Arithmetic Encoding of Markov Random Fields

Matthew G. Reyes

Joint work with David L. Neuhoff

University of Michigan

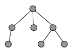
Motivation

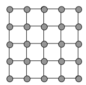
- Markov random fields (MRFs) used to model images, gene networks, more
- MRFs extensively studied, many tools available
- Little work done into compressing MRFs
- Recent (lossy) bilevel image coder based on MRFs

Overview

- Arithmetic Encoding (AC) of acyclic MRF using Belief Propagation (BP)
- For cyclic MRF, AC encode a loop cutset - subset of sites whose removal leaves acyclic graph – then encode remainder using acyclic MRF method
- Encoding of loop cutset done with Local Conditioning (LC), a variant of BP
- New extension of LC to undirected graphs

Background: MRFs





- graph $G=(V,E)$
- alphabet $\mathcal{X}$
- random field $X=\{X_i : i \in V\}$
- image $x=\{x_i : i \in V\}$
- for $U \subset V$, X_U and x_U
- Markov Property

$$\Pr(X_i = x_i | X_{V\setminus i} = x_{V\setminus i}) = \Pr(X_i = x_i | X_{\partial i} = x_{\partial i})$$

- potentials $\Phi_{ij} : \mathcal{X}\times\mathcal{X} \to \mathbb{R}_+$, $\Phi_i : \mathcal{X} \to \mathbb{R}_+$
- Gibbs distribution

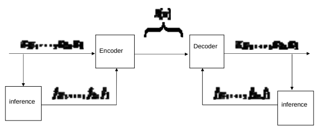
$$p(x) = \frac{1}{Z} \prod_{\{i,j\}\in E} \Phi_{ij}(x_i,x_j) \prod_{i\in V} \Phi_i(x_i)$$

$$= \frac{1}{Z} \Gamma(x)$$

Background: AC

- order sites into scan
- coding distributions $\{f_i\}$

$$f(x) \triangleq \prod_{i=1}^{N} f_i(x_i)$$



- bit string length

$$l(x) \approx -\prod_{i=1}^{N} \log f_i(x_i)$$

$$E_p\, l(X) = \sum_{x\in\mathcal{X}^N} p(x) l(x) \approx H(X) + D(p\|f)$$

- optimal coding

$$p_{i|\cdot}(x_i) \triangleq \Pr(X_i = x_i | X_1^{i-1} = x_1^{i-1})$$

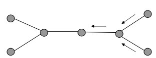
$$f_i = p_{i|\cdot} \implies E_p\, l(X) = H(X)$$

Background: BP

- messages

$$m_{i\to j} = \sum_{x_i} \Phi_i(x_i)\Phi_{ij}(x_i,x_j)$$

$$m_{i\to j}(x_j) = \sum_{x_i} \Phi_i(x_i)\Phi_{ij}(x_i,x_j) \prod_{k\in\partial i\setminus j} m_{k\to i}(x_i)$$



- beliefs

$$Z_A(x_A) = \sum_{x_{V\setminus A}} \Gamma(x_A, x_{V\setminus A})$$

$$p(x_A) \propto Z_A(x_A)$$

$$Z_A(x_A) = \prod_{i\in A} \Phi_i(x_i) \prod_{\{i,j\}\subset A} \Phi_{ij}(x_i,x_j) \prod_{j\in\partial A} m_{j\to i}(x_i)$$

- messages and beliefs are vectors
- complexity $O(|\mathcal{X}|^2)$

AC Encoding an Acyclic MRF

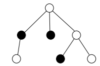
- connected scan – parent already encoded
- in tree, conditioning on parent optimal

$$p_{i|\cdot} = \frac{Z_{\{\pi(i),i\}}(x_{\pi(i)},x_i)}{Z_{\pi(i)}(x_{\pi(i)})}$$

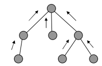
Proposition 0.1 *Let $x = (x_1,\cdots,x_N)$ be an image from an acyclic MRF. If the scan is connected, then*

$$p_{i|\cdot}(x_i) = \frac{\Phi_{\pi(i),i}(x_{\pi(i)},x_i)\Phi_i(x_i) \prod_{j\in\partial i\setminus\pi(i)} m_{j\to i}(x_i)}{\sum_{x_i'} \Phi_{\pi(i),i}(x_{\pi(i)},x_i')\Phi_i(x_i') \prod_{j\in\partial i\setminus\pi(i)} m_{j\to i}(x_i')}.$$

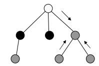
- two-pass algorithm: upward BP, downward AC



1. Realization of binary MRF



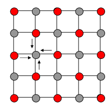
2. Upward BP, marginalization



3. Downward AC, conditioning

AC Encoding a Cyclic MRF

- encode loop cutset x_L
  - scan cutset nodes
  - use Local Conditioning [Diez] to compute $\{f_i : i \in L\}$
- encode (acyclic) remainder $x_{V\setminus L}$ conditioned on
  - x_L fixed, modify self-potentials $\hat{\Phi}_i(x_i) = \Phi(x_i, x_{\partial i})$, $i \in L$
  - run BP, compute $\{m_{i\to j}^{(x_L)}\}$, $Z_i^{(x_L)}$, $i \notin L$



Encoding a Loop Cutset

- Global Conditioning [Pearl]: condition on all possible values of X_L
  - conditional messages and beliefs (large arrays)

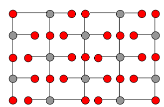
$$m_{i\to j}^{(L)} = \left[m_{i\to j}^{(x_L)} : x_L \in \mathcal{X}_L \right] \qquad Z_i^{(L)} = \left[Z_i^{(x_L)} : x_L \in \mathcal{X}_L \right]$$

  - compute beliefs $Z_i(x_i) = \sum Z_i^{(x_L)}(x_i)$

- unwrapped graph
  - for $l \in L$ multiple copies $\{l^{(a)} : a \subset \partial l\}$
  - modify self-potentials $\hat{\Phi}_{l^{(a)}}(x_l) = \Phi_l^{1/|\partial l|}(x_l)\Phi(x_l, x_a)$

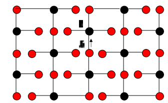
- complexity $O(|L||\mathcal{X}|^{|L|})$

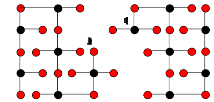
- use LC to reduce complexity



Local Conditiong for undirected graphs:

- consider message from k to i



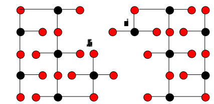


- downstream nodes

Lemma 0.1 [illegible]

$$m_{k\to i}^{(x_L)} = m_{k\to i}^{(x_{L'})}.$$

$L_{d(k)}$ = loop cutset nodes all copies of which downstream



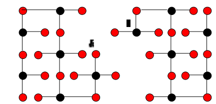
downstream of $M_{k\to i}$

- upstream nodes
  - “summed out” messages

$$m_{k\to i}^{(x_{L_{ki}})}(x_i) \triangleq \sum \sum \Phi_{ki}(x_k,x_i)\Phi_k(x_k) \prod_{j\in\partial k\setminus i} m_{j\to k}^{(\cdot)}(x_k).$$

Proposition 0.2 [illegible]

$L_{u(k)}$ = loop cutset nodes all copies of which upstream



upstream of $M_{k\to i}$

- relevant sets
  - edge $\{k,i\}$ $R_{ki} = L\setminus(L_{u(ki)} \cup L_{d(ki)})$ $L = (L_{u(ki)}, R_{ki}, L_{d(ki)})$
  - node i $R_i = \bigcup_{k\in\partial i} R_{ki}$

Theorem 0.1 *The LC formulas (in matrix notation) are*

$$Z_i^{(R_i)} = Z_i^{(R_i)}(x_i) \circ \prod_{k\in\partial i} m_{k\to i}^{(R_{ki})}$$

and

$$m_{k\to i}^{(R_{ki})} = A_{k\to i} \cdot \left[\Phi_k^{(\cdot)}(x_k) \circ \prod_{j\in\partial k\setminus i} m_{j\to k}^{(R_{jk})} \right] \cdot A_{ki}$$

where [illegible]

Theorem 0.2 [illegible]

$$c(L) = \max_{i\in V} |R_i|$$

Theorem 0.3 [illegible]

$$O(|\mathcal{X}|^{c(L)})$$

Conclusions/comments:

- Introduced problem of AC encoding MRF
- General method for AC encoding acyclic or cyclic MRF
- Extended LC to undirected graphs, give complexity formula

References:

- J. Pearl, “Probabilistic Reasoning in Intelligent Systems”, 1988
- F.J. Diez, “Local Conditioning in Bayesian Networks”, 1996
- M.G. Reyes, D.L. Neuhoff, “Arithmetic Encoding of Markov Random Fields”, ISIT 2009