* Methods
  + EM Algorithm

Expectation Maximum algorithm is an iterative method to find the local maximum likelihood. More specifically, it maximize the log likelihood of the data to find an optimal model for GMM.

There are two major steps in each iteration of EM algorithm. The E step calculate the gamma probability $\gamma$, which is the posterior probability of each class, given current model $\Lambda$. After this step, the current likelihood will be calculated to check if it has reached its local maximum. While the likelihood is still converging, the M step will use calculated gamma probability to re-estimate the GMM model for next iteration of maximization.

Evaluation

* + HMM
* Discussion

Comparing the result from mixed audio-visual with purely audio or visual, we find that at any chosen weight, the accuracy from audio+visual data is at least as high as the lower of the audio or the visual data, and many cases much higher than either of the two. This is because in the AV tests, the joint probability of audio and visual data rather than just one of them is calculated, providing more evidence of a given class/person/label.

The AV accuracy gains its maximum accuracy at weight 0.5, and it gradually becomes lower towards either end. This is because as weight increases, a larger weight/reliability is given onto audio samples, and it becomes closer to pure audio tests (and the same happens for the other direction). An equal weight considers audio and visual samples equally, and aligns perfectly with the Bayes rule of HMM, thereby producing the maximum accuracy.