Modified Radial Basis Functions Approximation Respecting Data Local Features

Jakub Vasta  
Department of Computer Science and Engineering  
Faculty of Applied Sciences, University of West Bohemia  
Plzen, Czech Republic  
vastja@students.zcu.cz

Michal Smolik  
Department of Computer Science and Engineering  
Faculty of Applied Sciences, University of West Bohemia  
Plzen, Czech Republic  
smolik@kiv.zcu.cz

Vaclav Skala  
Department of Computer Science and Engineering  
Faculty of Applied Sciences, University of West Bohemia  
Plzen, Czech Republic  
skala@kiv.zcu.cz

Martin Cervenka  
Department of Computer Science and Engineering  
Faculty of Applied Sciences, University of West Bohemia  
Plzen, Czech Republic  
cervemar@kiv.zcu.cz

Abstract—This paper presents new approaches for Radial basis function (RBF) approximation of 2D height data. The proposed approaches respect local properties of the input data, i.e. stationary points, inflection points, the curvature and other important features of the data. Positions of radial basis functions for RBF approximation are selected according to these features, as the placement of radial basis functions has significant impacts on the final approximation error. The proposed approaches were tested on several data sets. The tests proved significantly better approximation results than the standard RBF approximation with the random distribution of placements of radial basis functions.

Index Terms—Radial basis function, approximation, inflection points, stationary points, Canny edge detector, curvature

I. INTRODUCTION

The approximation is commonly used and well known technique in many computer science disciplines. This technique can be divided into two groups. The first one is the approximation that use the mesh and its connectivity. Some well known approaches that use the triangulation are [1]–[4]. However, all those approaches need the mesh connectivity, i.e. triangulation, which can be time consuming and difficult to compute for higher dimensions. On the opposite site, the second group are approximation techniques that does not require any mesh, i.e. they are called meshless methods. This paper focuses on this kind of approximation.

In the book [5] is provided an introduction for each of the most important and classic meshless methods along with the complete mathematical formulations. In total, it presents 19 meshless methods in detail with full mathematical formulations showing numerical properties such as convergence, consistency and stability. One example of approximation technique is Kriging [6]. It depends on expressing spatial variation of the property in terms of the variogram, and it minimizes the prediction errors which are themselves estimated. Extension and variations of this method are available in [7]–[9]. Another approximation technique is weighted least square method [10], [11] it is simple because it is based on the well-known standard least squares theory. It is attractive because it allows one to directly use the existing body of knowledge of the least squares theory and it is flexible because it can be used to a broad field of applications in the error-invariable models. Very similar approach is the LOWESS method [12] which is used for meshless smoothing and approximation of noisy data.

The Radial basis function (RBF) methods have been widely used for approximation of scattered data, recently. A brief introduction to this method is in [13]. Comparison of different radial basis functions is in [14], [15]. The paper [16] presents an approach for large scattered data interpolation. It uses the space subdivision to reduce the computation time and more importantly to reduce the needed memory for RBF approximation. A modification of this algorithm for 3D vector field data approximation is presented in [17]. Very important for the final RBF approximation quality is the distribution of radial basis functions. This problem is described and suggested solution in [18], [19]. Many papers also propose a solution to the selection of the best shape parameters of radial basis functions [20]–[24].

II. RADIAL BASIS FUNCTIONS

The Radial basis functions (RBF) are commonly used for n-dimensional scattered data approximation and interpolation. This approach is used in many areas, e.g. image reconstruction [25], [26], neural networks [27] and surface reconstruction [28], [29]. The task can be stated as follow. Find analytic function for given pairs of values \((x_i, h_i)\), where \(x_i\) is a point position in n-dimensional space and \(h_i\) is value in this point. For such data it is not possible to use standard approximation and interpolation techniques because lack of
knowledge about data connectivity and ordering. Therefore, the RBF approximation has the following attributes:

- Designed for scattered data approximation/interpolation
- Independent of the data dimension
- Not separable, i.e. it is not valid to approximate/interpolate data “dimension by dimension”
- Invariant with respect to euclidean transformations

Hardy [30] proposed RBF interpolation based on interpolation equation:

\[ f(\vec{x}) = \sum_{i=1}^{M} \lambda_i \theta(||\vec{x} - \vec{x}_i||), \]  

(1)

where \( \vec{x}_i \) is data point and \( \lambda_i \) is point weight. \( \theta(r_{ij}) \) is radial basis function, where \( r_{ij} = ||\vec{x} - \vec{x}_i|| \). Radial basis function can differ, but they can be divided into two main groups by theirs range of influence, i.e. global and local functions.

RBF approximation/interpolation leads to the linear equation system \( A\vec{x} = \vec{b} \), where approximation differ from interpolation only with form of matrix \( A \). Solvability and stability problems were solved for example in [31] [32]. Wright [32] extend original RBF interpolation with polynomial and added more conditions.

A. Radial Basis Functions approximation

RBF approximation is based on point distance in n-dimensional space and is derived from the same equation (2) as interpolation is.

\[ f(\vec{x}) = \sum_{i=1}^{M} \lambda_i \theta(||\vec{x} - \vec{e}_i||), \]  

(2)

where \( M \) is number of radial basis functions, \( \lambda_i \) is weight of radial basis function, \( \theta \) is radial basis function and \( \vec{e}_i \) is placement of radial basis function.

Given set of value pairs \( \{\vec{x}_i, h_i\}_N \), where \( \vec{x}_i \) is point position in n-dimensional space, \( h_i \) is value in this point, \( N \) is number of given points. \( N \ll M \) therefore we obtain over-determined system of linear equations.

\[ h_i = f(\vec{x}_i) = \sum_{i=1}^{M} \lambda_i \theta(||\vec{x}_i - \vec{e}_j||) \]  

(3)

It can be rewritten in matrix form

\[ A\vec{\lambda} = \vec{h} \]  

(4)

This over-determined system of linear equations can be solved by LSE or QR decomposition.

III. PROPOSED APPROACH

Radial basis function placement is important factor for approximation error. In this contribution, property of these good points are proposed with the way to find them. First group are extreme points i.e. local/global minimum or maximum. Next group are points of inflection. These points represents changes in data. Another proposed group of points are stationary points of curvature. These points represents extreme curvature values and in some case they are similar to points of inflection. Last group are edge points as known from image processing, because we can look at data as on image depending on data structure or sampling. Search for important points is amended with Halton sequence [13] sampling with special stress on border and corner sampling for covering whole data set. The last step is reduction of points number with nearest neighbour condition.

The Halton sequence is computed using the following formula:

\[ Halton(p)_k = \sum_{i=0}^{\lfloor \log_p k \rfloor} \frac{1}{p^{i+1}} \left( \left\lfloor \frac{k}{p^i} \right\rfloor \mod p \right), \]  

(5)

where \( p \) is a prime number, \( k \) is the order of the element of the Halton sequence, i.e. \( k \in \{1, \ldots, n\} \). For generation of random points with Halton distribution in higher dimension, different values of \( p \) are used for each dimension.

A. RBF Approximation with Stationary Points

It is known that stationary points are such points that hold equation

\[ \frac{\delta f}{\delta x} = 0 \land \frac{\delta f}{\delta y} = 0 \]  

(6)

This condition is not enough to determine whether the point is global/local extreme or just a saddle point. It is necessary to examine Hessian to determine point property. On the other hand it can be seen that saddle points are as important as points of extreme.

Evaluation of partial derivatives and comparing with zero is not optimal way to find stationary points. Better way is to compare given point with its surrounding i.e. masks for minimum, maximum and saddle point can be created.

B. RBF Approximation with Inflection Points

Points of inflection are such points where surface change from convex to concave or the other way round. For points of inflection in continuous space hold that Gauss curvature is equal to zero

\[ K_{gauss} = \frac{\frac{\delta^2 f}{\delta x^2} \frac{\delta^2 f}{\delta y^2} - \left( \frac{\delta^2 f}{\delta x \delta y} \right)^2}{\left( \frac{\delta^2 f}{\delta x^2} + \frac{\delta^2 f}{\delta y^2} \right)^2} \]  

(7)

It can be seen, from equations above, that Gauss curvature is zero only when numerator is equal zero, i.e. Hessian matrix determinant is equal to zero.
It is not necessary to compute exact curvature value to find point of inflection. We need to find just points where curvature sign change from negative to positive or vice versa. It can be seen that sign of Gauss curvature only depends on numerator sign because denominator is always positive.

C. RBF Approximation with Stationary Points of Curvature

From (7) we can compute curvature of given surface in every data point. Then it is possible to find stationary points in curvature similar to process described in the section III-A.

D. RBF Approximation with Edge Detection

In the simplified scenario, where the points are sampled in the grid pattern, we can look at the data as image i.e. value \( h \) on position \( [x, y] \) can be considered to be brightness intensity \( I \) in pixel \( (i, j) \). In case of scatter data there is need to use some special data structure to obtain points adjacency information e.g. kd-tree or adjust used algorithms.

We proposed to detect edges in image i.e. transitions between low and high values. Another suggested approach is to compute data gradient magnitude in each data point and then run edge detector over such field of gradient magnitudes. This approach will find transitions between low and high gradient magnitude areas. To detect edges we can use existing detectors from image processing e.g. Canny, Sobel, Prewitt etc.

IV. EXPERIMENTAL RESULTS

Proposed methods were tested on several test functions which were designed to represent special data set behaviour. Sampling step is 0.01 and functions are normalized to interval \((x, y, z) \in [-1, 1] \times [-1, 1] \times [0, 1]\) for comparison purposes. For all functions Gauss radial basis function \( \phi(r) = e^{-\epsilon r^2} \) was used.

For comparison was used square mean error per point as can be seen in Fig. 6, Fig. 7 and Fig. 8. In the 1st test function \((8)\) random distribution of placement with Halton sequence provide good results in comparison with other methods, see Fig. 3. This is caused by function shape which fill whole space.

In the 2nd test function \((8)\) can be seen improvement when proposed methods are used because of its special behaviour only in some areas of its domain, see Fig. 4.

The last test function \((8)\) is design to test RBF approximation in general and even with our improvements lot of methods fails, see Fig. 5. What is even more it was found that with proper placement it has no effect on precision to add more points from Halton sequence.

\[
\begin{align*}
 f_1(x, y) &= \frac{2}{\Pi} \left( \sin (4x^2 + 4y^2) - x + y - \frac{5}{2} \right) \\
 f_2(x, y) &= \frac{3}{4} e^{-\frac{1}{4}((9x-2)^2 + (9y-2)^2)} \\
 &\quad + \frac{3}{4} e^{-\frac{1}{49}(9x+1)^2 - \frac{1}{49}(9y+1)^2} \\
 &\quad + \frac{1}{2} e^{-\frac{1}{4}(9x-7)^2 + (9y-3)^2} - \frac{1}{5} e^{-(9x-4)^2 - (9y-7)^2} \\
 f_3(x, y) &= \frac{1}{9} \tanh(9y - 9x) + 1
\end{align*}
\]

(8)
V. CONCLUSION

The proposed methods were tested on several standard testing functions, however, only some representative functions are mentioned in this contribution. The above presented methods proved very good results in precision of approximation, even though some special types of functions, e.g. fast changes, are problematic for all approaches. The experiments also proved validity of the proposed methods for the Radial basis function approximation of scattered data, with regard to a low approximation error with high points reduction leading to a high compression ratio.

ACKNOWLEDGMENT

The authors would like to thank their colleagues at the University of West Bohemia, Plzen, especially to colleague Zuzana Majdisova, for their discussions and suggestions. This research was supported by the projects: Czech Science Foundation (GACR) No. GA17-05534S and partially by SGS 2019-016.
Fig. 6: Mean square error on 1st test function (8)

Fig. 7: Mean square error on 2nd test function (8)

Fig. 8: Mean square error on 3rd test function (8)

REFERENCES


