Natural Neighbor Concepts in Scattered Data Interpolation and Discrete Function Approximation

Tom Bobach and Georg Umlauf
University of Kaiserslautern
Geometric Algorithms Group
Department of Computer Sciences
D-67653 Kaiserslautern, Germany
http://www-umlauf.informatik.uni-kl.de/˜bobach | ˜umlauf

Abstract: The concept of natural neighbors employs the notion of distance to define local neighborhoods in discrete data. Especially when querying and accessing large scale data, it is important to limit the amount of data that has to be processed for an answer. Because of its implicit definition on distances, the natural neighbor concept is extremely well suited to provide meaningful neighborhoods in spatial data with a scattered, inhomogeneous distribution.

This paper revisits some unique properties of natural neighbor based methods and summarizes important findings for their successful application to scattered data interpolation, and the computation of discrete harmonic functions.

1 Introduction

Many scientific areas deal with phenomena in a spatial context when modeling or investigating real world problems. Computational approaches, be it in the endeavor of data generation or data analysis, inherently deal with discrete data representations. Based on the spatial structure of the data which can range from completely regular to inhomogeneous and scattered, the way how we access the data model has big impact on space and time requirements of algorithms.

In the first part of this work we are concerned with smooth, local interpolants in large, inhomogeneous, and unstructured data, a challenging problem for data representation and -access. We focus on the local reconstruction of a real-valued function from data in an adequate spatial neighborhood. Often – in absence of further knowledge about the modeled problem – this neighborhood is best expressed in terms of natural neighbors, a neighborhood relation defined on the Voronoi diagram of the data sites.

Another problem that is related to locally defined functions is the computation of discrete harmonic functions. We discuss implications of a discretization domain that itself is a function of time and how natural neighbors are beneficial in this context.

In the remainder we first revisit previous work while discussing under which circumstances natural neighbor concepts are feasible. We then turn to the algorithmic impact caused by
natural neighbor concepts in scattered data interpolation. The issues that arise in the course of applying $C^2$-continuous natural neighbor interpolation are presented along with a brief sketch of available solutions. Leaving the area of function interpolation we then look at the computation of discrete harmonic functions over scattered point sets under continuous deformation.

2 The Natural Neighbor Concept

In a set of geometric entities $X = (x_1, \ldots, x_m)$, we call $x_i$ and $x_j$ natural neighbors if there is a point that is closer to $x_i$ and $x_j$ than to any other entity in $X$. This notion of spatial proximity is formalized in the Voronoi diagram, also known as Dirichlet tessellation or Thiessen polygons [Aur91, OBSC00]. The discrete set $X$ of known geometric entities is called the set of Voronoi sites, and its Voronoi diagram is the partition of space into so-called Voronoi tiles ($T_1, \ldots, T_m$), such that

$$p \in T_i \iff d(p, x_i) \leq \min_{1 \leq j \leq n} d(p, x_j)$$

for every point $p$ and a distance measure $d$. In this structure, natural neighbors are exactly those Voronoi sites whose tiles have a non-empty intersection.

One of our goals is the reconstruction of an unknown function that is locally defined by nearby data, where proximity is defined via the distance measure. In this context natural neighbors provide an excellent notion of neighborhood.

Although (1) encompasses Voronoi sites of any shape and dimension, and is valid for arbitrary distance measures $d$, not all choices are algorithmically feasible. We present two of the most basic and elegant variants.

The traditional Voronoi diagram is uniquely defined by the point-shaped sites $x_i \in \mathbb{R}^n$ and the Euclidean metric $d_2(x, y) = \|x - y\|_2$. It comes with properties that are typically required in locally operating geometric algorithms: invariance under rigid transformations and uniform scaling. Moreover, its tiles $T_1, \ldots, T_m$ are convex polyhedra, allowing easy and robust geometric access. The most important fact about Voronoi diagrams, from a computational point of view, is their duality to Delaunay triangulations. These data structures are well-understood with a large support in terms of fast and robust algorithms and computation libraries, and allow easy and fast access to the entities of the Voronoi diagram.

The Voronoi diagram can be generalized e.g. in terms of different distance measures, or by the shape of sites. Each generalization usually leads to more complex algorithms and data structures and causes the loss of one or the other desirable property of the traditional Voronoi diagram. One noteworthy generalization that retains most of the geometric elegance of the traditional Voronoi diagram is the power- or Laguerre Voronoi diagram. Based on a non-uniform metric

$$(d_p(p, x_i))^2 = (d_2(p, x_i))^2 - w_i$$

that uses weights $w_i$ associated with the sites $x_i$, it still leads to convex Voronoi tiles and
is dual to the *regular triangulation* of a weighted point set, which is only slightly more complicated to maintain than the Delaunay triangulation.

Generalizations in the shape of the sites usually do not dualize as easily to a triangulation and require more general graph structures to represent the Voronoi diagram.

Depending on the characteristics of a data set, natural neighbor concepts can prove valuable for local function approximation or interpolation, the key indicators being

- a spatial setting,
- proximity that is based on a distance measure with a geometric interpretation,
- a sought-after function that locally depends on known data.

### 3 Related Work

Although the general concept of natural neighbors is not restricted to interpolation and local coordinates, the latter are their most prominent applications. The introduction of natural neighbor based local coordinates which possess $C^1$ continuity almost everywhere was done in the pioneering work of [Sib80] under the nowadays misleading term “natural neighbor coordinates” and later applied to globally $C^1$ scattered data interpolation in [Sib81]. Further results on the properties of the $C^1$ coordinates followed in [Far90, Pip92].

Following the concept underlying the $C^1$-continuous coordinates, $C^0$-continuous coordinates have been independently introduced by [CFL82] in the context of random lattices in nuclear physics, by [BIK+97] under the term “Non-Sibsonian” coordinates, and by [Sug99] as “Laplace natural neighbor” coordinates.

In [HS00b] an integral relation between $C^0$ and $C^1$ coordinates was shown and generalized to almost everywhere $C^k$-continuous natural neighbor coordinates.

The drawback of reduced ($C^0$) continuity of interpolants based on the coordinates was overcome in [Sib81, Far90] who devised globally $C^1$-continuous interpolants, and in [HS04] who devised a globally $C^2$ continuous interpolant.

A generalization of natural neighbor interpolation to line- and circle-shaped Voronoi sites was performed in [AMG98, GF99, HS00a].

The algorithmic implications of natural neighbor interpolation have been investigated in [BBU06, BS95, Hiy05], and acceleration approaches were presented in [FEK+05, PLK+06].

Support for the manipulation and access of generalized Voronoi diagrams by means of graphics hardware is due to [HCK+99].

The implicit definition of higher order of continuity based on data that lacked explicit derivative information has been the goal of [Sib81, Cla96, Flö03, BBU06].

Finally, the non-trivial definition of natural neighbor coordinates in the tangent space of manifolds was investigated in [BC00, Flö03].
4 Smooth Data Interpolation with Natural Neighbors

The scattered data interpolation problem can be stated as follows. We assume a scattered, unstructured set of data sites \( X \) and a partially defined function \( f : X \rightarrow \mathbb{R} \) together with its first \( k \) derivatives \( f^{(1)}, \ldots, f^{(k)} \) at each site. We seek to construct a function \( \Phi \) that agrees with \( f \) and its first \( k \) derivatives on \( X \) and otherwise satisfies properties including, but not limited to, smoothness and variation minimization. In the following we consider the interpolation of scalar values and derivatives up to order two in \( \mathbb{R}^2 \), i.e. gradients \( \nabla f(x_i) \) and Hessians \( \mathcal{H} f(x_i) \), and concentrate on the evaluation of \( \Phi(q) \) at a query position \( q \in \mathbb{R}^2 \).

In the corresponding research area of scattered data interpolation, many efficient local and global schemes have been proposed, able to deal with a large variety of input data. Interpolation schemes with global support lead to better results in general than schemes with local support, at the expense of considerably increased computational complexity. Local schemes, on the other hand, depend on a definition of “local” that often amounts to a user-provided parameter, making especially inhomogeneously distributed data hard to deal with.

Natural neighbor scattered data interpolation determines the local support for the reconstruction from the set of natural neighbors, thus coming with implicit and automatic control over the neighborhood. The evaluation of natural neighbor schemes at a point \( q \) operates on the set \( N_q := (x_1, \ldots, x_n) \) of natural neighbors in the Voronoi diagram of \( X \cup \{ q \} \), and involves the following steps:

1. computation of coordinates \( \lambda(q) = (\lambda_1(q), \ldots, \lambda_n(q)) \in \mathbb{R}^n \) of \( q \) with respect to the points in \( N_q \),
2. if not provided, estimation of the first \( k \) derivatives \( f^{(1)}, \ldots, f^{(k)} \) with respect to \( x \) at the data points, and
3. setup and evaluation of a multivariate function \( \varphi \) in the coordinates, giving the interpolant as \( \Phi = \varphi(\lambda(q)) \).

The rest of section is dedicated to the algorithmic efforts involved in the computation of \( C^0 \), \( C^1 \), and \( C^2 \)-continuous interpolants. After pointing out in what respect the implementation of the considered \( C^2 \) interpolant is more complex, we devise an algebraic rather than geometric approach for the computation of coordinates to simplify the algorithmic realization. A similar difficulty can be observed for the estimation of higher order derivatives, for which we refer to a recursive approach with limited complexity.

4.1 Natural neighbor schemes up to first order continuity

Natural neighbor coordinates as proposed in [Sib80, CFL82, Sug99, BIK+97] are computed from sizes of geometric entities in the Voronoi diagram of the data sites. Thanks to the duality between Voronoi diagram and Delaunay triangulation, operations on the
Voronoi diagram can in general be reduced to the traversal of adjacent elements in the Delaunay triangulation. We briefly sketch the steps involved in the computation of $C^0$ and $C^1$-continuous interpolants.

The set $N_q$ of natural neighbors is identical to the set of edge-adjacent neighbors of $q$ in the Delaunay triangulation of $X \cup \{q\}$. We can assume $N_q = (x_1, \ldots, x_n)$ to be ordered counter-clockwise around $q$. The vertices $(v_1, \ldots, v_n)$ of $q$'s Voronoi tile $T_q$ are the circumcenters of the triangles $\triangle(q,x_i,x_{i+1}), 1 \leq i \leq n$, where we assume $N_q$ to be cyclic, i.e. $x_{n+1} := x_1$.

Now, the $C^0$-continuous natural neighbor coordinates $\lambda^0(q) = (\lambda^0_1(q), \ldots, \lambda^0_n(q))$ of $q$ with respect to $(x_1, \ldots, x_n)$ are defined as

$$\lambda^0_i(q) = \lambda^0(q) / \sum_{x_j \in N_q} \lambda^0_j(q), \quad \tilde{\lambda}^0_i(q) = \|v_{i-1} - v_i\| / \|x_i - q\|.$$

From the above definitions it is obvious that the computation of $\lambda^0(q)$ amounts to the iteration of the one-ring of $q$ after it has been inserted into the Delaunay triangulation of $X$. In three-dimensions, the computation of $\lambda^0_i(q)$ requires the iteration around the Delaunay edge $(q,x_i)$ and the area computation of a convex polygon.

A $C^0$-continuous interpolant is now given by

$$\Phi^0(q) = \varphi(\lambda^0(q)) = \sum_{x_i \in N_q} \lambda^0_i(q)f(x_i).$$

An example of a height field interpolation based on $\Phi^0$ is shown in Figure 1.

Natural neighbor coordinates with $C^1$-continuity in $\mathbb{R}^2 \setminus X$, initially proposed by Sibson [Sib80], are defined by

$$\lambda^1(q) = |\mathcal{P}_i \cap \mathcal{T}_q| / |\mathcal{T}_q|,$$

where $\mathcal{P}_i$ is the tile of $x_i$ in the Voronoi diagram of $X$, and $\mathcal{T}_q$ is the tile of $q$ in the Voronoi diagram of $X \cup \{q\}$. In [Wat92] it was observed that $|\mathcal{P}_i \cap \mathcal{T}_q|$ is the sum of signed areas of “dual triangles” which are defined on circumcenters of triangles formed by $q$ and the vertices of Delaunay triangles that would locally be modified by the insertion of $q$. The beauty of this lies in the generality of the resulting formula for the intersection volumes, which is directly applicable to higher dimensions, operating on “dual simplexes”.

Figure 1: Discrete, scalar valued height field interpolated using the $C^0$-continuous interpolant $\Phi^0$. 
It should be noted, however, that this approach is numerically unstable near and on edges (or $k-1$-simplexes) of the Delaunay triangulation due to an infinite volume of the dual triangle (simplex). A remedy to this can be found in [Hiy05].

To achieve $C^1$ continuity also at the data sites $X$ by interpolating gradients, two constructions have been proposed in [Sib81] and [Far90] that define $\varphi$ as a polynomial in $\lambda^1(q)$ such that the gradient of the interpolant matches that defined at the data site. The construction in [Far90] builds on cubic Bézier simplexes $b^3$ in $n$ variables for which it is close to trivial to model derivatives at their vertices, which coincide with the data sites. For a thorough treatment of Bézier simplexes, see [dB87]. Consequently, the $C^1$ interpolant is

$$\Phi^1(q) = \varphi^1(\lambda^1(q)) = b^3(\lambda^1(q)).$$  \hfill (3)

Figure 2(b) shows a discrete height field with given gradients interpolated using $\Phi^1$.

In case of unknown gradients, [Sib81] proposed to estimate the gradient at $x_i$ from the weighted least squares plane through $N_i$. This, being a standard approach to derivative estimation, works remarkably well thanks to the utilization of $\lambda^1$ as weights in the least squares fit.

### 4.2 Natural neighbor coordinates with second order continuity

The evaluation of $C^0$ and $C^1$-continuous interpolants requires only the Delaunay triangulation of $X$ and a simple traversal of vertex-adjacent elements. This section deals with the first step in globally $C^2$-continuous interpolation, the computation of local coordinates.

In [HS00b], a framework for the computation of natural neighbor coordinates with $C^k$-continuity in $\mathbb{R}^2 \setminus X$ was proposed that contained above mentioned $C^0$ and $C^1$ coordinates as special cases. The definition of these is based on the concept of power diagrams, which differ from the ordinary Voronoi diagram in the use of the custom distance measure (2). The power diagram shares all properties of the Voronoi diagram with additional control.
of the Voronoi tile sizes by means of the site weights $w_i$, where tiles can vanish for small enough values of $w_i$.

Interestingly, the bisectors bounding a tile $\mathcal{T}_i$ in the power diagram are linearly displaced depending on $w_i$. Thus, for a uniform choice of $w_i = 0$ at all sites except $q$, the corresponding weight $w_q \in [-w_{\text{max}}, 0]$, where $\mathcal{T}_q(-w_{\text{max}}) = \emptyset$, continuously blends between the ordinary Voronoi diagram of $X$ and that of $X \cup \{q\}$, as illustrated in Figure 2(a).

These last facts led to the observation that the areas $\mathcal{T}_q \cap \mathcal{T}_i$ used in the computation of Sibson’s coordinates are swept by the edges of the weight-dependent tile $\mathcal{T}_q(w_q)$ as $w_q$ runs from $-w_{\text{max}}$ to 0. This constitutes an integral relation between the variable length of the Voronoi edge and the area of overlap $\mathcal{T}_q \cap \mathcal{T}_i$ that was generalized to

$$\lambda^k_i(q) = \lambda^k_i(q)/\sum_j \hat{\lambda}^k_j(q), \quad \hat{\lambda}^k_i(q) = \lambda^k_i(q, 0),$$

$$\hat{\lambda}^k_i(q, u) = \int_{-\infty}^{0} \hat{\lambda}^{k-1}_i(q, v)dv, \quad \hat{\lambda}^0_i(q, u) = l_i(q, u)/r_i,$$

where $l_i(q, u)$ is the length of the tile edge separating $\mathcal{T}_q(u)$ and $\mathcal{T}_i(u)$, and $r_i = \|x_i - q\|$.

Figure 2(a) shows that $l_i(q, u)$ is a piecewise linear function, making $\hat{\lambda}^k_i(q, u)$ a piecewise polynomial on support intervals that are determined by the geometry of the Voronoi diagram. The implementation of the integral expression (4) involves rather complex geometric operations and requires careful treatment of degenerate cases.

In [BBU06] the author devised an algebraic approach to the determination of $l(q, u)$ that naturally deals with degenerate situations and generalizes more easily to 3D. A sketch of the approach follows.

The weight-dependent tile $\mathcal{T}_q(u)$ in our consideration is always a convex polytope which has an alternate representation as an intersection of half-spaces whose representation in Hessian normal form can easily be derived from the Delaunay triangulation,

$$\mathcal{T}_q(u) = \bigcap_{x_i \in N_q} H_i(u), \quad H_i(u) = \{p \mid (p - q)(x_i - q) \leq b_i(u)\},$$

where $b_i(u)$ is a linear function.

In [Las83] a recursive algorithm was proposed for the computation of the volume of convex polytopes in half-space representation such as (5). Each recursion expresses the $d$-dimensional volume as a function of $d - 1$-dimensional volumes, until only intersections of real-valued intervals are to be computed at the deepest level. We exploit this by applying the recursive algorithm to (3) while keeping track of the modifications done to the linear functions $b_i(u)$. After $d - 1$ recursions, the 1-dimensional volumes correspond to $l_i(q, u)$.

For $\mathcal{T}_q(u) \in \mathbb{R}^2$, this leads to a linear programming problem whose solution corresponds to the piecewise linear function required for the computation of $\lambda^k_i(q)$ in (4), and can still be solved in a decent fashion for $\mathcal{T}_q(u) \in \mathbb{R}^3$, where this time roots of second order polynomials are involved.
4.3 Natural neighbor interpolant with second order continuity

Based on the $C^k$ coordinate construction and the Bézier simplex idea introduced in [Far90], a globally $C^2$ continuous interpolation scheme was proposed in [HS04], representing $\varphi$ by a quintic Bézier simplex $b_5$ over $\lambda^2(q)$. While the $C^1$ approach in [Far90] interpolates given gradients at the nodes, the $C^2$ approach in [HS04] additionally interpolates Hessians. The evaluation of the interpolant is rather costly in terms of floating point operations, yet the quality of the results is very high.

However, if only function values are provided as input, $\nabla f(x_i)$ and $\mathcal{H}f(x_i)$ must be estimated from the data. To estimate $\nabla f(x_i)$, [Sib81] used Sibson’s coordinates as weights for the least squares plane through the natural neighbors, yielding good results and the reproduction of spherical quadratics (see also [Flö03]). If also second order derivative information shall be extracted from the data, the set $N_q$ is of insufficient size, which renders the estimation unstable most of the time as shown in Figure 3(b). We now sketch the recursive scheme for the estimation of higher order derivatives devised by the author in [BBU06].

Starting with a first iteration over the data, intermediate gradients $\hat{\nabla} f(x_i)$ are fitted based on each site’s natural neighborhood. The second iteration fits $\nabla f(x_i)$ and $\mathcal{H}f(x_i)$, which are the first two terms of the Taylor expansion of $f$ at $x_i$, to the $f(x_i)$ and $\hat{\nabla} f(x_i)$ such that both are approximated as closely as possible. The result of this improved method is shown in Figure 3(c). The approach readily generalizes to higher order derivative estimation.

5 Discrete Harmonic Functions in Time-Dependent Point Sets

We now turn our attention to the approximation of harmonic functions. By definition, a function $f$ is harmonic on a domain $\Omega$ if it satisfies the Laplace equation $\Delta f|_\Omega = 0$, where $\Delta f = \nabla^2 f$. In case of a boundary value problem with Dirichlet boundary conditions of the form $f|_{\partial\Omega} = u$ for a given $u$, a fundamental result from harmonic function theory states that there exists a unique $f$ satisfying the Laplace equation, c.f. [ABR01].
5.1 Discrete Harmonic Functions

In case of a discrete domain \( X = \{ x_1, \ldots, x_m \} \subset \Omega \), the characterization of a discrete harmonic function \( f : X \to \mathbb{R} \) utilizes an approximation of \( \Delta f(x_i) \), which is commonly modeled as a weighted sum of differences

\[
\Delta f_i := \Delta f(x_i) = \sum_{x_j \in N_{x_i}} \lambda_{ij} (x_i - x_j),
\]

where \( N_{x_i} \) is some nearby neighborhood and \( \lambda_{ij} \in \mathbb{R}^+ \) reflect the approximation of the metric. This is a generalization of the sum of second derivatives in finite differences.

This definition directly corresponds to that of the graph Laplacian on a directed, weighted graph \( G = (X, E, \Lambda) \) over the nodes \( X \), with edges \( E = \{ e_{ij} \}_{ij} \subset X \times X \) and edge weights \( \Lambda : E \to \mathbb{R} \), where \( \lambda_{ij} = \Lambda(e_{ij}) \). Under the assumption that the weights \( \Lambda \) are positive and \( G \) is connected, a similar statement about existence and uniqueness of a discrete harmonic function exists for a set of fixed function values.

In order to approximate the continuous Laplacian, a reasonable choice of \( E \) are the edges in the Delaunay triangulation of \( X \). This leaves the choice of \( \lambda_{ij} \), which can be chosen such that the discrete Laplacian reproduces certain properties of the continuous Laplacian, namely

\[
\Delta \text{id} = \Delta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
\]

This is trivially fulfilled if \( \lambda_{ij} \) are chosen to be generalized barycentric coordinates of \( x_i \) in the one-ring neighborhood of the Delaunay triangulation, since by definition of barycentric coordinates,

\[
0 = \sum_{e_{ij} \in E} \lambda_{ij} (x_i - x_j), \quad 1 = \sum_{e_{ij} \in E} \lambda_{ij}, \quad 0 \leq \lambda_{ij}.
\]

Obviously, \( \lambda_{ij} \) is only defined in the interior of the convex hull \( \mathcal{C}(X) \) of \( X \). Therefore, the Dirichlet conditions of the boundary value problem are automatically required at convex hull vertices of \( X \).

With this in hand, the discrete harmonic function that is the solution to the boundary value problem

\[
f |_{\partial \mathcal{C}(X)} = u, \quad \Delta f_i = 0
\]

is found as the solution of a sparse linear system.

When ignoring the limit behavior of the above introduced Laplacian approximation the author confirmed in [BHFU07] that the computed discrete harmonic functions approximate continuous harmonic functions very well.
5.2 Deforming Domain

If the domain discretization $X$ is now a continuous function of time $X(t)$, say in the course of a dynamic simulation, then we intuitively expect the Laplacian approximation and with it the computed discrete harmonic function to continuously reflect the deformation that occurs with time.

The author investigated this aspect in [BHFU07]. It turns out that among the many choices of barycentric coordinates that render acceptable approximations of the Laplacian in the static case, only natural neighbor coordinates lead to a Laplacian approximation that continuously depends on the deformation $X(t)$.

The reason in here is the continuity of local coordinates with respect to their defining neighborhood. The majority of generalized barycentric coordinates are defined with respect to an enclosing polygon, which is derived from the connectivity of the Delaunay triangulation. The connectivity, however, must change at some point of an arbitrary deformation for any triangulation to remain valid. These changes lead to discontinuous changes of the polygonal neighborhood, and consequently to discontinuous jumps in generalized polygonal barycentric coordinates, which is illustrated in Figure 4(a).

Natural neighbor coordinates, on the other hand, are zero exactly when edge flips occur in the Delaunay triangulation, which comes from their relation to sizes of entities in the Voronoi diagram which themselves continuously depend on the positions of the Voronoi sites. This behavior is illustrated in Figure 4(b). Furthermore, the continuity of $\lambda_{ij}$ with respect to $q$, as discussed in Section 4, carries over to the Laplacian approximation.

6 Conclusion

We motivated that for scattered data methods in a spatial context, the definition of “neighborhood” plays an important role. Without prior knowledge about the data, spatial proximity is a valid criterion for neighborhood definition. A completely automatic determination of such a neighborhood is given in terms of natural neighbors in the Voronoi diagram of a
set of data sites.

By pointing out the advantages and algorithmic implications of natural neighbor based methods we provided insight into key indicators for their application.

This has been further supported by sketching the robust implementation of $C^2$ natural neighbor interpolation, derivative estimation and the application of natural neighbor coordinates to the computation of discrete harmonic functions.

7 Acknowledgements

Financial support by the German Research Foundation DFG within the International Research Training Group 1131 ‘Visualisation of Large and Unstructured Data Sets. Applications in Geospatial Planning, Modeling, and Engineering’ is gratefully acknowledged by the authors.

References


