Gordon Surface Modeling in a Network Environment

Jan Smid
SKS Enterprises
Finksburg, MD 21048
Email: jsmid@sks99.com

Petr Volf
Institute of Information Theory and Automation
Prague, Czech Republic
Email: volf@utia.cas.cz

***Abstract*—With the advent of the Internet, we have entered the age of ubiquitous systems that will need to be able to communicate, establish criteria for measurement and then perform these measurements. These future systems, built for engineering and personal applications, will have components that have their own identities. Such dialog systems will be able to communicate among themselves and with the user.This means that individual components or nodes must be based on much richer semantics than is currently available in contemporary systems. Engineering devices of the future will have to perform traditional tasks such as monitoring and control in mutual collaborationn with users. In this paper we present a framework and some specifications for one of the spectrum of possible tasks, namely temperature modeling using fixed and mobile network of sensors.**

The goal of the presented temperature modeling procedure is to investigate and describe the process of these measurements. The mathematical model extends from one dimensional generalized linear networks to a Gordon surface approximant. The Gordon surface interpolates a network of curves [2]. The network of curves can even be degenerated (one side collapsed into a point) [7]. The role of the entire dialog system is to provide interactive approach to solving a variety or ordinary task and to evolve the system by absorbing declarative and procedural knowledge in a way that does not rely entirely on user's interpretation (semantics).

I. INTRODUCTION

The world of engineering and personal experience have become increasingly distributed and interactive. The need for being able to talk to our devices and network nodes becomes almost obvious. In the past, the focus has been on deepening individual methods. What the network approach brings is collaboration and dialogs. This requires new methods and views that support this effort. This view is not considered mainstream but it exists, see e.g. [10]. Most simply, communicating networks of a variety of nodes are ultimately stronger then one super strong node. The biological world is our paradigm in many ways.

II. PROBLEM DEFINITION

Temperature sources of unknown locations in an area create a temperature field. The system has a set of fixed sensors and a set of mobile sensors. The task is to develop a model and a measurement plan. The number of measurements is unknown ahead of time, the measurement procedure does not stop until the model is satisfactory. First, we will consider a one dimensional model and then a class of two dimensional models, specifically the Gordon surface. There are a number of regression procedures for regression type of analysis that can be used as examples to show a conversion from standard to network conscious approach. We experiment with models that we worked with in the past [7], [9], [8].

III. DIALOG ALGORITHM

There are three types of algorithms. Two of them are standard and distributed. The new type is the algorithm in a network and dialog environment. Standard fitting algorithm consists of finding a model when a data set is known. Distributed algorithms work with distributed resources. Dialog algorithms work with negotiating entities, such as agents, nodes. In the context of temperature modeling, the main task component is to make decisions whether to obtain more measurements. The algorithm has two main steps:

1) Compute the model for a given measurement set
2) Evaluate the model for regions where no data is provided, if satisfactory goto 5, else goto 3
3) measure to add data
4) goto 1.
5) done

IV. ONE DIMENSIONAL MODEL AND STOPPING CRITERION

The data is a set of pairs x^n, t^n. One dimensional model of a function that fits a given set of data can be written as follows:

$$y(x; w) = \sum_{j=0}^{M} w_j \phi_j(x)$$

The goal is to find the weights w_j, and uncertainties such that

$$y(x^n; w) = t^n$$

in the least square sense. Under simplifying assumptions the distribution of the predicted outputs of the model can be calculated explicitely, which is convenient for mathematical analysis and for gaining insight. Bishop providees the formula for Gaussian noise and one dimensional case. The variance of posterior distribution of predicted values y is equal to

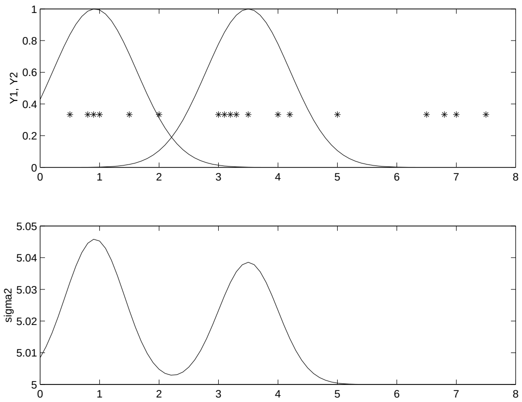

Fig. 1. Variance Two Basis Functions

$$\sigma^2(x) = 1/\beta + \phi(x)^T A^{-1}\phi(x)$$

$$A = \beta \sum_n \phi(x^n)\phi(x^n)^T + \alpha I$$

α, β are hyperparameters (pp.390-394 [1]), I is the identity matrix.

The variance σ is a function of the Hessian matrix, and the behavior of basis functions. The variance varies with the measurement points distribution. The fact that we can control the width of the variance can be utilize for the stopping criterion definition as well.

The stopping criterium: The width of the variance band can be made smaller for a sufficient number of measurement points and basis functions of the generalized linear model. When the variance band becomes smaller then a heuristically estimated threshold value the model is done by definition. This concept is based on the matrix norm which tends to go to zero for a given set of basis function and s sufficient large set of measurements. This is a one dimensional concept that can be extended to two dimension using the Gordon surface interpolant. The main idea of this generalization is to view variance as a Gordon surface constructed from one dimensional network of curves that are calculated using Bishop like formulas.

Figure 1 shows variance $\sigma^2(x)$ distribution at different x, for a system of two basis functions. Figure 2 and 3 show the effect of measurement points on the magnitude of variance σ.

V. GORDON SURFACE

The basic Gordon surface [2],[7] is a mapping

$$S(s,t) = G + F + corrections$$

$$G = \sum_{j=1}^{m+2} G_j(s)\Psi_j(t)$$

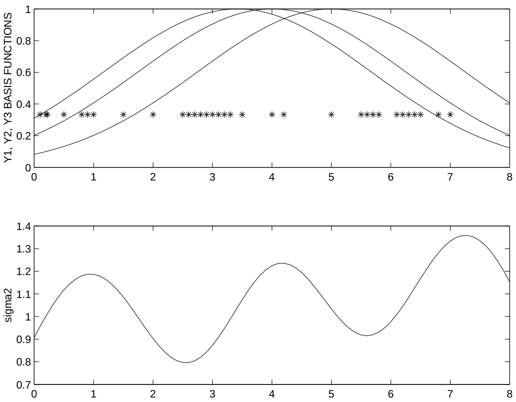

Fig. 2. Low variance, added data

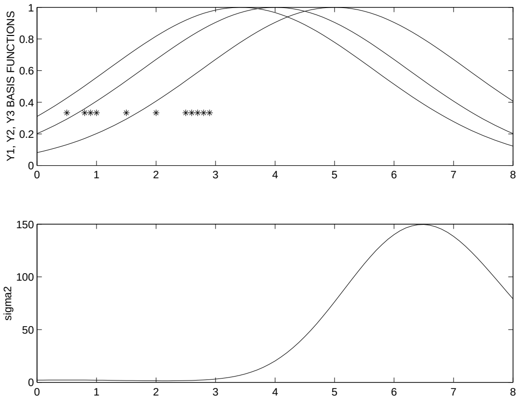

Fig. 3. No data, large variance

$$F = \sum_{i=1}^{n+2} F_i(t)\Phi_i(s)$$

$$corrections = -\sum_{i=1}^{n+2}\sum_{j=1}^{m+2} S_{ij}\Phi_i(s)\Psi_j(t).$$

from a rectangle $[s_1, s_n] \times [t_1, t_m]$ in the (s,t) plane to R^3. We can assume that $s_1 = 0$, $s_n = 1$, $t_1 = 0$, $t_n = 1$.

We are given a network of intersecting curves G_j and F_i in R^3, homeomorphic to a planar rectangular grid. Each curve is parametrized on the interval $[0,1]$. Here, $\{F_i\}_{i=1}^{n}$ and $\{G_j\}_{j=1}^{m}$ are given 3-dimensional vector functions defining curves in R^3, while $\{F_i\}_{i=n+1}^{n+2}$ and $\{G_j\}_{j=m+1}^{m+2}$ are cross-boundary derivatives. The fixed vectors S_{ij} represent either the curve intersections or the derivatives at the boundary grid points. Scalar-valued blending functions are denoted by

$$\Psi_j(t), \Phi_i(s), \quad (i = 1, 2, \cdots, n+2;\ j = 1, 2, \cdots, m+2).$$

For a degenerate case where one boundary curve is reduced to a single point see appendix.

Individual meshlines can be created by applying the one dimensional fitting process. In the two step procedure, the first step consists of projecting the data points on a few planes cut

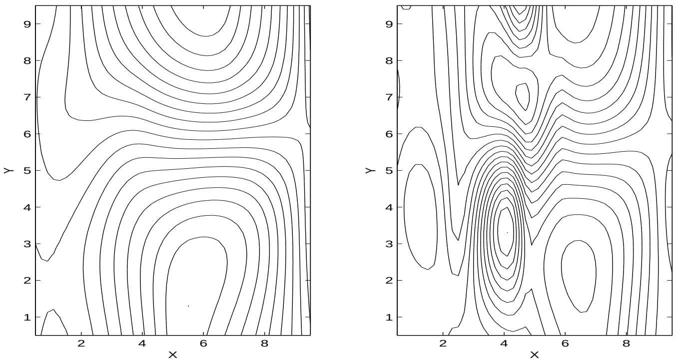

Fig. 4. Gordon surface, Left: uniform points, Right: missing points

through the area of interest. The second step then creates a curve by connecting points that belong to the same group of points. One way to create a group of points is to collect points close to a plane cut through the area.

The Gordon surface techniques can be viewed as a method of multi-dimensional smoothing. When data are projected to the planes, simultaneously we can assess the weights of data points relative to these planes, e.g. with the help of a kernel function, as in the kernel smoothing approach. Then, along the planes, the optimal curves are fitted with the aid of weighted least squares. Simultaneously, these kernel function are also a convenient choice for the blending. Example of those kernel functions are the triangular kernel and the Gaussian kernel.

VI. VARIANCE FOR GORDON SURFACE

The exact form of variance band (or confidence band for actual surface, or prediction band for new data) is difficult to obtain in two-dimensional case. However the blending Gordon technique can be used to obtain estimates. The substantial advantage of this approach is that all calculations are reduced to one dimension.

For each curve of the network for which we wish to interpolate a smooth surface, we can compute a variance function and thus we obtain second network of curves. We can measure uniformity of this surface and if needed modify it by obtaining more measurements or increasing the number basic functions of the model.

The other approach is based on representing the variability of possible surfaces by a sample of surfaces obtained by repeated fitting the Gordon surface to data with randomized selection of cut planes and also of fitted 1-dimensional curves along them (i.e. the centers of Radial Basis functions or nodes of splines are randomized). This technique is actually the Metropolis algorithm that produces the representation of Bayes posterior distribution of possible surfaces. This distribution can be measured either by sample variances or by the inter-quartile range. The disavantage of this method is high computational time, even in one-dimensional case. Therefore the Gordon method is clearly superior.

Figure 4 shows contour graphs of two Gordon surfaces. One surface was generated using a uniform grid of measured points and the other was generated where the sub-region $x \in (3,5) \times y \in (3,5)$ was empty. One-dimensional curves along cut planes were approximated by cubic splines. Gaussian kernels were used both for weighting and blending. Uneven set of measurement point generate a Gordon surface with more local minima and maxima, as was expected.

VII. DATA COMMUNICATION WITH ROBOTS

Special protocols are used to control electronic toys, industrial and scientific devices. In situations where no configuration is possible or covenient, the internet protocols (TCP/IP, HTTP) must take over. In one setup of our experiments, a robot receives wirelessly strings that are interpreted as commands to robots servos. The base sending signals is a miniserver that listens to a user or an agent. No further configuration is reguired to operate the robot.

1) The easiest communication of data is via the HTTP protocol whenever possible. The HTTP protocol based commnunication sends parameters using the GET or POST methods. These parameters can start programs or retrieve data from registers of microprocessors.
2) Socket based TCP/IP communication, is more complicated to implement compare to the HTTP communication, but it is needed to run permanent processes and different threads. Our language of choice for running multiple processes is Java.
3) An example of network processes can be illustrated using image analysis for robot navigation (please see section IX).

VIII. NETWORK TOPOLOGY EVALUATION

When we deal with a large number of deployed sensors one aspect which might be often of interest is to know whether the network coverage is adequate or not.

Computational homology [5] has been proposed as a tool for analyzing sensor data and images collected from recording devices. The Proposed methods are based on the methods of algebraic topology. In case of sensor networks, good coverage by sensors can be defined as the coverage where every subregion is observed by three nodes or sensors that can see or communicate with each other.

The network can be symbolically represented by a graph, where nodes nodes represent sensor locations and the edges indicate wheather or not there is a communication of a sort (e.g. visual or radio communication).

A good coverage network then can be represented as a triangulated area. Incompletely covered area is such where some triangles are missing. To detect holes or missing triangles a simpler tool might be used at times instead of more powerful algebraic topology tools. To convey the idea we will use the expand on a particular figure. This approach is a well established geometric tradition. Imagine a triangulated triangle

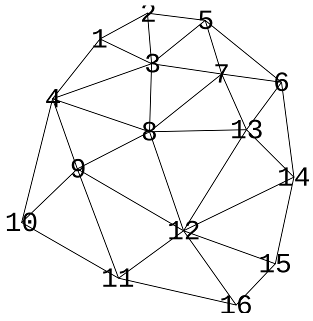

Fig. 5. Sensor setwork nodes

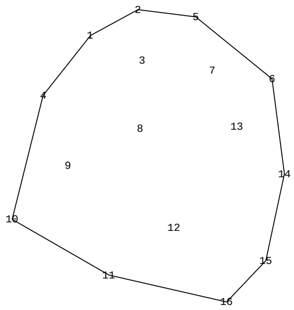

Fig. 6. Boundary edges, no Holes

[4], (see appendix for definition). The famous Euler theorem relates the number of faces F, the number of vertices V, and the number of edges E for a sphere (convex polyhedron)

$$F + V - E = 2$$

or more generaly

$$F + V - E = 2 - 2k$$

for convex polyhedron with k holes. The Euler theorem cannot be applied directly as the simplest criterion unless we know the number of faces. However we can still considerably simplify the general approch by computational homology.

Boundary edges will produce another graph and we can find out algoritmically wheather an edge belongs to the boundary or not. The boundary edges belong to only one triangle, and this feature can be checked easily. If this graph is disconnected or if this graph has more then one cycle then we know there are holes in the network coverage. This method is much faster then straightforward approach using computational homology.

IX. VISUAL NAVIGATION AND DOCUMENTATION

Locations of robots are clearly critical for correct implementation of the measurement process. IPcam are mounted on the robot and at differend nodes in the area of measurement. At times IPcams might be the only navigational tools available.

Arguably the simplest method of identifying an object for the purpose of navigation is to use the concept of bounding box. This method is implemented in CMUcam in the micro-controller environment [3]. A system of distributed IPcams and temperature sensors provide data that are processed by a system of webprocessors (servers). We use the standard techniques for object tracking to find a robot marked by a color button. [3]. More sophisticated approach involves standard background modeling algorithm that is described for example in [11].

Change detection for numerical data (such as temperature distribution can processed by another class of techniques. For concrete examples see for example analysis in [8], [9]. Regardless of technique is used it can be treated as a dialog process where one agent or more agents can provide results provided that input data set is known.

Regardless whether the sophisticated [6] or simplest techniques (such as the CMUcam color tracking algorithm) are used for specialized tasks, the dialog domain encapsulates and hides those differences. The user or an agent asks the same (or almost the same) question, command or a request, for example: Is there a change in behavior of a temperature distribution? Or perhaps a question: Can you show me typical distribution for February and March last year? The ability to do that does not come for free and a dialog enironment must be added.

1) The robot navigation camera takes an image that is analyzed. Location of a mark allows to calculate the robot position.
2) The IPcam takes a jpeg of a robot. The difference image is the location of a robot. The system finds the best positioned robot and then sends this robot to a desired location using elementary commands, whatever they are, by sending a sequence of characters to a base webprocessor that passes this string to the robot. This is the simplest scenarion with minimal requirements on the robot's abilities. However the robot can be very autonomous when so equipped.

X. LANGUAGE COMMUNICATION OF NODES AND USERS

Communication in networks does occur at different levels. The top communication level resembles a dialog using language. With increasing heterogeneity of individual network nodes no fixed protocol will do. We initialize the system with dialogs that take on a form of commands (c), requests (r) and declarative information (di).

The goal is to develop a dialog system based on a distributed system of databases and webprocessors. Databases provide a rich content of information. Some content is declarative and logical inferencing is appropriate. Some content is procedural (programs).

The system will have manipulation functions. An example of a manipulation function is to combine objects with attributes into a new object. A new object can be a graph or a constructor function. An example is a function that creates a new table either permanent or temporary, such as a spectrum of two series.

The system will be able to list atrributes of objects, e.g. what quantities and concepts are available: graphs, monotone segments, averages, an so on.

The simplest dialog scenarion will be based on restricted menus and keywords.

The dialog will allow the user and agents to input declarative facts. The system will have inferencing capabilities such as the resolution for ambiguous semantic pairs (utterance, semantic expression)

Resolved semantic mapping determines the group of applicable programs.

1) semantic mapping definition:
   Given a set of pairs (utterance, semantic expression) find a function f() or a class of functions from a set of words U to a set of semantic expressions M, so that each utterance is mapped to a corresponding expression.
2) Semantics generation:
   For a given utterance find a group of semantic expressions that are compatible with the utteration and an existing context.
3) Action:
   Based on a semantic group provide the required action.
4) Basic actions are: command, declarative query, information.

Our experimental implementation environment consists of PHP, MySQL, and Java.
An example of robot navigation and control

1) Someone needs to make a measurement at a position A=(100, 20).(c)
2) Go to the back left corner of 213f.(c)
3) Take a picture of IP=192.168.1.22 robot (in order to update a position of that robot).(c)
4) There is a new robot IP= 192.168.121. (di)

An example of IPcamera navigation and control

1) Turn the camera to the right .(c)
2) Find in database B a picture from yesterday. (c)
3) Take a picture of IP and find out robot IP position. (c)
4) The IPcam IP=192.168.1.117 recognizes only commands ”move left”, ”move right”. (di)

These commands, and other forms of language, are backed by programs and functions (semantics). Structure of semantics can vary from node to node, or in other words the same utterance can point to different groups of programs.

XI. CONCLUSION

Gordon surface algorithm is one example of a general approach to distributed dialog based systems. The presented example illustrates a broad class of algorithms designed for standalone applications, in which the user is the interpreter of the environment. This environment can then can be adapted into a networked environment in which individual nodes will grow their identities for the benefit of the entire network.

XII. APPENDIX

1) A triangulation of a triangle D (or any other polygon) is a division of D into a finite number of triangles so that each boundary edge of D is the edge of just one triangle of the subdivision and each edge in the interior of D is an edge of exactly two triangles of the subdivision. A triangulation is a type of complex.
2) The paper by W.J.Gordon [2] provides a solution of passing a ’smooth’ free-form surface through a network of intersecting curves, with derivatives prescribed across the outer boundaries. Occasionally, the network of curves is irregular and one of the boundary curves reduces to a single point. We will call the surface that evolves from such boundary pathology the degenerate Gordon surface. The degenerate Gordon surface can have, but typically does not have, unidirectional normals at the degenerate side. In this case, the normals must be multidirectional if the ”transverse” meshlines that terminate in the single-point boundary do not have coplanar tangents where they meet. On the other hand, if the tangents are coplanar, then by suitably interpolating a tangent distribution between the concurrent meshlines, one can create a degenerate Gordon surface of desired degree of differentiability throughout [7]. This construction is important because it provides an algorithm for creating smooth models for irregular regions.

ACKNOWLEDGMENT

Jan Smid would like to thank the Department of Computer Science, Bowie State University, Bowie, MD for support of this project.

REFERENCES

[1] Ch. Bishop: Neural Networks for Pattern Recognition. Cambridge University Press.

[2] W.J.Gordon: Spline-Blended Surface Interpolation Through Curve Networks. Journal of Mathematics and Mechanics. Vol. 18, No. 10, (1969).

[3] A. Rowe, Ch. Rosenberg, I. Nourbakhsh: A Low Cost Embedded Color Vision System. http://www.cs.cmu.edu/ cmucam/.

[4] M. Henle: Combinatirial Introduction to Topology. Dover Publications 1994.

[5] T. Kaczynski, K. Mischaikow, M.Mrozek, 2004.*Computational Homology*, Springer, Applied Mathematical Sciences, Vol.157.

[6] T. Serre, L.Wolf, S. Bileschi, M. Riesenhuber and T. Poggio: Robust Objesct Recognition with Cortex-like Mechanisms. IEEE Transaction On Pattern Analysis And Machine Intelligence. August 2006.

[7] J. Smid: Triangular Gordon Surface with Unique Normal. TR EDS-GM. 1991.SIAM Conference on Geometric Modeling, Tampe, AZ, November 1991.

[8] J.Smid, P.Volf, B. Markham, J. Seiferth. Modeling and Change Point Detection for Satellite Instrumentation Calibration Coefficient Time Series: An Example Based on Landsat-5 Thematic Mapper Data. The SPIE paper 3117-28, Earth Observing Systems II, San Diego, 1997.

[9] J.Smid, P.Volf. Dynamics Approximation and Diagnosis from a Neural Network Model. Mathematics of Neural Networks, Eds. S.Ellacott, J.Mason. I.Anderson, Kluwer Academic Publishers,1996.

[10] T. Toffoli: A Knowledge Home. Personal knowledge structuring in a computer world. January 2002. http:/kh.bu.edu/.

[11] K.Toyama, J.Krumm, B.Brumitt, B. Meyers: Wallflower:Principles and Practice of Background Maintenance. IEEE 2001.