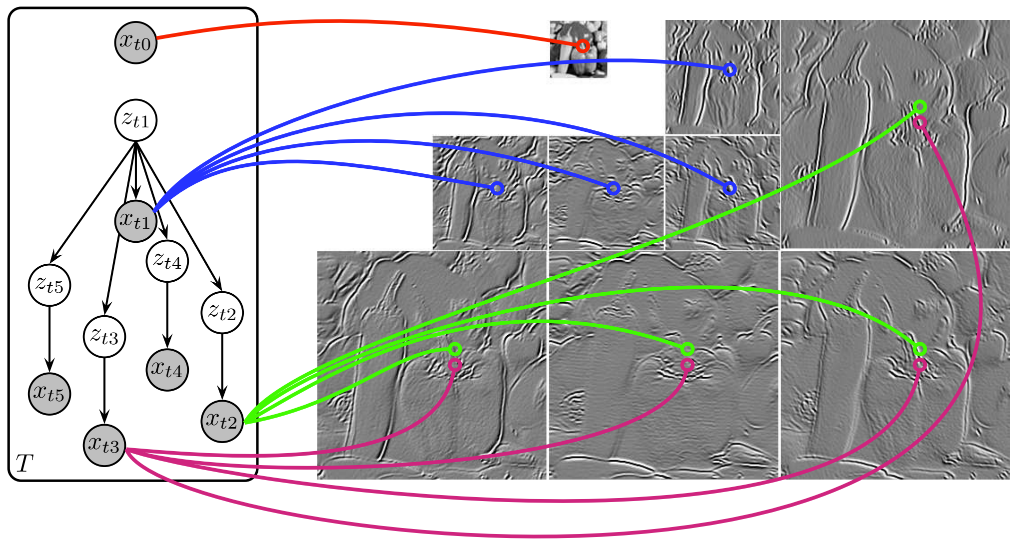
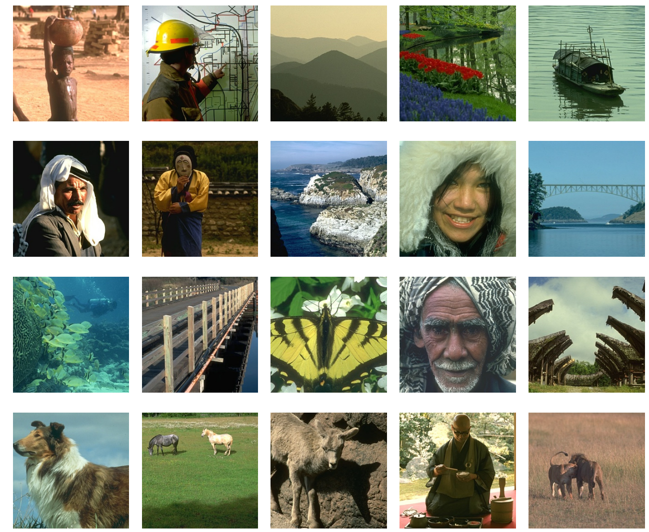
Transfer Denoising with Hierarchical DP Hidden Markov Trees

Jyri Kivinen, Erik Sudderth, and Michael Jordan





- *Capture* *multiscale dependencies* *using a tree of latent variables*

  $z_{ti} \mid z_{\mathrm{Pa(ti)}} \sim \pi_{z_{\mathrm{Pa(ti)}}}$ $x_{ti} \mid z_{ti} \sim \mathcal{N}\left(0, \Lambda_{z_{ti}}\right)$

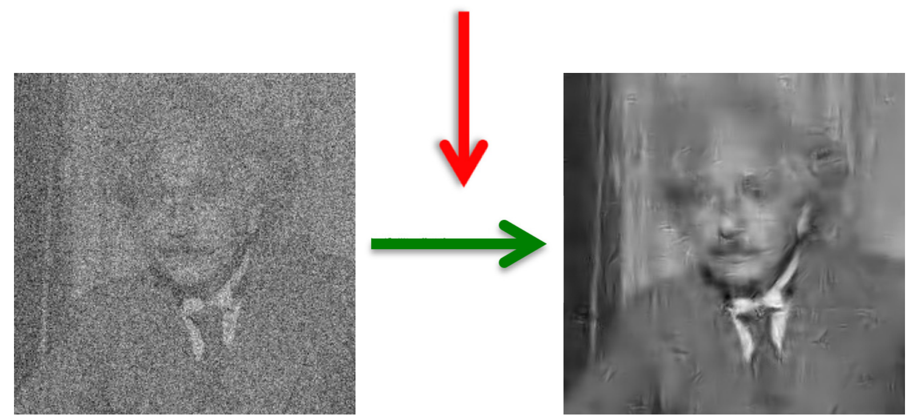
  State assignments   *Observed wavelet coefficients*

- *Automatically* *adapt the number of states* *to the statistics of observed data*

$$p(x_{ti} \mid \beta, \Lambda_1, \Lambda_2, \ldots) = \sum_{k=1}^{\infty} \beta_k \mathcal{N}(x_{ti}; 0, \Lambda_k) \qquad \beta \sim \mathrm{Stick}(\gamma)$$

Marginal distributions of wavelet coefficients

- *Reuse* *multiscale hidden state patterns from* *clean* *images to make robust predictions*



$$\mathbb{E}[x_{ti} \mid \mathbf{w}, \theta^{(s)}] = \sum_{k=1}^{K_s} p(z_{ti} = k \mid \mathbf{w}, \theta^{(s)})\, \mathbb{E}[x_{ti} \mid w_{ti}, \Lambda_k^{(s)}]$$

From belief propagation   *Linear least-squares smoothing*