Thin Plate Spline Interpolation on Large 2D Grids

Barrodale Computing Services Ltd. (BCS)

www.barrodale.com

Abstract

In the early 1990s BCS was involved in some image processing and detection projects arising in both natural resource and defense settings. We were convinced that thin plate spline (TPS) functions could play an important role in developing effective software for image registration and change detection, but initially we encountered some computational challenges when using TPS functions on mid-size to large problems. Fortunately, we were able to enlist the assistance of Professor M. J. D. Powell of Cambridge University; the end result is that BCS has very efficient, numerically stable software for solving the underlying problem of TPS evaluation on large 2D grids (such as raster images). This note provides background to, and some mathematical details of, the algorithms involved.

Introduction

Given a set of distinct control points \( \{(x_i, y_i), i = 1, 2, \ldots, n\} \) that are not all collinear, with associated function values \( \{z_i\} \), it is often desirable to estimate corresponding function values at the node points of a large rectangular grid. While many approaches to this problem are possible (e.g., bilinear interpolation and triangulation techniques), here we seek the smoothest possible real-valued function \( F \) satisfying

\[
F(x_i, y_i) = z_i, \quad \text{for} \quad i = 1, 2, \ldots, n \tag{1}
\]

where our measure of smoothness is the integral

\[
I(F) = \iint_{\mathbb{R}^2} \left( \frac{\partial^2 F}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 F}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 F}{\partial y^2} \right)^2 \, dx \, dy. \tag{2}
\]

In [3] it is proved that the variational problem of minimizing (2) subject to the interpolation conditions (1) is solved uniquely by a thin plate spline (TPS) of the form

\[
F(x, y) = a_0 + a_1 x + a_2 y + \sum_{i=1}^n \lambda_i r_i^2 \log r_i \tag{3}
\]

where \( a_0, a_1 \) and \( a_2 \) are the coefficients of the planar term of the spline, \( \lambda_i \) is the coefficient for the \( i \)th spline term, and

\[
r_i^2 = (x - x_i)^2 + (y - y_i)^2.
\]
In addition, the integral $I$ in (2) is finite if and only if the $\lambda_i$’s have the property that

$$
\sum_{i=1}^{n} \lambda_i = 0, \quad \sum_{i=1}^{n} \lambda_i x_i = 0, \quad \text{and} \quad \sum_{i=1}^{n} \lambda_i y_i = 0. \quad (4)
$$

An alternative “engineering” motivation for the use of TPSs as smooth interpolation functions is provided in [4], where the small deflection equation for a thin metal plate of infinite extent is considered. If deflections $z_i$ are specified at each of the control points $(x_i, y_i)$, the plate will require point loads at each control point in order to achieve the desired deflections. Again, it turns out that the mathematical model for this problem is a TPS of the form (3) whose coefficients determine the point loads. In addition, if the plate is required to flatten out away from the point loads, this is achieved by imposing equilibrium conditions on the plate so that the sum of the loads and their moments is zero (i.e., the conditions (4) pertain).

Thus, a TPS has the desirable properties that it asymptotically approaches a global planar model outside the region of the control points, and the effects at each control point are predominantly local, i.e., largely confined to an immediate region around each $(x_i, y_i)$.

A particularly appropriate application for TPS gridding is in image warping ([1], [2]). Here, $n$ pairs of control points are defined that link corresponding features in a reference image and a new image (sometimes called a source or sensed image), and two TPS functions are generated for shifting the $x$- and the $y$-components, respectively, of all the pixels of the new image onto the reference image.

Previous attempts to apply TPSs to the generation of large grids with substantial numbers of control points have encountered two major difficulties: (1) the system of equations defining the coefficients can be very ill-conditioned, and (2) the evaluation of TPS functions is very time-consuming. In the BCS implementation these difficulties have been overcome (see below), allowing large rectangular grids to be generated stably and accurately using TPS functions with, for example, hundreds of control points.

TPSs were originally derived for the physical case where $x$ and $y$ are spatial variables and their units are the same. However, our approach can be applied to any $x$ and $y$ variables provided that the grid intervals in each dimension are viewed as being equivalent. The current implementation allows TPS grids to be generated for arbitrary values of $x$ and $y$, and forms the TPS corresponding to a grid with unit spacing. While this does not conform strictly to the physical model for which TPSs were originally defined, it does allow the desirable properties of TPSs to be used in gridding functions of any two variables.

Because of the numerical sensitivity of the calculations, in practice all computations are carried out in double precision arithmetic.
TPS Coefficient Generation

Let \( \{(x_i, y_i), \ i = 1, \ldots, n\} \) be a set of \( n \) control points (distinct and not all collinear), each with associated \( z_i \), for which a TPS function \( F \) satisfying the interpolation conditions (1) is sought. A unique TPS with \((n+3)\) coefficients exists in the form

\[
z_i = F(x_i, y_i) = a_0 + a_1 x_i + a_2 y_i + \sum_{j=1}^{n} \lambda_j r_{ij}^2 \log r_{ij},
\]

where \( r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 \), and the conditions (4) also apply.

The \( \lambda_j \)'s and \( \lambda_j \)'s in (5) can be computed from the matrix equation \( \mathbf{Hv} = \mathbf{b} \), where

\[
\mathbf{H} = \begin{bmatrix}
0 & r_{21}^2 \log r_{21} & \cdots & r_{n1}^2 \log r_{n1} & 1 & x_1 & y_1 \\
r_{12}^2 \log r_{12} & 0 & \cdots & r_{n2}^2 \log r_{n2} & 1 & x_2 & y_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
r_{1n}^2 \log r_{1n} & r_{2n}^2 \log r_{2n} & \cdots & 0 & 1 & x_n & y_n \\
1 & 1 & \cdots & 1 & 0 & 0 & 0 \\
x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\
y_1 & y_2 & \cdots & y_n & 0 & 0 & 0
\end{bmatrix},
\]

\[
\mathbf{v}^T = [\lambda_1, \lambda_2, \ldots, \lambda_n, a_0, a_1, a_2] \quad \text{and} \quad \mathbf{b}^T = [z_1, z_2, \ldots, z_n, 0, 0, 0].
\]

This system of equations can be extremely ill-conditioned, particularly for larger numbers of control points, and application of a numerically stable algorithm is essential when computing TPS coefficients. A review of available TPS algorithms is given in [6], where direct (i.e., finite) matrix methods are advocated for smaller values of \( n \), and iterative algorithms are recommended for larger problems (say, when \( n \) exceeds 1000). In the BCS implementation we assume that \( n \) does not exceed several hundred control points\(^1\), and the direct method used here is to construct an orthogonal matrix (which is a product of \((3n-6)\) Givens rotations) that reduces the interpolation problem to a square positive definite system of \((n-3)\) equations. The resulting system is solved to eventually provide stable computed values for all \((n+3)\) TPS coefficients.

---

\(^1\)We have occasionally used this coefficient generation software on much larger problems, but it is not computationally efficient on very large problems when compared to iterative methods.
Direct evaluation of a TPS via the functional form (3) can be prohibitively time-consuming on a large grid because each node point requires the evaluation of $n$ distinct terms involving the logarithm function. In image warping problems, the grids can easily be of size $2000 \times 2000$ points (pixels) or larger, and involve hundreds of control points. The fast TPS evaluation used in our implementation, which is based on a tabulation scheme given in [5], replaces the direct evaluation form (3) with an approximation strategy that yields estimates accurate to within a user-specified tolerance. The use of this approximation technique can result in evaluation times that are more than 1000 times faster than those for the direct TPS form (3), with minimal effects on the accuracy of the gridded values. Table 1 shows the relative improvement in evaluation times (for various grid sizes and numbers of control points) that resulted from a numerical comparison of the fast versus direct TPS evaluation procedures. Notice that the last line of this table provides an example where use of the fast evaluation procedure reduces the computer evaluation time by three orders of magnitude.

Table 1: Ratio of evaluation times for fast (approximate) and direct TPS gridding algorithms.

<table>
<thead>
<tr>
<th>Number of Control Points</th>
<th>Ratio of Evaluation Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid size 300x300</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>51</td>
</tr>
<tr>
<td>50</td>
<td>59</td>
</tr>
<tr>
<td>100</td>
<td>68</td>
</tr>
<tr>
<td>200</td>
<td>70</td>
</tr>
<tr>
<td>400</td>
<td>72</td>
</tr>
<tr>
<td>Grid size 1000x1000:</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>143</td>
</tr>
<tr>
<td>50</td>
<td>233</td>
</tr>
<tr>
<td>100</td>
<td>342</td>
</tr>
<tr>
<td>200</td>
<td>445</td>
</tr>
<tr>
<td>400</td>
<td>525</td>
</tr>
<tr>
<td>Grid size 2000x2000</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>156</td>
</tr>
<tr>
<td>50</td>
<td>319</td>
</tr>
<tr>
<td>100</td>
<td>525</td>
</tr>
<tr>
<td>200</td>
<td>910</td>
</tr>
<tr>
<td>400</td>
<td>1300</td>
</tr>
</tbody>
</table>
References


