Proceedings of the IEEE Workshop on Statistical and Computational Theories of Vision, Fort Collins, CO, June, 1999. (Published on Web).

An Equivalence Theorem for Mixed Markov Fields

Artur Fridman[1]

Abstract

We develop a framework for a new class of probability distributions on graphical structures, to be used as prior models particularly in computer vision. Mixed Markov Fields, as they are called, generalize Markov Random Fields by allowing dynamic links between vertices. We state and prove an analog of the Hammersley-Clifford equivalence theorem for Mixed Markov Fields. We then express several standard image analysis problems in the Mixed Markov Field framework.

1. Department of Mathematics, Brown University, Providence, RI 02912

1 Introduction

A considerable effort in computer vision, as well as in other areas of pattern recognition, has been directed toward developing computationally feasible statistical models that appropriately represent available prior knowledge. Markov Random Fields (MRF's) have been used extensively for that purpose in computer vision for image restoration (GEMAN and GEMAN 1984), boundary detection (GEMAN 1987), texture discrimination (GEMAN and GRAFFIGNE 1987), as well as in medical expert systems, language parsing, and beyond.

Sometimes, however, a Markov Random Field is not the most appropriate as a prior model. Consider, for example, a texture discrimination problem discussed in GEMAN and GRAFFIGNE (1987). It can be solved within the Bayesian framework by introducing, in addition to the pixel process $(x_i)_{i \in I}$, an auxiliary label process $(x_l)_{l \in L}$, where x_l is the texture of a pixel block $l \in L$, encoded as a label. A random process on (x_I, x_L) is built from the local interaction terms between neighboring pixels and texture blocks, and the texture discrimination problem is solved by MAP estimation.

Now suppose the image contains *only one* texture block of a certain shape and size, and the background is modeled with a different texture. Performing MAP estimation on the combined pixel-label process in this case is not only time-consuming because of the presence of label to label interactions, it is actually unnecessary, since all we have to do is find *where* in the image the differently textured block is located. It is thus better to introduce, in place of the label process, a single *pixel*-valued random variable, which will point to the location of the texture block. It is this need for combining pixel-valued *address* variables with the standard pixel process that leads us to the concept of a Mixed Markov field.

The idea of constructing probability distributions from local interactions of intensity variables and address variables belongs to MUMFORD (1995, 1997), who proposed a class of probability models on graphs which have two types of vertices. Random variables that "live" on address vertices are vertex-valued. Assigning a value to an address variable produces a "dynamic" edge connecting the address vertex on which this variable lives and the vertex to which it is pointing. This edge is in addition to the set of "static" edges that exist before the address variables have been assigned values. The edge structure is completely determined after all the address variables have been assigned values and is a union of dynamic and static edges. A probability distribution is then defined as the product over the cliques in the full graph.

2 The Definition of a Mixed Markov Field

The simple texture discrimination example above can be generalized considerably by allowing multiple templates of various shapes and textures in the image, by allowing template to template interactions (i.e., we may require that clumping together a group of texture blocks produce a new, distinct texture), etc. To adequately represent the observed and unobserved variables in this case, a new model with the following components is proposed:

- a graph $G = (I \cup A, E)$ where I is the image plane with a neighborhood relation N determined by the set of edges E, and A is a collection of templates;

- a random vector $\mathbf{x} = (x_I; i_A)$ where $x_I = (x_i)_{i \in I}$ is a vector of intensity variables with values in $\mathbf{R}$ and $i_A = (i_a)_{a \in A}$ is a vector of address variables with values in $I \cup \{\text{nil}\}$. i_a holds the location of template a in the image I. If a is not present in the image, then $i_a = \text{nil}$.

- for each clique $C \in \mathfrak{C}$, for each subset D of the set of templates A, and for each vector i_D of pixels inside C, a neighbor potential $V_C(x_C; i_D)$ determining the distribution of the intensity variables x_C needed to produce the texture created by clumping together the templates from D at positions i_D inside C. In particular, the set of neighbor potentials $V_C(x_C, \emptyset)$ determines the background texture on I, and $V_C(x_C; i_a)$, $a \in A$ determines the "pure a" distribution of x_C produced when template a is placed at pixel i_a inside C.

Note that a pixel-valued *vector* i_A is a distinct entity from a *subset* $\{i_A\}$ of the image plane. Given i_A, we'll denote by $i_{\uparrow C}$ a subvector of i_A consisting of those i_a satisfying: $i_a \in C$. That is, $\uparrow C = \{a \in A\colon\ i_a \in C\}$.

A Mixed Markov Field (MMF) on $I \cup A$ is now defined as the product over the cliques in I of the neighbor potentials:

$$P(x_I; i_A) = \frac{1}{Z} \prod_{C \in \mathfrak{C}} V_C(x_C; i_{\uparrow C})$$

This definition leads naturally to the idea of *dynamic cliques*. Once the values of the address variables have been assigned, a neighborhood relation on the set of vertices in A (templates) is determined as follows:

a subset D of A is a dynamic clique if the set $\{i_D\}$ is a clique in I

(in particular, $i_a \neq \text{nil} \quad \forall a \in D$). Changing the values of the address variables produces different sets of dynamic cliques. We might say that the vertices in A get temporarily linked into groups based on the proximity of the values of the address variables in the image plane.

We will return to the concept of dynamic templates in Section 4, when discussing the face template of FISCHLER and ELSCHLAGER (1973). A typical Mixed Markov Field is shown in Figure 1.

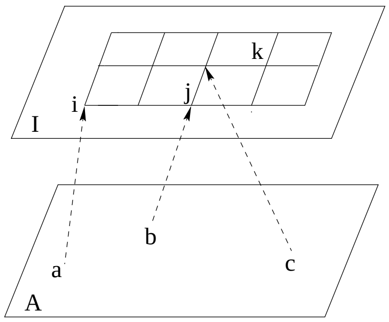

Figure 1: An Example of a Mixed Markov Field

3 An Equivalence Theorem

The utility of Mixed Markov Fields comes the following Equivalence theorem, which establishes that

1. the local characteristics of a Mixed Markov Field have a Markov property similar to that of the local characteristics of a Markov Random Field, and

2. that any random process on $(x_I; i_A)$ whose local characteristics have this Markov property is a Mixed Markov Field.

Two special cases of this theorem have been previously analyzed. In the absence of address variables, a Mixed Markov Field becomes a Markov Random Field, for which the above result is known as the Hammersley-Clifford theorem (cf. HAMMERSLEY and CLIFFORD 1968, GRIMMETT 1973). A Mixed Markov Field with indistinguishable templates (i.e., $P(x_C; i_A) = P(x_C; j_A)$ whenever $\{i_A\} = \{j_A\}$) and a blank background is essentially a Markov Point Process (cf. RIPLEY 1977), for which the equivalence of Markov and Gibbs characterizations was shown in RIPLEY and KELLY (1977).

The exact statement of the main theorem is given below. We denote by $N(i)$ the set of neighbor pixels of i, and by $\overline{N(i)}$ the set $N(i) \cup \{i\}$. By agreement, $\overline{N(\text{nil})} = \emptyset$.

Theorem 3.1 *Suppose $P(x_I; i_A) > 0$. P is a Mixed Markov Field on $I \cup A$ if and only if*

- *$\forall j \in I$, the ratio $\dfrac{P(x_j, x_{I\setminus\{j\}}; i_A)}{P(x'_j, x_{I\setminus\{j\}}; i_A)}$ is a function only of $x_{N(j)}$ and $i_{\uparrow\overline{N(j)}}$*

- *$\forall a \in A$, the ratio $\dfrac{P(i_a, i_{A\setminus\{a\}}; x_I)}{P(nil_a, i_{A\setminus\{a\}}; x_I)}$ is a function only of $x_{\overline{N(i_a)}}$ and $i_{\uparrow\overline{N(i_a)}}$.*

The proof of this Theorem is given in the Appendix.

4 Examples of Mixed Markov Fields

4.1 The Kelly-Ripley Point process

Consider first the Markov point process with values in I, described in RIPLEY and KELLY (1977). Its probability density, conditioned on the number of points (say, n), is given by

$$p(x) = p(x_1, \ldots, x_n) = \alpha \cdot \gamma^{r(x)}, \qquad \text{where}$$

$$r(x) = \#\{(k,l) : k < l, |x_k - x_l| < r\}; \quad r > 0.$$

This point process can be expressed as a Mixed Markov Field on $I \cup A$, where $A = \{a_1, \ldots, a_n\}$ is a set of indistinguishable templates. The neighborhood system on I is determined by setting $N(i) = \{j \in I : |i - j| < r\}$. Neighbor potentials V_C are set as follows:

$$V_C(x_C; i_A) = \begin{cases} \gamma & \text{if } |C| = 2 \text{ and } C \subseteq \{i_A\} \\ 1 & \text{otherwise} \end{cases}$$

Setting V_C as above attaches penalty γ to every pair of points in $\{i_A\}$ closer that r to each other.

4.2 The Face Template of Fischler and Elschlager

We now look at the face model discussed in FISCHLER and ELSCHLAGER (1973); see also YUILLE (1991), cf. Figure 2.

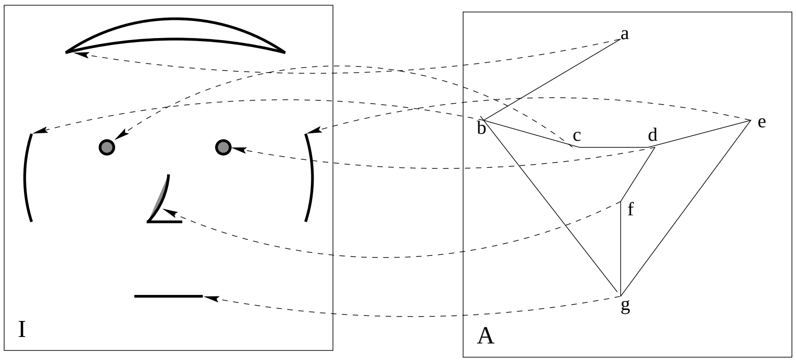

Figure 2: The face template of Fischler and Elschlager

An important difference between this and the earlier example is that here, some of the address variables are connected by "springs" with associated functions $g_{ab}(i,j)$ which determine the cost of placing template $a \in A$ at position $i \in I$ and $b \in A$ at $j \in I$. The posterior energy function of the process is given by

$$E(i_A|x_I) = \sum_{a \in A} l_a(i_a) + \sum_{<a,b>} g_{ab}(i_a, i_b)$$

where $l_a(i)$ measures how well template a fits at i and $g_{ab}(i,j)$ are the cost functions.

If all the $g_{ab}(i,j)$ are constant whenever $i \notin \overline{N(j)}$ then we can incorporate them inside the neighbor potentials V_C. In general, however, we would like to expand the Mixed Markov Field framework to allow for A to have a neighborhood structure. Certain vertices in A may be permanently linked into groups, called *static* cliques, to distinguish them from the dynamic cliques formed after the values of the address variables have been assigned. Potentials $G_D(i_D)$, where D is a static clique in A will encode long-range geometric constraints on the values of the variables in D, that is, on the mutual locations of the templates in D. We then define a Mixed Markov Field to be a distribution on $I \cup A$ of the following kind:

$$P(x_I; i_A) = \frac{1}{Z} \prod_{C \in \mathfrak{C}} V_C(x_C; i_{\uparrow C}) \prod_{D \in \mathfrak{D}} G_D(i_D). \qquad (1)$$

An Equivalence theorem for this type of Mixed Markov Field reads as follows:

Theorem 4.1 *Suppose $P(x_I; i_A) > 0$. P has a Gibbs characterization as in (1) if and only if*

- *$\forall j \in I$, the ratio $\dfrac{P(x_j, x_{I\setminus\{j\}}; i_A)}{P(x'_j, x_{I\setminus\{j\}}; i_A)}$ is a function only of $x_{N(j)}$ and $i_{\uparrow\overline{N(j)}}$*

- $\forall a \in A$, *the ratio* $\dfrac{P(i_a, i_{A\backslash\{a\}}; x_I)}{P(nil_a, i_{A\backslash\{a\}}; x_I)}$ *is a function only of* $x_{\overline{N(i_a)}}$, $i_{N(a)}$, *and* $i_{\uparrow\overline{N(i_a)}}$.

The proof is analogous to that of Theorem 3.1 and is not included.

The face model can now be expressed in the Mixed Markov Field framework by setting potentials V_C and G_D as follows:

- for each of the seven templates $a \in A$, set an auxiliary variable $W_C(x_C; i_a)$ so that

$$\prod_{C:\ i_a \in C} W_C(x_C; i_a) = \exp(l_a(i_a));$$

- set $V_C(x_C; i_{\uparrow C}) = \prod_{a \in \uparrow C} W_C(x_C; i_a)$ or 1 if $\uparrow C = \emptyset$.
- set $G_{\{a,b\}}(i, j) = \exp(g_{ab}(i, j))$.

Then

$$\log\left(\prod_{C \in \mathfrak{C}} V_C(x_C; i_{\uparrow C}) \cdot \prod_{D \in \mathfrak{D}} G_D(i_D)\right) = \frac{1}{Z_{(x_I)}}\left(\sum_{a \in A} l_a(i_a) + \sum_{<a,b> \in \mathfrak{D}} g_{ab}(i_a, i_b)\right),$$

which shows that this process can be represented in the Mixed Markov Field framework.

5 Conclusion

Constructing priors out of local interactions of address and intensity variables is a promising new approach in statistical modeling. The power and utility of the Mixed Markov Field framework comes from the equivalence of Markov and Gibbs characterizations of Mixed Markov Fields, established in Section 3.

Appendix

Proof of Theorem 3.1: If

$$P(x_I; i_A) = \frac{1}{Z} \prod_{C \in \mathfrak{C}} V_C(x_C; i_{\uparrow C})$$

then

$$\begin{aligned}
\forall j \in I, \quad \frac{P(x_j, x_{I-j}; i_A)}{P(x'_j, x_{I-j}; i_A)} &= \prod_{C:\ j \in C} \frac{V_C(x_j, x_{C-j}; i_{\uparrow C})}{V_C(x'_j, x_{C-j}; i_{\uparrow C})} = f(x_{N(j)}, i_{\uparrow\overline{N(j)}}),\\
\forall a \in A, \quad \frac{P(i_a, i_{A-a}; x_I)}{P(\mathrm{nil}_a, i_{A-a}; x_I)} &= \prod_{C:\ i_a \in C} \frac{V_C(x_C; i_a i_{(\uparrow C)-a})}{V_C(x_C; \quad i_{(\uparrow C)-a})} = f(x_{\overline{N(i_a)}}, i_{\uparrow\overline{N(i_a)}}).
\end{aligned}$$

Turning to the "if" part, let $\mathbf{o} = (o_I; \mathrm{nil}_A)$ be the reference configuration. From the Möbius inversion formula,

$$\begin{aligned}
\frac{P(x_I; i_A)}{P(\mathbf{o})} &\equiv \frac{P(x_I; i_{\uparrow I}, \mathrm{nil}_{A-\uparrow I})}{P(\mathbf{o})} = \prod_{D \subset I} V_D(x_D; i_{\uparrow D}), \quad \text{where} \\
V_D(x_D; i_{\uparrow D}) &= \prod_{E \subset D} \left[\frac{P(x_E o_{I-E}; i_{\uparrow E} \mathrm{nil}_{A-\uparrow E})}{P(\mathbf{o})} \right]^{(-1)^{|D-E|}}
\end{aligned}$$

We will show $V_D = 1$ unless D is a clique. In fact, if $\exists u \neq v \in D$ such that $u \notin N(v)$ then

$$\begin{aligned}
V_D(x_D; i_{\uparrow D}) &= \prod_{E \subset D - \{u,v\}} \left[\frac{P(x_u x_v y_R; i_{\uparrow u} i_{\uparrow v} j_S)}{P(o_u x_v y_R; \mathrm{nil}_{\uparrow u} i_{\uparrow v} j_S)} : \frac{P(x_u o_v y_R; i_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)}{P(o_u o_v y_R; \mathrm{nil}_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)} \right]^{(-1)^{|D-E|}} \\
\text{where} \quad R &= I - \{u,v\}, \quad y_R = (x_E o_{R-E}) \\
S &= A - (\uparrow u \cup \uparrow v), \quad j_S = (i_{\uparrow E} \mathrm{nil}_{S-\uparrow E})
\end{aligned}$$

The first fraction inside square brackets is equal to

$$\begin{aligned}
& \frac{P(x_u x_v y_R; i_{\uparrow u} i_{\uparrow v} j_S)}{P(o_u x_v y_R; i_{\uparrow u} i_{\uparrow v} j_S)} \cdot \frac{P(i_{\uparrow u} i_{\uparrow v} j_S; o_u x_v y_R)}{P(\mathrm{nil}_{\uparrow u} i_{\uparrow v} j_S; o_u x_v y_R)} \\
= & \frac{P(x_u o_v y_R; i_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)}{P(o_u o_v y_R; i_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)} \cdot \frac{P(i_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S; o_u o_v y_R)}{P(\mathrm{nil}_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S; o_u o_v y_R)} \quad \text{due to the Markov properties} \\
= & \frac{P(x_u o_v y_R; i_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)}{P(o_u o_v y_R; \mathrm{nil}_{\uparrow u} \mathrm{nil}_{\uparrow v} j_S)},
\end{aligned}$$

which is the second fraction inside square brackets. Thus

$$P(x_I; i_A) = P(\mathbf{o}) \cdot \prod_{C \in \mathfrak{C}} V_C(x_C; i_{\uparrow C}), \qquad \text{Q.E.D.}$$

References

[1] D. Geman, *Stochastic Model for Boundary Detection*, Image and Vision Computing **5**, 61–65, 1987

[2] S. Geman and D. Geman, *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*, IEEE Trans. PAMI **6**, 721–741, 1984

[3] S. Geman and Chr. Graffigne, *Markov Random Field Models and Their Applications to Computer Vision*, In: Proc. International Congress of Mathematicians (1986), Amer. Math. Soc., 1496–1517, 1987

[4] G. Grimmett, *A Theorem about Markov Fields*, Bull. Lond. Math. Soc., **5**, 81–84, 1973

[5] J. Hammersley and P. Clifford, *Markov Fields on Finite Graphs and Lattices*, Preprint Univ. of Cal., Berkeley, 1968

[6] D. Mumford, *The Statistical Description of Visual Signals*, in “ICIAM 95”, Akademie Verlag, 1996

[7] D. Mumford, *The Mathematical Modeling of Cortical Functioning and Thought*, Proc. Norbert Wiener Centennial Conference, Amer. Math. Soc., 1997

[8] B. Ripley, *Modeling Spatial Patterns* (with discussion), J. R. Statist. Soc. B **39**, 172–212, 1977

[9] B. Ripley and F. Kelly, *Markov Point Processes*, J. Lond. Math. Soc., **15**, 188–192, 1977

[10] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer, 1995

[11] A. Yuille, *Deformable Templates for Face Recognition*, J. Cog. Neuroscience, Vol. 3, Num. 1, 59–70, 1991