Gibbs Random Fields

- A set of random variables F is said to be a Gibbs random field (GRF) on S with respect to $\mathcal{N}$ if and only if its configurations obey a Gibbs distribution
- A Gibbs distribution takes the following form

$$P(f) = Z^{-1} \times e^{-\frac{U(f)}{T}} \qquad Z = \sum_{f \in \mathbb{F}} e^{-\frac{U(f)}{T}}$$

is a normalizing constant called the partition function, T is a constant called the temperature which shall be assumed to be 1 and $U(f)$ is the energy function.

- The energy

$$U(f) = \sum_{c \in C} V_c(f) \tag{1}$$

is a sum of clique potentials $V_c(f)$ over all possible cliques C

- The value of $V_c(f)$ depends on the local configuration on the clique c

Gibbs Random Fields: Property

- The Gaussian distribution is a special member of this Gibbs distribution family
- A GRF is said to be homogeneous if $V_c(f)$ is independent of the relative position of the clique c in S
- It is said to be isotropic if V_c is independent of the orientation of c
- To calculate a Gibbs distribution, it is necessary to evaluate the partition function Z which is the sum over all possible configuration in $\mathbb{F}$
- There are a combinatorial number of elements in $\mathbb{F}$ for a discrete $\mathcal{L}$, the evaluation is prohibitive even for problems of moderate sizes

Gibbs Random Fields: Configuration

- $P(f)$ measures the probability of the occurrence of a particular configuration or pattern, f
- The more probable configurations are those with lower energies
- The temperature T controls the sharpness of the distribution
- When the temperature is high, all configurations tend to be equally distributed
- Near the zero temperature, the distribution concentrates around the global energy minima
- Given T and $U(f)$, a class of patterns can be generated by sampling the configuration space $\mathbb{F}$ according to $P(f)$
- For discrete labeling problem, a clique potential $V_c(f)$ can be specified by a number of parameters
- For continuous labeling problem, $V_c(f)$ is continuous function of f_c

Gibbs Random Fields: Energy

- The energy of a Gibbs distribution is expressed as the sum of several terms, each related to cliques of a certain size, i.e.,

$$U(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i,i'\} \in C_2} V_2(f_i, f_{i'}) + \sum_{\{i,i',i''\} \in C_3} V_2(f_i, f_{i'}, f_{i''}) + \dots$$

- An important special case is when only cliques of size up to two are considered. In this case, the energy can also be written as

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'})$$

- $\{i, i'\}$ and $\{i', i\}$ are two distinct cliques in C_2 because the sites in a clique are ordered

GRF: Equivalence

- The conditional probability can be written as

$$P(f_i|f_{\mathcal{N}_i}) = \frac{e^{-\left[V_1(f_i)+\sum_{i'\in\mathcal{N}_i} V_2(f_i,f_{i'})\right]}}{\sum_{f_i\in\mathcal{C}} e^{-\left[V_1(f_i)+\sum_{i'\in\mathcal{N}_i} V_2(f_i,f_{i'})\right]}}$$

- Equivalence between Markov Random Field and Gibbs Random Field
- An MRF is characterized by its local property (the Markovianity)
- A GRF is characterized by its global property (the Gibbs distribution)
- The Hammersley-Clifford theorem establishes the equivalence of these two types of properties
- The theorem states that F is an MRF on S with respect to $\mathcal{N}$ if and only if F is a GRF on S with respect to $\mathcal{N}$.

Markov-Gibbs Equivalence

- This theorem provides a simple way of specifying the joint probability
- *One can specify the joint probability $P(F = f)$ by specifying the clique potential functions $V_c(f)$ and choose appropriate potential functions for desired system behavior. It encodes the a priori knowledge or preference about interactions between labels*
- *The MRF-Gibbs distribution has been used in solving optimization problems. In optimization problems, an objective function is in the form of an energy function and is to be minimized*
- *As the quantitative cost measure, an energy function defines the minimal solution as its minimum, usually a global one*
- *Formulate an energy function so that the correct solution is embedded as the minimum*

MRF Image Modeling: Example

- image modeled as a random field $Y = \{Y_s : s \in L\}$
- $L = \{(i,j)|0 \leq i \leq N_1 - 1, 0 \leq j \leq N_2 - 1\}$ index set, a set of site indices on 2-D discrete $N_1 \times N_2$ rectangular integer lattice
- for each lattice point or pixel $s = (i,j) \in S$ Y_s is a real-valued random variable
- random field Y is characterized by a joint probability distribution P_Y which may be characterized by an associated parameter set θ_Y
- random variables Y_s will take on sample values or realizations y_s from a common finite set of integers $\{0, 1, 2, \ldots, L_Y - 1\}$

MRF Image Modeling: Example

- let random field $X = \{X_1, \ldots, X_m\}$ be a Markov Random Field defined on L and m is the total number of the classes
- sites in L are related to each other via a neighborhood system $\Psi = \{N_l, l \in L\}$ N_l is the set of neighbor of site
- site is not a neighbor of itself
- clique is a subset of sites in N_l $c \in N_l$ is a clique of every pair of distinct sites in c are neighbors
- random field X is considered to be an MRF on S if and only if $P(X = x) > 0$ and $P(X_l = x_l | X_r = x_r, r \neq s) = P(X_l = x_l | X_r = x_r, r \in N_l)$
- difficult to determine the above characteristics in practice

MRF Image Modeling: Example

- MRF has the form

$$P(X = x) = \frac{1}{Z} e^{-U(X)/T}$$

- $X = x$ is a realization from $X = \{X_1, \ldots, X_m\}$, i.e., $x = \{x_1, \ldots, x_m\}$ is a set of random field X and

$$U(X) = \sum_{L}^{\text{I}} V_c(x)$$

  is global energy function and it is given by the sum of clique potentials $V_c(x)$, over all possible cliques
- choice of energy function is arbitrary

MRF Image Modeling: Example

- general expression for the energy function

$$U(X) = \sum_{L} V_c(ci) + \sum_{c \in N_i} V_c(ci, cj)$$

- known as Potts model $V_c(ci)$ external field that weighs the relative importance of different classes
- simplified Potts model with no external energy $V_c(ci) = 0$
- local spatial transitions are taken into account and all the classes in the label segmentation, X_{opt}, as near as possible to the real image X^*

$$V_c(ci, cj) = \begin{cases} -\beta & \text{if } ci = cj \\ 0 & \text{otherwise} \end{cases}$$

$$V_c(ci, cj) = \begin{cases} -\frac{\beta\sigma_i^2}{(\sigma_i^2 + (y_{ci} - y_{cj})^2 \times d_{cicj})} & \text{if } ci \neq cj \\ -\beta & \text{if } ci = cj \end{cases}$$

MRF-MAP Segmentation

- y_{ci} and y_{cj} are the pixel intensities of ci and cj, $d_{cicj} = 1$ or $\sqrt{2}$ represents the distance between the two pixels
- β is constant that controls the classification
- decreasing of d_{cicj} and $(y_{ci} - y_{cj})^2$, $V_c(ci, cj)$ decreased to $-\beta$, i.e., ci and cj are right one pixel or their intensities are same
- image Y as rectangular lattice L y_l denotes the intensity of the pixel at l and it correspond to the label x_l in X
- Bayes theorem yields a complete model coupling intensities and labels

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

MRF-MAP Segmentation

- the prior probability of the image $P(Y)$ is independent of the labelling X, using MAP

$$X_{opt} = \max \arg_{X \in L} \{P(Y|X)P(X)\}$$

- assume that image data is obtained by adding an identical independently distributed (i.i.d) Gaussian noise

$$p(y_I|x_I) = \frac{1}{\sqrt{2\pi\sigma_I^2}} \exp\left\{-\frac{(y_I - \mu_{xI})^2}{2\sigma_I^2}\right\}$$

- based on conditional independent assumption of Y the conditional density $P(Y|X) = \prod_L P(y_I|x_I)$

MRF-MAP Segmentation

$$P(Y|X) = \prod_L \left[\frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left\{ -\frac{(y_l - \mu_{xl})^2}{2\sigma_l^2} \right\} \right] \propto \exp\left[-\sum_L \frac{(y_l - \mu_{xl})^2}{2\sigma_l^2} \right]$$

- potential function of the conditional probability can be written as

$$U(Y|X) = \sum_L \frac{(y_l - \mu_{xl})^2}{2\sigma_l^2}$$

- similarly prior density of MRF takes the form of

$$P(X) = \frac{1}{Z} \exp\left[-\sum_L^1 \sum_{c \in N_l} \frac{V_c(ci, cj)}{T} \right]$$

$$P(X|Y) \propto \exp\left(-\sum_L \left[\frac{(y_l - \mu_{xl})^2}{2\sigma^2} + \sum_{c \in N_l} \frac{V_c(ci, cj)}{T} \right] \right)$$

MRF-MAP Segmentation

- image Y is then segmented by finding the field of labels X

$$
\begin{aligned}
X_{opt} &= \min \arg_{x \in L} U(X|Y) \\
&= \min \arg \left(\sum_{L} \left[\frac{(y_I - \mu_I)^2}{2\sigma^2} + \frac{1}{T} \sum_{c \in N_I} V_c(ci, cj) \right] \right)
\end{aligned}
$$

- optimization method: simulated annealing (SA), iterated conditional model (ICM) used to find the solution
- SA is slow but guarantee a global minimum solution
- ICM likely to reach local minima and no guarantee that a global minimum of energy function can be obtained, provides much faster convergence
- ICM iteratively decrease the energy by visiting and updating the pixel

MRF-MAP Segmentation

- for each pixel l, given the observed image Y and current labels of all the pixels in the neighborhood, the label of X_l is replaced with one that can maximize the probability as

$$X_l^{(k+1)} = \arg\max P(X_l^{(k)} | Y, X_r^{(k)}, r = l)$$

- starting from the initial state, keep on running on the procedure above until either the predefined number of iterations is reached or the label of X does not change

Thank You