Synchronous Image restoration

Laurent Younes

CMLA-DIAM, Ecole Normale Supérieure de Cachan, 61 avenue du Président Wilson, 94 235 Cachan CEDEX.

Abstract. We analyse a class of random fields invariant by stochastic synchronous updating of all sites, subject to a generalized reversibility assumption. We give a formal definition and properties of the model, study the problem of posterior simulation, parameter estimation, and then present experimental results in image restoration.

1 Introduction.

We utilize in this work a new class of random fields on some given set of sites, which are invariant under the action of a transition probability which synchronously updates the states of all sites. We illustrate, with an issue of image restoration, the usefulness of these models, and their feasibility for practical applications.

We begin by a summary of the definition of p-periodic synchronous random fields. Let S be the finite set of sites, F a finite state space. The set of all configurations on S is F^S, and will be denoted by Ω_S (or Ω if no confusion is possible). A random field (r.f.) on S is a probability distribution on Ω.

Definition 1 *A synchronous kernel of order q is a family $\mathcal{P}$ of transition probabilities from Ω^q to F, denoted*

$$p_s(x(q),\ldots,x(1)\ ;\ y_s),$$

($x(l)\in\Omega$, $l=1,\ldots,q$, $y_s\in F$).

We say that $\mathcal{P}$ is positive.

Definition 2 *If $\mathcal{P}$ is a synchronous kernel, its associated (synchronous) transition probability from Ω^q to Ω is*

$$P(x(q),\ldots,x(1)\ ;\ y)=\prod_{s\in S}p_s(x(q),\ldots,x(1)\ ;\ y_s).$$

It corresponds to a simultaneous updating of all sites according to the local transitions p_s.

For a function f defined on Ω^q, define

$$Pf(x(q),\ldots,x(1))=\int f(x(q-1),\ldots,x(0))P(x(q),\ldots,x(1)\ \ dx(0)). \tag{1}$$

Then, there exists a unique distribution $\bar{\pi}$ on Ω^q which satisfies :

$$\int Pf d\bar{\pi}=\int f d\bar{\pi} \tag{2}$$

for all f. We finally define a probability distribution on Ω, (rather than Ω^q), by

Definition 3 *We say that a law π on Ω is associated to the synchronous kernel $\mathcal{P}$ if π is one of the marginal distributions of $\bar{\pi}$, where $\bar{\pi}$ is defined in (2).*

It is easy to deduce from equation (2) that all the marginals of $\bar{\pi}$ are identical.

We now define p-periodicity :

Definition 4 *Let $p=q+1$. The transition probability $P(x(q),\ldots,x(1)\ ;\ x(0))$ is p-periodic with respect to a distribution $\bar{\pi}$ on Ω^q if, for any functions $f_q,\ldots,f_0$ defined on Ω,*

$$\int f_q[x(q)]\ldots f_1[x(1)]Pf_0[x(q),\ldots,x(1)]d\bar{\pi}=\int Pf_q[x(q-1),\ldots,x(0)]f_{q-1}[x(q-1)]\ldots f_0[x(0)]d\bar{\pi} \tag{3}$$

Lecture Notes in Computer Science, Vol. 801
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94
© Springer-Verlag Berlin Heidelberg 1994

In other terms, the compound distribution *on* Ω^p*, defined by*

$$\mu[x(q),\ldots,x(0)] = \bar{\pi}[x(q),\ldots,x(1)]P[x(q),\ldots,x(1)\ \ x(0)] \tag{4}$$

is invariant by circular permutation of $x(q),\ldots,x(0)$.

We shall say that a synchronous field is p-periodic if it is the marginal distribution of $\bar{\pi}$ (or μ) with respect to which P is p-periodic.

In *(younes 1993)*, the following characterization theorem is proved. We identify Ω^p with $\Omega_{S_1}\otimes\ldots\otimes\Omega_{S_p}$ where $S_i = S\times\{i\}$, $i = 1,\ldots,p$, are copies of S and add a new element to S, which can be considered as a "void site", and which we will denote by 0 ($0\notin S$).

Theorem 1 *Let* π *be a probability measure on* Ω*. The following properties are equivalent :*

1. π *is a synchronous p-periodic r.f. associated to a positive synchronous kernel.*
2. *There exists a distribution* μ *on* Ω^p *such that*
   - *i-* μ *is invariant by circular permutation of the coordinates.*
   - *ii- For all* $i = 1,\ldots,p$*, the variables* X_s*,* $s\in S_i$ *are* μ*-conditionally independent given the other variables* X_t*,* $t\notin S_i$*.*
   - *iii- The conditional distribution*
   $$\mu(X_s \mid X_t, t\neq s)$$
   for $s\in S_i$ *is positive.*
   - *iv-* π *is the marginal distribution of* μ *over* Ω_{S_1}*.*

In that case, the associated synchronous kernel may be described as follows.

For each p-uple in $(S\cup\{0\})^p$*, where 0 is the void site, of the kind* $\bar{s} = (s_1,\ldots,s_p)$*, there exists a function*

$$h_{\bar{s}}[x(1),\ldots,x(p)] = h_{\bar{s}}[x_{s_1}(1),\ldots,x_{s_p}(p)]$$

defined on Ω^p*, which only depends on variables* $x_{s_k}(k)$ *for indices* k *such that* $s_k\neq 0$*, such that* $p_u[x(1),\ldots,x(q)\ ;\ x_u(p)]$ *takes the form :*

$$\exp\left\{\sum_{\bar{s}=(s_1,\ldots,s_{p-1},u)} h_{\bar{s}}[x(1),\ldots,x(p)]\right\} Z_u[x(1),\ldots,x(p-1)]. \tag{5}$$

Moreover, the functions $h_{\bar{s}}$ *are invariant by circular permutation of the indices, in the sense that, for any* $x(1),\ldots,x(p)$*,*

$$h_{\bar{s}}[x(1),\ldots,x(p)] = h_{\bar{s}_\tau}[x(p),x(1),\ldots,x(p-1)],$$

where $\bar{s}_\tau = (s_p, s_1\ldots,s_{p-1})$.

The p-step distribution μ *can be expressed in terms of the functions* $h_{\bar{s}}$ *:*

$$\mu[x(1),\ldots,x(p)] = \frac{1}{Z}\exp\left\{\sum_{\bar{s}} h_{\bar{s}}[x(1),\ldots,x(p)]\right\}. \tag{6}$$

We shall also refer to the following notion :

Definition 5 *A compound (or joint) partially synchronous distribution of order p is a Gibbs distribution* μ *on* $\Omega_{\bar{S}}$ *such that the variables* X_s*,* $s\in S_i$ *are* μ*-conditionally independent given the other variables* X_t*,* $t\notin S_i$*.*

If μ *is a compound synchronous distribution, its marginal over* $\Omega_{S_1}\equiv\Omega_S$ *will be called a partially synchronous distribution of order p.*

Remark. To sample from a partially synchronous distribution, one needs to sample from μ, which involves $p-1$ auxilliary variables, each of which being synchronously sampled conditionally to the others. The efficacy of such a parallel sampling therefore decreases when p is large.

From the preceding definition, we see that p-periodic fields are partially synchronous distributions associated to μ which is invariant by circular permutation of the coordinates. There is no such loss of efficiency in this case, since all auxilliary variables have the same marginals.

2 Synchronous sampling of a posterior distribution.

2.1 Sitewise degradation.

This issue can be considered as critical to measure the usefulness of a model for applications.

To formalize the situation, let Ω denote the configuration space of S, with state space F, and Ω' denote some other configuration space on some state space G. Assume that we are given a family of functions $(b_s,\ s \in S)$ from Ω to G, yielding a function b from Ω to Ω' defined by $b(x) = \xi$ with $\xi_s = b_s(x)$ for all s. In a statistical interpretation, Ω is the set of "original" configurations, and Ω' the set of "observed" configurations.

For a given probability distribution π on Ω, and a given configuration $\xi \in \Omega'$, we are concerned with the issue of sampling from the *posterior distribution* $\pi(\ . \ |\xi)$.

We shall consider the following important particular case in which the computation of ξ from x is performed coordinatewise, i.e. ξ_s only depends on x_s (so that b_s is a function from F to G). Examples in which this is satisfied are

- $F = G \times G'$ and only the G-component of x is observed (partial observations).
- Let $(x^0_s,\ s \in S)$ be an unobserved random field, and (ϵ_s) be a noise which is independent of x^0. Assume that x^0 and ϵ both follow a synchronous distribution, or more generally that the joint distribution of (x^0, ϵ) is synchronous. Assume finally that the observation takes the form : $\xi_s = b_s(x^0_s, \epsilon_s)$.

Proposition 1 *Assume that π is p-periodic and that for all s, $b_s(x) = b_s(x_s)$ only depends on x_s. Then, the posterior distribution $\pi(\ . \ |\xi)$. is partially synchronous of order p.*

This is not valid anymore when the condition $b_s(x) = b_s(x_s)$ is relaxed. Although this condition is true for a large range of applications, there remain some significant cases for which it is not satisfied. The most important among these, especially in the context of image restoration is the case of blurring. The next section addresses the case of linear blurring with additive Gaussian white noise.

2.2 Restoration of blurred pictures.

In this section, the state space F is no more finite, but equal to the real line $\mathbf{R}$. The random fields are assumed to have densities with respect to Lebesgue measure on Ω, which are given by the same kind of formulae as in the finite case, with the implicit integrability assumptions.

We assume that the observation is obtained through the equation

$$\xi_s = \sum_t \eta_{st} x_t + \epsilon_s, \tag{7}$$

where ϵ is some Gaussian white noise of variance σ^2, and η_{st} are the coefficients of a point-spread function around s. A particular case is when $\sigma^2 = 0$, in which the restoration problem reduces to deblurring the picture.

In order to restore the original picture from the observed ξ, the problem is still to devise an efficient sampling algorithm of the posterior distribution. The previous methods cannot be applied, unless $\eta_{st} = 0$ for $s \neq t$. The difficulty comes from the fact that, when expressing the energy of the conditional distribution of x given ξ, there appears a term $\sum(\xi_s - \sum_t \eta_{st}x_t)^2$, therefore yielding interactions between x_s and x_t for $t \neq s$. In the following proposition, we show how a very simple trick can be used to solve this problem.

Proposition 2 *Assume that π is p-periodic and that ξ is given by equation (7) above. Then, there exists a compound partially synchronous distribution μ, of order $p+1$, of the kind*

$$\mu(z, \bar{z}, x(2), \ldots, x(p)),$$

such that the distribution of $(z + \bar{z})/2$ is $\pi(\ . \ |\xi)$.

Proof: Let $\mu_0[x(1), \ldots, x(p)]$ be the density of the p-step distribution associated to π ; it is of the kind

$$\exp[-Q(x(1), \ldots, x(p))]/Z,$$

Q being the associated energy function. To simplify notations, we set $x = x(1)$.

The energy function associated to the joint field $(x(1),\dots,x(p),\xi)$ is

$$\frac{1}{2\sigma^2}\sum_s(\xi_s-\sum_t\eta_{st}x_t)^2+Q(x,x(2),\dots,x(p))$$

Introduce a new r.f., (u_s), which is Gaussian with energy $\frac{1}{2\sigma^2}(a\sum_s u_s^2-\sum_s(\sum_t\eta_{st}u_t)^2)$,, a being large enough for this quadratic form to be positive. Assume that u is independent of the other fields ($x(i)$, $i=1,\dots.p$ and ξ). After a simple transformation, the joint energy of $x, x(2),\dots,x(p),\xi,u$ can be written

$$\frac{1}{2\sigma^2}\{\sum_s[\xi_s-\sum_t\eta_{st}(x_t+u_t)][\xi_s-\sum_t\eta_{st}(x_t-u_t)]+a\sum_s u_s^2\}+Q(x,x(2),\dots,x(p))$$

It suffices now to set $z_t=x_t-u_t$ and $\bar z_t=x_t+u_t$. The distribution of z, $\bar z$, $x(2),\dots,x(p)$ and ξ has for energy

$$\frac{1}{2\sigma^2}\{\sum_s[\xi_s-\sum_t\eta_{st}z_t][\xi_s-\sum_t\eta_{st}\bar z_t]+\frac{a}{4}\sum_s(z_s-\bar z_s)^2\}+Q((z+\bar z)/2,x(2),\dots,x(p)).$$

This provides the desired structure for the conditionnal distribution of z, $\bar z$, $x(2),\dots,x(p)$ given ξ, which is the distribution μ we were looking for. □

3 Experiments.

We now present examples to illustrate the restoration of noisy pictures in this last case.

3.1 Modeling.

We include, as it is standard in image modeling, edge elements within the prior distribution and introduce two hidden fields, $(h_s, s\in S)$ and $(v_s, s\in S)$, with values in $\{0,1\}$, respectively indicating the presence of a horizontal edge ($h_s=1$) or vertical edge ($v_s=1$). More precisely, if $s=(i,j)$ is the representation of s on the image grid, h_s indicates an edge between (i,j) and $(i-1,j)$, and v_s between (i,j) and $(i,j-1)$. To shorten notation, we set $(i-1,j)=s.h$ and $(i,j-1)=s.v$.

The prior distribution is therefore defined on the set of all configurations of (x,h,v). To model a synchronous r.f., we must introduce auxilliary fields $(y,\bar h,\bar v)$, and model a compound distribution $\mu\left((x,h,v),(y,\bar h,\bar v)\right)$ with the property that, given $(y,\bar h,\bar v)$, all component $(x_s,h_s,v_s), s\in S$ are independent, and conversely ; we do not impose mutual independence of x_s, h_s and v_s. Up to a scaling factor, the density of this distribution μ is given by

$$\exp\{-\tfrac{1}{2\tau^2}[\,\delta\sum_s(x_s^2+y_s^2)+\sum_s(x_s-y_s)^2+\kappa\sum_s((x_s-y_{s.v})^2+(y_s-x_{s.v})^2-2\theta_0)(1-v_s) \tag{8}$$
$$+\kappa\sum_s((x_s-y_{s.h})^2+(y_s-x_{s.h})^2-2\theta_0)(1-h_s)\,]+Q(h,v,\bar h,\bar v)\,\}$$

This density is with respect to the product of Lebesgue measures at each site for x_s and y_s variables, and counting measures on $\{0,1\}$ for edge variables. The parameter δ is an arbitrary, very small number ensuring the integrability of the above expression. In fact, when we will be considering the posterior, for which this problem disappears, we will let δ tend to 0.

The second sum forces gray-level variables x_s and y_s to have values which are not too far apart. This allows us to interpret the third and fourth terms, which are weighted by a positive parameter κ, as terms forcing the differences of gray-level at neighboring pixels to be small unless an edge separates them. The parameter θ_0 appears like a threshold below which this difference should be in the absence of edge. For the experiments, we have heuristically fixed the values of τ, κ and θ_0. This is more or less made possible by the simplicity of this part of the model.

The "edge energy", Q, is quadratic in its variables. It has the form

$$Q(w^1,w^2,w^3,w^4)=\beta_0\sum_{i,s}w_s^i\beta_1\sum\nolimits_1 w_s^iw_t^j+\dots+\beta_k\sum\nolimits_k w_s^iw_t^j \tag{9}$$

The sums $\sum_1, \ldots, \sum_k$ are made over specified families of indices i, s, j, t. For example, one of them represents self-relation (like the α parameter in the synchronous Ising model), and the sum is made over i, s, j, t such that $s = t$ and either $i = 1$ and $j = 3$ (horizontal edges) or $i = 2$ and $j = 4$ (vertical edges). Other contain interaction between adjacent aligned edge elements, of edge elements making a right angle, and so on. The constraint is that there may not be an interaction within (h, v) nor within $(\bar{h}, \bar{v})$, which means that one may not have $i = 1$ together with $j = 1$ or 2, nor $i = j = 2$, and similarly for 3 and 4. In our experiments, we used $k = 15$. It is clear that it is not possible to work with heuristics, nor by trial-and-error to set the value of such a number of parameters. This has been done with a help of a learning procedure which is summarized in remark R16.

Using the method given in Proposition 3, we introduce two additional auxilliary fields, z and $\bar{z}$, such that $x_s = (z_s + \bar{z}_s)/2$, and one obtains the distribution of $(z, \bar{z}, y, h, \bar{h}, v, \bar{v})$ given ξ.

This distribution can be sampled by iterating the following sequence of *synchronous* steps ; assume that a current configuration of $(z, \bar{z}, y, h, \bar{h}, v, \bar{v})$ is given. Then, a global updating of this configuration can be done by

1. Update z, given $\bar{z}$, y, h, v and $\bar{h}$, $\bar{v}$ given h and v.
2. Set $x = (z + \bar{z})/2$
3. Update y given x, h, v.
4. Update h, v given x, y, $\bar{h}$, $\bar{v}$ and set $\bar{z} = z$.

We give some results of experiments in this context. The noise is Gaussian, additive, with variance 200, the image being coded in gray levels between 0 and 256. The blur is obtained through a 5 by 5 Gaussian filter given by $\eta_{st} = c.\exp(-\|s - t\|^2/2\zeta)$ if $\max(|s_1 - t_1|, |s_2 - t_2|) \leq 2$, and 0 if not, c being a normalization ensuring that the sum of the η_{st} is 1.

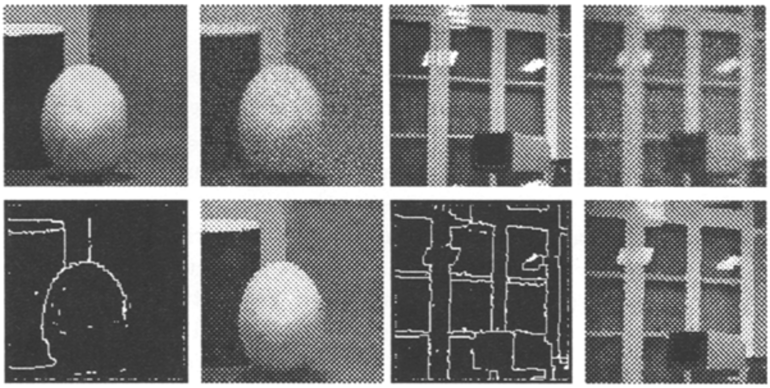

Fig. 1. Restoration : blurred pictures ($\zeta = 2$) with additive noise of variance 200 (Upper left : Original, upper right : Noisy, lower left : Estimated edges, lower right : restored picture).

References

J. Besag (1974) : Spatial Interaction and the Statistical Analysis of Lattice Systems. *J. of Roy. Stat. Soc.* B-36 pp 192-236.

D.A. Dawson (1975) : Synchronous and asynchronous reversible Markov systems *Canad. Math. Bull.* 17 633-649.

D. Geman (1991) :*Random Fields and Inverse Problems in Imaging,* In *Proceedings of the Ecole d'été de Saint-Flour,* Lecture Notes in Mathematics, Springer Verlag, New York.

D. and S. Geman (1984): Stochastic Relaxation, Gibbs Distribution and Bayesian Restoration of Images. *IEEE TPAMI.* Vol PAMI-6 pp 721-741.

O. Koslov and N. Vasilyev (1980) : Reversible Markov chains with local interactions. In *Multicomponent Random Systems*, R.L. Dobrushin and Ya. G. Sinai Editors. (Dekker New York). 415-469.

L. Younes (1993) : Synchronous Random Fields and Image restoration (preprint).