# 第十一章：基于概率图模型的图像分析

中国科学技术大学
电子工程与信息科学系

# 图像建模

- 图像处理和分析：观测数据 $\xrightarrow{\text{推测}}$ 未知数据
- 不确定性
  - 观测不确定性
  - 预测不确定性
- 概率模型
  - 链式法则： $P(x_1, \cdots, x_N) = P(x_1)\prod_{i=2}^{N} P(x_i | x_1, \cdots, x_{i-1})$
  - 条件独立性假设
- 概率图模型
  - 概率模型
  - 引入图作为表示工具

# 图像建模

- 概率无向图模型（马尔可夫随机场）
- 条件随机场

# 概率无向图模型（马尔可夫随机场）

- 模型定义
- 概率无向图模型因子分解
- 例子：图像去噪

# 概率无向图模型

## 模型定义

- 图（graph）是由结点（node）和边（edge）组成的集合
  - ✓ 结点：$v$，结点的集合：$V$
  - ✓ 边：$e$，边的集合记：$E$
  - ✓ 图：$G = (V, E)$
  - ✓ 无向图是指边没有方向的图
- 概率图模型
  - ✓ 结点$v$，随机变量$Y_v$
  - ✓ 边$e$：随机变量之间的概率依赖关系
  - ✓ 用无向图$G = (V, E)$表示联合概率分布$P(Y)$
- 无向图表示的随机变量之间存在的性质
  - ✓ 成对马尔可夫性（pairwise Markov property）
  - ✓ 局部马尔可夫性（local Markov property）
  - ✓ 全局马尔可夫性（global Markov property）

# 模型定义

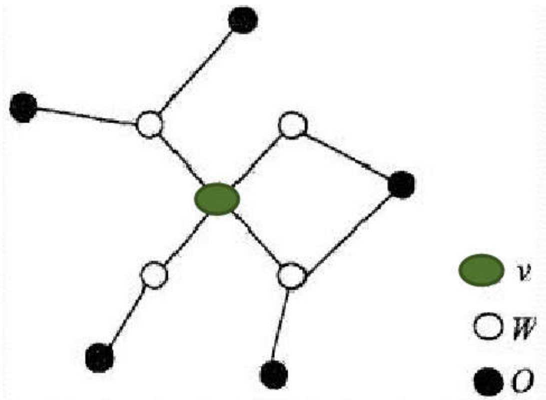
- 成对马尔可夫性（Pairwise Markov property）
  - 设$u$和$v$是无向图$G$中任意两个没有边连接的结点，结点$u$和$v$分别对应随机变量$Y_u$和$Y_v$
  - 其他所有结点为$O$，对应的随机变量组是$Y_O$
  - 定义：给定随机变量组$Y_O$的条件下，随机变量$Y_u$和$Y_v$是条件独立的

$$P(Y_u, Y_v | Y_O) = P(Y_u | Y_O) P(Y_v | Y_O)$$

# 模型定义

- 局部马尔可夫性（Local Markov property）
  - $v \in V$是任意结点
  - $W$是与$v$有边相连的所有结点
  - $O$是其他所有结点
  - 对应的随机变量$Y$
    - ✓ $v$：$Y_v$，$W$：$Y_W$，$O$：$Y_O$
  - 定义：给定随机变量组$Y_W$的条件下随机变量$Y_v$与随机变量$Y_O$是独立的

$$P(Y_v, Y_O | Y_W) = P(Y_v | Y_W) P(Y_O | Y_W)$$




  - 在$P(Y_O | Y_W) > 0$时，等价于

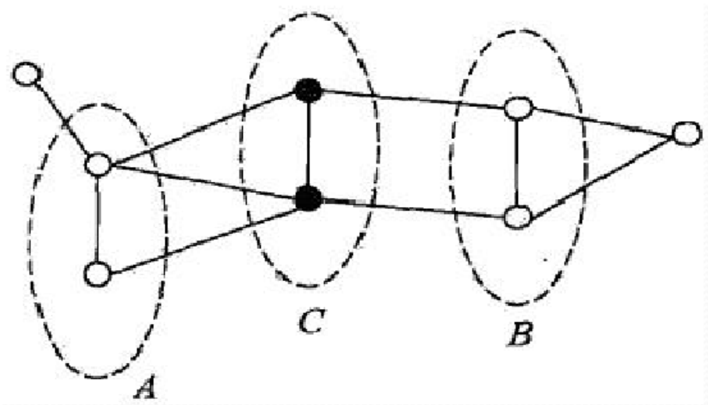
$$P(Y_v | Y_W, Y_O) = P(Y_v | Y_W)$$

  - 马尔可夫毯：与一个结点相邻的结点构成的集合

# 模型定义

- 全局马尔可夫性（Global Markov property）
  - 结点集合$A$，$B$是在无向图$G$中被结点集合$C$分开的任意结点集合
  - $A$，$B$，$C$对应的随机变量组分别为$Y_A$，$Y_B$，$Y_C$
  - 定义：给定随机变量组$Y_C$的条件下随机变量组$Y_A$与随机变量组$Y_B$是条件独立的

$$P(Y_A, Y_B | Y_C) = P(Y_A|Y_C)P(Y_B|Y_C)$$

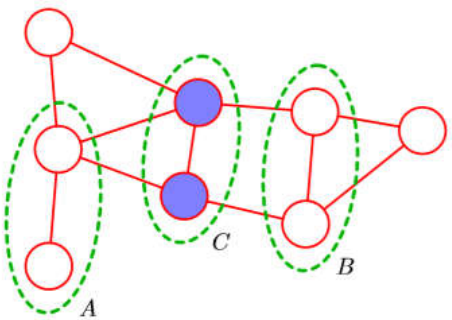


# 模型定义

## □ 条件独立性质

在无向图中，有三个结点集合$A$、$B$、$C$

- 考虑连接集合$A$的结点和集合$B$的结点的所有可能路径
- 如果所有这些路径都通过了集合$C$中的一个或多个结点，那么所有这样的路径都被阻隔 → 条件独立
- 把集合$C$中的结点以及与这些结点相连的链接全部删除，考察是否存在一条从$A$中任意结点到$B$中任意结点的路径，如果没有，那么条件独立性质成立




$C$为条件，$A$与$B$条件独立

# 模型定义

- 概率无向图模型
  - 设有联合概率分布 $P(Y)$，由无向图 $G=(V,E)$ 表示，在图 $G$ 中，结点表示随机变量，边表示随机变量之间的依赖关系
  - 如果联合概率分布$P(Y)$满足成对、局部或全局马尔可夫性，就称此联合概率分布为概率无向图模型（probabilistic undirected graphical model），或马尔可夫随机场（Markov random field）
  - 问题关键
    - ✓ 求解联合概率，引申为对联合概率进行因子分解

# 概率无向图模型（马尔可夫随机场）

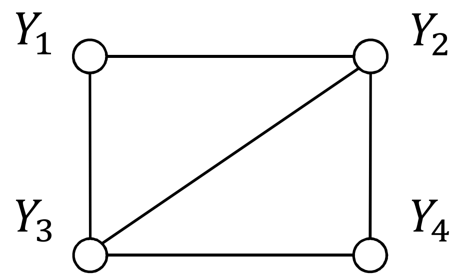
- 模型定义
- 概率无向图模型因子分解
- 例子：图像去噪

# 概率无向图模型的因子分解

## 团与最大团

- 无向图 $G$ 中任何两个结点均有边连接的结点子集称为团（clique）
- 若$C$是无向图$G$的一个团，并且不能再加进任何一个$G$的结点使其成为一个更大的团，则称此$C$为最大团（maximal clique）
- 联合概率分布分解的因子定义为团块中变量的函数
- 定义在最大团上的函数包含最大团的子集

- 2个结点的团：
  - ✓ $\{Y_1, Y_2\}$，$\{Y_2, Y_3\}$，$\{Y_3, Y_4\}$，$\{Y_4, Y_2\}$，$\{Y_1, Y_3\}$
- 3个结点的团
  - ✓ $\{Y_1, Y_2, Y_3\}$，$\{Y_2, Y_3, Y_4\}$
- 最大团
  - ✓ $\{Y_1, Y_2, Y_3\}$，$\{Y_2, Y_3, Y_4\}$



# 概率无向图模型的因子分解

- 概率无向图模型的因子分解（Factorization）
  - 将概率无向图模型的联合概率分布表示成最大团上的随机变量的函数的乘积形式的操作
- 给定概率无向图模型
  - 设其无向图为$G$，$C$为$G$上的最大团，$Y_C$表示$C$对应的随机变量
  - 概率无向图模型的联合概率分布$P(Y)$可写作图中所有最大团$C$上的函数$\psi_C(Y_C)$的乘积形式

$$P(Y) = \frac{1}{Z}\prod_C \psi_C(Y_C)$$

联合概率分布

$$Z = \sum_Y \prod_C \psi_C(Y_C)$$

规范化因子$Z$（划分函数）

✓ 若模型中有$M$个离散点，每个结点有$K$个状态，涉及到$K^M$个状态求和

✓ 最坏情况下，计算量是模型大小的指数形式

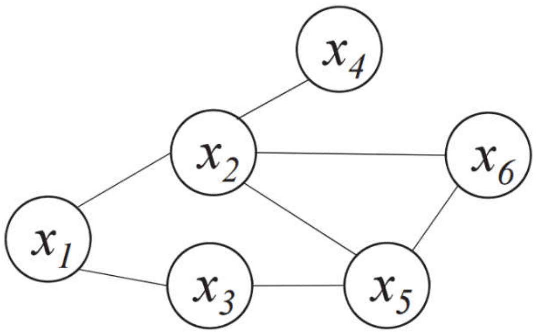
✓ 对于参数学习来说，划分函数是必要的；对于局部条件概率分布的计算，划分函数是不需要的

# 概率无向图模型的因子分解

- 联合概率分布定义
  - 联合概率分布可以使用最大团定义
  - 假设所有最大团构成的集合为$C^*$

$$P(\mathrm{x}) = \frac{1}{Z}\prod_{C\in C^*}\psi_C(Y_C)$$

- 图模型




- 联合概率分布

$$P(\mathbf{x}) = \frac{1}{Z}\psi_{12}(x_1,x_2)\psi_{13}(x_1,x_3)\psi_{24}(x_2,x_4)\psi_{35}(x_3,x_5)\psi_{256}(x_2,x_5,x_6)$$

# 概率无向图模型的因子分解

- 势函数

$$\psi_C(Y_C) = \exp\{-E(Y_C)\}$$

  - ✓ $E(Y_C)$被称为能量函数（energy function）
  - ✓ 指数表示被称为玻尔兹曼分布（Boltzmann distribution）
  - ✓ 联合概率分布被定义为势函数的乘积
  - ✓ 总能量可以通过将每个最大团的能量相加的方法得到

- Hammersley-Clifford定理
  - 概率无向图模型的联合概率分布$P(Y)$可以表示为如下形式

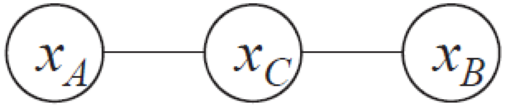
$$P(Y) = \frac{1}{Z}\prod_C \psi_C(Y_C)$$

$C$是无向图的最大团
$Y_C$是$C$的结点对应的随机变量
$\psi_C(Y_C)$是$C$上定义的严格正函数

$$Z = \sum_Y \prod_C \psi_C(Y_C)$$

# 概率无向图模型的因子分解

- 势函数$\psi_C(Y_C)$的作用是定量刻画变量集$Y_C$中变量的相关关系，应为非负函数，且在所偏好的变量取值上有较大的函数值




- 上图中，假定变量均为二值变量，定义势函数

$$\psi_{AC}(x_A, x_C) = \begin{cases} 1.5, & \text{if } x_A = x_C; \\ 0.1, & \text{otherwise} , \end{cases}$$

$$\psi_{BC}(x_B, x_C) = \begin{cases} 0.2, & \text{if } x_B = x_C; \\ 1.3, & \text{otherwise} , \end{cases}$$

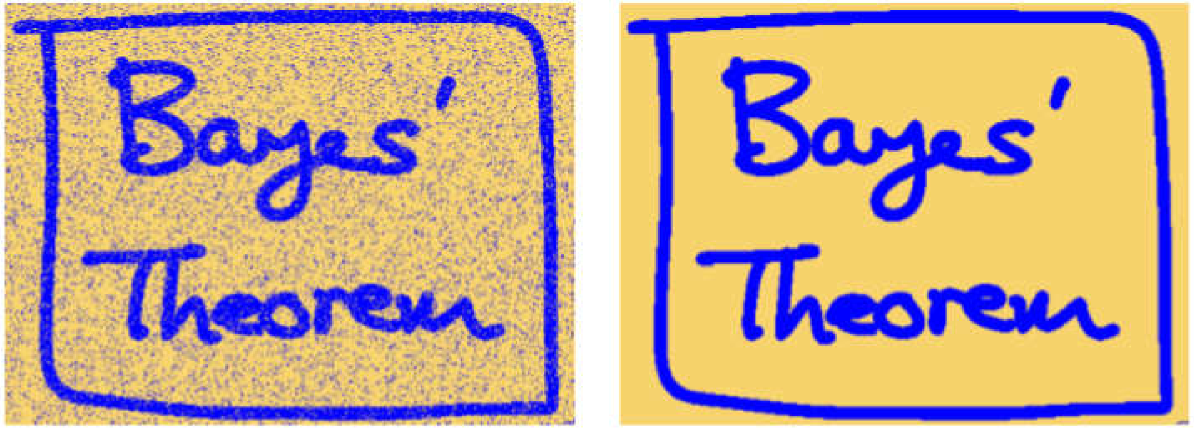
- 说明模型偏好$x_A$与$x_C$有相同的取值，$x_B$与$x_C$有不同的取值；换言之，$x_A$与$x_C$正相关，$x_B$与$x_C$负相关。所以令$x_A$与$x_C$相同且$x_B$与$x_C$不同的变量值指派将有较高的联合概率

# 概率无向图模型（马尔可夫随机场）

- 模型定义
- 概率无向图模型因子分解
- 例子：图像去噪

# 马尔可夫随机场

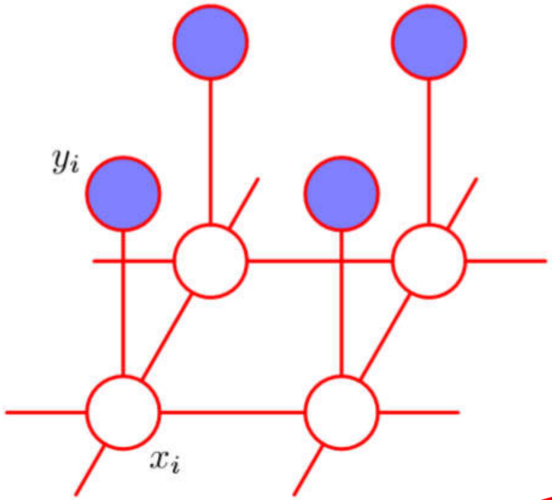
## □ 图像去噪




- 带噪图像：$y_i \in \{-1,+1\}$，$i=1,\ldots,D$覆盖所有像素
  - ✓ 由无噪声图像，以一个较小概率随机翻转像素值符号得到
- 未知的无噪声图像：$x_i \in \{-1,+1\}$
- 目标：给定带有噪声的图像，恢复原始的无噪声图像

# 马尔可夫随机场

## □ 图像去噪




能量函数

两种类型的团

$$E(x,y) = h\sum_i x_i - \beta\sum_{\{i,j\}} x_i x_j - \eta\sum_i x_i y_i$$

- $-\beta x_i x_j$
  - ✓ $x_i$和$x_j$符号相同时，能量函数会给出一个较低的能量（即较高的概率）
  - ✓ $x_i$和$x_j$符号相反时，能量函数会给出一个较高的能量（即较低的概率）
- $-\eta x_i y_i$
  - ✓ $x_i$和$y_i$符号相同时，能量函数会给出一个较低的能量
  - ✓ $x_i$和$y_i$符号相反时，能量函数会给出一个较高的能量

# 马尔可夫随机场

- 图像去噪

  联合分布

$$p(x,y)=\frac{1}{Z}\exp\{-E(x,y)\}$$

  固定噪声图像的像素给出的观测值$y$的元素，噪声图像隐式地定义了一个无噪声图像上的条件概率分布$p(x|y)$
  使用迭代条件模式（ICM）算法，求解获得无噪图像$x$

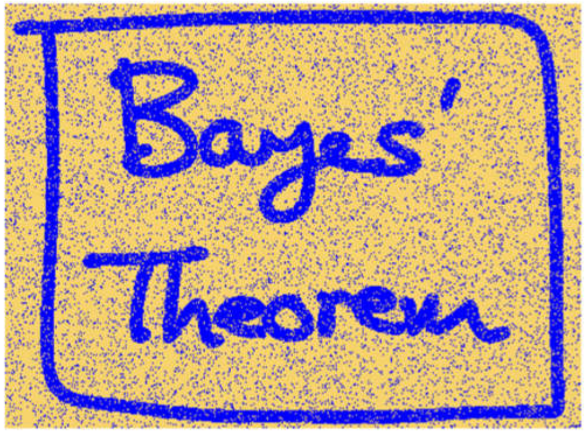
  - 对所有像素，初始化变量$x_i=y_i$
  - 每次取一个$x_j$结点，计算两个可能状态$x_j=+1$和$x_j=-1$的总能量，保持其他所有结点变量固定
  - 将$x_j$设置为能量较低的状态
  - 对其他结点重复更新过程，直到满足某个合适的停止条件

# 马尔可夫随机场

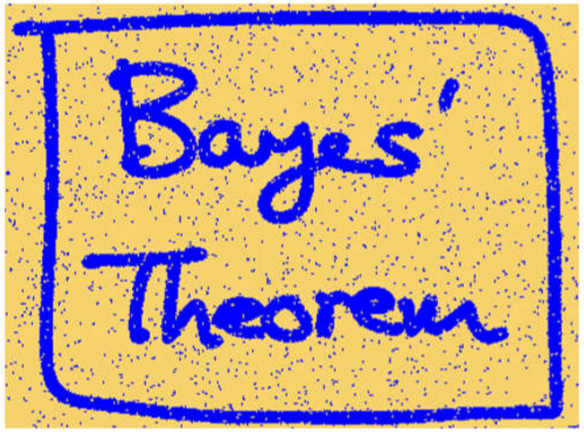
- 图像去噪

原始二值图像

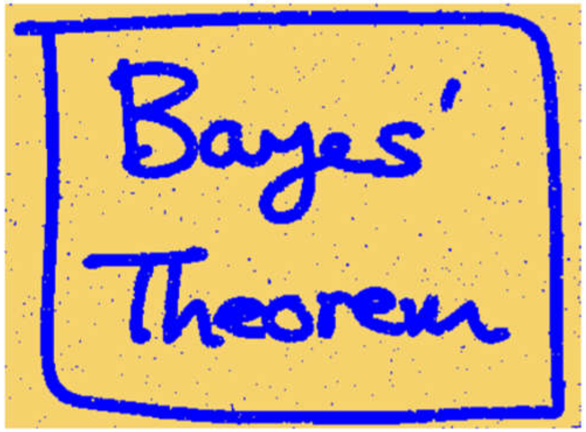

带噪图像




迭代条件模型
恢复的结果




图割算法恢
复的结果



# 生成式图像建模

- 概率无向图模型（马尔可夫随机场）
- 条件随机场

# 条件随机场

- 条件随机场定义与形式
- 条件随机场概率计算问题
- 条件随机场预测算法
- 例子：图像语义分割

# 条件随机场定义

- 条件随机场
  - 给定随机变量$X$条件下，随机变量$Y$的马尔可夫随机场
- 条件随机场的定义
  - 设$X$与$Y$是随机变量，$P(Y|X)$是在给定$X$的条件下$Y$的条件概率分布，若随机变量$Y$构成一个由无向图$G=(V,E)$表示的马尔可夫随机场，即满足马尔可夫性

$$P(Y_v|X, Y_w, w \neq v) = P(Y_v|X, Y_w, w \sim v)$$

对任意结点$v$成立，则称概率分布$P(Y|X)$为条件随机场，式中$w \sim v$表示在图$G=(V,E)$中与结点$v$有边连接的所有结点$w$，$w \neq v$表示结点$v$以外的所有结点。

# 条件随机场定义

- 定义在线性链上的特殊的条件随机场
  - 线性链条件随机场（linear chain conditional random field）
  - 线性链条件随机场可以用于标注等问题
  - 在条件概率模型$P(Y|X)$中，$Y$是输出变量，表示标记序列，$X$是输入变量，表述需要标注的观测序列，也把标记序列称为状态序列

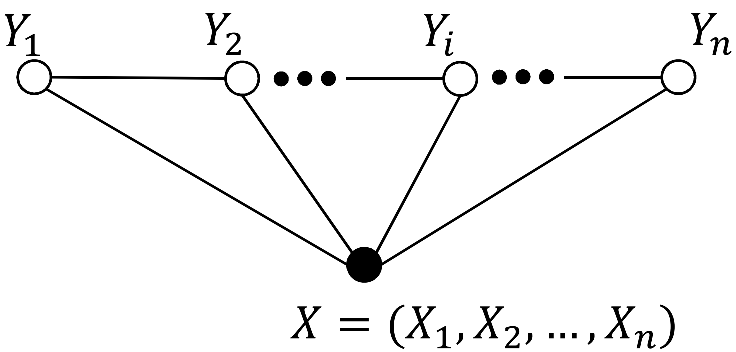
# 条件随机场定义

- 线性链情况

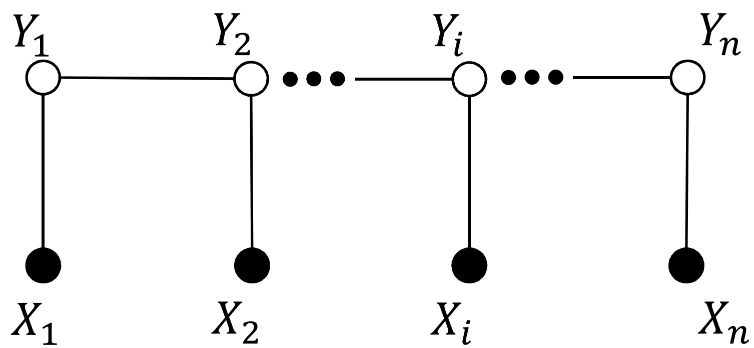
$$G = (V = \{1,2,\dots,n\}, E = \{(i,i+1)\}), i = 1,2,\dots,n-1$$

$$X = (X_1, X_2, \dots, X_n), \qquad Y = (Y_1, Y_2, \dots, Y_n)$$

- 最大团是相邻两个结点的集合，线性链条件随机场




线性链条件随机场




$X$和$Y$有相同的图结构的线性链条件随机场

# 条件随机场定义

- 线性链条件随机场
  - 设$X=(X_1,X_2,\dots,X_n)$，$Y=(Y_1,Y_2,\dots,Y_n)$均为线性链表示的随机变量序列，若在给定随机变量序列$X$的条件下，随机变量序列$Y$的条件概率分布$P(Y|X)$构成条件随机场，即满足马尔可夫性

$$P(Y_i|X,Y_1,\dots,Y_{i-1},Y_{i+1},\dots,Y_n)=P(Y_i|X,Y_{i-1},Y_{i+1})$$

$$i=1,2,\dots,n$$（在$i=1$和$n$时只考虑单边）

  - 称$P(Y|X)$为线性链条件随机场
  - 在标注问题中，$X$表示输入观测序列，$Y$表示对应的输出标记序列或状态序列

# 条件随机场参数化形式

- 定理：线性链条件随机场的参数化形式
  - 设$P(Y|X)$为线性链条件随机场，则在随机变量$X$取值为$x$的条件下，随机变量$Y$取值为$y$的条件概率具有如下形式：

$$P(y|x)=\frac{1}{Z(x)}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1},y_i,x,i)+\sum_{i,l}\mu_l s_l(y_i,x,i)\right)$$

$$Z(x)=\sum_{y}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1},y_i,x,i)+\sum_{i,l}\mu_l s_l(y_i,x,i)\right)$$

- ✓ $Z(x)$：规范化因子
- ✓ $t_k$是定义在边上的特征函数，转移特征，依赖于前一个和当前位置
- ✓ $s_l$是定义在结点上的特征函数，状态特征，依赖于当前位置
- ✓ $\lambda_k$和$\mu_l$是$t_k$和$s_l$对应的权值

# 条件随机场的简化形式

- 条件随机场的简化形式

  条件随机场中同一特征在各个位置都有定义，可以对同一个特征在各个位置求和，将局部特征函数转化为一个全局特征函数，这样就可以将条件随机场写成权值向量和特征向量的内积形式，即条件随机场的简化形式

简化步骤：

- 首先将转移特征和状态特征及其权值用统一的符号表示，设有$K_1$个转移特征，$K_2$个状态特征，$K = K_1 + K_2$，记

$$f_k(y_{i-1}, y_i, x, i) = \begin{cases} t_k(y_{i-1}, y_i, x, i), & k = 1,2,\dots,K_1 \\ s_l(y_i, x, i), & k = K_1 + l;\ l = 1,2,\dots,K_2 \end{cases}$$

# 条件随机场的简化形式

- 对转移特征与状态特征在各个位置$i$求和

$$f_k(y,x)=\sum_{i=1}^{n} f_k(y_{i-1},y_i,x,i), \qquad k=1,2,\dots,K$$

- 权值：

$$w_k=\begin{cases}\lambda_k, & k=1,2,\dots,K_1 \\ \mu_l, & k=K_1+l;\, l=1,2,\dots,K_2\end{cases}$$

- 条件随机场可表示为

$$P(y|x)=\frac{1}{Z(x)}\exp\sum_{k=1}^{K} w_k f_k(y,x) \qquad Z(x)=\sum_{y}\exp\sum_{k=1}^{K} w_k f_k(y,x)$$

# 条件随机场的简化形式

- $w$表示权值向量

$$w = (w_1, w_2, \dots, w_K)^T$$

- 以$F(y,x)$表示全局特征向量，即

$$F(y,x) = \left(f_1(y,x), f_2(y,x), \dots, f_K(y,x)\right)^T$$

- 条件随机场内积形式

$$P_w(y|x) = \frac{\exp(w \cdot F(y,x))}{Z_w(x)}$$

$$Z_w(x) = \sum_y \exp(w \cdot F(y,x))$$

# 条件随机场的矩阵形式

- 线性链条件随机场，引进特殊的起点和终点状态标记 $Y_0 = start$，$Y_{n+1} = stop$，这时$P_w(y|x)$可以通过矩阵形式表示
- 对观测序列$x$的每一个位置$i = 1,2,\ldots,n+1$，定义一个$m$阶矩阵（$m$是标记$Y_i$取值的个数）

$$M_i(x) = [M_i(y_{i-1}, y_i|x)]$$

$$M_i(y_{i-1}, y_i) = \exp(W_i(y_{i-1}, y_i|x))$$

$$W_i(y_{i-1}, y_i|x) = \sum_{i=1}^{K} w_k f_k(y_{i-1}, y_i, x, i)$$

# 条件随机场的矩阵形式

- 线性链条件随机场

$$P_w(y|x) = \frac{1}{Z(x)} \exp\left(\sum_i \sum_k w_k f_k(y_{i-1}, y_i, x, i)\right)$$

$M_1(y_0, y_1|x)$ $M_2(y_1, y_2|x)$

$$= \frac{1}{Z(x)} \exp\left(\sum_k w_k f_k(y_0, y_1, x, 1)\right) \exp\left(\sum_k w_k f_k(y_1, y_2, x, 2)\right) \cdots$$

$$\exp\left(\sum_k w_k f_k(y_n, y_{n+1}, x, n+1)\right)$$

$M_{n+1}(y_n, y_{n+1}|x)$

# 条件随机场的矩阵形式

- 给定观测序列$x$，标记序列$y$的非规范化概率可以通过$n+1$个适当的矩阵元素的乘积表示：

$$\prod_{i=1}^{n+1} M_i(y_{i-1}, y_i|x)$$

- 条件概率$P_w(y|x)$：

$$P_w(y|x) = \frac{1}{Z_w(x)} \prod_{i=1}^{n+1} M_i(y_{i-1}, y_i|x)$$

- $Z_w(x)$为规范化因子，是$n+1$个矩阵的乘积

$$Z_w(x) = [M_1(x)M_2(x) \cdots M_{n+1}(x)]_{\text{start,stop}}$$

# 条件随机场的矩阵形式

- 例 1：给定一个线性链条件随机场，观测序列 $x$，状态序列 $y$，$i=1,2,3, n=3$，标记 $y_i \in \{1,2\}$，假设 $y_0 = \text{start} = 1, y_4 = \text{stop} = 1$，各个位置的随机矩阵$M_1(x), M_2(x), M_3(x), M_4(x)$分别是

$$M_1(x) = \begin{bmatrix} a_{01} & a_{02} \\ 0 & 0 \end{bmatrix},\ M_2(x) = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

$$M_3(x) = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},\ M_4(x) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},$$

试求状态序列 $y$ 以 start 为起点 stop 为终点所有路径的非规范化概率及规范化因子。

# 条件随机场

- □ 条件随机场定义与形式
- □ 条件随机场概率计算问题
- □ 条件随机场预测算法
- □ 例子：图像语义分割

# 条件随机场的概率计算问题

## □ 概率计算问题

- 给定条件随机场$P(Y|X)$，输入序列$x$和输出序列$y$
- 计算条件概率：$P(Y_i = y_i|x)$，$P(Y_{i-1} = y_{i-1}, Y_i = y_i|x)$
- 以及相应的数学期望

$$P(Y_i = y_i|x)$$

$$= \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} \frac{1}{Z_w(x)} \prod_{j=1}^{n+1} M_j(y_{j-1}, y_j|x)$$

$$= \frac{1}{Z_w(x)} \left( \sum_{y_1} \cdots \sum_{y_{i-1}} \prod_{j=1}^{i} M_j(y_{j-1}, y_j|x) \right) \times \left( \sum_{y_{i+1}} \cdots \sum_{y_n} \prod_{j=i+1}^{n+1} M_j(y_{j-1}, y_j|x) \right)$$

$$= \frac{1}{Z_w(x)} \left( \sum_{y_{i-1}} M_i(y_{i-1}, y_i|x) \cdots \sum_{y_2} M_3(y_2, y_3|x) \sum_{y_1} M_1(y_0, y_1|x) M_2(y_1, y_2|x) \right)$$

$$\times \left( \sum_{y_{i+1}} M_{i+1}(y_i, y_{i+1}|x) \cdots \sum_{y_{n-1}} M_{n-1}(y_{n-2}, y_{n-1}|x) \sum_{y_n} M_n(y_{n-1}, y_n|x) M_{n+1}(y_n, y_{n+1}|x) \right)$$

# 条件随机场的概率计算问题

- 前向-后向算法：前向
  - 对每个指标$i = 0,1,\ldots,n+1$，定义前向向量$\alpha_i(x)$

$$\alpha_0(y_0|x) = \begin{cases} 1, & y_0 = start \\ 0, & \text{否则} \end{cases}$$

  - 递推公式

$$\alpha_i^T(x) = \alpha_{i-1}^T(x) M_i(x)$$

  - $\alpha_i(y_i|x)$表示在位置$i$的标记是$y_i$，且到位置$i$的前部分标记序列的非规范化概率，$y_i$可取的值$m$个，所以$\alpha_i(x)$是$m$维列向量

# 条件随机场的概率计算问题

- 前向-后向算法：后向
  - 对每个指标$i = 0,1,\dots,n+1$，定义后向向量$\beta_i(x)$

$$\beta_{n+1}(y_{n+1}|x) = \begin{cases} 1, & y_{n+1} = stop \\ 0, & \text{否则} \end{cases}$$

  - 递推公式

$$\beta_i(x) = M_{i+1}(x)\beta_{i+1}(x)$$

  - $\beta_i(y_i|x)$表示在位置$i$的标记是$y_i$，且从位置$i+1$到$n$的后部分标记序列的非规范化概率

- 由前向-后向向量得

$$Z(x) = \alpha_n^T(x) \cdot \mathbf{1} = \mathbf{1}^T \cdot \beta_1(x)$$

# 条件随机场的概率计算问题

## □ 概率计算

- 按照前向-后向向量的定义，计算标记序列在位置$i$是标记$y_i$的条件概率和在位置$i-1$与$i$是标记$y_{i-1}$和$y_i$的条件概率

$$P(Y_i = y_i|x) = \frac{\alpha_i(y_i|x)\beta_i(y_i|x)}{Z(x)}$$

$$P(Y_{i-1} = y_{i-1}, Y_i = y_i|x) = \frac{\alpha_{i-1}(y_{i-1}|x)M_i(y_{i-1}, y_i)\beta_i(y_i|x)}{Z(x)}$$

$$\boxed{Z(x) = \alpha_n^T(x) \cdot \mathbf{1}}$$

# 条件随机场的概率计算问题

- 期望值的计算
  - 利用前向-后向向量，可以计算特征函数关于联合分布$P(X,Y)$和条件分布$P(Y|X)$的数学期望
  - 特征函数$f_k$关于条件分布$P(Y|X)$的数学期望是：

$$E_{P(Y|X)}[f_k] = \sum_y P(y|x) f_k(y,x)$$

$$= \sum_{i=1}^{n+1} \sum_{y_{i-1} y_i} f_k(y_{i-1}, y_i, x, i) \frac{\alpha_{i-1}(y_{i-1}|x) M_i(y_{i-1}, y_i) \beta_i(y_i|x)}{Z(x)}$$

$$k = 1,2,\dots,K$$

$$\boxed{Z(x) = \alpha_n^T(x) \cdot \mathbf{1}}$$

# 条件随机场的概率计算问题

- 假设经验分布为$\tilde{P}(X)$特征函数$f_k$关于联合分布$P(X,Y)$的数学期望是：

$$E_{P(Y|X)}[f_k] = \sum_{x,y} P(x,y) \sum_{i=1}^{n+1} f_k(y_{i-1}, y_i, x, i)$$

$$= \sum_{x} \tilde{P}(x) \sum_{y} P(y|x) \sum_{i=1}^{n+1} f_k(y_{i-1}, y_i, x, i)$$

$$= \sum_{x} \tilde{P}(x) \sum_{i=1}^{n+1} \sum_{y_{i-1}y_i} f_k(y_{i-1}, y_i, x, i) \frac{\alpha_{i-1}(y_{i-1}|x) M_i(y_{i-1}, y_i) \beta_i(y_i|x)}{Z(x)}$$

$$k = 1,2, \dots, K$$

$$\boxed{Z(x) = \alpha_n^T(x) \cdot \mathbf{1}}$$

# 条件随机场

- 条件随机场定义与形式
- 条件随机场概率计算问题
- 条件随机场预测算法
- 例子：图像语义分割

# 条件随机场的预测算法

- 预测算法：
  - 给定条件随机场$P(Y|X)$和输入序列（观测序列）$x$
  - 求：条件概率最大的输出序列（标记序列）$y^*$

- 维特比算法

$$P_w(y|x) = \frac{\exp(w \cdot F(y,x))}{Z_w(x)} \quad \Rightarrow \quad \begin{aligned} y^* &= \arg\max_y P_w(y|x) \\ &= \arg\max_y \frac{\exp(w \cdot F(y,x))}{Z_w(x)} \\ &= \arg\max_y \exp(w \cdot F(y,x)) \\ &= \arg\max_y w \cdot F(y,x) \end{aligned}$$

非规范化概率最大的最优路径

$$\max_y w \cdot F(y,x)$$

# 条件随机场的预测算法

- 路径表示标记序列

$$w = (w_1, w_2, \ldots, w_K)^T$$

$$F(y, x) = \left(f_1(y, x), f_2(y, x), \ldots, f_K(y, x)\right)^T$$

$$f_k(y, x) = \sum_{i=1}^{n} f_k(y_{i-1}, y_i, x, i)\,, k = 1,2, \ldots, K$$

- 只计算非规范化概率

$$\max_y \sum_{i=1}^{n} w \cdot F_i(y_{i-1}, y_i, x)$$

$$F_i(y_{i-1}, y_i, x) = \left(f_1(y_{i-1}, y_i, x, i), f_2(y_{i-1}, y_i, x, i), \ldots, f_K(y_{i-1}, y_i, x, i)\right)^T$$

# 条件随机场的预测算法

- 维特比算法
  - 首先求出位置1的各个标记$j = 1,2,\ldots,m$的非规范化概率

$$\delta_1(j) = w \cdot F_1(y_0 = start, y_1 = j, x), j = 1,2,\ldots,m$$

  - 由递推公式，求出到位置$i$的各个标记$l = 1,2,\ldots,m$的非规范化概率的最大值，同时记录最大值路径

$$\delta_i(l) = \max_{1 \le j \le m} \delta_{i-1}(j) + w \cdot F_i(y_{i-1} = j, y_i = l, x), j = 1,2,\ldots,m$$

$$\psi_i(l) = \arg\max_{1 \le j \le m} \delta_{i-1}(j) + w \cdot F_i(y_{i-1} = j, y_i = l, x), j = 1,2,\ldots,m$$

# 条件随机场的预测算法

## 维特比算法

- 直到$i=n$时终止，这时求得非规范概率的最大值为

$$\max_y \left(w\cdot F(y,x)\right) = \max_{1\le j\le m} \delta_n(j)$$

- 最优路径的终点

$$y_n^* = \arg\max_{1\le j\le m} \delta_n(j)$$

- 由此最优路径终点返回

$$y_i^* = \psi_{i+1}(y_{i+1}^*), i = n-1, n-2, \ldots, 1$$

- 得最优路径

$$y^* = (y_1^*, y_2^*, .., y_n^*)^T$$

# 条件随机场的预测算法

- 例 2: 假设随机矩阵$M_1(x), M_2(x), M_3(x), M_4(x)$分别是

$$M_1(x) = \begin{bmatrix} 0 & 0 \\ 0.5 & 0.5 \end{bmatrix},\ M_2(x) = \begin{bmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{bmatrix},$$

$$M_3(x) = \begin{bmatrix} 0.5 & 0.5 \\ 0.6 & 0.4 \end{bmatrix},\ M_4(x) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

求以$\mathrm{start} = 2$, $\mathrm{stop} = 2$为终点的所有路径的状态序列$y$中概率最大的序列。

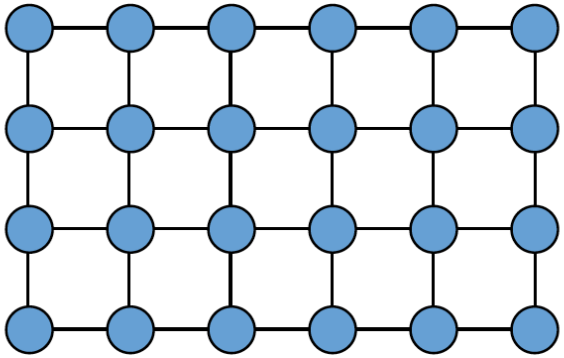
# 条件随机场

- 条件随机场定义与形式
- 条件随机场概率计算问题
- 条件随机场预测算法
- 例子：图像语义分割

# 基于条件随机场的图像分割

- 能量函数

$$E(x) = \sum_i \psi_u(x_i) + \sum_i \sum_{j \in \mathcal{N}_i} \psi_p(x_i, x_j)$$



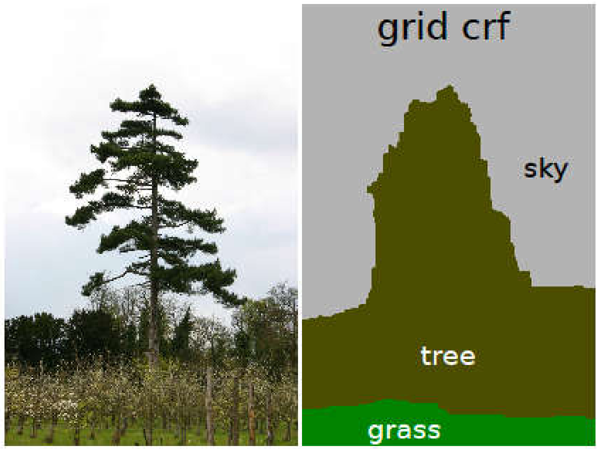
- $\psi_u(x_i)$
  - 来自分类器（像素）
  - 利用颜色、纹理、位置等特征
- $\psi_p(x_i, x_j)$
  - 考虑相邻像素

$$\psi_p(x_i, x_j) = 1_{[x_i \neq x_j]}\left(w^{(1)} \exp\left(-\frac{|\mathrm{I}_i - \mathrm{I}_j|^2}{2\theta_\beta^2}\right) + w^{(2)}\right)$$

# 基于条件随机场的图像分割

- 能量函数

$$E(x) = \sum_{i} \psi_u(x_i) + \sum_{i} \sum_{j \in \mathcal{N}_i} \psi_p(x_i, x_j)$$

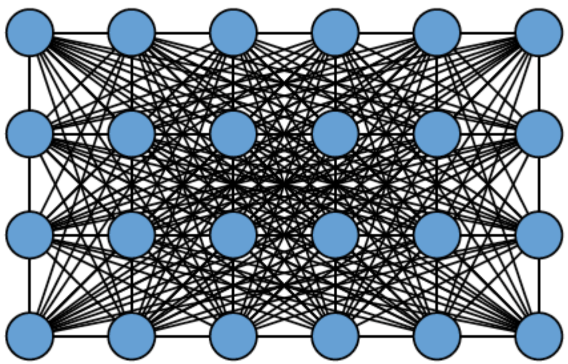



- 推断高效
- 仅考虑了局部关系
- 物体边界过度平滑

# 全连接条件随机场

- 能量函数

$$E(x)=\sum_i \psi_u(x_i)+\sum_i \sum_{j>i} \psi_p(x_i,x_j)$$



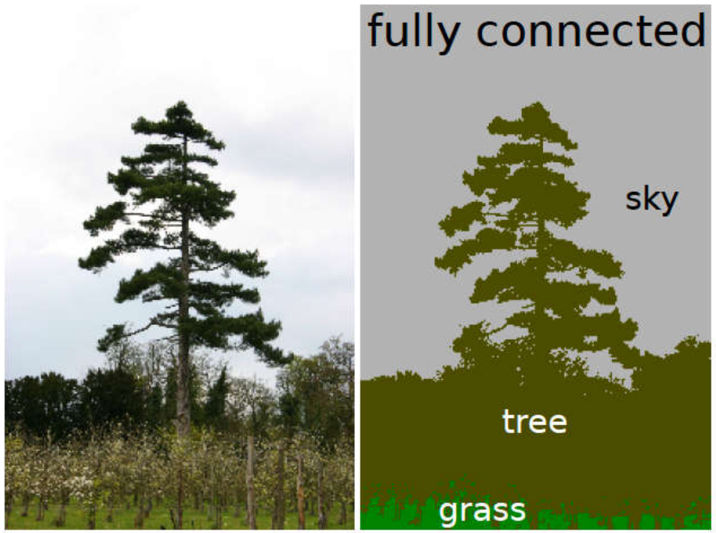
- 每个结点和所有其他结点相连
- 不同的连接具有不同的权重
- 体现远距离相关性

P. Krähenbühl and V. Koltun, Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials, in NIPS 2011.

# 全连接条件随机场

- 能量函数

$$E(x) = \sum_i \psi_u(x_i) + \sum_i \sum_{j>i} \psi_p(x_i, x_j)$$




- 缓解了边界过度平滑问题
- 推断效率低
  - 以像素为结点
  - 结点数和边数巨大

# 全连接条件随机场

$$E(x) = \sum_i \psi_u(x_i) + \sum_i \sum_{j>i} \psi_p(x_i, x_j)$$

- 高斯边势能

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)$$

  - 标记相容性函数 $\mu$
  - 高斯核函数线性组合
  - 任意特征空间

# 全连接条件随机场

$$\psi_p(x_i, x_j) = \mu(x_i, x_j)\left(w^{(1)}\exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}\right) + w^{(2)}\exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}\right)\right)$$

- 标记相容性
  - $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - $\mu(x_i, x_j)$ 从训练数据中学习
- 外观核函数
  - 颜色
- 局部平滑性
  - 抑制像素级的噪声

# 全连接条件随机场

- 推断问题
  - 寻找最可能的标记

$$\hat{x} = \underset{x}{\operatorname{argmax}} P(x) \quad \text{其中} \quad P(x) = \exp(-E(x))$$

- 平均场近似
  - 寻找近似分布$Q(x) = \prod_i Q_i(x_i)$使其在KL散度$D(Q||P)$下与$P(x)$接近
  - $\hat{x}_i \approx \underset{x_i}{\operatorname{argmax}} Q_i(x_i)$

# 全连接条件随机场

- 平均场近似

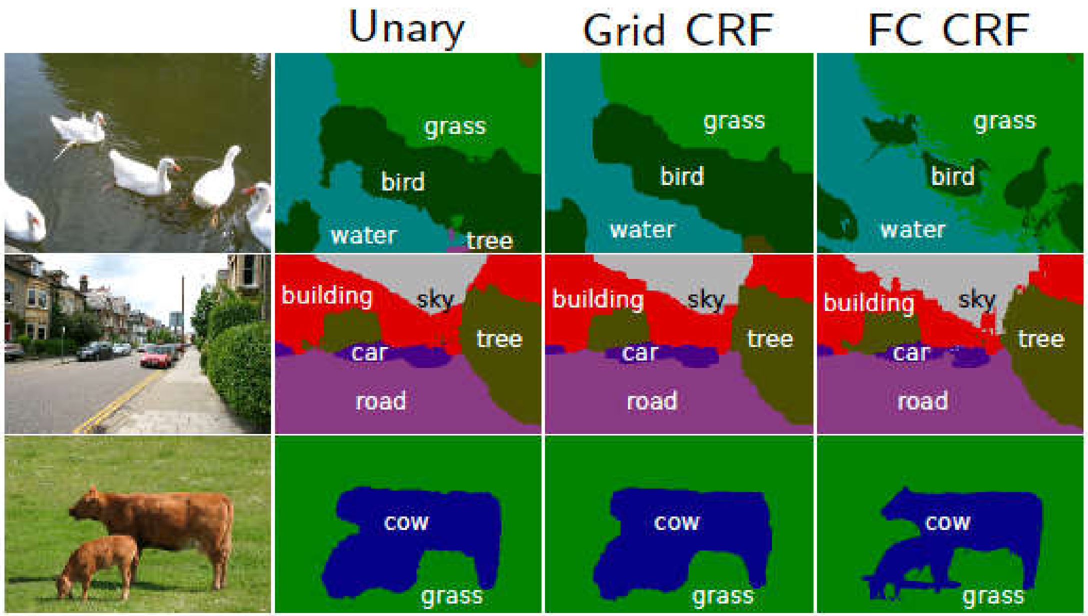
$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{ -\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)\, Q_j(l') \right\}$$

- 初始化$Q_i(x_i) = \frac{1}{Z_i} \exp\{-\psi_u(x_i)\}$
- 迭代直至收敛
  - 消息传递: $\boxed{\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)\, Q_j(l)}$
  - 相容性变换： $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - 局部更新： $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - 归一化： $Q_i(x_i)$

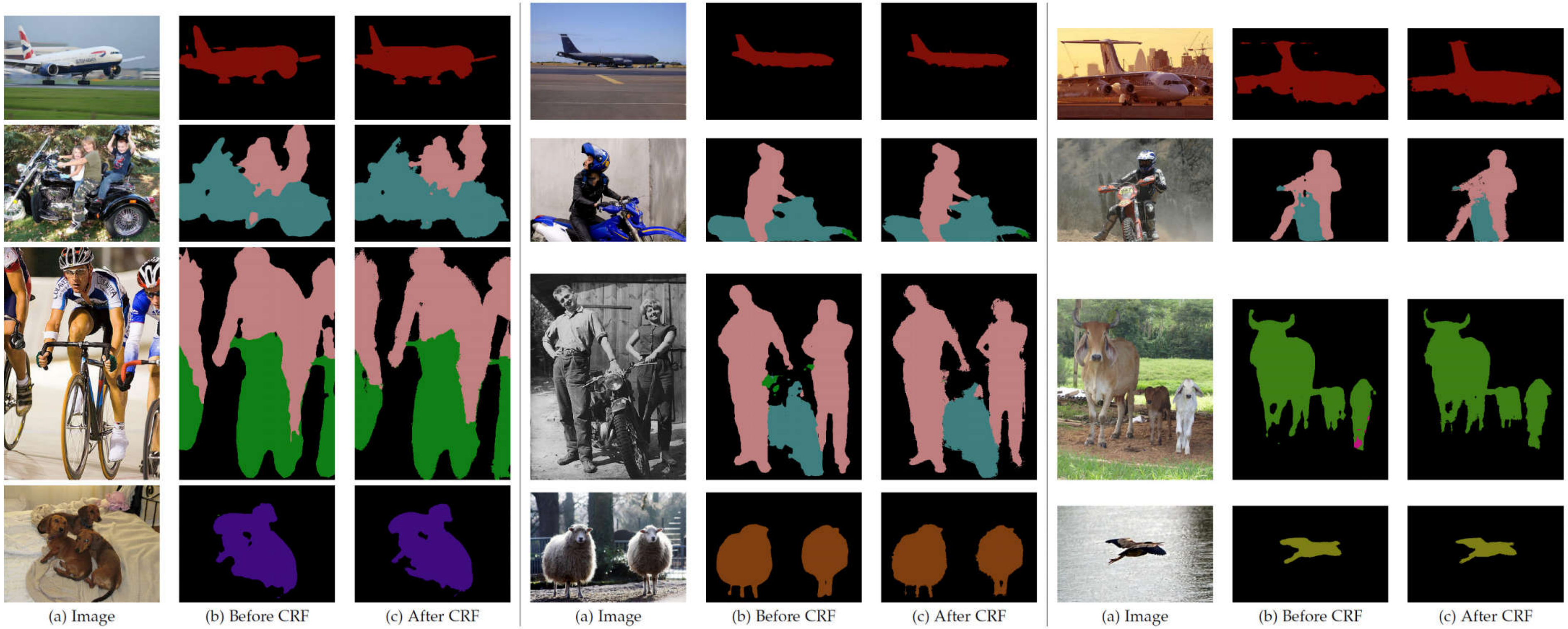
# 全连接条件随机场

- 实验结果
  - MSRC 数据集

| | Time | Global | Avg |
|---|---|---|---|
| Unary | - | 84.0 | 76.6 |
| Grid CRF | 1s | 84.6 | 77.2 |
| **FC CRF** | **0.2s** | **86.0** | **78.3** |



# 全连接条件随机场




| Method | before CRF | after CRF |
|---|---|---|
| LargeFOV | 65.76 | 69.84 |
| ASPP-S | 66.98 | 69.73 |
| ASPP-L | 68.96 | 71.57 |

*L.-C. Chen**, G. Papandreou*, I. Kokkinos, K. Murphy, and A. L. Yuille, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, PAMI 2018.