

Bayes Filter for Gesture Recognition

Variables

Hidden State Space

$$\mathbb{Z} = \{z_{\emptyset}^t, z_1^t, \dots, z_m^t\}$$

Where there are m objects and t represents the timestamp on the state.

A state consists of an object being referenced (or no object being referenced) and a timestamp.

Each state z_i^t consists of a series of 3D points.

NOTE: Consider combinations of objects? This would allow a good model of "The yellow objects". Initial model is for single objects only, but that is definitely a worthwhile extension.

Observations

$$\mathbb{O} = \{o_1, \dots, o_t\}$$

Where the subscript represents the timestamp for each observation.

Where $o_i = \{l_i, r_i, h_i, s_i\}$

Each observation consists of four parts:

l_i , the left arm vector at time i

r_i , the right arm vector

h_i , the head vector

s_i , speech

Since not all components of the observation tuple are guaranteed to be present at the same time, any of the four can take on a null value

Transition Function

The transition function $\mathbb{T}(z_i^t, z_k^{t+1})$ returns the probability of state z_i^t transitioning to z_k^{t+1} .

Transition probabilities should be high for $\mathbb{T}(z_i^t, z_k^{t+1})$ when $i = k$ and low otherwise. (Or we can start with uniform, but this makes more sense to me intuitively)

We could also weight the transition functions based on shared properties.

Let \mathbb{N} represent the probability of seeing a specific sample with a normal distribution of the specified parameters, θ represent the angle between the input vector and the sample point, and w, x , and y be exponents for weighting each sample appropriately.

Equations

We wish to know the most likely state (object being referenced) given our observations and previous state estimation, namely:

$$\operatorname{argmax}_{z_i^t} [P(z_i^t | o_{t-1}) * \sum_{z_k^{t-1} \in \mathbb{Z}} P(z_i^t | z_k^{t-1})]$$

Where:

$$P(z_i^t | z_k^{t-1}) = \mathbb{T}(z_k^{t-1}, z_i^t) * P(z_k^{t-1})$$

$$P(z_k^0) = \frac{1}{m+1}$$

$$P(z_i^t | o_{t-1}) = P(z_i^t | l_{t-1})^u * P(z_i^t | r_{t-1})^w * P(z_i^t | h_{t-1})^x * P(z_i^t | s_{t-1})^y$$

$$P(z_i^t | l_{t-1}) = \prod_{p \in z_i^t} \mathbb{N}(\mu_l = 0, \sigma_l, \theta(l_{t-1}, p))$$

$$P(z_i^t | r_{t-1}) = \prod_{p \in z_i^t} \mathbb{N}(\mu_r = 0, \sigma_r, \theta(r_{t-1}, p))$$

$$P(z_i^t | h_{t-1}) = \prod_{p \in z_i^t} \mathbb{N}(\mu_h = 0, \sigma_h, \theta(h_{t-1}, p))$$

$$P(z_i^t | s_{t-1}) = \text{TBD}$$