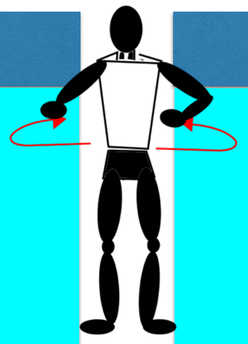
Complexity-Aware Assignment of Latent Values in Discriminative Models for Accurate Gesture Recognition

Manoel Horta Ribeiro, Bruno Teixeira, Antonio Otávio Fernandes, Wagner Meira Jr., Erickson R. Nascimento

Context



Problem: gesture and activity recognition on video streams

Solution: probabilistic graphical models based on conditional random fields using latent variables.

State-of-the-Art

Use **latent variables** to encode complex gestures/activities in a one-to-one mapping. They model gestures with that have many variations or complicated interactions in the time domain.

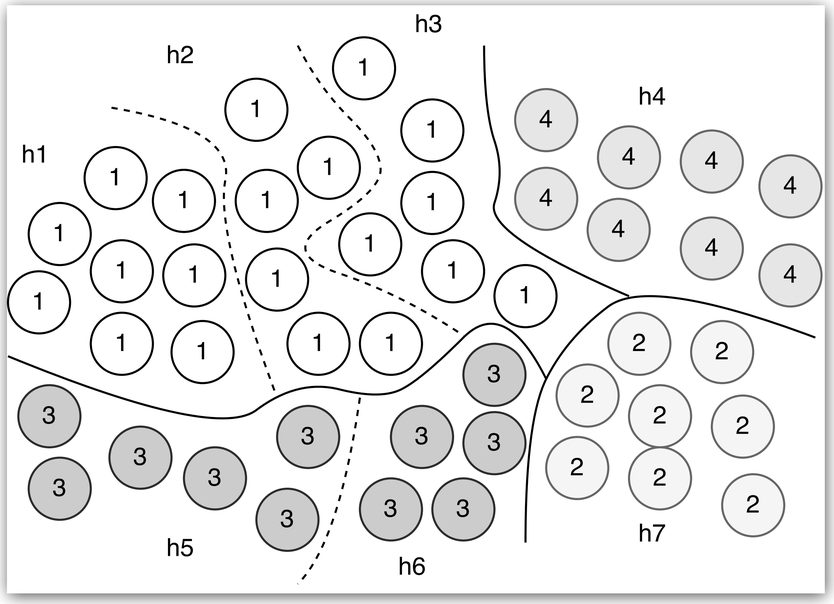


Figure 1. Label space divided by latent variables

Problem and Proposed Solution

✦ Latent values are arbitrarily are arbitrarily distributed, **N** for each label

✤ We propose distributing according to a measurement of complexity of each label.

$$P(\mathbf{y}|\mathbf{x}) = \sum_{\mathbf{h}:\forall \mathbf{h_j} \in \mathbf{H_{Y_j}}} P(\mathbf{h}|\mathbf{x}), \quad P(\mathbf{h}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp\left\{ \sum_{t=1}^{T} \sum_{k=1}^{K} \theta_k f_k(h_t, h_{t-1}, \mathbf{x}) \right\}$$

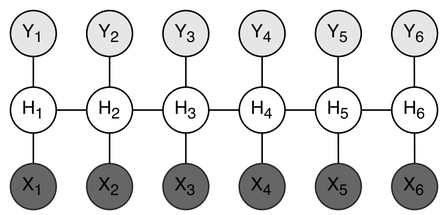


Figure 3. Example of a Latent-Dynamic Conditional Random Field.

Methodology

- Calculate complexity meas. for each label

1. For each two attributes of two instances of a given sample:
2. Normalize the two samples in the time domain.
3. Calculates point-to-point distance.
4. CM for each label is the normalized sum of this value for every instance and every attr.

- Distribute latent values accordingly

1. Give one latent value for each label.
2. While there are still latent values left, always give the label whose number latent values differs the most from its complexity measurement.

(Selected) Results

- Binarized datasets (NATOPS, ArmsGesture)

Dataset	Many Gestures			
	Ours c=1	Ours c=0.75	Normal	
AG0−12345	97.3 ± 1.3	97.5 ± 0.9	97.7 ± 1.1	
AG01−2345	85.0 ± 24.8	97.2 ± 0.8	97.3 ± 1.3	
AG012−345	89.2 ± 2.9	90.7 ± 0.6	90.6 ± 0.9	
NT0−12345	91.5 ± 1.2	92.6 ± 0.8	92.0 ± 1.1	
NT01−2345	86.2 ± 3.6	$\mathbf{93.8 \pm 0.4}$	83.8 ± 22.0	+ 10%
NT012−345	74.2 ± 1.3	$\mathbf{79.5 \pm 17.6}$	74.2 ± 1.3	+ 5%
AG01−2345	85.0 ± 24.8	97.2 ± 0.8	97.3 ± 1.3	
AG12−0345	90.1 ± 1.6	90.4 ± 1.3	90.5 ± 1.9	
AG23−0145	96.9 ± 1.7	$\mathbf{97.2 \pm 1.1}$	94.7 ± 4.8	+ 2%
AG34−0125	90.8 ± 1.6	90.8 ± 1.6	90.8 ± 1.6	
AG45−0123	$\mathbf{96.3 \pm 1.2}$	95.4 ± 0.7	94.6 ± 0.3	+ 2%
NT01−2345	86.2 ± 3.6	$\mathbf{93.8 \pm 0.4}$	83.8 ± 22.0	+ 10%
NT12−0345	70.5 ± 1.6	72.4 ± 3.4	72.0 ± 2.3	
NT23−0145	$\mathbf{89.7 \pm 1.8}$	81.1 ± 1.4	85.7 ± 12.0	+ 4%
NT34−0125	69.3 ± 1.6	$\mathbf{69.7 \pm 1.6}$	66.8 ± 1.8	+ 2%
NT45−0123	94.4 ± 1.0	$\mathbf{94.7 \pm 1.2}$	87.8 ± 10.4	+ 7%