

Problem:

• Predicting activities or imminent events from observed actions or events in the video.

Potential Applications

- Surveillance system
- Sports video analysis
- Smart room (elder caring)

Idea:

Reasoning-based activity prediction system should understand underlying causality and predictability of activities.

Our contribution:

- Actionlet: Encodes a temporal decomposition of long activities into a sequence of meaningful action units.
- **PST** (Probabilistic Suffix Tree): modeling causality
- **PAF** (Predictive Accumulative Function): modeling predictability





Definition:

- Let Σ be the finite set of actionlets, learned from activity temporal decomposition and encoding.
- $D_{training} = \{r^1, r^2, ..., r^m\}$ be the training set of m sequences over the actionlet alphabet Σ , where the length of the *i*th (i = 1, ..., m)sequence is l_i .

Goal:

• Learn a model P that provides a probability assignment p(t) for an ongoing actionlet sequence $t = t_1, t_2, ..., t_{||t||}$.



Predictive Accumulative Function (PAF)--- Characterizing Predictability

- We want to depict the predictable characteristic of a particular activity "tennis game" is a late-predictable problem.
- "drinking water" is an early predictable problem.
- PAF learned from information entropy changes along every stage of activity progress.

Formulation:

- $k \in [0,1]$, fraction of begninning portion (prefix) of any sequence.
- *D*_{trainning} set of trainning actionlet sequnces.
- D_k , set of sequences, where each element is first k percentage of training sample.

formation gain:
$$y_k = \frac{H(D) - H(D|D)}{H(D)}$$

Where,
$$H(D) = -\sum_{r \in D} p^{\overline{T}}(r) \log p^{\overline{T}}(r),$$

 $H(D|D_k) = -\sum_{r_{\text{pre}(k)} \in D_k} \sum_{r \in D} p^{\overline{T}}(r, r_{\text{pre}(k)}) \log p^{\overline{T}}(r|r_{\text{pre}(k)})$

From trained PST model \overline{T} , we write:

$$p^{\overline{T}}(r) = \prod_{j=1}^{\|r\|} \gamma_{s^{j-1}}(r_j) \text{ and } p^{\overline{T}}(r_{\text{pre}(k)}) = \prod_{j=1}^{\|r_{\text{pre}(k)}\|} \gamma_{s^{j-1}}(r_{\text{pre}(k)_j})$$
Early Predictable

From sequence of data pair (k, y_k) , we can fit PAF:

$$\mathbf{y} = \mathbf{f}_{\mathbf{p}}(\mathbf{k})$$



Modeling Complex Temporal Composition of Actionlets for Activity Prediction

Kang Li, Jie Hu and Yun Fu



