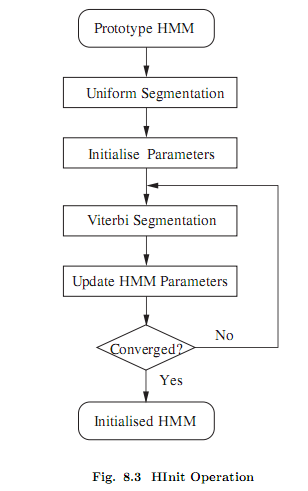
HTK from HTKBook.pdf to HInit.c Code,

Verification using real data

/RESEARCHS/neural\_network\_shmm/logbook\_hinit\_calculation\_verification.doc



# Hinit

# The basic principle of HInit depends on the concept of an HMM as a generator of speech vectors. Every training example can be viewed as the output of the HMM whose parameters are to be estimated. [HTKBOOK,sect. 8.2].

# @todo inside htk

## @todo uniform segmentation on hinit

Well actually it’seems only one segment

### scripts/log\_0001.txt

Just like using only 1 segment contain only 322 observations …

@TODO

- system is PLAIN

Aligning Segment Number 1

1 -87.79 ---- ----

2 -176.60 -173.12 ----

3 -272.34 -263.89 -279.29

4 -380.97 -366.06 -388.23

5 -497.45 -479.17 -501.99

The first line contain only a column ? why ?

@todo whats in a column ? whats is aligning …. Column is an emission state[2,3,4], in column is state probability ? Column contain OutP which is negative by gaussian order.

1 -87.79 ---- ----

### @todo, Hinit.c

EstimateModel()

🡪ViterbiAlign()

* OutP()

#### @todo Estimatemodel() explained

IntVec States has size 322 ?why

#### @todo ViterbiAlign() Explained

hmmLink🡪transP in ViterbiAligned() already contain log e values. For example

#### @todo

#### @todo gconst calc.

Hmodel.c: FixDiagGConst

Gconst = n\*log(2PI)+sigma(log var)

/\* EXPORT->FixDiagGConst: Sets gConst for given MixPDF in DIAGC case \*/

**void** FixDiagGConst(MixPDF \*mp)

{

**float** sum;

**int** i,n;

LogFloat z;

Vector v;

v = mp->cov.var; n=VectorSize(v); sum = n\*log(TPI);

**for** (i=1; i<=n; i++){

z = (v[i]<=MINLARG)?LZERO:log(v[i]);

sum += z;

}

mp->gConst = sum;

}

/\* DOutP: Log prob of x in given mixture - Diagonal Case \*/

**static** LogFloat DOutP(Vector x, **int** vecSize, MixPDF \*mp)

{

**int** i;

**float** sum,xmm;

sum = mp->gConst;

**for** (i=1;i<=vecSize;i++) {

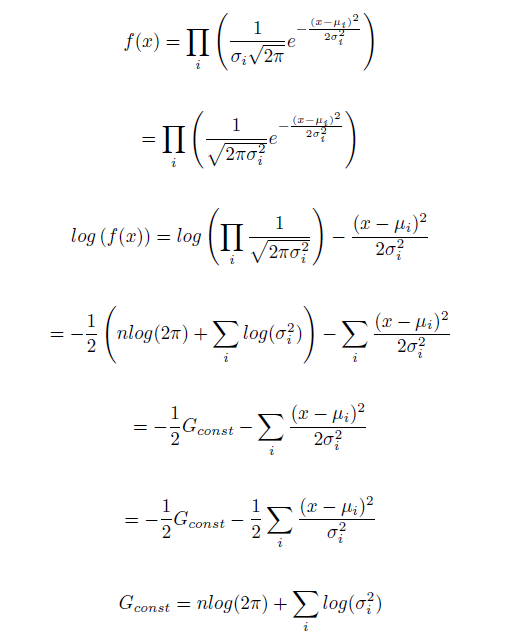
xmm=x[i] - mp->mean[i];

sum += xmm\*xmm/mp->cov.var[i];

}

**return** -0.5\*sum;

}



#### @todo when hmm parameter updated

#### @todo traceback matrix,

Is traceback matrix has relationship with multi viterbi path ? or multi solution ? how many path htk save ?

#### @todo convergence criterion htk hinit?

The convergence of HInit is depend on difference between total Probability every iteration.

HInit.c:1227

Delta=newP-totalP

#### @todo MFCC observations/vectors not clear GetSegObs compare to octave

….

Must load mfcc from octave/matlab …. And compare to getsegobs

Using hello.m 🡪 arctic\_0001.txt we got

-6.6966 4.1143 -2.6649 14.079 10.799 4.6172 12.076 ……..

-7.1163 2.0789 0.3962 12.47 5.7629 8.6881 17.707 ……….

-5.1815 4.9537 1.2972 9.3509 5.0981 0.93263 15.483 ………..

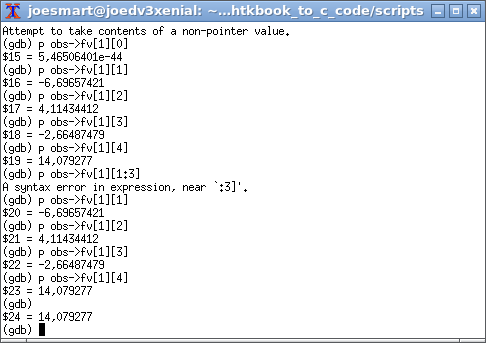
-3.3229 5.59 0.74351 13.567 11.653 5.0787 13.797

….

As We see in HParm.h:pringobservations(), Observations->fv is a matrix with two indexes,

obs->fv[s][j];

in our case we have only one mixture 🡪 s=1, and j from 1 (not 0) ,remember htk vector start at 1, 0 for its size, …



#### @todo mfcc file

#### type,

#### @data type

##### HMMSet \*/HLink, HMMDef

HMMDef /HLink

HMMSet

owner\*

##### Observations data type

**typedef** **struct** {

Boolean eSep; /\* Energy is in separate stream \*/

**short** swidth[SMAX]; /\* [0]=num streams,[i]=width of stream i \*/

ParmKind bk; /\* parm kind of the parm buffer \*/

ParmKind pk; /\* parm kind of this obs (bk or DISCRETE) \*/

**short** vq[SMAX]; /\* array[1..swidth[0]] of VQ index \*/

Vector fv[SMAX]; /\* array[1..swidth[0]] of Vector \*/

} Observation;

##### OutP

Outp return log of probs

###### IDOutP

DOutP

For diagonal (non inverse covariance)

LogFloat DOutP(Vector x, **int** vecSize, MixPDF \*mp)

### @todo what is traceback matrix

## @todo segmental k –means ? where in code

## @todo Debug observations into teks ….

In top of htk tools there is tracing level … for example

**#define** T\_VIT 0020 /\* Detailed Viterbi Alignment \*/

Which when executed … using –T switch as in scripts/all.sh

cmd="$HTKTOOLSDIR/HInit -C configtrain.txt -T 20 -m 1 -M hinitoutput VOI arctic\_a0001.mfc"

echo $cmd ; eval $cmd

# @todo, read tutorial uniform segmentation, viterbi training at speech.zone

# @todo, viterbi training?

$pwd

/home/joesmart/RESEARCHS/htkbook\_to\_c\_code/scripts

../../neural\_network\_shmm/htk\_35/HTKTools/HInit -T 3 -m 1 -M hinitoutput VOI arctic\_a0001.mfc

Initialising HMM VOI . . .

States : 2 3 4 (width)

Mixes s1: 1 1 1 ( 39 )

Num Using: 0 0 0

Parm Kind: MFCC\_D\_A\_Z\_0

Number of owners = 1

SegLab : NONE

maxIter : 20

epsilon : 0.000100

minSeg : 1

Updating : Means Variances MixWeights/DProbs TransProbs

- system is PLAIN

322 observations loaded from arctic\_a0001.mfc

1 Observation Sequences Loaded

Starting Estimation Process

Iteration 1: Average LogP =-29415.41797

Iteration 2: Average LogP =-29252.76953 Change = 162.64844

Iteration 3: Average LogP =-29247.36133 Change = 5.40820

Iteration 4: Average LogP =-29247.36133 Change = 0.00000

Estimation converged at iteration 5

Output written to directory hinitoutput

## arctic\_a0001.wav properties

Dump of "D:\RESEARCHS\htkbook\_to\_c\_code\scripts\arctic\_a0001.sfs" on Sat Mar 03 20:39:52 2018

Item 1

Data Type : 1.01 SPEECH

History : slink(file=d:/researchs/htkbook\_to\_c\_code/scripts/arctic\_a0001.wav,headerlen=44,start=0,end=51761,freq=16000,channels=1/1,dc=0,mult=0)

Parameters :

Process Date : Fri Mar 02 21:10:51 2018

Format : 2 byte integer

Frame size : 1 Frame count : 51761

Total Length : 512 Frame Duration : 6.25e-005 (16000 Hz)

Window size : 1 Overlap : 0

Offset : 0 Last Position : 0

Comment :

Linked item :

Filename : d:/researchs/htkbook\_to\_c\_code/scripts/arctic\_a0001.wav

Filepath :

Filetype : 0

Item no. : 0.00

Offset : 44

Multiplex : 0

Byte swap : no

DC offset : 0

Bit shift : 0

Date : Fri Mar 02 21:08:58 2018

lab\_match: out of date labels file 'C:\htk\SFS/data/labels'

file time=1513586305 internal time=1288920030

Total time = 51761 x 6.25e-6 = 3.24

## @todo 322 vs 434 observations

arctic\_a0001.wav

configcopy.txt;TARGETRATE =100 000 =1e5

target rate in s = 1e5x1e-7 = 1e-2 s=10 mili seconds;

jumlah frame = 3.24 s / 1e-2 = 3.24 x 1e2 = 324 obs – 2 (first and last ) = 322 obs.

2 observations is thrown out because of windowing …250

## @todo ron matlab/octave check … on Average LogP

## @todo how htk count Average LogP

# journals

## 1990 juang rabiner

State transition probability:

# Installation setup

## Hinit Gdb from shell

Must tell gdb to directory search:

$pwd

/home/joesmart/RESEARCHS/htkbook\_to\_c\_code/scripts

./gdb.sh

In gdb prompt:

gdb>b main

gdb> r `cat arg.txt`

## Eclipse

### Use F3 to see implementation alt 🡨 and alt 🡪 to go back

### Eclipse cdt installed over eclipse

No debugger in eclipse indigo …

Sudo apt-get install eclipse-cdt

Now create project htk\_35

Debugger failed no HInit.c

Using eclipse only for editing … using gdb in bash for debugging …

## Octave and gnuplot installed, ron htk matlab downloaded

## Matlab r2015

~/ bin/matlab\_2015

Failed, segmentation fault cor dumped

Using arg “-nodesktop”, ~/bin/matlab\_r2015 –nodesktop –nojvm

To able to use plot , “remove -nojvm”, only use –nodesktop options …

### Reading mfcc htkbook\_to\_c\_code/mscripts/hello.m

### Command line mode getplot ?

### Commandl line mode path

Use rmpath, addpath, path to remove, add, and show search path.

## Initial

3/2/2018 9:41:21 PM

htkbook\_to\_c\_code/scripts/all.sh

3/1/2018 10:13:29 PM  
Project dir. Created ~/RESEARCHS/htkbook\_to\_c\_code/  
HInit.c