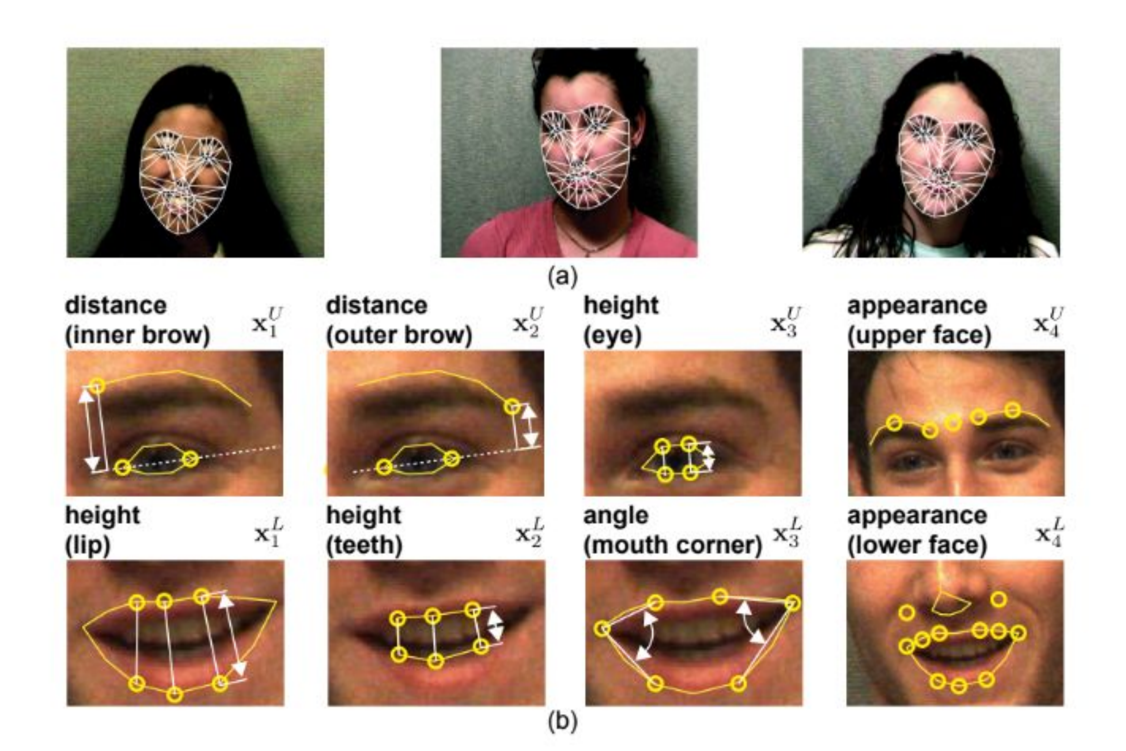
Unsupervised Discovery of Facial Events

Outline

- Features
- ACA
  - DTAK
  - KKM
- Optimizing

Features



(a)
distance (inner brow) $\mathbf{x}_1^U$
distance (outer brow) $\mathbf{x}_2^U$
height (eye) $\mathbf{x}_3^U$
appearance (upper face) $\mathbf{x}_4^U$
height (lip) $\mathbf{x}_1^L$
height (teeth) $\mathbf{x}_2^L$
angle (mouth corner) $\mathbf{x}_3^L$
appearance (lower face) $\mathbf{x}_4^L$
(b)

ACA (Aligned Cluster Analysis)

- KKM : clustering
- DTAK : distance between sequence data

KKM (kernel k-means)

minimize $$J_{kkm}(\mathbf{M}, \mathbf{G}) = ||\phi(\mathbf{X}) - \mathbf{M}\mathbf{G}||_F^2,$$ $$||\mathbf{X}||_F^2 = tr(\mathbf{X}^T\mathbf{X})$$

optimal $$\mathbf{M} = \phi(\mathbf{X})\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}$$ (cluster centroid)

$$J_{kkm}(\mathbf{G}) = tr\left(\mathbf{L}\mathbf{K}\right) \quad \mathbf{L} = \mathbf{I}_n - \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}$$

$$\mathbf{K} = \phi(\mathbf{X})^T\phi(\mathbf{X})$$

DTAK (Dynamic time alignment kernel)

find path maximizes the **weighted sum of similarity** between sequences

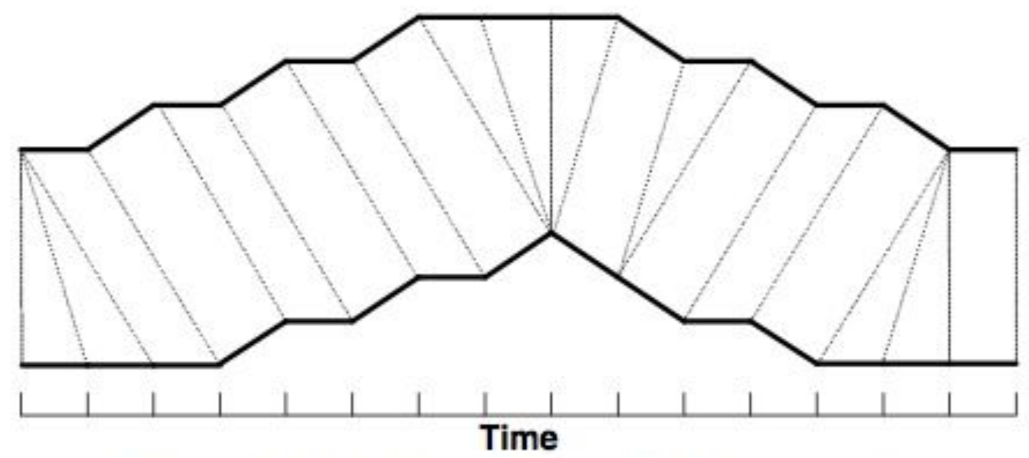


Figure 1. A warping between two time series.

DTAK (Dynamic time alignment kernel)

sequence $\mathbf{X} \doteq [\mathbf{x}_1, \cdots, \mathbf{x}_{n_x}]$ $\mathbf{Y} \doteq [\mathbf{y}_1, \cdots, \mathbf{y}_{n_y}]$

alignment matrix $\mathbf{Q} \in \mathbb{R}^{2 \times l}$ $\begin{bmatrix} \cdots & q_{1c} & \cdots \\ \cdots & q_{2c} & \cdots \end{bmatrix}$

step

correspondence matrix $\mathbf{W} \in \mathbb{R}^{n_x \times n_y}$

$w_{ij} = \frac{1}{n_x + n_y}(q_{1c} - q_{1c-1} + q_{2c} - q_{2c-1})$ if $q_{1c} = i$ and $q_{2c} = j$

else = 0

DTAK (Dynamic time alignment kernel)

$$\tau = \max_{\mathbf{Q}} \sum_{c=1}^{l} \frac{1}{n_x + n_y} (q_{1c} - q_{1c-1} + q_{2c} - q_{2c-1}) \kappa_{q_{1c} q_{2c}}$$

$$\kappa_{ij}(\mathbf{x}_i, \mathbf{y}_j) = \exp(\tfrac{1}{2\sigma^2} ||\mathbf{x}_i - \mathbf{y}_j||_2^2)$$

$$\tau(\mathbf{X}, \mathbf{Y}) = tr(\mathbf{K}^T \mathbf{W}) = \psi(\mathbf{X})^T \psi(\mathbf{Y})$$

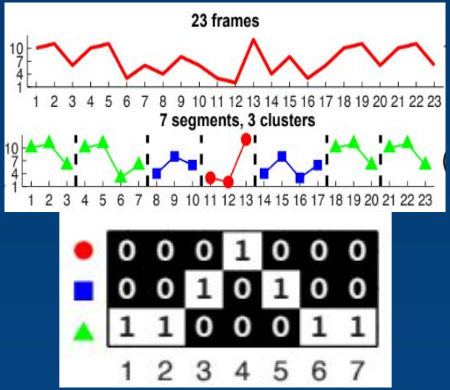
$\psi(\cdot)$ denotes a mapping of the sequence into a feature space, and $\mathbf{K} \in \mathbb{R}^{n_x \times n_y}$

ACA (Aligned Cluster Analysis)

minimize $J_{aca}(\mathbf{G}, \mathbf{M}, \mathbf{s}) = ||[\psi(\mathbf{Z}_1)\ \cdots\ \psi(\mathbf{Z}_m)] - \mathbf{M}\mathbf{G}||_F^2$

$\mathbf{Z}_i$: segment **i**, total **m** segment $\in \mathbb{R}^{d \times n_i}$

$\psi(\cdot)$: mapping s.t. $\psi(\mathbf{Z}_i)^T \psi(\mathbf{Z}_j) = tr(\mathbf{K}_{ij}^T \mathbf{W}_{ij})$



indicator matrix $\mathbf{G} \in \{0, 1\}^{k \times m}$ **k** clusters, **m** segment

optimal $\mathbf{M} = \phi(\mathbf{X})\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}$ (cluster centroid)

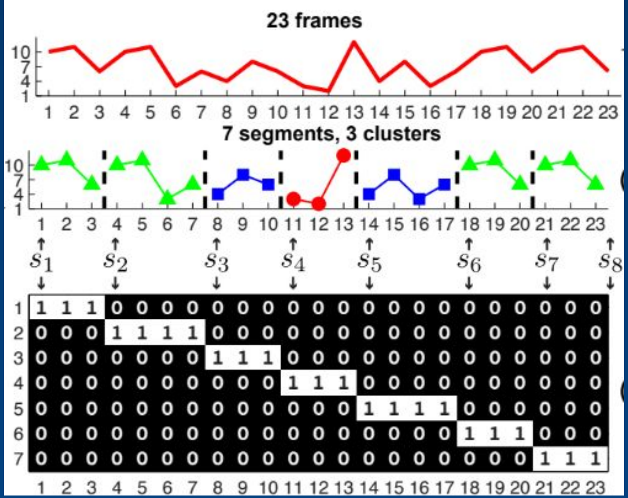
$J_{aca}(\mathbf{G}, \mathbf{s}) = tr\left((\mathbf{L} \circ \mathbf{W})\mathbf{K}\right)$ $\mathbf{L} = \mathbf{I}_n - \mathbf{H}^T\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}\mathbf{H}$

ACA (Aligned Cluster Analysis)

minimize $J_{aca}(\mathbf{G}, \mathbf{s}) = tr\Big((\mathbf{L} \circ \mathbf{W})\mathbf{K}\Big)$

$$\mathbf{L} = \mathbf{I}_n - \mathbf{H}^T\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}\mathbf{H}$$

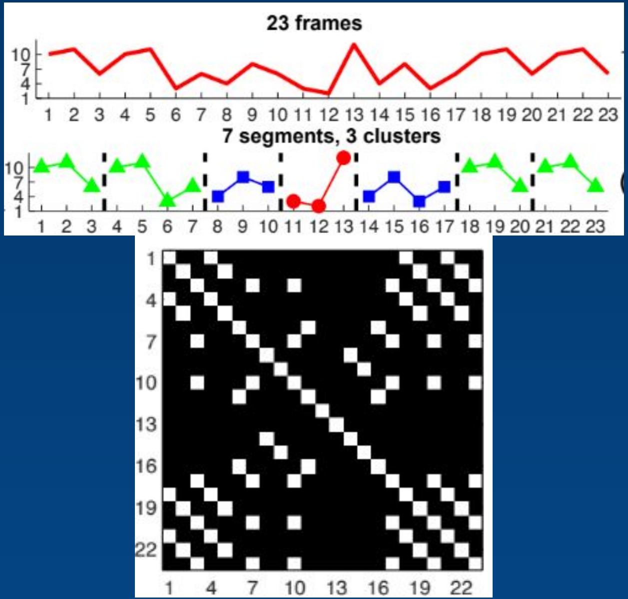
frame-segment indicator $\mathbf{H} \in \{0,1\}^{m \times n}$



ACA (Aligned Cluster Analysis)

minimize $J_{aca}(\mathbf{G}, \mathbf{s}) = tr\Big((\mathbf{L} \circ \mathbf{W})\mathbf{K}\Big)$

$$\mathbf{L} = \mathbf{I}_n - \mathbf{H}^T\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}\mathbf{H}$$



self-similarity matrix
(sample kernel matrix)

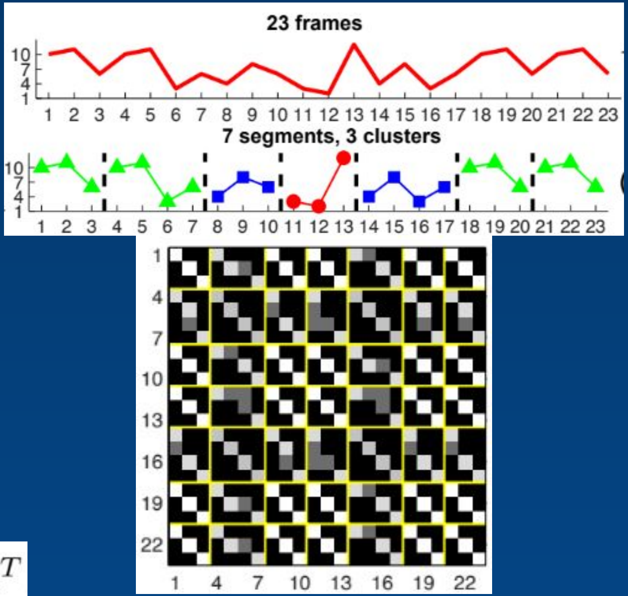
$\mathbf{K} \in \mathbb{R}^{n \times n}$

$\mathbf{K} = \phi(\mathbf{X})^T\phi(\mathbf{X})$

ACA (Aligned Cluster Analysis)

minimize $J_{aca}(\mathbf{G}, \mathbf{s}) = tr\Big((\mathbf{L} \circ \mathbf{W})\mathbf{K}\Big)$

$$\mathbf{L} = \mathbf{I}_n - \mathbf{H}^T\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}\mathbf{H}$$



correspondence matrix $\mathbf{W} \in \mathbb{R}^{n \times n}$

$m \times m$ blocks of $\mathbf{W}_{ij} \in \mathbb{R}^{n_i \times n_j}$

DTAK for each two segments

kernel segment matrix $\mathbf{T} \in \mathbb{R}^{m \times m}$

$$\mathbf{T} = [\tau_{ij}]_{m \times m} = [tr(\mathbf{K}_{ij}^T \mathbf{W}_{ij})]_{m \times m} = \mathbf{H}(\mathbf{K} \circ \mathbf{W})\mathbf{H}^T$$

Optimizing ACA

- minimize $J_{aca}(\mathbf{G}, \mathbf{s})$
- Using coordinate descent strategy alternates between optimizing $\mathbf{G}, \mathbf{s}$ while computing $\mathbf{M}$
- rewrite k-means equation:

$$J_{aca}(\mathbf{G}, \mathbf{s}) = \sum_{c=1}^{k} \sum_{i=1}^{m} g_{ci} dist_{\psi}^{2}(\mathbf{Z}_i, \mathbf{m}_c) = \sum_{i=1}^{m} dist_{\psi}^{2}(\mathbf{Z}_i, \mathbf{m}_{c_i^*})$$

Optimizing ACA

$$J(v) = \min_{\mathbf{G},\mathbf{s}} J_{aca}(\mathbf{G},\mathbf{s})|_{\mathbf{X}_{[1,v]}}$$

$$J(v) = \min_{1<i\leq v} \left(J(i-1) + \min_{\mathbf{G},\mathbf{s}} J_{aca}(\mathbf{G},\mathbf{s})|_{\mathbf{X}_{[i,v]}} \right)$$

$$J(v) = \min_{v-n_{\max}<i\leq v} \left(J(i-1) + \min_{\mathbf{g}} \sum_{c=1}^{k} g_c dist_{\psi}^2(\mathbf{X}_{[i,v]}, \dot{\mathbf{m}}_c) \right) \tag{9}$$

$$i_v^*, \mathbf{g}_v^* = \arg\min_{i,\mathbf{g}} J(v)$$

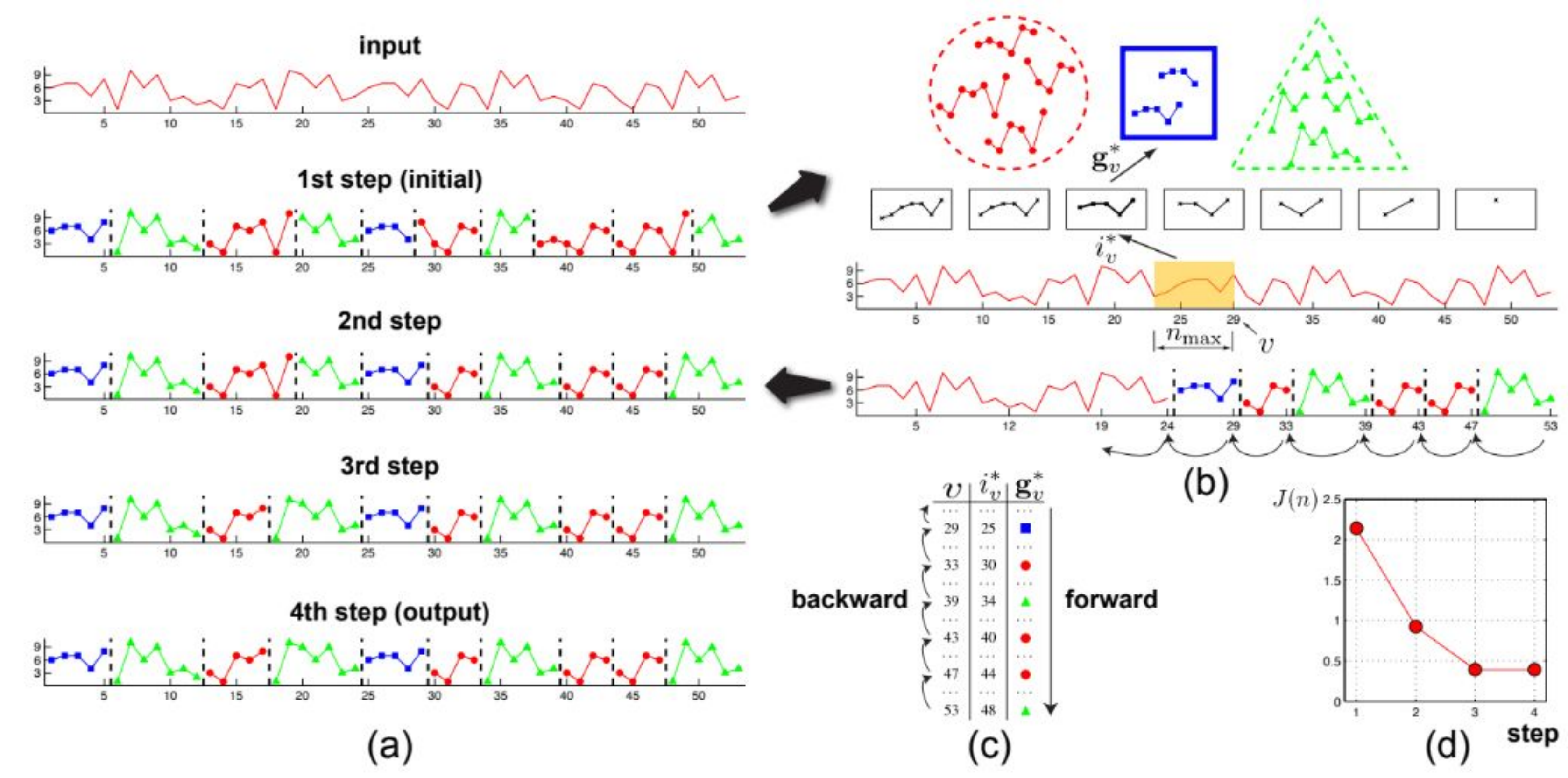
head position, **label** of right segment

$$dist_{\psi}^2(\mathbf{X}_{[i,v]}, \dot{\mathbf{m}}_c) = \tau(\mathbf{X}_{[i,v]}, \mathbf{X}_{[i,v]}) - \frac{2}{\dot{m}_c} \sum_{j=1}^{\dot{m}} \dot{g}_{cj} \tau(\mathbf{X}_{[i,v]}, \dot{\mathbf{Z}}_j) + \frac{1}{\dot{m}_c^2} \sum_{j_1,j_2=1}^{\dot{m}} \dot{g}_{cj_1} \dot{g}_{cj_2} \tau(\dot{\mathbf{Z}}_{j_1}, \dot{\mathbf{Z}}_{j_2})$$

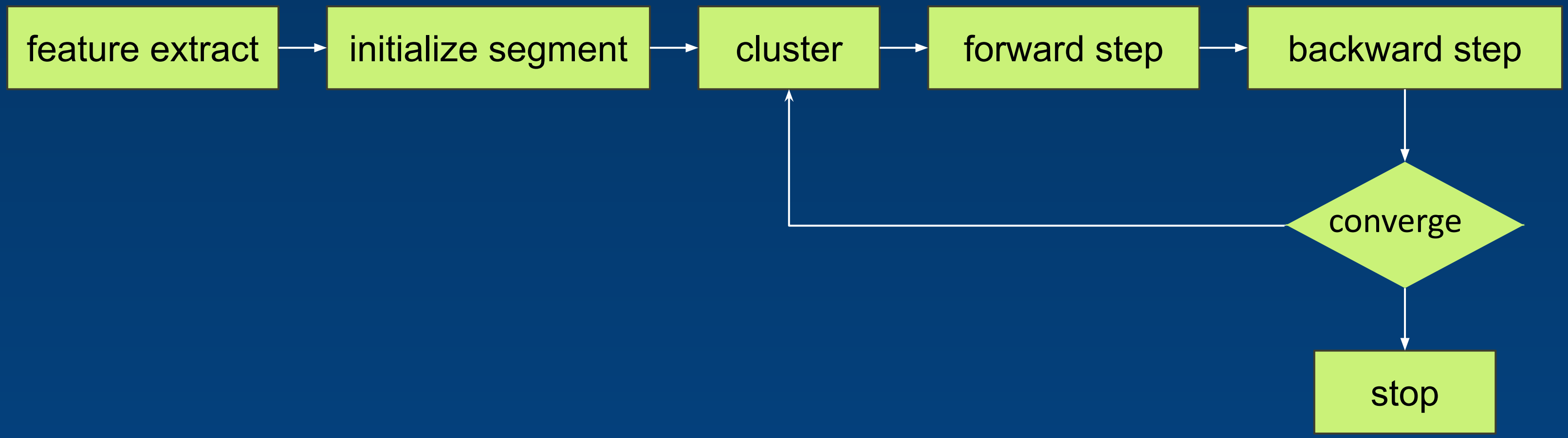
Optimizing ACA

- initialize segment
- Forward step
  - from v = 1 to n, compute $J(v)$ $i_v^*, \mathbf{g}_v^*$
- Backward step
  - back from v=n, cutoff the segment whose $s = i_v^*$ $\mathbf{g} = \mathbf{g}_v^*$
  - repeat on $v = i_v^* - 1$
- Repeat until J(n) converge

Optimizing ACA



Summary


feature extract
initialize segment
cluster
forward step
backward step
converge
stop