Visual Task Inference Using Hidden Markov Models

CS365 Project Proposal, Group 6
P.Yaswanth Kumar (yashkp@iitk.ac.in)
Jitendra Kumar Dhakar(jdhakar@iitk.ac.in)
Guide: Prof. Amitabha Mukerjee

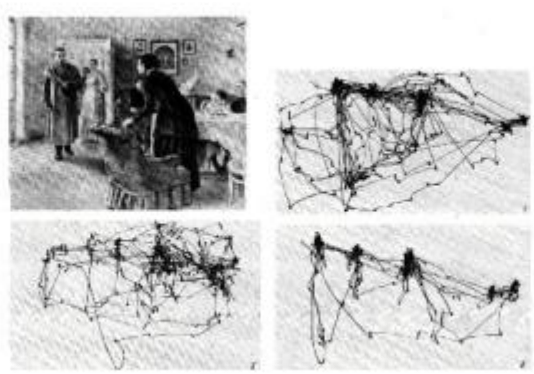
MOTIVATION:

[1]Visual tasks such as reading, counting and searching, greatly influence eye movement patterns. Yarbus showed that different eye movement trajectories emerge depending on the visual task that the viewers are given. Our main aim is to build an inverse Yarbus process whereby we can infer the visual task by observing the measurements of a viewer's eye movements while executing the visual task. The method we are using is Hidden Markov Models (HMMs) to create a probabilistic framework to infer the viewer's task from eye movements.

WORK BEFORE:

Two major models that were used to do task inference are Bottom-up Models where we assume *Naïve Bayes Assumption*[1] that the observations are conditionally independent of each other and are task independent. Another model is Top-down model which is task dependent but we still assume *Naïve Bayes Assumption* . So we use Hidden Markov Models by slightly modifying these two models in this project.

METHOD:



[1]In the above picture, the black lines represent the eye trajectories based on different tasks for a given image.

Learning of the HMMs: (**training**)

For each task, we take some observation sequences. Let { $\mathbf{O}^1$, $\mathbf{O}^2$,….,$\mathbf{O}^N$} be the set of sequences for learning where 'N' is the number of observations taken for each task. Its element is a sequence of eye positions $\mathbf{O}^n = \mathbf{o}_1^n\ \mathbf{o}_2^n \ldots\ \mathbf{o}_T^n$[2].

Using *Baum-Welch algorithm* [3] we can find the parameters $\boldsymbol{\pi}$, $\mathbf{A}$, $\mathbf{B}$ of a HMM $\boldsymbol{\lambda}$. So in this way, we can construct an action model for each task '**y**' denoted by $\boldsymbol{\lambda}_{\mathbf{y}}$.

Finding Likelihood term for each task:

For a new given input sequence of eye positions **O**, we now find the likelihood term for each task **y** by using *forward algorithm* [3].

Finding Posterior Probabilities and task inference:

Using Bayes Rule, we can find the posterior probability i.e probability that the task is '**y**' for a given input sequence **O** as follows.

$$P(\lambda_y|\mathbf{O}) = \frac{P(\mathbf{O}|\lambda_y)P(\lambda_y)}{P(\mathbf{O})}. \qquad [1]$$

So the task with maximum value of $\mathbf{P(\lambda_y|O)}$ has the more probability that the input sequence corresponds to that task and so is the required task.

DATA SET:

We will be creating our own dataset for training and testing purposes. (using eye gaze tracking)

REFERENCES:

[1] Haji-Abolhassani, A**.** and Clark, J.J., "Visual Task Inference Using Hidden Markov Models", proceedings of International Joint Conference on Artificial Intelligence (IJCAI), pp. 1678--1683, 2011

[2] Daiki KAWANAKA, Nonmember, Takayuki OKATANI, and Koichiro DEGUCHI, "Hierarchical-HMM Based Recognition of Human Activity"

[3] [Rabiner, 1990] L.R. Rabiner. A tutorial on hidden Markov mod-els and selected applications in speech recognition. Readings in speech recognition, 53(3):267–296, 1990.