TREIZIÈME COLLOQUE GRETSI - JUAN-LES-PINS DU 16 AU 20 SEPTEMBRE 1991

A multilevel GMRF model for image restoration and segmentation

Fabio Arduini, and Carlo S. Regazzoni

Department of Biophysical and Electronic Engineering (DIBE) - University of Genoa -
Via All'Opera Pia 11A- I16145 Genova-ITALY

RÉSUMÉ

La méthode de Gibbs et Markov est étendue à l'intégration des observations fournies par capteurs virtuels et organisées selon une taxnomie hiérarchique. Cette méthode multiniveau couple le procédés de retauration et segmentation à des niveaux d'abstraction multiplees en modelant les observations (acquises avec les capteurs virtuels) par une fanction d'énergie différente pour chaque capteur virtuel.

ABSTRACT

An extension of Gibbs-Markov approach to integration of observations provided by virtual sensors and organized according to a hierarchical taxonomy is presented. The multi-level approach couples restoration and segmentation processes at multiple abstraction levels by modeling observations provided by related virtual sensors by means of energy functions which couple a labeling process to the observation process.

1. INTRODUCTION

The Gibbs-Markov model is a powerful method [1] for representing a-priori knowledge about different classes of images. The main advantage of this method is the capability of representing a global constraint on the characteristics of a signal as the sum of local terms, which can be separetly computed on different set of samples. This aspect makes this method attractive for the inherent high degree of parallelism. Moreover, it has been shown [2,3] that the global constraint can be interpreted as a tool for regularizing a-priori knowledge necessary to solve ill-posed problems [2], such as segmentation, edge-extraction and restoration.

Such regularization is possible, as the constraints to be used for the restoration and segmentation processes can be associated with the solution probability, which depends on the set of acquired observations. Stochastic [4] and deterministic [5,6] optimization algorithms can be used to maximize this a- posteriori probability. The availability of appropriate optimization algorithms makes the method particularly interesting.

In this paper, the Gibbs-Markov approach is extended to integration of observations provided by virtual sensors and organized according to a hierarchical taxonomy. The extension consists in using of Gibbs-Markov Random Fields (GMRFs) for higher-level processes. The classical approach is strictly used for models of pixel-level processes. Other hierarchical models [1,7] have been proposed in the literature: they couple restoration of original colours of image sites (low-level) with placement of labels (high-level). The multilevel GMRF approach is an extension of the hierarchical one in that coupled processes are performed at multiple abstraction levels corresponding to different virtual sensors. In particular, the single level restoration and edge extraction processes proposed in [1] are extended by using a two-level distributed representation. To this end, two coupled GMRFs models are associated with two specialized virtual sensors.

A modified version, (F^0,R^0), of the coupled field (F,L) proposed in [1] is associated with the first level sensor. F is the set of variables related to the image to be restored; the values of such variables are included in the set of digitized luminance values of a pixel. L (line process) is a dual field assuming binary values related to the presence of discontinuities between adjacent image pixels. A labeling process R is here used instead of L; a pixel is uniquely identified by a label chosen from a discrete set, and region properties are taken into account when describing problem constraints. Using region properties instead of discontinuities allows one to compute a probability measure which takes into account more information about a scene. In other problems, such as depth determination, the use of discontinuities may be sufficient.

A similar coupled field, (F^1,R^1), is associated with the higher level virtual sensor. Coupled solutions provided as observations by lower-level virtual sensors are considered as input. Input observations are organized into a graph whose nodes are regions (i.e., groups of pixels with the same label). The modelled process consists in labelling regions in order to merge them into complex groups. Consequently, coarser discontinuities arise at a higher level, as local differences of labels.

In section 2, the proposed extension is explained, and in section 3, models of multilevel coupled fields are described. In section 4 results on real and synthetic images are reported .

2. GENERAL DEFINITIONS

In the GMRF model [1], an image I is represented as a square lattice $S_0=\{s_m, m=1..M^2\}$ whose sites (i.e., pixels) are associated with variables, f_m, which assume random values included in a discrete set F^0. Similarly, a dual lattice D_0 is introduced, whose sites assume binary values representing the presence of edges. In this paper, we perform a pixel- labelling process at the lower level, segmenting the image into sets of homogeneous regions. A label r_m is assigned to each pixel, which takes on a random value chosen from a set of M labels. Restoration and segmentation at the pixel level are obtained as the estimation of the joint best configuration (Maximum A-Posteriori, MAP) of two sets of variables, i.e.,$\{F^0,R^0\}$ depending on input observations $G^0=\{G(i,j):(i,j) \in S^0\}$, provided by the camera.

The obtained regions are considered as input observations G^1 provided by virtual sensors at level 2, and are represented by a graph, defined as the lattice S^1, where $S^1=\{k:k=1..K\}$, and k is a generic node. Each node of the lattice S^1 corresponds to a different region r_m0, and is associated with a vector of variables $f(k)^1$. The i-th element of f(k) indicates the i-th attribute of region k (e.g., area, perimeter, average gray level, shape factor, etc.).

The neighbourhood system W1 associated with the virtual sensor at level 2 is obtained by representing observations as a region-adjacency graph (RAG) [8]; pairs of nodes are considered as neighborhoods if the related regions share a perimeter portion. Relational attributes refer to links between nodes (e.g., length of the shared perimeter). Given a lattice S^1 and its neighbourhood system W^1, an a-priori model of the attributes of a region can be defined by letting f(k) to be a GMRF with respect to (S^1, W^1). The following properties must be verified: first, the effect of all regions, say h, on a node k can be expressed by means of the only effect of the neighbouring nodes of k, i.e., $N_k=\{j \in S^1: k \in N_j\}$. This is the Markov property for region level, and can be written as:

$P(f(k)/\{f(h):h=1..H\})=P(f(k)/\{f(j):j \in N_k\})$(1).

Moreover, the probability that an attribute will assume a value included in a discrete set must be greater than zero, that is:

$P(f(k))>0$ (2)

A first order neighbourhood system has been selected, that is, only the regions that are directly adjacent to a node k belong to N_k. The neighbourhood system at level 2 is variant in respect to graph nodes.

Given the pair (S^1,W^1), a global constraint can be fixed to model a-priori knowledge necessary to restore and segment observations of the virtual sensor. The optimal configuration of

$(F^1=\{f(k)^1:k=1.K\}, R^1=\{r(k)^1:k=1..K\})$,

can be estimated, provided that an a-priori model of F^1 and R^1 on S^1 is given by a Gibbs distribution, and that an observation model of the virtual sensor is available. The estimation depends on the solution (F^{*0}, R^{*0}) provided by the lower level result.

The estimation criterion usually followed when dealing with single-level GMRF problems is the MAP (Maximum A-Posteriori) criterion. According to this criterion, the f-value is chosen from a set of variables F which maximizes the conditioned probability of f given a set of observed variables G, that is:

$\max\_F\ P(F/G) = \max\_F\ P(G/F) * P(F)$ (3).

The two terms to be jointly maximized in (1) can be respectively interpreted [3] as a model of how a sensor transforms the real world represented by variables F into observations G (i.e., P(G/F)), and an a-priori expectation about the values assumed by variables F (i.e., P(F)). The Hammersley- Clifford theorem [5] allows one to express the a-priori probability of a MRF F in terms of a Gibbs distribution, i.e.:

$P(F)=1/Z \exp(-U(F)/T)$ (4),

where Z is the partition function, T is the temperature of the system, and U(F) is a data-dependent energy function that can be computed in a local way. Under assumptions about the observation model it is possible [1], to express $U_{F/G}$ as the sum of U_F, and $U_{G/F}$, that is:

$U_{F/G} = U_{G/F} + U_F$ (5).

The Hammersley-Clifford Theorem makes the maximization of the probability equivalent to the minimization of the related energy function. It has been shown [2,3] that $U_{G/F}$ can be regarded as a term that charges for the differences between sensor observations and the solution, and that U_F is a stabilizer that makes the extimation problem well-posed.

3 MULTILEVEL GMRFs

The multilevel model for restoration and segmentation processes is defined is defined by the energy functions at the image and region levels.

3.1 IMAGE LEVEL

At this level, the GMRF model is given in terms of the function $P(F^0,R^0/G^0,F^1,R^1)$ (6)

where F^0 is the observation process of the image level, R^0 is the labelling process which assignes a label to each pixel of the input image, and G^0 is an initial measure of the observation related to the estimated process data by the relation

$G^0=F^0+N$ (7), where N is a white Gaussian noise with zero mean and variance σ^2. Image blurring is considered negligible. A hierarchical GMRF model of (6) defined by two coupled processes has been considered: the observation process is performed at the image level on a regular lattice, where each site corresponds to a pixel in the digitized image; the labelling process consists in assigning labels to all pixels on the basis of a p.d.f., as defined below. The p.d.f. (6) can be decomposed into three terms

$P(F^0,R^0/G^0,F^1,R^1)=P_{ic}(F^0,R^0)P_\lambda(F^0,R^0/G^0)P_\pi(F^0,R^0/G^0 \cdot F^1,R^1)$ (8)

where P_{ic} is a term related to regularization constraints at the image level, P_λ is a term derived from the evidence provided by the physical sensor, and P_π is an expectation term supplied by the region level. Using the Hammersley Clifford Theorem, (8) can be rewritten as:

$U(F^0,R^0/G^0,F^1,R^1)=U_{ic}(F^0,R^0)U_\lambda(F^0,R^0/G^0)U_\pi(F^0,R^0/G^0 \cdot F^1,R^1)$ (9)

The problem of performing the segmentation of a scene starting from data coming from a visual sensor is generally ill-posed, and requires a regularization process. A weak smoothness constraint similar to the one presented in [1], has been imposed on possible solutions, together with a term which favours the expansion of a label in the related process:

$$U_{ic}(F^0,R^0)=\sum_{i}\sum_{L j C_{Fi}}(f_i-f_j)^2(1-\delta_{r_{ij}})+\sum V_c(R^0) \quad (10)$$

In this way, data are approximated by a piecewise constant function, as in [1]. The first sum is extended to all sites of the image process, and, for each site, a first-order neighbourhood system with the usually associated cliques is selected . The second sum is extended to all sites of the lattice by considering the neighbourhing pixels of each site. The term $\delta_{r_{ij}}$ is a Kronecker delta which couples the labelling and the observation processes, and acts as the line process in [1].

Compatibility of the observation process with the measured data is reached through the evidence term U_λ of the energy function

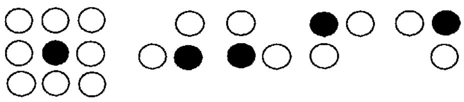
$$U_\lambda(F^0,R^0/G^0)=\frac{1}{2\sigma^2}\sum_i (f_i-g_i)^2+U_r(R^0/G^0) \quad (11)$$

The term U_R can be neglected in the minimization process, as all the configurations of R0 are assumed to do not depend from G0. The third term U_π is an expectaton term which takes into account suggestions coming from the higher level. U_π can be expressed as a sum of potentials:

$$U_\pi(F^0,R^0/F^1,R^1)=\sum_i U_\pi(R^0/R^1)+\sum_i (f_i-f_{r_i})^2 \quad (12)$$

This sum is extended to all sites of the lattice. The first term relates the area of a region to its expansion probability: regions of locally larger size are favoured to grow, as the related potential cost becomes lower. The choice of a region expansion also depends on the values of the intensity gradients between adjacent regions. In the following table, a set of potential scores of cliques based on a second order neighbourhood system are shown; it is worth-noting that these potential depend on the area of neighbouring regions, S(.).

The second term refers to the average grey level of a region. This term aims at favouring the aggregation of a pixel into a region characterized by an average grey level closer to the luminance of the pixel itself.



F(r) = area of region r

Ⓐ $F(A)<F(C)$ and $F(B)<F(C)$
Ⓒ Ⓑ $Vc(C) = 1.2$ else $V(C) = 2.0$

Ⓐ $F(A)<F(C)$
Ⓒ Ⓒ $Vc(A) = 0.4$ else $Vc(C) = 1.2$

Ⓐ $F(A)<F(C)$
Ⓒ Ⓐ $Vc(A) = 0.4$ else $Vc(C) = 1.2$

Table 1 : Potential Cost of clique configurations for the segmentation process at the Image Level

3.2 REGION LEVEL

At the region level, a set of measures for each region found at level 0 is considered as input. Regions are organized into a graph whose nodes contain a set of attribute values which characterize the regions; a branch between two nodes of the graph indicates that the related regions are adjacent. The probability of being to be maximized at this level can be expressed as:

$$P(F^1,R^1/G^1) \quad (12)$$

where F^1 and R^1 are the fields to be estimated at this level and G^1 are the observations provided by the image-level process, that is:

$$G^1=(F^0,R^0) \quad (13)$$

As at the image-level, it is possible to write the local a-priori probability model as:

$$U(F^1,R^1/G^1)=U^1_{ic}(F^1,R^1)U^1_\lambda(F^1,R^1/G^1) \quad (14)$$

The ic part consists of two different terms which take into account relational properties between adjacent regions, the first term is:

$$U_{ic}(F^1,R^1)=\sum_i\sum_j \{(f_i-f_j)^2+k\,U_{per}(P_{r_{ij}},P_T)\} \quad (15)$$

where f_j is the average grey level of a region j belonging to the set of neighbourhoods of i, N_i; P_{rij} is the length of the common perimeter between regions i and j; and P_T is the total perimeter of region i. Consequently, to assign the same label to two adjacent regions these must have both a similar average grey level and a long enough shared perimeter. In general each property of a region can be used to drive the merging process, that is:

$$U_{ic}(F^1,R^1)=\sum_i\sum_k \{C_k(F_i,F_j)\}$$

where C_k is a generic cost function for two neighbouring sites, and f(i), f(j) are vectors of attributes of regions i and j.This term acts as a stabilizer at the region level.

The second term of the cost function is:

$$U_\lambda(F_1,R^1/G^1)=\sum_i \frac{1}{2\sigma_i}(f_i-g_i)^2$$

The cost function is higher for each site whose restored average gray level f_{i1}, (which is an estimation of the actual mean gray level), is far from the average gray level, g_i, which is originally assigned to regions by considering the image-level output. In this way, the final result depends on intermediate outputs provided by the lower-level virtual sensor. The value of sigma is computed as the variance of the pixel luminance inside region i, and is used as an uncertainty dependent weight.Regions characterized by a spread histogram are considered less effective in the computation of the global cost. The term P_π is not included, as it implies the presence of hierarchically higher levels which drive segmentation on the basis of the specific application domain considered.

4 RESULTS

Results are reported on a real image representing an airplane displayed in Fig1.a. After 100 iterations of the image-level process, the image is splitted into several regions (see Fig.1.c); however, such regions do not affect the accuracy of the contour of the object contained in the image. After the region level process, the coarser final segmentation shown in figure 1.d is obtained, which maintains a clear discrimination between the airplane and the background; in addition the plane shape contains small, though significant, regions (e.g., the region representing letters and symbols by which the plane can be identified). The restored image, shown in figure 1.b, obtained at the end of the process appears very promising too. One can easily notice that the boundaries between the regions identified by the proposed approach have been preserved. Finally, the graphs in figure 2 show the behaviour of the energy function, both in the case of the image and region level processes (Fig. 2.b) and in the case of the image level (Fig.2.a) process only. It appears evident that the convergence at a correct segmentation result is obtained when both levels are used in a reduced number of iterations.

5 CONCLUSIONS

A multilevel GMRF model of image segmentation and restoration has been presented. The model is suitable for a hierarchical network of virtual sensors, and implies the development of distributed a- priori models of observations at each level, and of top-down and bottom-up dependences between adjacent levels. Results are comparable with those obtained by a single-level approach, but require a lower computational load. Future work will concern the assessment of distributed optimization algorithms for networks of the proposed type and an extension [9] of the proposed approach segmenatation of image sequences.

Fig1 from top to bottom: Original Image (Fig 1.a), Restored Image (Fig 1.b), Segmented Image after the Image level Process (Fig 1.c), and after the Region Level Process (Fig 1.d)

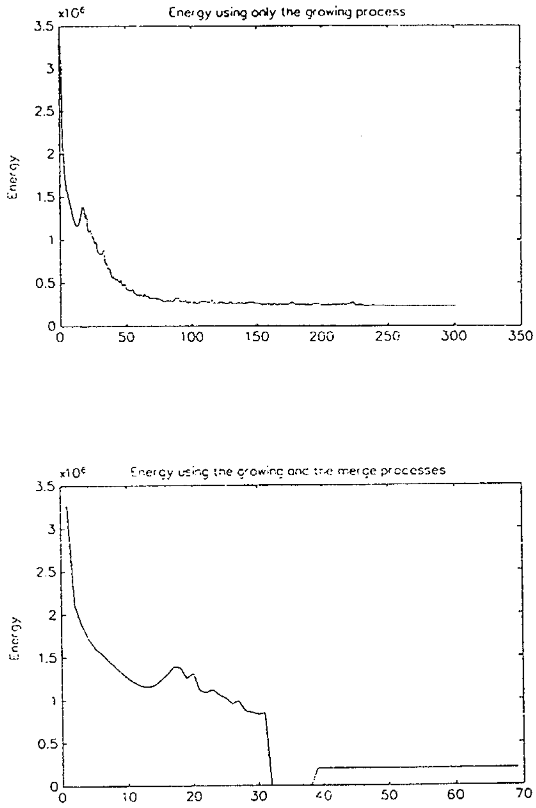


Fig 2 (From top to Bottom): behaviour of the Energy Fuction when only the Image Level is applied (Fig.2.a); behaviour of the Energy Function when both levels are applied (Fig.2.b).

REFERENCES

[1] S.Geman, D.Geman, "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images", IEEE Trans. on PAMI, Vol.PAMI-6, No.6, pp.721-741, 1984.

[2] M. Bertero, T.Poggio, and V.Torre"Ill posed problems in early vision", MIT A.I.Memo 924, May 1987.

[3] R.Szelinski, "Bayesian Modeling of Uncertainty in Low-Level Vision", Kluwer Academic Publishers, Norwell, Massachusetts, 1989.

[4] S.Kirkpatrick, C.D.Gelatt, M.P.Vecchi, "Optimization by Simulated Annealing", Science, Vol.220, No.4598, pp.671-680.

[5] J.Besag, "Spatial Interaction and the Statistical Analysis of Lattice Systems", J.Royal Stat. Soc.B 34, pp.75-83, 1972.

[6]J. Zerubia and R.Chelappa, "Mean Field Approximation using Compound-Markov Random Fields for Edge Detection and Image Restoration, Proc. 1990 IEEE ICASSP, pp.2193-2196, 1990.

[7]H.Derin and H.Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields", IEEE Trans. on PAMI, PAMI-6, pp.707-720, 1987.

[8] D.Ballard and C.Brown, "Computer Vision", Prentice Hall, Englewood Cliffs, 1982.

[9] F.Arduini, R.Cabri, G.Foresti, V.Murino and C.Regazzoni, "A numerical and symbolic fusion method for interpretation of image sequence" , Proc. 1991 IEEE ICASSP, pp. 2477-2480, May 1991.