Supplementary Material
Latent Trees for Estimating Intensity of Facial Action Units

Sebastian Kaltwang
Imperial College London
sk2608@imperial.ac.uk

Sinisa Todorovic
Oregon State University
sinisa@eecs.oregonstate.edu

Maja Pantic
Imperial College London
m.pantic@imperial.ac.uk

7. Additional Qualitative Results

Fig. 6 shows the face instances probabilistically sampled from LT, which is learned on the DISFA data, for different FAUs. For every FAU, the corresponding face figure depicts the mean landmark locations for both the zero intensity (black crosses) and highest intensity levels (red crosses). Also, every face figure shows the standard deviation of the landmark locations along x and y axes for the highest intensity level (red ellipses). As can be seen, our LT structure learning is capable of correctly capturing the underlying higher-order dependencies among the landmark locations for different FAUs and their intensities.

As in the main paper, we can clearly see the correspondence between the facial regions influenced by the related FAU, e.g. FAUs 1 mainly influences the eyebrow landmarks while FAU 12 mainly influences the landmarks around the mouth.

Regarding FAU15, the learned distribution is almost identical the neutral face and thus is not sufficiently discriminative to achieve good results.

FAU5 influences the points around the right eyelid, which is expected. Additionally, the distributions of almost all other points increase their standard deviation, which indicates that FAU5 is accompanied by large head movements, which in turn lower the precision of the points.

FAU9 is not detected by points around the nose, but rather through the eyebrows. This indicates that the brow lowering is often co-curring with the nose wrinkle and thus the model detects FAU9 indirectly through FAU4. However, the co-occurrence is not strict and thus the LT performance is lower than SVR.

8. LT Parameter Updates

This section provides detailed formulas for the parameter updates of the distributions defined in (2), (3) and (4). Solving (13) regarding the parameters $(\boldsymbol{\mu}_{k,l})^{\text{new}}$ for the distributions defined in (2) leads to the update:

$$\mu_{k,l}^{\text{new}}(k1) = \frac{\sum_n q^{(n)}(\mathbf{h}_l = k1, \mathbf{h}_{P(l)} = k)}{N_{k,P(l)}}, \tag{14}$$

with $N_{k,P(l)} = \sum_n q^{(n)}(\mathbf{h}_{P(l)} = k)$.

Furthermore, solving (13) regarding the parameters $(\boldsymbol{\mu}_{k,m}, \boldsymbol{\Sigma}_{k,m})^{\text{new}}$ for the distributions defined in (3) leads to the update:

$$\boldsymbol{\mu}_{k,m}^{\text{new}} = \frac{\sum_n q^{(n)}(\mathbf{h}_{P(m)} = k)\mathbf{x}_m^{(n)}}{N_{k,P(m)}} \tag{15}$$

$$\boldsymbol{\Sigma}_{k,m}^{\text{new}} = \frac{\sum_n q^{(n)}(\mathbf{h}_{P(m)} = k)(\mathbf{x}_m^{(n)} - \boldsymbol{\mu}_{k,m}^{\text{new}})(\mathbf{x}_m^{(n)} - \boldsymbol{\mu}_{k,m}^{\text{new}})^\top}{N_{k,P(m)}}, \tag{16}$$

with $N_{k,P(m)} = \sum_n q^{(n)}(\mathbf{h}_{P(m)} = k)$.

Finally, solving (13) regarding the parameters $(\boldsymbol{\mu}_r)^{\text{new}}$ for the distributions defined in (4) leads to the update:

$$\mu_r^{(\text{new})}(k) = \frac{\sum_n q^{(n)}(\mathbf{h}_r = k)}{N}. \tag{17}$$

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

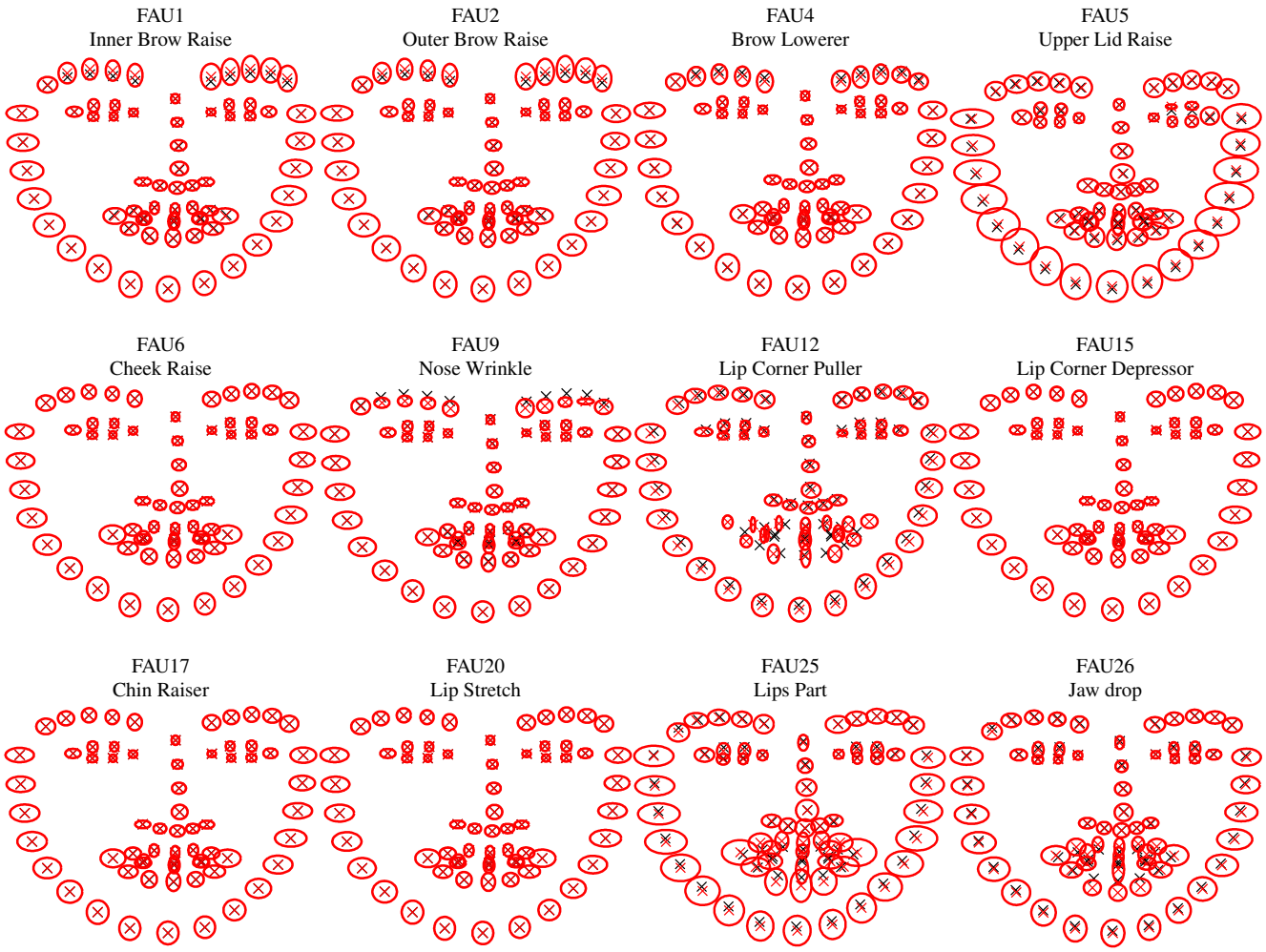

Figure 6. Landmark locations probabilistically sampled from an LT model trained on the DISFA data. For every FAU, the corresponding face figure depicts the mean landmark locations for both the zero intensity (black crosses) and highest intensity levels (red crosses). Also, every face figure shows the standard deviation of the landmark locations along x and y axes for the highest intensity level (red ellipses). The crosses cover each other, i.e. if the distribution does not change, then just a red cross is visible.