Simplification of Scalar Data via Monotone-Light Factorizations

Martin Brooks* Liam Watson†

Abstract

Using the monotone-light factorization of continuous functions [7, 11, 12], we outline a new approach to data simplification. This approach puts very few restrictions on the sampling domain, and provides a mathematical framework for computational methods simplifying scientific and geometric data.

1 Introduction

This paper outlines *extremal simplification* [2], a new approach to data simplification that utilizes the monotone-light factorization of a continuous function [7, 12, 11]. A finite data set on a topological space X must be interpolated to a continuous function in order to make use of the monotone-light factorization. As such, extremal simplification is comprised of two key steps: Interpolation of a discrete function by patches (section 4), and simplification of the continuous interpolation (section 3). In particular, extremal simplification provides the mathematics required by algorithms that move from discrete data to continuous interpolation to topological structure, simplify the topological structure, and then generate simplified data. Extremal simplification is correct in the sense that continuous interpolation of the simplified data has the intended simplified topological structure.

Various topological structures have been used to guide simplification of scalar data, including the Reeb graph or contour tree [3, 4, 10], the Morse-Smale complex [1], and the persistence diagram [6, 5]. Extremal simplification may be thought of as an enriched variation of Reeb graph methods; the additional information provided by the monotone-light factorization provides tools for analysis (section 5) and a foundation for further algorithmic extensions. In addition, extremal simplification provides a general theory of data simplification that is well suited to a range of practical applications. These generalities include flexibility in both input and output. Indeed, extremal simplification applies to data sampled on spaces of any dimension, as well as *glued together* spaces such as networks and geometric models. Data is simplified by removing local extrema; the extrema to be removed may be selected by arbitrary preference, by persistence [6], or via local geometric measures [3]. Further, extremal simplification does not uniquely specify simplified data values, thereby allowing the simplified data to be engineered to meet specific application requirements.

*Institute for Information Technology, National Research Council of Canada, martin.brooks@nrc.ca

†Département de Mathématiques, Université du Québec à Montréal, liam.watson@cirget.ca

2 Some requisite point-set topology

We give a brief overview of the definitions from point-set topology that will be required in the sequel. All of these are standard and there are many references, for example [8]. We assume that the reader is familiar with the definition of a metric space, as well as the basic language of set theory.

A topological space X is a set together with a collection of open subsets $\mathcal{U}$; the only requirements are that the the empty set $\emptyset$ and the entire set X be included in this collection. Further, for open sets U_1 and U_2 the open set $U_1 \cap U_2 \in \mathcal{U}$, as well as $\cup_{i\in I} U_i \in \mathcal{U}$ for an arbitrary collection of open sets $U_i \in \mathcal{U}$ (the index set I may be infinite). A basic example is the Real line $\mathbb{R}$, where the open sets are open intervals of the form (a, b) for real numbers $a < b$ (together with $\emptyset$ and $\mathbb{R}$). These open intervals form a basis of this topology on $\mathbb{R}$. The complement of an open set is a closed set.

A function $f : X \to Y$ between topological spaces is called continuous if for every open set $U \in Y$, the set $f^{-1}(U) \in X$ is open also. In the case where the spaces X and Y are additionally metric spaces (as is, for example, $\mathbb{R}$), it can be verified that this definition coincides with the standard epsilon-delta definition of continuity. If $g : Y \to X$ is continuous with $f \circ g$ and $g \circ f$ identity functions then X and Y are said to be homeomorphic.

Given topological spaces X and Y, the product $X \times Y$ may be given the product topology where the open sets are generated by $U \times V$ for U and V open sets of X and Y respectively. The plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is one such example. Another standard topology is the subspace topology: Any subset of X may be viewed a topological space by intersecting open sets with the given subspace. In this way, the graph of a continuous function $f : X \to Y$ may be viewed as a topological space as a subspace of $X \times Y$.

We will make use of the quotient topology construction throughout. Given a topological space X and a closed subset K (that is, the complement of K is open in X), we define the quotient X/K as the space $(X \smallsetminus K) \cup \{*\}$ where $*$ is a single point.[1] It remains only to describe the open sets for this new topological space; the quotient topology. There is a natural map associated to a quotient given by

$$q(x) = \begin{cases} x & x \notin K \\ * & x \in K. \end{cases}$$

A set in X/K is declared open whenever the inverse image $q^{-1}(K)$ is an open set of X. For example a circle S^1, as a topological space, may be obtained as a quotient of $\mathbb{R}$ by any closed set of the form $\mathbb{R} \smallsetminus (a, b)$. In this setting the quotient operation identifies two ends of a closed interval.

More generally, given any equivalence relation $\sim$ on the topological space X we may consider the quotient $X/\sim$ endowed with the quotient topology (above, the equivalence relation is simply $x \sim y$ whenever $x, y \in K$). In this case the points of $X/\sim$ are the equivalence classes $[x] = \{y \in X | x \sim y\}$. Again there is a natural quotient map q, and the open sets are determined by there inverse images.

The definitions given above are general, and examples of topological spaces abound. In practice however, we need to make some additional assumptions on the spaces used. Indeed our example $\mathbb{R}$ clearly has many additional properties (it is a metric space, for example). Extremal simplification applies to a finite data set drawn from a Peano space X; that is, a compact, connected and locally connected metric space. We recall these definitions presently.

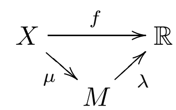
A topological space X is said to be compact when every cover of X by open sets contains a finite subcover and connected whenever $X = U_1 \cup U_2$ and $U_1, U_2 \neq \emptyset$ implies that $U_1 \cap U_2 \neq \emptyset$. Note that S^1 is compact and connected, as is any closed interval in $\mathbb{R}$.

It is easy to see that the curve given by $\sin\left(\frac{1}{x}\right) \cup \{0\}$, as a subspace of $\mathbb{R}^2$, is connected. However, there are open sets containing 0 that are disconnected. This is an example of a space that is connected, but not locally connected. A locally connected space X has the property that the open sets of X may be described by a basis of open sets which are connected.

It should be noted that for a topological space to be given a metric structure, it suffices that the topological space be regular (given a point $x \in X$ and an open set U containing x there is a closed set K such that $x \in K \subset U$) and second countable (X has a countable basis of open sets).

Peano spaces include bounded subsets of $\mathbb{R}^n$ (in particular, closed intervals in $\mathbb{R}$ as above, and graphs as in section 3), and compact manifolds in any dimension (S^1, for example). However, Peano spaces may be more complicated; they can be of of variable dimension or have fractal characteristics.

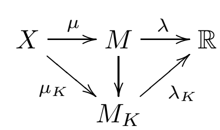
3 Monotone-Light factorizations and simplification



Suppose $f : X \to \mathbb{R}$ is a continuous function defined on a Peano space X. A classical result due to Eilenberg [7] and Whyburn [11, 12] states that f admits a unique monotone-light factorization.[2] That is, $f = \lambda \circ \mu$ where $\mu : X \to M$ is *monotone* and $\lambda : M \to \mathbb{R}$ is *light*. A function is called monotone whenever the inverse image of a connected set is connected, and light whenever the inverse image of a point is discrete (that is, a finite collection of points). M is called the *middle space* for the function f.

It is instructive to see how such a decomposition may be constructed (cf. [12]). To each point y in the range of f we may consider the level set $f^{-1}(y) \subset X$. We take as equivalence classes the connected components of every such level set, and form the quotient, $M = X/\sim$. That is, $[x]$ is the connected component of $f^{-1} \circ f(x)$ containing x. The associated quotient map is the desired monotone factor μ, since the equivalence classes are connected by construction. Now define the light factor as $\lambda([x]) = f \circ \mu^{-1}([x]) = y$.

From this decomposition, the function f factors through the middle space M. In this setting M is a directed graph. The light factor λ identifies each arc [3] of M with a closed interval in $\mathbb{R}$; the arc's direction is from smaller to larger. In general, this directed graph may be infinite, but we restrict our attention to functions having middle space M that is a finite graph (this is the case for admissible interpolations, see section 4). As a result, the middle space is the function's Reeb graph [9].



We now describe extremal simplification. Let $K \subset M$ be a closed, connected subset of the middle space M such that the restriction to the boundary $\lambda|_{\partial K}$ is constant. Suppose that K has the additional property that if all arcs intersecting ∂K are outgoing (resp. incoming), then ∂K is comprised entirely of minima (resp. maxima) of f. Such a K is referred to as an *extremal collapse set*. Generally, the middle space M has many extremal collapse sets K; each choice of K results in a distinct simplification. The key step for extremal simplification is to choose an extremal collapse set K and to collapse K to a point, resulting in a (simplified) quotient graph $M_K = M/K$, having fewer extrema M. In this case the graph M_K

[1] Notation $X \smallsetminus K$ indicates *set subtraction*: all elements of X not in K.

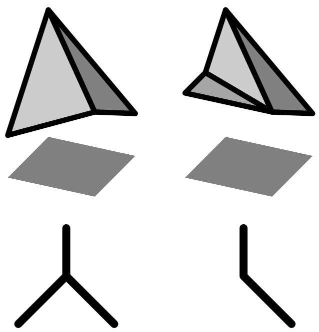
[2] In fact, the result is much more general; the range of f need not be restricted to $\mathbb{R}$.

[3] Arcs include their endpoints.

determines a new monotone factor μ_K by composition with the quotient. The new light factor is given by

$$\lambda_K(x) = \begin{cases} \lambda(x) & x \notin K \\ c & x \in K \end{cases}$$

where c is the constant value of $\lambda|_{\partial K}$. Thus, the simplified function $f_K = \lambda_K \circ \mu_K$ has middle space M_K.



Note that this new function differs from f by flattening the region $\mu^{-1}K$, and is familiar from the computational topology literature [3, 10]. We call f_K a *flat* simplification. An example of a piecewise linear function and one of its flat simplifications is shown on the right, together with their respective Reeb graphs.

The flat simplification f_K is only one of many possible extremal simplifications of f that collapse K; however, we will not discuss the others in this paper, except to mention that this flexibility will be useful in applications. It is shown in [2] that extremal simplification's use of quotients is more general than contour tree pruning rules [3]. Extremal simplification's freedom to choose the collapse set K allows successive simplifications to proceed in the order most appropriate to the application.

4 Interpolation and patching

Suppose we have a finite set $D \subset X$ and real-valued measurements $F : D \to \mathbb{R}$; we refer to F as *scalar data*. A continuous interpolation f is obtained by covering X with patches, interpolating on each patch. There may be many ways that X can be covered,[4] and many ways in which the individual patches may be interpolated. Approaches in the literature include linear interpolation on a triangular mesh, as well as bi-linear and tri-linear interpolation on square and cubic meshes respectively. In contrast, extremal simplification uses an axiomatic approach to patch interpolation, as follows.

A finite collection $\mathcal{P} = \{P_1, \dots, P_n\}$ of subsets of X is a *patch geometry* whenever the following five conditions are satisfied:

(1) $P_1 \cup \cdots \cup P_n = X$;

(2) each P_i is connected;

(3) each $P_i = \overline{P_i^\circ}$ (that is, each P_i is the closure of its interior);

(4) $P_i^\circ \cap P_j^\circ = \emptyset$ whenever $i \neq j$ (that is, intersection of patches occurs only on the boundary); and

(5) each intersection $P_i \cap P_j$ has finitely many connected components.

[4] Of course, we assume that at least one such covering exists.

Functions $h_i : P_i \to \mathbb{R}$ (for $i \in \{1, \dots, n\}$) constitute *patch interpolants* for F when $h_i|_{(P_i \cap D)} = F|_{(P_i \cap D)}$ and the h_i agree on the patch intersections, i.e. $h_i|_{P_j} = h_j|_{P_i}$ for all $i, j \in \{1, \dots, n\}$. Thus, the individual patch interpolants combine to provide a *patch interpolation* $f : X \to \mathbb{R}$ for F by defining $f|_{P_i} = h_i$.

Suppose patch interpolants $h_1, \dots, h_n$ yield the patch interpolation f. Choose any pair of indices $i, j \in \{1, \dots, n\}$ (these need not be distinct), consider a connected component $C \subset P_i \cap P_j$. Let f's maximum and minimum values on C be $z_{\max} = \max_{c \in C}\{f(c)\}$ and $z_{\min} = \min_{c \in C}\{f(c)\}$, and now consider the level sets $H_{\max} = f^{-1}(z_{\max})$ and $H_{\min} = f^{-1}(z_{\min})$. We say that D *witnesses* C whenever $D \cap C \cap H_{\max} \neq \emptyset$ and $D \cap C \cap H_{\min} \neq \emptyset$, and that D witnesses f when D witnesses every component C of every intersection $P_i \cap P_j$.

We now impose an admissibility requirement on f: The patch interpolation f is *admissible* if D witnesses f and each patch interpolant h_i is a monotone function.

Example. Barycentric interpolation on simplexes having data at the vertices constitutes admissible interpolation for any data.

Example. n-linear interpolation on n-cubes having the data at the vertices is admissible only if the data F is such that each cube's interpolant is monotone. However, when some cells' interpolants are not monotone, then those cells can be subdivided into linearly interpolated simplexes, providing a hybrid patch geometry having admissible interpolation.

Current research includes investigation of NURBS and other commonly used interpolations with respect to admissibility.

The following theorems show the importance of admissible interpolations.

Theorem 1 *Suppose f and f' are two admissible patch interpolations of F, both defined on patch geometry $\mathcal{P} = P_1 \dots P_n$. Then the monotone-light factorizations of f and f' have identical middle spaces (Reeb graphs) and identical light factors.*

For patch geometry $\mathcal{P}$, this result allows us to assign a unique Reeb graph and light factor to the scalar data F, since the choice of interpolant is irrelevant. This in turn allows us to define simplification of scalar data: Suppose f is any admissible interpolation of F, and suppose f has monotone-light factorization $f = \lambda \circ \mu$ with Reeb graph M. Choose any extremal collapse set $K \subset M$, and let f_K be the flat simplification, having Reeb graph M_K and light factor λ_K. We now define simplified scalar data as $F_K = f_K|_D$.

Theorem 2 *f_K is an admissible interpolation of F_K for patch geometry $\mathcal{P}$.*

Theorem 2 tells us that F_K has Reeb graph M_K and light factor λ_K, as intended. Theorem 1 tells us that every admissible interpolation of F_K defined on $\mathcal{P}$ has this same Reeb graph and light factor.

5 Approximation error

A simple analysis of f's Reeb graph M yields the persistence [6] of each extremum; this is a scale measurement that may be interpreted as the significance of the extremum, defined as follows. Given a vertex $p \in M$ corresponding to an extremum of f, we will consider real closed intervals I such that $\max(I) = \lambda(p)$ when p is a maximum and $\min(I) = \lambda(p)$ when p is a minimum. Then there is a smallest such interval I_p such that either (1) there is a point $q \neq p$ with $\lambda(p) = \lambda(q)$ in the connected component of $\lambda^{-1}(I_p)$ containing p, or (2) $I_p = \text{image}(f)$. The length of I_p is the persistence [6] of the extremum p.

The monotone-light factorization yields a lower bound on approximation error for simplification. Suppose that $f, g : X \to \mathbb{R}$ are continuous functions.[5] Let $\sigma(f)$ denote the minimal persistence among f's extrema.

Theorem 3 *If g has fewer extrema than f then*

$$\max_{x \in X} |f(x) - g(x)| \geq \frac{\sigma(f)}{2}.$$

This bound also applies to scalar data. Suppose $F, G : D \to \mathbb{R}$ are scalar data and $\mathcal{P}$ is a patch geometry such that both F and G have admissible interpolations for $\mathcal{P}$. Define $\sigma(F) = \sigma(f)$ for any admissible interpolation f of F, noting that this quantity is independent of the choice of f.

Theorem 4 *If G has fewer extrema than F then*

$$\max_{x \in D} |F(x) - G(x)| \geq \frac{\sigma(F)}{2}.$$

6 Conclusion

This paper has introduced extremal simplification, a useful consequence of the monotone-light factorization of classical point-set topology. The monotone-light factorization provides a mathematical foundation for rigorous development of data simplification algorithms; details and proofs are found in [2]. Applicable to scalar data defined on spaces of any dimension, extremal simplification allows extrema to be removed in application-specific order. Moreover, admissibility ensures that the simplification is independent of the choice of interpolation. We have also shown that approximation error bounds can be proved for both scalar data and continuous functions.

[5] Assume f and g have middle spaces that are finite graphs.This is true, for example, when each function is an admissible interpolation of scalar data as in section 4.

Acknowledgements

This work was carried out in 2005 while the second author was a visiting researcher at the National Research Council.

References

[1] P.-T. Bremer, B. Hamann, H. Edelsbrunner, and V. Pascucci. A topological hierarchy for functions on triangulated surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 10(4):385 – 396, 2004.

[2] M. Brooks and L. Watson. Simplification of sampled scalar fields by extremal simplification. *NRC Technical Report ERB-1147 / NRC 49359*, 2007.

[3] H. Carr. Topological manipulation of isosurfaces. PhD Thesis, University of British Columbia, 2004.

[4] H. Carr, J. Snoeyink, and M. van de Panne. Simplifying flexible isosurfaces using local geometric measures. In *IEEE Visualization*, pages 497–504. IEEE Computer Society, 2004.

[5] H. Edelsbrunner, D. Morozov, and V. Pascucci. Persistence-sensitive simplification of functions on 2-manifolds. *Proc. 22nd ACM Symp. Computational Geometry*, 2006.

[6] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comp. Geo.*, 28:511 – 533, 2002.

[7] S. Eilenberg. Sur les transformations continues d'espaces métriques compacts. *Fundam. Math.*, 22:292–296, 1934.

[8] John L. Kelley *General Topology.* Graduate Texts in Mathematics, v. 27. Springer-Verlag, 1975. (first published 1955)

[9] G. Reeb. Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique. *C. R. Acad. Sci. Paris*, 222:847–849, 1946.

[10] G.H. Weber, S.E. Dillard, H. Carr, V. Pascucci, and B. Hamann. Topology-controlled volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 13(2), 2007.

[11] G.T. Whyburn and E. Duda. *Dynamic topology.* Springer-Verlag, New York, 1979. Undergraduate Texts in Mathematics, Kelley.

[12] G.T. Whyburn. *Analytic Topology.* American Mathematical Society Colloquium Publications, v. 28. American Mathematical Society, New York, 1942.