

Unsupervised learning of image transformations

Roland Memisevic

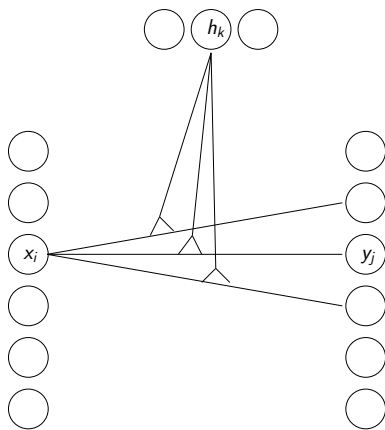
University of Toronto

CIAR summer school 2006

Unsupervised learning of transformations

- ▶ Unsupervised learning: exploit regularities in data.
- ▶ Ex. statistics of natural images...
- ▶ What about *transformations* of data?
- ▶ Ex.: Optical flow, view-point invariance...

Modulator units



$$E = \sum_{ijk} w_{ijk} x_i y_j h_k$$

Modulator units

- ▶ Using the score

$$E = \sum_{ijk} w_{ijk} x_i y_j h_k$$

model the *joint conditional* as

$$p(\mathbf{y}, \mathbf{h} | \mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp(E(\mathbf{y}, \mathbf{h}; \mathbf{x}))$$

with

$$Z(\mathbf{x}) = \sum_{\mathbf{y}, \mathbf{h}} \exp(E(\mathbf{y}, \mathbf{h}; \mathbf{x}))$$

- ▶ From this we get the *conditional marginal*

$$p(\mathbf{y} | \mathbf{x}) = \sum_{\mathbf{h}} p(\mathbf{y}, \mathbf{h} | \mathbf{x})$$

- ▶ (Hinton, Lang 1985), (He, Zemel 2004)

Two views

- ▶ Conditional RBM: Potentials dependent on inputs.
 - ▶ A CRF with hidden variables.
 - ▶ A CRF that *learns* its features.
- ▶ Modulated regression: Functions modulated by hidden units.
 - ▶ Exponential mixture of experts (with weight sharing).
 - ▶ Hidden units provide a *factorial* code for transformations.

Training

- ▶ For training maximize $L = \sum_{\alpha} \log p(y^{\alpha} | x^{\alpha})$
- ▶ Gradient:

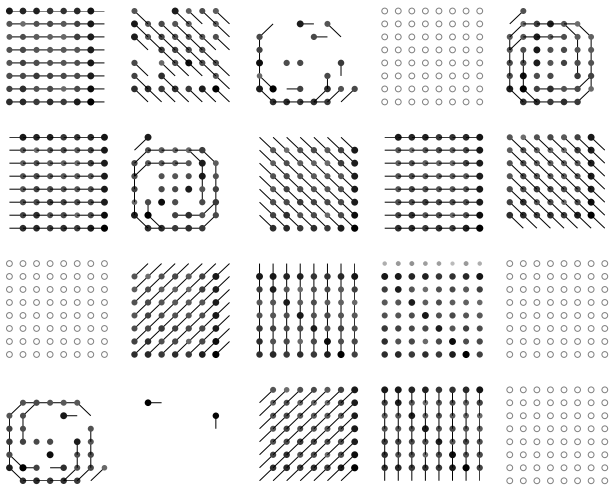
$$\frac{\partial L}{\partial w} = \sum_{\alpha} \left[\sum_{h^{\alpha}} p(h^{\alpha} | y^{\alpha}, x^{\alpha}) \frac{\partial E(h^{\alpha}, y^{\alpha})}{\partial w} - \sum_{h^{\alpha}, y^{\alpha}} p(y^{\alpha}, h^{\alpha} | x^{\alpha}) \frac{\partial E(h^{\alpha}, y^{\alpha})}{\partial w} \right]$$

- ▶ Can use contrastive divergence

Example: Simple transformations

- ▶ Show video1.
- ▶ Flow represented as binary vector.
- ▶ Can use supervised learning, e.g. to classify egomotion.
- ▶ How can we get a dense flowmap?

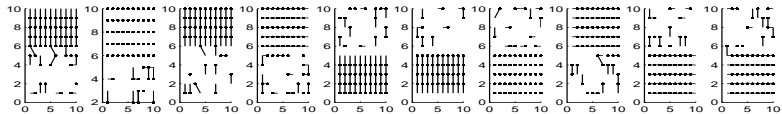
Flowfields



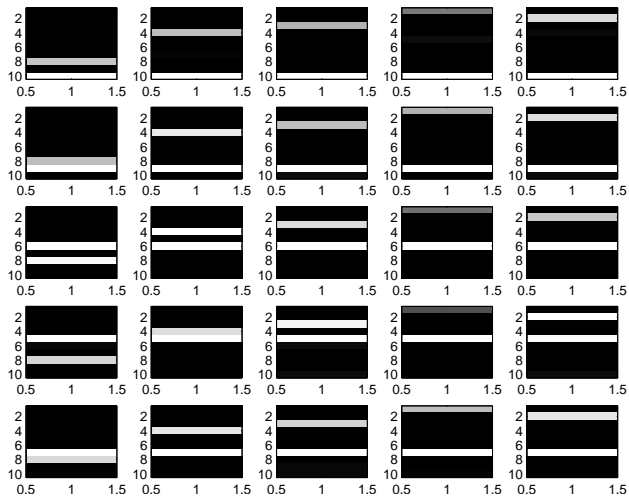
Example: Factorial transformations

- ▶ Show video2.

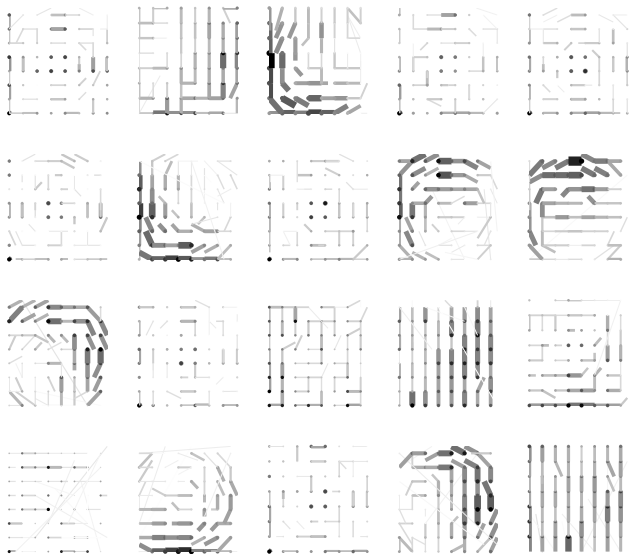
Factorial flowfields



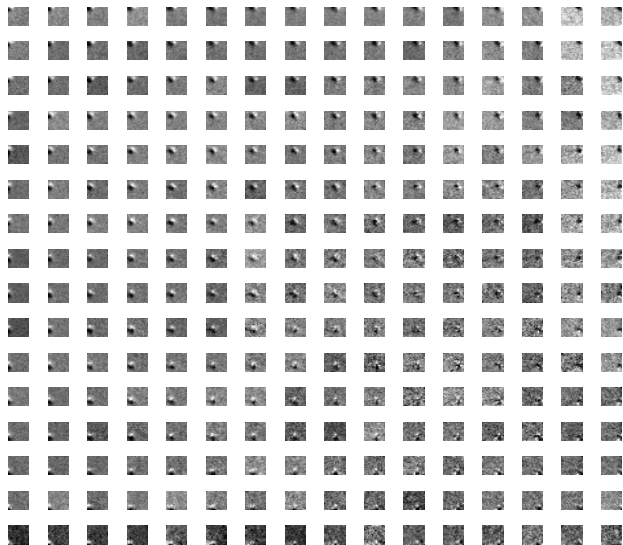
Factorial flowfields



Flowfields on natural images



One hidden unit



Example: Factorial transformations

- ▶ Show video3,4.

Discussion

- ▶ Local/ convolutional versions
- ▶ Layers
- ▶ Inference for optical flow fast
- ▶ Flow represented as binary vector. Can continue with *supervised learning* from here.
- ▶ Viewpoint invariance, stereo, depth estimation similarly
- ▶ Occlusion, noise