Lab 2 – Hidden Markov Models with Gaussian Emissions

Presentation of the data

The main data

We are working on the 44 utterances of the tidigits set. We will try to find which digit is said by using a special model for each model.

All the models are in the variable models and there is a section for the Gaussian Mixture Model (GMM) and an other one for the Hidden Markov Models (HMM).

Analyze of the HMM models

1. Start propability

Always start with the first state since the probability is 1. This is logical since the model can't go back to the previous state and we don't want to miss a state

2. Transition probabilites

The biggest coefficients are on the diagonal of the matrix. This means that the model has most of the chance to stay on the same state. The coefficients just after the diagonal (except first and last state) is bigger than 0. Thus, we can go to the next state but the probability is really small. All the other coefficients are set to 0. We can conclude, we are using the first assumption of Markov.

Example

In order to verify each steps, we base the test on the mfcc of the example variable :

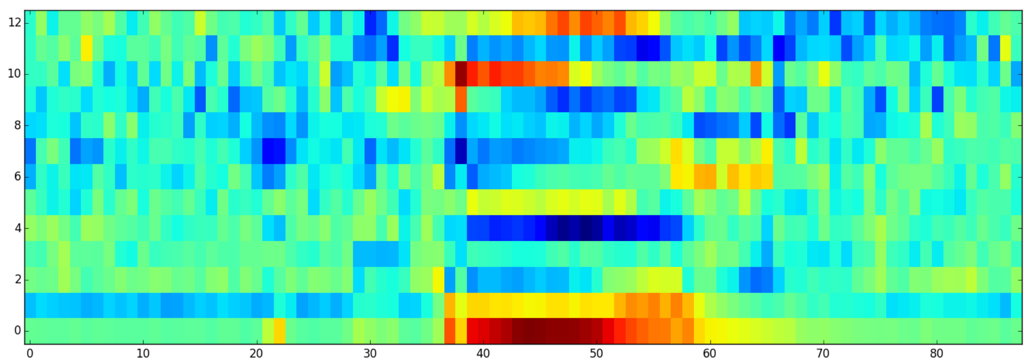

Illustration 1: MFCC of example

Multivariate Gaussian Density

The first third and the last third are the representation of the silence. The likelihood is homogeneous strong. However, in the middle we can see that the likelihood is weaker. There are a yellow stain and a blue one. The middle represents a phoneme.

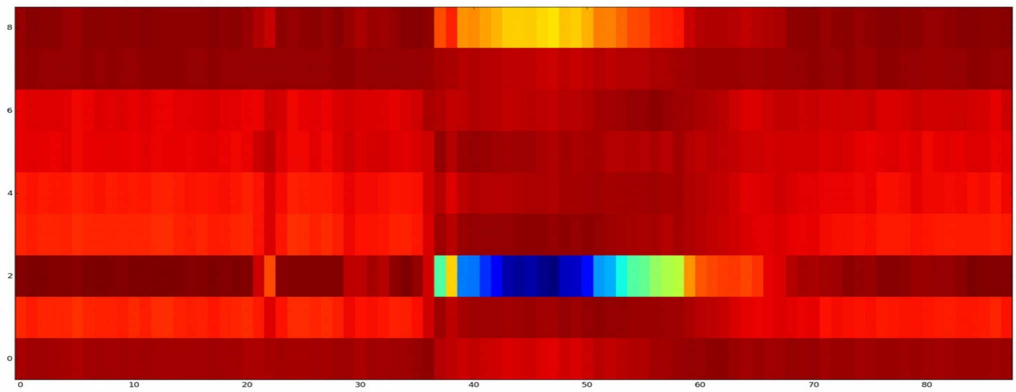

Illustration 2: GMM observations log likelihood

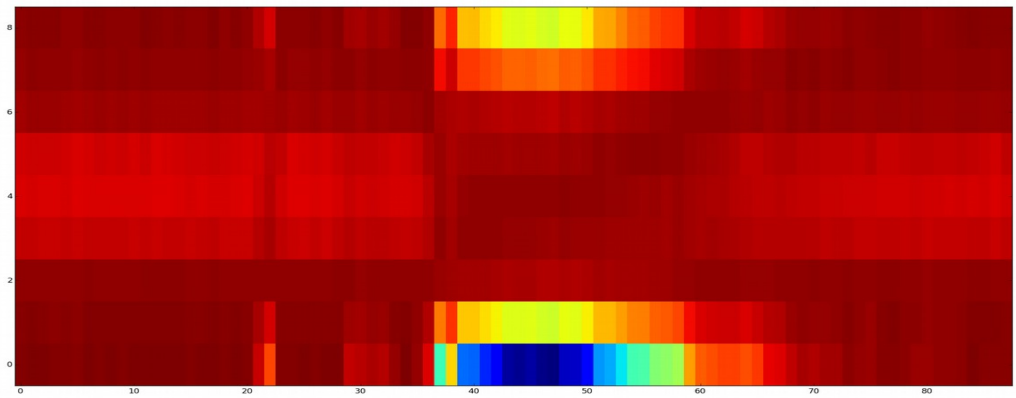

Illustration 3: HMM observation log likelihood

GMM Likelihood and Recognition

For getting the result, we choose the biggest value in the array. The index tells which model is the best one. A good result appears when the digit recognized is the same that was used for training the model.
In this case, we have only good results.

HMM Likelihood and Recognition

Forward Algorithm

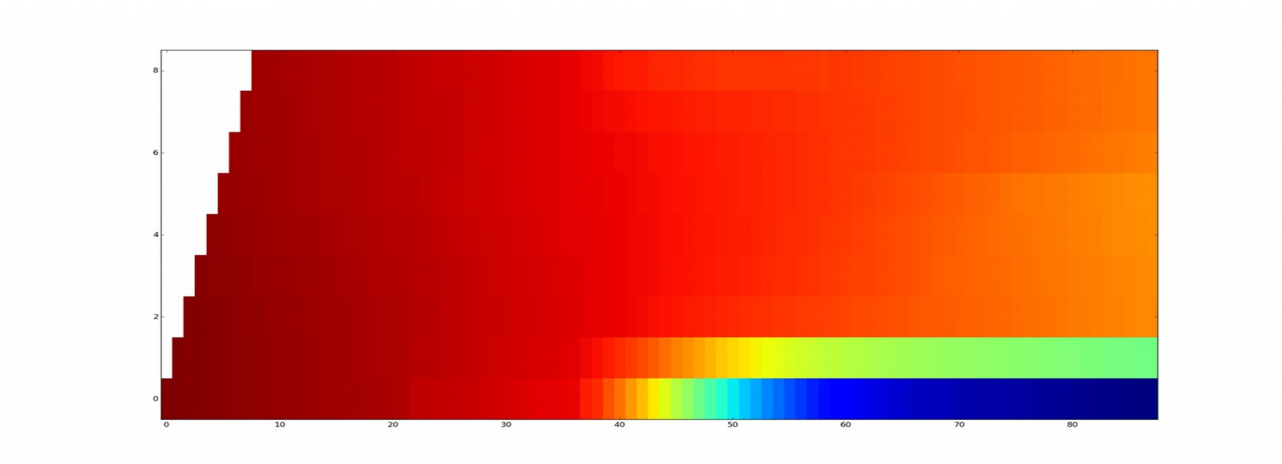

Illustration 4: Log Alpha for HMM

If we take the maximum likelihood for each utterance, we find two mistakes. The number 22 and 40 are wrong. This method is not as good as the one before.

The recognition using the Gaussian distribution in the HMM models gives the same result as the GMM likelihood. All the utterances are correctly classified.

The influence of the HMM transition model is that, it is trying to reproduce the impact of the weights for knowing what has the most of chance of happening.

On the first steps, the HMM topology is spreading but keep a strong probability. When continuing in the last steps, the probability decrease homogeneously except for the first two states that have a very low probability.

Viterbi Approximation

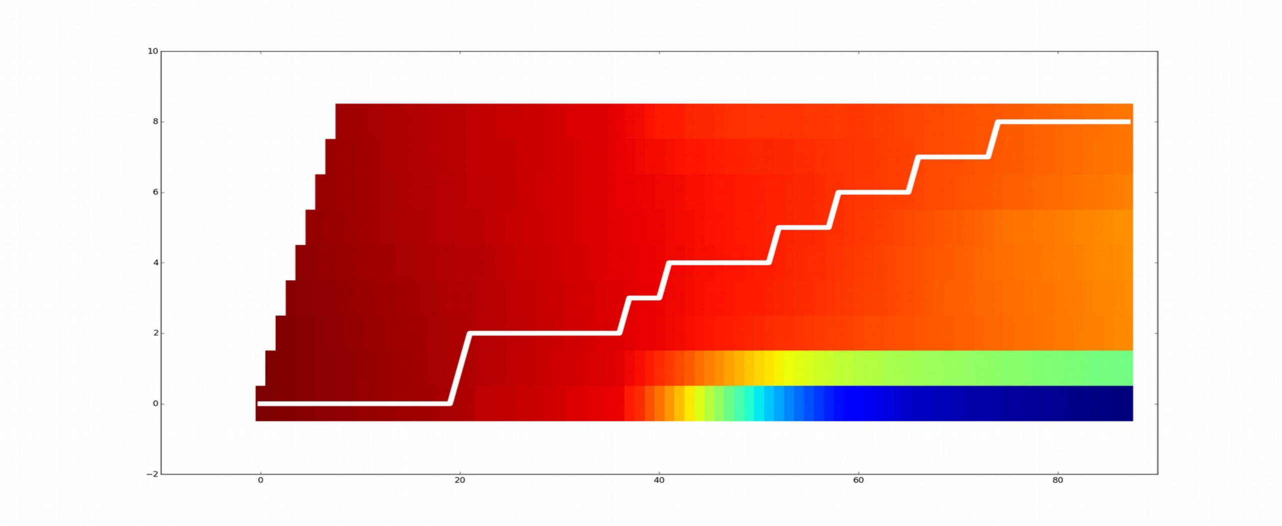

Illustration 5: Viterbi path on the log alpha from HMM

The path rises slowly since the biggest probability is to stay at the same state.

We retrieve the same result as for the HMM likelihood. This algorithm is not enough good for correcting the failure of HMM likelihood.
(I have checked the loglik and the vloglik are different for all utterances)

Backward Algorithm

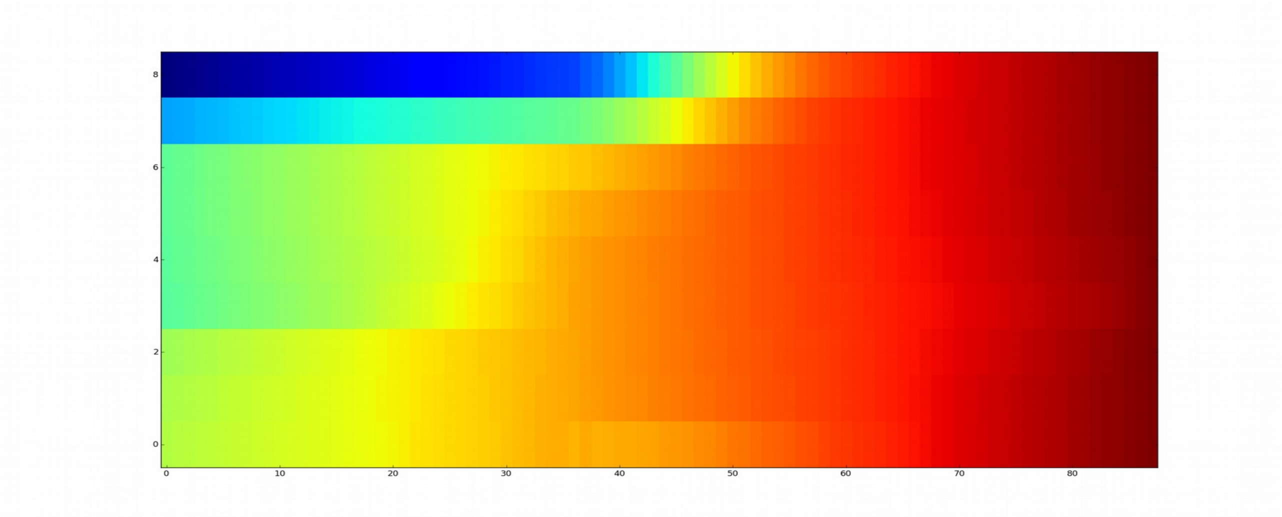

Illustration 6: Log beta from HMM

Without any surprise, we retrieve the same result as for HMM forward algorithm.

Results

UT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44

GMM																																												
HMM																					6																			4				
H/G																																												
VIT																					6																			4				
Bck																					6																			4				