Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

JAN REININGHAUS, JENS KASTEN, TINO WEINKAUF, AND
INGRID HOTZ

Combinatorial Feature Flow Fields: Tracking Critical Points in Discrete Scalar Fields

ZIB-Report 11-02 (February 2011)

Combinatorial Feature Flow Fields: Tracking Critical Points in Discrete Scalar Fields

Jan Reininghaus, Jens Kasten, Tino Weinkauf, and Ingrid Hotz

February 10, 2011

Abstract

We propose a combinatorial algorithm to track critical points of 2D time-dependent scalar fields. Existing tracking algorithms such as Feature Flow Fields apply numerical schemes utilizing derivatives of the data, which makes them prone to noise and involve a large number of computational parameters. In contrast, our method is robust against noise since it does not require derivatives, interpolation, and numerical integration. Furthermore, we propose an importance measure that combines the spatial persistence of a critical point with its temporal evolution. This leads to a time-aware feature hierarchy, which allows us to discriminate important from spurious features. Our method requires only a single, easy-to-tune computational parameter and is naturally formulated in an out-of-core fashion, which enables the analysis of large data sets. We apply our method to a number of data sets and compare it to the stabilized continuous Feature Flow Field tracking algorithm.

1 Introduction

Time-dependent 2D scalar data arises in many scientific disciplines. For the analysis of such data, the extraction of minima, saddles, and maxima of each individual time step has been proven useful. These point features of the data are often called critical points. To understand the dynamic behavior of time-dependent data, it can be beneficial to analyze the temporal evolution of these critical points.

To enable an efficient quantification of the temporal evolution of the critical points, we can track them over time. In this paper, we call such a tracked critical point a critical line of the data. Many different algorithms that extract critical lines have been proposed, see Section 2 for a small overview.

For smooth data, the Feature Flow Field method [TS03] provides a particularly sound mathematical foundation. Given a smooth time-dependent scalar field, the critical lines are implicitly defined by streamlines in a higher dimensional derived vector field. While this method works well for smooth data, its application to data that is only continuous is problematic as derivatives have to be computed. To circumvent this problem, derivative free algorithms employing concepts from algebraic topology have been developed recently, see Section 2.

The main remaining weakness of the available algorithms is their inability to handle noisy data in a meaningful way. Such data usually contains an overwhelming number of critical lines that hinder meaningful visual data analysis. To reduce the number of critical lines, one typically smooths the data or discards short critical lines. Both approaches can be problematic. A simple smoothing of the data may remove important critical lines and affect the spatial position of the critical lines, see Figure 7 for an example. Discarding short critical lines may remove an important and stable, but short lived feature. See Figure 9 for an example of such a short but important critical line.

This paper proposes a combinatorial algorithm that is able to track critical points in noisy data. This robustness is achieved by combining Forman's notion of a combinatorial gradient field [For98a] with the notion of Persistence proposed by Edelsbrunner et al.[EHZ01]. Persistence is a well founded importance measure for critical points. Together, these concepts enable a robust and consistent combinatorial representation of the gradient of a scalar function. Both fundamental concepts will be briefly introduced in a graph theoretic formulation in Section 3.1.

A definition for a critical line of a sequence of combinatorial gradient fields was recently proposed by King et al. [KKM08]. The basic idea is similar to the continuous Feature Flow Field method - a higher dimensional field is constructed in which the critical lines are given by combinatorial streamlines. We therefore refer to the higher dimensional field as a Combinatorial Feature Flow Field in this paper. We formulate King's definition of critical lines in combinatorial gradient fields using a graph theoretic formulation in Section 3.2.

Our main contribution is the introduction of the first efficient algorithm that extracts the critical minima, saddle, and maxima lines in 2D discrete scalar fields using Combinatorial Feature Flow Fields.

The proposed algorithm has many valuable properties. It has a reasonable running time and is naturally formulated in a out-of-core fashion enabling the analysis of large data sets as only two subsequent time steps have to be kept in memory. The input consists of a regular cell complex, so the algorithm can deal with many widely used representations of discrete data like triangulations, quadrangulations, or a mixture of these. It contains only one easily-tuned computational parameter, the persistence threshold σ, used to construct the combinatorial representation of the gradient fields.

Due to the combinatorial nature of our algorithm, we can formulate a natural spatio-temporal importance measure for the resulting critical lines called Integrated Persistence (see Section 3.3).

2 Related Work

Many algorithms that track features in time-dependent data have been proposed in many different scientific communities. A lot of this work has been partially inspired by object tracking methods in the area of computer vision, see [YJS06] for a survey. In the context of visual data analysis, tracking approaches can roughly be categorized into three classes depending on the treatment of the temporal dimension [Pos03].

The first class considers feature tracking as a two-step process: feature extraction for each time slice and subsequent feature matching solving a correspondence. Such methods do not rely on a temporal interpolation. Event analysis mostly happens implicitly during tracking defined by event functions. Common tracked features are volumes or areas, boundaries or contours and points. Correspondence criteria use distance metrics of the domain and the attribute space, which are in general based on application specific heuristics. Typical attributes comprise feature size, shape descriptors or also texture characteristics [CJR07]. Features are linked, if their distance falls below a given threshold [SSZC94, RPS99, LBM+06, dLvL01]. Improvements using feature overlap instead of Euclidean distances are used in [SW97]. A more global approach is followed in [Ji06] employing a best matching algorithm. Improved tracking can be achieved by utilizing additional information for motion prediction [RSVP02]. [BSS02] proposes a progressive tracking of isosurfaces using the isosurface at time t as an initial guess for the next time-step $t+1$. An extension to tracking of the entire contour tree using volume overlap has been proposed in [SB06].

The second class of algorithms considers time as additional dimension, treated equally to spatial dimensions. Features are extracted from space-time directly, thus increasing the dimension of the domain and the features by one. Tracking is accurate with respect to the chosen temporal interpolation. No explicit distance metrics for features are needed. Event analysis is mostly a subsequent step after tracking and is based on well-founded theory. Methods extracting isosurfaces in space-time have been proposed in [WB98, JSW03]. A topological event analysis based on the Reeb-Graph of the surface resulting from sweeping contours has been performed in [WBD+ar, BWP+10]. The development of topological structures in 2D and 3D flow fields has been analyzed in [TWSH02, GTS04]. These algorithms consider vector fields composed of space-time cells with linear interpolation, for which events are restricted to cell boundaries. Critical point tracking thus reduces to the computation of entry and exit points for each cell. Similarly, [BP02] introduces an algorithm to track vortex core lines over time and scale space searching for features, represented as parallel vectors, on all boundary cells of the space-time cell [BP02]. While giving accurate results, these methods are prone to noisy data and a high feature density. To reduce the number of extracted features and events, a common practice is to delete short living features. A combinatorial approach to track critical points is based on the definition of Jacobi sets [EH04]. It consists of Jacobi edges, which are extracted from a spatial-temporal simplicial complex assuming a linear interpolant. The decision whether an edge belongs to the Jacobi curve involves the topological analysis of the lower link of vertices and edges of the simplicial complex. While providing a nice theoretical framework, the resulting Jacobi curves of real data sets are often very complex and hard to analyze. Based on this work it is also possible to track the evolution of the Reeb-graph of a scalar function [EHM+08].

The third class of algorithm combines aspects of both above-mentioned types. They represent the dynamic behavior of features implicitly as streamlines of a higher dimensional derived vector field in space-time. Critical points can then be tracked by computing certain streamlines in this vector field, referred to as a Feature Flow Field [TS03]. Recently, a combinatorial version of the Feature Flow Field method has been proposed [KKM08]. This method is discussed in detail in Section 3.2 and provides the mathematical foundation for our novel tracking algorithm presented in Section 4.

3 Fundamental Concepts

The purpose of this section is to introduce the reader to the main concepts that build the mathematical foundation for our combinatorial tracking algorithm described in detail in Section 4. We first introduce the reader to the well known concept of combinatorial gradient vector fields (CGF) in Section 3.1 using a graph theoretic formulation. Using this concept we can define the notion of a combinatorial feature flow field (CFFF) in Section 3.2. We conclude this Section with a definition of a time-aware importance measure for the tracked critical points that is based on the notion of persistence.

3.1 Combinatorial Gradient Fields

For simplicity, we restrict ourselves to 2D manifolds while the mathematical theory for combinatorial gradient fields is defined in a far more general setting [For98a]. Let C denote a finite regular cell complex of a 2D manifold. Examples of such cell complexes are triangulations or quadrangular meshes. Given C, we first define its cell graph $G_C = (S, L)$ that encodes the combinatorial information contained in C in a graph theoretic setting.

The nodes S of the graph consist of the cells C of the complex and each node u^p is labeled with the dimension p of the cell it represents. For a triangulation, the nodes of the cell graph therefore consist of the vertices (0-cells), edges (1-cells), and triangles (2-cells).

The links L of the graph encode the neighborhood relation of the cells in C: if the cell represented by node u^p is in the boundary of the cell represented by node w^{p+1} then $\ell^p = \{u^p, w^{p+1}\}$ is a link in the graph. Note that we label each link with the dimension of its lower dimensional node.