

Supplementary Material

for Facial Expression Transfer with Input-Output Temporal Restricted Boltzmann Machines
NIPS 2011 Submission 932

6 Parameter updates

In this section we present the parameter updates that were not included in the main text due to space constraints.

6.1 Input-Output Temporal Restricted Boltzmann Machines

The CD updates for the IOTRBM are:

$$\Delta W_{ij} \propto \sum_{t=N+1}^T \langle v_{i,t} h_{j,t} \rangle_{\text{data}} - \langle v_{i,t} h_{j,t} \rangle_{\text{recon}} \quad (17)$$

$$\Delta A_{ki} \propto \sum_{t=N+1}^T \langle v_{i,t} v_{k,<t} \rangle_{\text{data}} - \langle v_{i,t} v_{k,<t} \rangle_{\text{recon}} \quad (18)$$

$$\Delta B_{kj} \propto \sum_{t=N+1}^T \langle h_{j,t} v_{k,<t} \rangle_{\text{data}} - \langle h_{j,t} v_{k,<t} \rangle_{\text{recon}} \quad (19)$$

$$\Delta P_{li} \propto \sum_{t=N+1}^T \langle v_{i,t} s_{l,<=t} \rangle_{\text{data}} - \langle v_{i,t} s_{l,<=t} \rangle_{\text{recon}} \quad (20)$$

$$\Delta Q_{lj} \propto \sum_{t=N+1}^T \langle h_{j,t} s_{l,<=t} \rangle_{\text{data}} - \langle h_{j,t} s_{l,<=t} \rangle_{\text{recon}} \quad (21)$$

$$\Delta a_i \propto \sum_{t=N+1}^T \langle v_{i,t} \rangle_{\text{data}} - \langle v_{i,t} \rangle_{\text{recon}} \quad (22)$$

$$\Delta b_j \propto \sum_{t=N+1}^T \langle h_{j,t} \rangle_{\text{data}} - \langle h_{j,t} \rangle_{\text{recon}} \quad (23)$$

where $\langle \cdot \rangle_{\text{data}}$ is an expectation with respect to the training data distribution, and $\langle \cdot \rangle_{\text{recon}}$ is the M -step reconstruction distribution as obtained by alternating Gibbs sampling, starting with the visible units clamped to the training data.

054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107

6.2 Factored, Third-order Input-Output Temporal Restricted Boltzmann Machines

The CD updates for the factored, third-order IOTRBM are:

$$\Delta W_{if}^v \propto \sum_{t=N+1}^T \left\langle v_{i,t} \sum_j W_{jf}^h h_{j,t} \sum_l W_{lf}^s s_{l, \leq t} \right\rangle_{\text{data}} - \left\langle v_{i,t} \sum_j W_{jf}^h h_{j,t} \sum_l W_{lf}^s s_{l, \leq t} \right\rangle_{\text{recon}} \quad (24)$$

$$\Delta W_{jf}^h \propto \sum_{t=N+1}^T \left\langle h_{j,t} \sum_i W_{if}^v v_{i,t} \sum_l W_{lf}^s s_{l, \leq t} \right\rangle_{\text{data}} - \left\langle h_{j,t} \sum_i W_{if}^v v_{i,t} \sum_l W_{lf}^s s_{l, \leq t} \right\rangle_{\text{recon}} \quad (25)$$

$$\Delta W_{lf}^s \propto \sum_{t=N+1}^T \left\langle s_{l, \leq t} \sum_i W_{if}^v v_{i,t} \sum_j W_{jf}^h h_{j,t} \right\rangle_{\text{data}} - \left\langle s_{l, \leq t} \sum_i W_{if}^v v_{i,t} \sum_j W_{jf}^h h_{j,t} \right\rangle_{\text{recon}} \quad (26)$$

The parameter updates for A , B , \mathbf{a} , and \mathbf{b} are given by Eq. 18, 19, 22, and 23 respectively.