since for all of them we will consider the element's coordinates in the 2D or 3D space as attributes.

The main objective of this work is to develop a simple interactive visual tool that facilitates multivariate spatial object exploration and analysis.

Contributions. We introduce a new approach that allows the user to explore visually multivariate spatial objects. A simple and intuitive interface prototype is proposed. In this interface, the user operates a *star coordinates*' visualization widget in order to extract meaningful structures. Once a user identifies a feature on this coordinate system, he/she defines a transfer function by mapping a selected feature region on such widget to a color and opacity. At the end, the feature is rendered on the object's space domain by the use of these maps.

We also illustrate the potential of this interface to different applications, having as input: a color image, a geological simulation volume, a color video and MRI slices.

Paper outline. In the next section, we discuss some previous and related works. After that, in section 3 we introduce formally *star coordinates* and *multidimensional transfer functions*. We present in section 4 our proposal for multivariate data exploration and visualization and the interface prototype. In section 5 we illustrate its power to some applications. Finally, in section 6 we analyze the proposal and suggest future works.

2 Previous and related works

Multidimensional visualization on 2D screen. Several methodologies have been proposed to visualize high-dimensional data on a 2D computer screen. Inselberg and Dimsdale proposed *Parallel Coordinates* [1] to visualize data in $\mathbb{R}^N$. In their framework, the N-dimensional space is represented on the 2D screen by the use on N parallel copies of the y-axis. There, each point in $\mathbb{R}^N$ of the data set is mapped to a polyline that connects the project of the point on each parallel axis. The *Radial Visualization* (*RadViz*) method [5] maps each point in $\mathbb{R}^N$ of the data set in a 2D disk. The axes are represented as a radial segment from the center to the boundary circle of the disk. *Star coordinates* [6] extends this idea allowing the user to control the size of the radial axis in order to facilitate the exploration of a suitable mapping. Finally, Artero and Oliveira in [10] extend the use of *RadViz* to 3D.

Transfer functions. The study of transfer functions are receiving a lot of attention caused by the advance in the graphics hardware. Its objective is to make the volume easier to explore by assigning rendering parameters to the voxels such as color and opacity reflecting their attributes. However it has been recognized by several authors that the problem of identifying a good transfer function is generally accomplished by trial and error. With a rough interface, it may result in a difficult and boring task for the user.

Pfister et al. proposed in [11] a genetic algorithm to construct transfer functions. The work [12] proposed an intuitive interface for the space of all possible transfer functions. The technique called *Contour Spectrum* [13] visually define the isosurface space by the use of a metric, like area and the mean gradient magnitude to guide an isovalue choice, and it also generates a transfer function. In [2] an interface is proposed to facilitate the definition of a transfer function. Their approach realizes the mapping definition independently for each property like color, opacity, and data range.

Levoy introduced in [7] the concept of 2D transfer function, where the second dimension is the gradient magnitude. The use of multidimensional transfer function appears in [3]. In [4] the use of Gaussian transfer functions are used to build multidimensional transfer functions. Tory et al. in [8] propose the use of an interface based on parallel coordinates to explicit represent the visualization parameter space of a transfer function.

Tzeng et al. [**?**] proposed a paint-based interface that allows the user to paint regions in an image (a plane on the volume) with different colors in order to mark some samples and to give labels to them. Those painted samples points are used as an input data for a learning process. Together with an isosurface visualization, this approach shows to be very useful for exploratory visualization.

3 Preliminaries

In this section we formalize mathematically the concepts used on the *star coordinates* and on the *multidimensional transfer function* techniques.

Star coordinates mapping. *Star coordinates* coordinate system generalizes the concept of scatter plot to a higher dimension. We are to present its formal definition. The basic idea is to arrange N axis on a circle in $\mathbb{R}^2$.

Consider a family of functions $\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}$ that maps an n-dimensional axis aligned box

$$\mathbf{\Xi} = [min_1, max_1] \times \ldots \times [min_N, max_N] \in \mathbb{R}^N$$

to the plane $\mathbb{R}^2$. Such family of functions is parameterized by a set $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ of N unitary vectors of $\mathbb{R}^2$ and

The corresponding work was published in the proceedings of the Sibgrapi 2006, pp. 273–280. IEEE Press, 2006.

by a set $\mathcal{S} = \{s_1, \ldots, s_N\}$ of N non-negative real numbers. The vector a_i in $\mathcal{A}$ represents the i^{th}-axis of $\mathbb{R}^N$ and the scalar s_i corresponds to an scale factor that multiplies a_i on star coordinate system.

Given a point $\mathbf{p} = (x_1, \ldots, x_N) \in \Xi$, the value of $\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}(\mathbf{p})$, called the *star coordinates of* $\mathbf{p}$, is computed by the use of the following expression:

$$\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}(\mathbf{p}) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - min_i)}{(max_i - min_i)} s_i \mathbf{a}_i$$

Figure 2 illustrates a star coordinates mapping, evaluated for point $\mathbf{p} = (1, 1, 0.5, 0.5) \in \Xi = [0, 2]^4$.

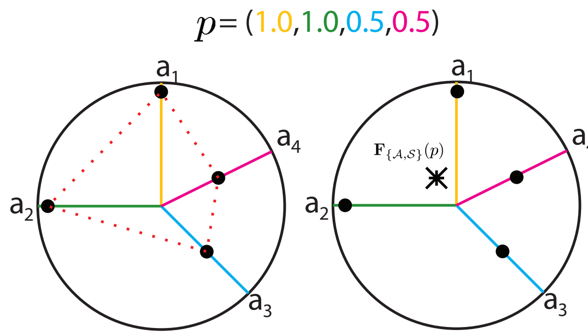

Figure 2: *Star coordinates mapping of a point in* $\mathbb{R}^4$.

Multidimensional transfer functions. The main role of multidimensional transfer functions in volumetric visualization is to map the voxel with multiple attributes to a color and opacity.

Formally, a *multidimensional transfer function* is a map $\mathbf{T} : \Xi \to \mathcal{C}$, where $\mathcal{C}$ is the space of all possible renderable properties such as color and opacity. Usually, a point on $\mathcal{C}$ has four coordinates: (R, G, B, A), where R, G, and B correspond, respectively, to the red, green and blue components of the color and A represents the opacity.

4 Our approach for exploratory visualization

Proposal overview. In this work we propose a simple and intuitive interface for the visual construction of a *multidimensional transfer function*.

The main idea is to define a two stage interactive process to build a *multidimensional transfer function*. In the first one, the user operates a star coordinates' widget to isolate a structure of the data. The interactive control are the parameters $\mathcal{A}$ and $\mathcal{S}$ of $\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}$. In the second stage, the user defines a mapping from the star coordinate system in $\mathbb{R}^2$ to the space $\mathcal{C}$ of renderable properties by the use of a transfer function $\mathbf{G}$.

Mathematically speaking, we substitute the hard task of defining a multidimensional transfer function $\mathbf{T} : \Xi \to \mathcal{C}$ by the function composition $\mathbf{G} \circ \mathbf{F}_{\{\mathcal{A},\mathcal{S}\}} : \Xi \to \mathcal{C}$.

The power of this approach relies in simplifying and accelerating the construction of meaningful multidimensional transfer functions by the use of visual data mining techniques.

Original star coordinates interface. Kondogan in [6] proposed a natural interface for data exploration using star coordinates. In his interface the main two operations are: rotation and scaling. With this two operations the user controls the $\mathcal{A}$ and $\mathcal{S}$ parameters for the mapping.

The objective of the rotation operations is to change the direction of a vector $\mathbf{a}_i \in \mathcal{A}$. To perform this operation, the user picks the axis $\mathbf{a}_i$ from any of its point and drag to set the new direction, whose direction is the vector from the origin to the drag point, Scale operations are used to change the scale factors $s_i \in \mathcal{S}$. To perform this scale operation, the user picks the end point of an axis $\mathbf{a}_i$ and push or pull towards or away from the origin. Figures 3 and 4 illustrate how the user operates the interface to define interactively the choice of $\mathcal{A}$ and the choice of $\mathcal{S}$.

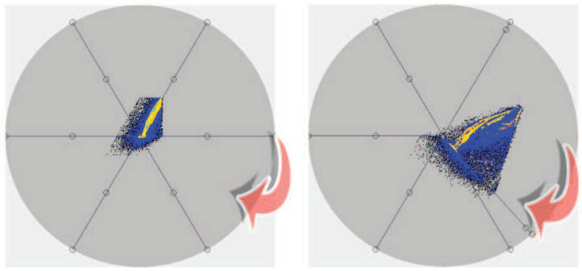

Figure 3: *Rotation operation: original mapping (left) and resulted mapping (right).*

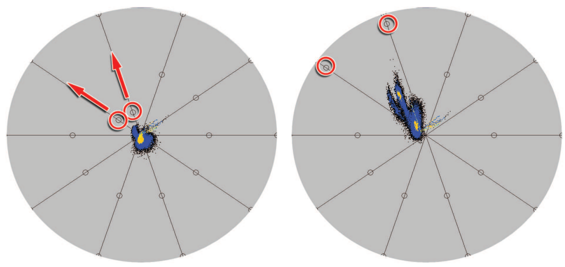

Figure 4: *Scale operation: original mapping (left) and resulted mapping (right).*

Improved star coordinates interface. We improve on the original star coordinates interface to facilitate the visual data exploration by the use of colors. The idea is to subdivide the space of the start coordinate system in small regions and to assign a color to each one according to the density of points mapped on it.

To subdivide the star coordinate domain, we propose two strategies: a regular square subdivision or a radial subdivision. In both cases, we assign a color to each region according to a threshold on the number of mapped points mapped on the region. Suppose that a vector of q thresholds $\mathbf{t} = (t_1, t_2, \ldots, t_q)$ and a vector of $q+1$ colors $\mathbf{c} = (c_1, c_2, \ldots, c_{q+1})$ are given. The color c_1 is assigned to a region that has less than t_1 points mapped on it. The color c_2 is assigned to the region when the number of mapped points is greater or equal to t_1 but less than t_2, and so on. The color c_{q+1} is assigned to the regions that have more than t_q points.

Figures 3 and 6 illustrate this idea using a regular square subdivision or a radial subdivision. And figures 3 and 4 shows a practical example using $q = 3$, i.e. only four colors: c_1 = black, c_2 = blue, c_3 = green and c_4 = yellow.

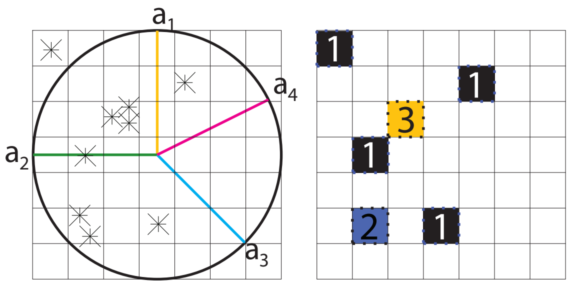

Figure 5: *Regular square subdivision.*

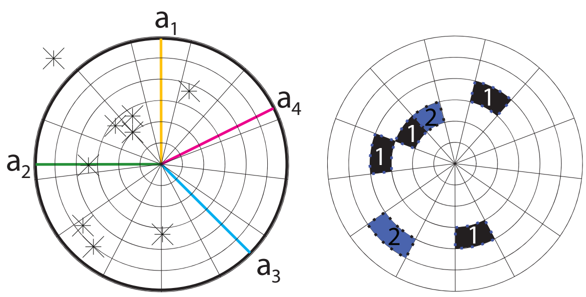

Figure 6: *Radial subdivision.*

The radial subdivision is very natural for star coordinates, but the computation of density in this case is more expensive when compared to regular squared subdivision.

The transfer function construction approach. Our proposal is to substitute the hard task of finding a good transfer function to recognize visually structures on multivariate spatial objects by the following process. First, the user identifies a structure on the star coordinates' space. After that, he/she defines a transfer function $\mathbf{G}$ by associating a color and opacity to a selected region on this space. Therefore, the points of the object mapped on this region are visualized using the associated rendering attributes. Next we will give a mathematical description of our approach.

Let consider that the data of the multivariate spatial object is contained in an N dimensional box $\mathbf{\Xi} \in \mathbb{R}^N$. Suppose that this box $\mathbf{\Xi}$ is a product of two sets: $\mathbf{\Xi} = \mathbf{\Omega} \times \mathbf{\Psi}$, where the set $\mathbf{\Omega}$ corresponds to the spatial attributes (e.g., the (X,Y,Z) coordinates of the voxel or (X,Y) for a pixel) and $\mathbf{\Psi}$ corresponds to other attributes of the data (e.g., value of a scalar function and the coordinates of this gradient).

Suppose also that the range of $\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}(\mathbf{\Xi})$ is contained in a disk of radius r $\mathbf{D}^2_r \subset \mathbb{R}^2$ centered at the origin.

The transfer function $\mathbf{G} : \mathbf{D}^2_r \to \mathcal{C}$ associates to each point in $\mathbf{D}^2_r$ a color and an opacity, where $\mathcal{C}$ is the space of all rendering attributes. Then we use the composition of $\mathbf{G}$ with $\mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}$, denoted by $\mathbf{G} \circ \mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}$, as a multivariate transfer function $\mathbf{T} : \mathbf{\Xi} \to \mathcal{C}$.

Once we have the function $\mathbf{G} \circ \mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}$, we visualize a point $\mathbf{p} = (\omega, \psi) \in \mathbf{\Omega} \times \mathbf{\Psi}$ using its spatial coordinates ω and the rendering attributes $\mathbf{G} \circ \mathbf{F}_{\{\mathcal{A},\mathcal{S}\}}(\mathbf{p})$.

Interface for the* G *function constructions. We propose a *pen tool* to define the function $\mathbf{G}$. The user selects a color and an opacity in the rendering attributes widget (see figure 7) and then paints a region (connected or not) on $\mathbf{D}^2_r$ by the use of this graphical tool (see figure 8). The user can repeat several times this operation to paint another region with a different color and opacity (see figure 1).

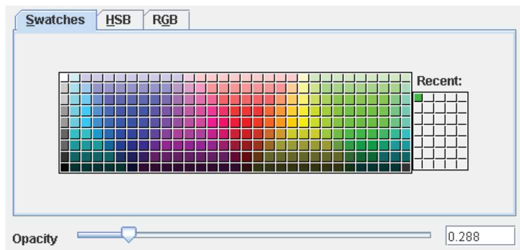

Figure 7: *Widget for color and opacity selection.*

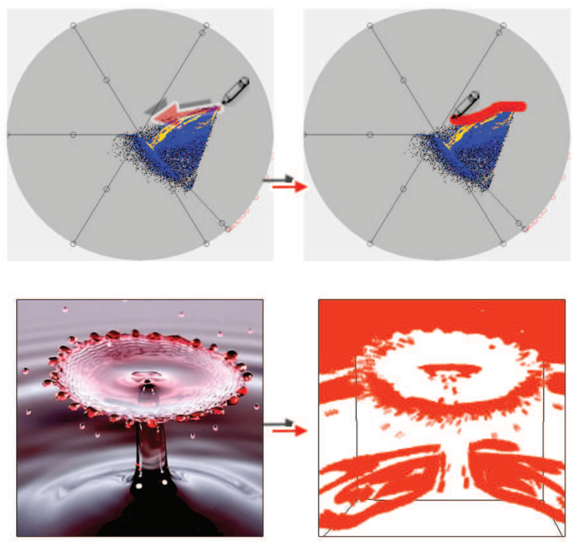

Figure 8: *Paint operation.*

A suitable procedure is to paint the regions with a highest density, and for that the color density scheme proposed above

The corresponding work was published in the proceedings of the Sibgrapi 2006, pp. 273–280. IEEE Press, 2006.

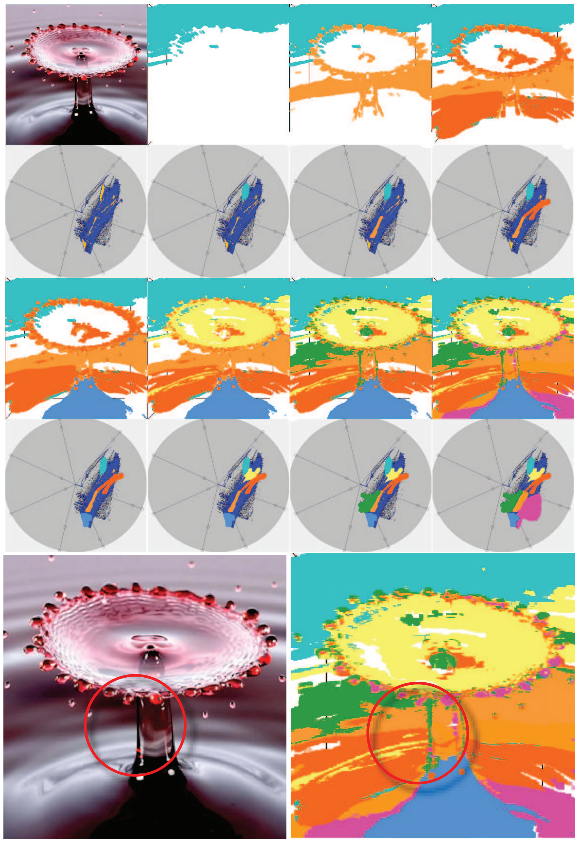

Figure 9: *Application in edge detection.*

serves as a guide to the user.

We init $\mathbf{G}$ with white color and opacity 0 for all points for all points of $\mathbf{D}_r^2$. When the user changes parameters $\mathcal{A}$ or $\mathcal{S}$, the function $\mathbf{G}$ is cleared, i.e. we assign the default value for all points in $\mathbf{D}_r^2$.

This interface can be improved by the addition of new parameters for the pen tool like size, shape, edge density, or pattern.

Next section exemplifies the potential of our approach in several applications.

5 Applications

Our approach was tested in a wide set of applications, like in multivariate images, volumes, MRI images, and videos.

Figure 9 illustrates the application of our method in color image segmentation. That figure shows the original image and, below it, the mapped points in the star coordinates system. Each pixel in this example has eight attributes: pixel X-coordinate, pixel Y-coordinate, red channel, green channel, blue channel, derivative of the red channel, derivative of the green channel, derivative of the blue channel. Using these attributes we could detect an edge (see the green part on the last picture) in this colored image that is very difficult to be identified by an automatic approach.

Figure 10 shows another example of segmentation using a 12-variate image. On the top-left picture we can see the star

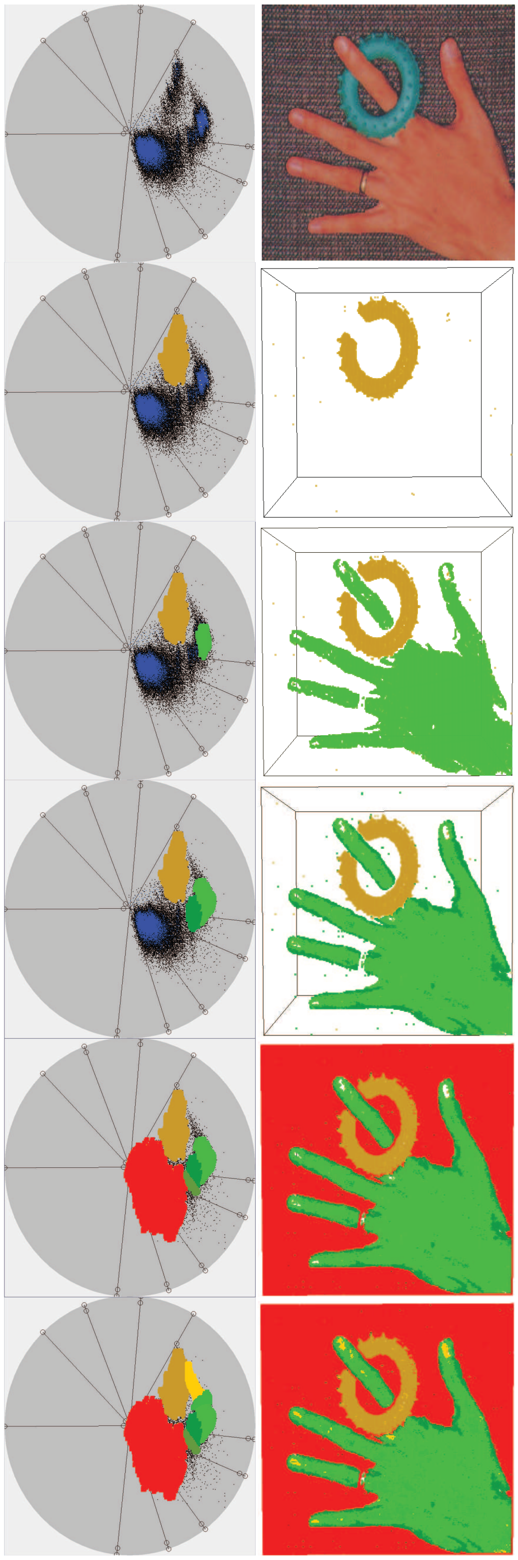

Figure 10: *Application in image segmentation.*

coordinates mapping for the image points. Choosing suitable parameters for the star coordinates, one can identify clearly three clusters on the image. In the sequence of pictures, the user marked the three clusters by the use of the pen tool.

Figure 11 shows the use of our approach on a sequence of color images (2D video). The data has seven attributes: the pixel coordinate (X,Y), the time T, the color (R,G,B), and the absolute value maximum variation in R, G or B. On the left, it shows the use of star coordinates mappings, which allows the user to visually identify moving parts on the video.

Figure 12 shows an application on an MRI image of resolution $80 \times 80 \times 44$. Although the resolution is very small, it was possible to the skin and a portion of the bones (in red).

Finally figure 13 shows a simulated probability volume of a 3D petroleum reservoir. Each voxel has as attribute the (X,Y,Z) coordinate, the probability of being a reservoir (on the bottom one can observe the XY-planes of probability), and the variation of the probability function in relation to the Z direction. It was possible to detect four clusters that are marked using different colors.

6 Conclusion and future works

In this work we introduced a simple and intuitive approach for visual exploration of multivariate spatial objects. We also proposed a very simple interface that facilitates the definition of a multivariate transfer functions by the use of star coordinates. The interface has only three operations: axis rotation, axis scale, and paint. The first two operations are used to find clusters using a star coordinate mapping. The paint is used to construct by composition a multivariate transfer function. The applications presented in this paper illustrate the great potential of this interface for visual data mining. To our knowledge, this is the first work that builds transfer functions based on star coordinates.

The authors plan to increment this interface by introducing automatic suggestion for the initial axis position and automatic region selection for painting. We plan also to improve the painting tool by introducing smooth variation of colors and to extend the work proposed by Tzeng et al. [?] to produce an intelligent system that identify a good choice of axis and scales.

References

[1] A. Inselberg and B. Dimsdale. Parallel coordinates: A tool for visualizing multidimensional geometry. In *IEEE Visualization'90*, pages 361–378, 1990.

[2] A. Konig and E. Groller. Mastering transfer function specification by using volumepro tecnology. In *Spring Conference on Computer Graphics 2001*, volume 17, pages 279–286, 2001.

[3] J. Kniss, G. Kindlmann and C. Hansen. Interactive translucent volume rendering and procedural modeling. In *IEEE Visualization'02*, 2002.

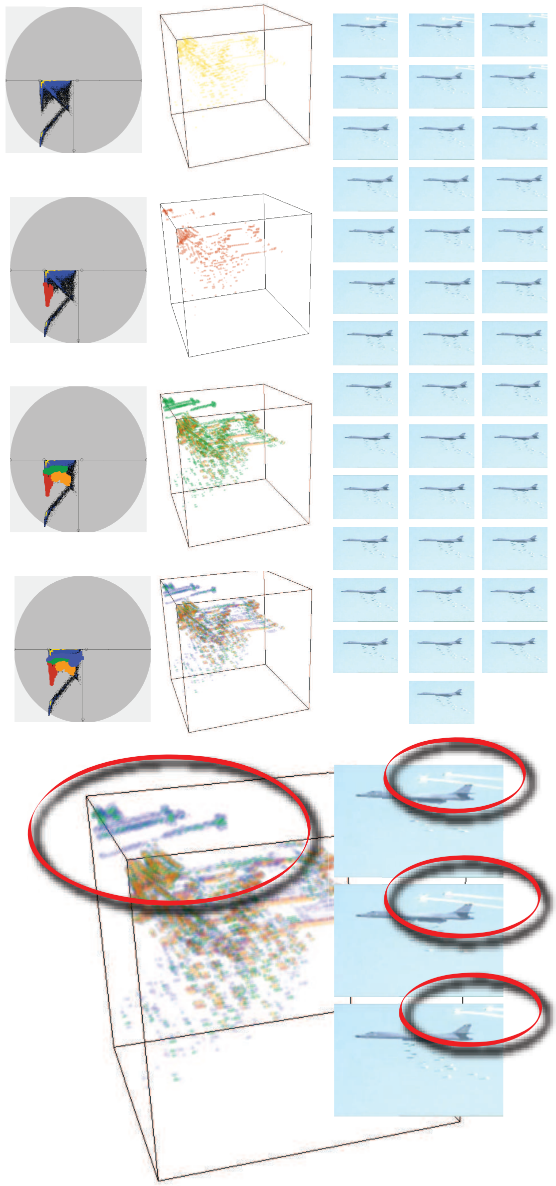

Figure 11: *Application in movement detection.*

The corresponding work was published in the proceedings of the Sibgrapi 2006, pp. 273–280. IEEE Press, 2006.

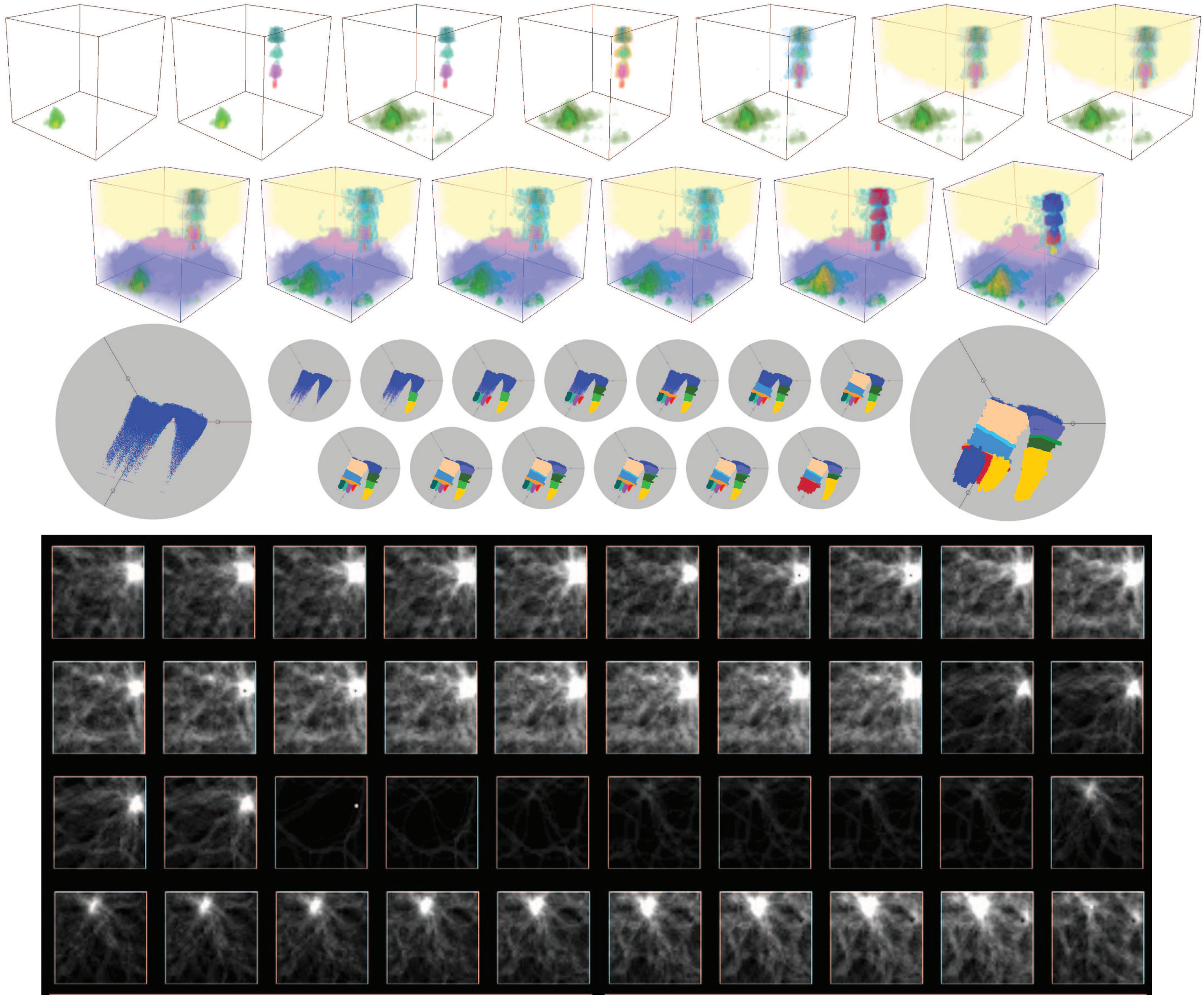

Figure 13: *Application in reservoir detection.*

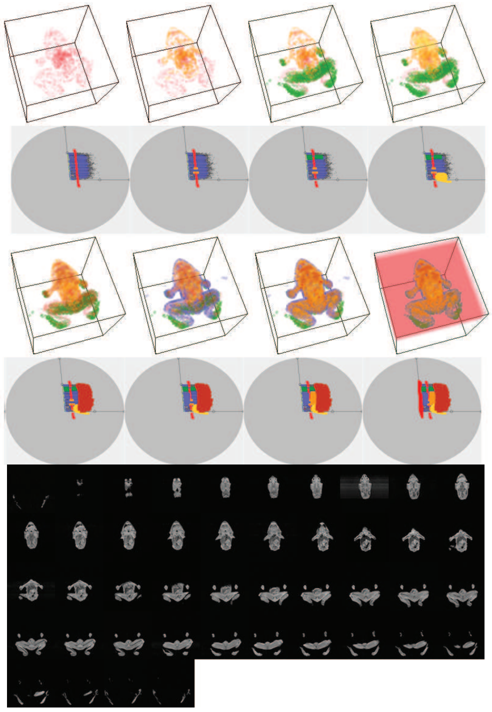

Figure 12: *Application in MRI.*

[4] J. Kniss, S. Premoze, M. Ikits, A. Lefohn, C. Hansen and E. Praun. Gaussian transfer functions for multi-field volume visualization. In *IEEE Visualization*, volume 11, page 65, 2003.

[5] P. Hoffman. Table visualizations: a formal model and its aplications. PhD thesis, *University of Massachusetts*, 1999.

[6] E. Kandogan. Visualizing multi-dimensional clusters, trends, and outliers using star coordinates. In *Proceedings of the 7th ACM SIGKDD*, pages 107–116, 2001.

[7] M. Levoy. Display on surfaces for volume data. *IEEE Computer Graphics and Aplications*, 8:29–37, 1988.

[8] M. Tory, S. Potts and T. Moller. A parallel coordinates style interface for exploratory volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 11(1):71–80, 2005.

[9] M. C. F. de Oliveira and L. Levkowitz. From visualization to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphcis*, 9:378–394, 2003.

[10] A. O. Artero and M. C. F. de Oliveira. Viz3d: effective exploratory visualization of large multidimensional data sets. In *Proceedings of the 17th Brazilian Symposium on Computer Graphics and Image Processing*, pages 340–347, 2004.

[11] T. He, L. Hong, A. Kaufman and H. Pfister. Generetion of transfer functions with stochastic search techniques. In *IEEE Visualization'96*, pages 227–234, 1996.

[12] J. Marks, B. Andalman, P.A.Beardsley and H.Pfister. Generetion of transfer functions with stochastic search techniques. In *IEEE Visualization'97*, pages 227–234, 1997.

[13] C. L. Bajaj, V. Pascucci and D. R. Schikore. The countor spectrum. In *IEEE Visualization'97*, pages 167–173, 1997.

The corresponding work was published in the proceedings of the Sibgrapi 2006, pp. 273–280. IEEE Press, 2006.