Variational Methods in Signal and Image Analysis

Vadim Mottl, Alexander Blinov,
Andrey Kopylov, Alexey Kostin
Tula State University
Lenin Ave. 92, 300600 Tula, Russia
mottl@atm.tsu.tula.ru

Ilya Muchnik
DIMACS, Center for Discrete Mathematics and
Theoretical Computer Science, Rutgers University
P.O. Box 1179, Piscataway, NJ 08855, USA
muchnik@dimacs.rutgers.edu

Abstract

A class of signal and image processing problems is considered from the standpoint of treating them as those of coordinating the local data-dependent information and model-based smoothness constraints. Such a generalized problem is set as the formal problem of minimization of a separable objective function defined on an appropriate neighborhood graph of data array elements.

1. Introduction

It was even Leonard Euler who marked that everything what happens in the world bears the sense of some minimum or maximum. Concerning data analysis, the verity of this observation, which procured the name of the variational principle, manifests itself in the fact that all the kinds of decision making rest on the intent to minimize an appropriate measure of discrepancy $J(X|Y)$ between the observed data set $Y \in \mathsf{Y}$ and result of its processing within the permissible range $X \in \mathsf{X}$. The variational principle turns a practical problem into the formal one of minimizing the respective objective function and shifts off the burden of heuristic constructing of the data analysis algorithm into the tilled field of the optimization theory.

We will consider a signal or image as a function y_t of one $t \in T$ or two variables $t = (t_1, t_2) \in T$ that takes values from an appropriate set $y \in \mathcal{Y}$ and is defined on the discrete axis of the variable forming the signal $T = \{t\} = \{1, \ldots, N\}$, usually, time, or on the image pixel grid $T = \{(t_1, t_2) \colon t_1 = 1, \ldots, N_1, \, t_2 = 1, \ldots, N_2\}$.

Both in signals and images, each element of the data array $t \in T$ is naturally associated with a number of its neighbors. It is convenient to express the neighborhood relation in the form of an undirected graph G which is understood as set of pairs of neighboring elements $G \subset T \times T$. The evident neighborhood graph for signals is chain (Fig.1,a), and the simplest neighborhood relation usually considered to be appropriate for images is rectangular lattice (Fig.1,b), but we will decompose it into a succession of trees (Fig.1,c).

In this work, we restrict our concern to only those of signal and image processing problems which can be represented as problems of transforming the original image or signal $Y = (y_t, t \in T)$, $y \in \mathcal{Y}$, into another function $X = (x_t, t \in T)$, that would be defined on the same argument set $t \in T$ and take values $x \in \mathcal{X}$ from a set specific for each particular problem.

To this class belong, in particular, the problems of smoothing, segmentation, time-frequency analysis of signals, texture analysis of images, matching of signals or images, directional analysis of texture images [2]. For these and kindred problems we propose here a unified variational approach essentially based on the fact that the objective functions intrinsic for ordered data sets are separable

$$J(X|Y) = \sum_{t \in T} \psi_t(x_t|Y) + \sum_{(t',t'') \in G} \gamma_{t',t''}(x_{t'}, x_{t''}), \quad (1)$$

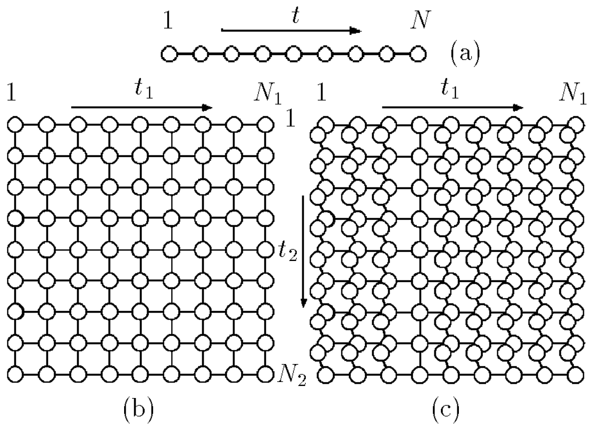

Figure 1. Neighborhood graphs on the set of data array elements: chain (a), rectangular lattice (b), and a tree (c) obtained from the lattice by splitting the nodes beyond the stem without edges in the resulting pairs.

Graph-Based Representations in Pattern Recognition (J.-M. Jolion and W.G. Kropatsch, ed.). Computing, Supplement 12. Springer-Verlag/Wien, 1998, pp. 135-145.

i.e. are sums of elementary functions of only one $\psi_t(x_t|Y)$ or two variables $\gamma_{t',t''}(x_{t'}, x_{t''})$ associated, respectively, with nodes $t \in T$ and edges $(t', t'') \in G \subset T \times T$ of the neighborhood graph.

It is assumed that the data-dependent node functions carry elementary units of information mined immediately from a vicinity of the current point t in the original signal or image, whereas the model-based edge functions are meant to express the a priori smoothness constraints of that or other kind. The processing of the data array boils down to global minimization of the separable objective function (1) as a way of finding a compromise between the mutually contradictory intents to comply with both of these two kinds of information. The minimum point $\hat{X}(Y)$ is considered as the result of processing.

However, the objective function's separability itself does not contribute essentially to an improvement of the computational solvability of the global optimization problem [1], until the variable adjacency graph is assumed to have no cycles, i.e. to be a tree [2]. It is just the latter assumption which leads to quick resulting algorithms and is systematically exploited here.

2. The basic optimization procedure

Let the tree G on the set of nodes $T = \{t\}$ define the variable adjacency in a separable objective function, and an arbitrary node t^* is assigned the root of the tree (Fig. 2). Then the node set T gets partitioned into a number of hierarchy levels T^j. Each variable x_t and the subtree of its descendants G_t will naturally produce a partial objective function $J_t(X_t)$ of absolutely the same structure as the full objective function (1) with the only difference that the sum is taken over sets $t \in T_t$ and $(t', t'') \in G_t$. It is obvious that the partial functions at the nodes (variables) of the lowest level, as well as at all the leaves, coincide with the node functions.

If we fix the value of an arbitrary variable x_t and mentally find the global minimum of the partial function over all the descendants of this variable, we will obtain function of one variable $\tilde{J}_t(x_t) = \min J_t(X_t)$ with respect to all x_s, $s \neq t$. It is clear that $\tilde{J}_t(x_t) = \psi_t(x_t)$ for the leaves, and at the root $\tilde{J}_{t^*}(x_{t^*}) = \min J(X)$ over x_s, $s \neq t^*$. As in the classical dynamic programming procedure, where the variables of the separable objective function are sequentially adjacent to each other, we will call such functions the Bellman functions. The global optimization procedure runs twice through the nodes of the tree. On the upward pass, it recalculates and stores the Bellman functions as result of minimization, at each node, of an untervening function of only one variable, and on the downward pass it finds the optimal values of all the variables.

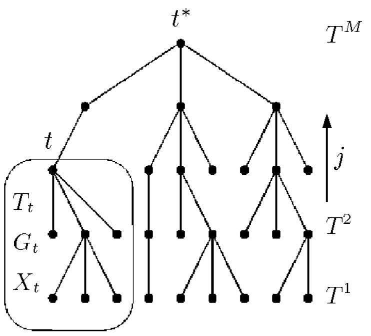

Figure 2. An arbitrary tree.

A numerical realization of the basic procedure for a separable function supported by a tree, just like the classical dynamic programming procedure for chains, is possible only in case there exists a finitely parametrized function family $\tilde{J}(x; \mathbf{a})$ concordant with node functions $\psi_t(x)$ and edge functions $\gamma_{t',t''}(x', x'')$ in the sense that the Bellman functions belong to this family at each step. In this case, the upward pass of the procedure consists in a recurrent re-evaluation of parameters $\tilde{\mathbf{a}}_t$ that completely represent the Bellman functions.

The evident case of a trivial parametrization is the case of a discrete range of node variables $\mathcal{X} = \{1, \ldots, m\}$. In this case, Bellman functions are m-dimensional vectors, and the partial optimization problems can be solved by elementary enumeration.

Another extreme case is the assumption that the range of node variables is n-dimensional real space $\mathbf{x} \in \mathcal{X} = \mathsf{R}^n$, and both node and edge functions are quadratic. It can be shown that the Bellman functions will be likewise quadratic and, so, completely representable by their minimum values, minimum points, and positive definite Hessian matrices.

The quadratic version of the dynamic programming algorithm for chains is a slight extension of the famous Kalman-Bucy filter and interpolator. The extension is more essential in the case of tree-like adjacency of variables in the separable objective function.

3. The generalized signal processing technique

Since the neighborhood relation between elements of a signal is completely expressed by a chain-like graph (Fig. 1,a), the natural instrument of solving a generalized signal processing problem is the classical dynamic programming procedure. To create the algorithm for a particular application, it remains to adequately choose the node and edge functions and to parametrize the Bellman functions in accordance with this choice.

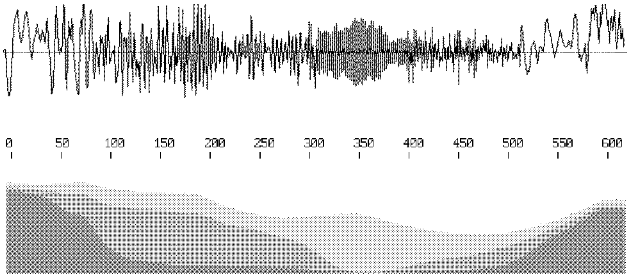

Figure 3. Spectral representation of the estimated autoregression model of a nonstationary signal.

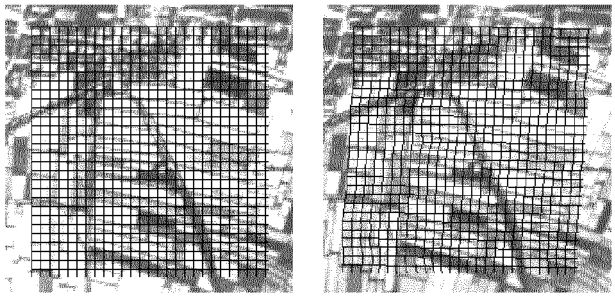

Figure 4. Disparity map of a stereo pair of satellite photographs.

As example of a signal processing algorithm built after such a scheme, we consider here the algorithm of constrained time-frequency analysis of nonstationary signals with calculation of instantaneous spectra from previously estimated parameters $\mathbf{x}_t = (x_{1,t}, \ldots, x_{n,t})^T \in \mathsf{R}^n$ of the nonstationary autoregression model $y_t \cong \sum_{i=1}^{n} x_{i,t} y_{t-i}$. For estimating autoregression coefficients, we used quadratic node and edge functions. All the Bellman functions remain quadratic, and the algorithm retains basic hereditary traits of the Kalman-Bucy filtration-interpolation procedure.

Fig. 3 shows the result of constrained time frequency analysis of a nonstationary signal. The diagram under the signal represents the instantaneous partial variances in the low (dark-gray), medium (gray) and high (light-gray) frequency bands as the summands of the full current variance which was estimated along with the autoregression coefficients.

4. Tree approximation of the image pixel lattice and the generalized image processing technique

It is clear that there is no way to replace the lattice-like neighborhood graph (Fig. 1,b) by a tree-like one without loss in the ability to provide smoothness of the secondary data field in all the directions from each point of the image plane.

To avoid this obstacle, for finding the values of the hidden field at each vertical row of the picture, we use a separate pixel neighborhood tree which is defined, nevertheless, on the whole pixel set and has the same horizontal branches as the others (Fig. 1,c). The generalized image processing procedure is aimed at finding the optimal values only for the goal variables at the stem nodes in each tree.

A typical image processing problem is the problem of stereo image matching. The first and algorithmically most difficult stage of the analysis of a stereo pair is matching, i.e. establishing a correspondence in the two images between the point pairs which could be treated as projections of the same point on the surface of the original three-dimensional object. It is required, for each point $\mathbf{t} = (t_1, t_2)$ of one of the images, let it be the left image conventionally taken as the reference one, to determine the vector $\mathbf{x_t} = (x_{1,\mathbf{t}}, x_{2,\mathbf{t}}) \in \mathsf{R}^2$ indicating the spatial shift between $\mathbf{t}$ and its match $\mathbf{t}' = \mathbf{t} + \mathbf{x_t}$ in the second (search) image with eventual interpixel interpolation [2]. Taken together, the shifts form the so-called disparity map (Fig. 4) to be recalculated into the geodesic elevation map.

As in the time-frequency analysis algorithm, we use here quadratic image-dependent node functions with parameters providing the best approximation of the discrepancy between the current point $\mathbf{t}$ of the reference image and its tentative match $\mathbf{t}' = \mathbf{t} + \mathbf{x}$ in the search one within the admissible shift set. The edge functions are taken also quadratic but with a theoretically infinite penalty for crossing of neighboring references in the search image. In this case, each Bellman function will be quadratic only within a quadrant in R^2, nevertheless, there exists a parametric family which contains all of them.

References

[1] S. Geman, and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. on PAMI*, 1984, 6, pp. 721-741.

[2] I. Muchnik, and V. Mottl. *Bellman Functions on Trees for Segmentation, Generalized Smoothing, Matching and Multi-Alignment in Massive Data Sets.* DIMACS Technical Report 98-15, February 1998. Center for Discrete Mathematics and Theoretical Computer Science. Rutgers, the State University of New Jersey, USA.