

TP BIOMETRIE : Réalisation d'un premier système de reconnaissance de locuteur avec des outils GMM (toolkit ALIZE)

1) INSTALLATION DES OUTILS

-telecharger l'outil de paramétrisation spro4.0 sur <http://www.irisa.fr/metiss/guig/spro/download.html>

-en cas de problèmes d'installation voir le repertoire <http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/tools/CommentsLBInstall/>

-une doc en ligne (en anglais !) se trouve sur <http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/tools/CommentsLBInstall/doc.pdf>

!!NOUVEAU : des versions pre-compilées et plus récentes sont disponibles sur <http://mistral.univ-avignon.fr/>

!!!!SI PROBLEMES D INSTALLATION, RECUPEREZ LE FICHER <http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/tools/TP-Biometrie.tar.gz>

et travaillez sur la machine **goedel** !!!!!

2) Reconnaissance de genre (semaines 1 et 2)

a) le corpus de parole se trouve dans <http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/tools/corpus/> (fichiers raw, PCM 16khz)

les listes de signaux sont dans <http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/tools/lst>

Placez vous dans le repertoire *TP-Biometrie* sous linux pour effectuer ce tp.

Générez tous les vecteurs de paramètres pour tous les signaux

Aide : utilisez la commande sfbcep de spro

exemple : sfbcep -F PCM16 -f16000 -p 19 -e -D -A inputfile.raw inputfile.prm

b) normalisez les paramètres en utilisant la commande suivante

NormFeat.exe --config ./cfg/NormFeat_energy.cfg --inputFeatureFilename ./lst/all.lst --debug false --verbose true

c) lancer l'exécutable de détection d'énergie sur tous les signaux

EnergyDetector.exe --config ./cfg/EnergyDetector.cfg --inputFeatureFilename ./lst/all.lst --verbose true --debug false

Vérifiez que les labels obtenus dans le repertoire ./lbl sont corrects (pour 2 ou 3 signaux)

c) re-normalisez les paramètres en utilisant la commande suivante

NormFeat.exe --config ./cfg/NormFeat.cfg --inputFeatureFilename ./lst/all.lst

d) Apprendre les modèles de genre (male / female) avec les étapes suivantes :

-création d'un modèle du monde :

TrainWorld.exe --config ./cfg/TrainWorldInit.cfg --inputStreamList ./lst/world.lst --weightStreamList ./lst/world.weight --outputWorldFilename world_init --debug false --verbose true

TrainWorld.exe --config ./cfg/TrainWorldFinal.cfg --inputStreamList ./lst/world.lst --weightStreamList ./lst/world.weight --outputWorldFilename world --inputWorldFilename world_init --debug false --verbose true

-création des modèles de genre

TrainTarget.exe --config ./cfg/target_male.cfg --targetIdList ./ndx/male.ndx --inputWorldFilename world --debug false --verbose true

TrainTarget.exe --config ./cfg/target_male.cfg --targetIdList ./ndx/female.ndx --inputWorldFilename world --debug false --verbose true

5) Testez les modèles ainsi obtenus sur les fichiers de tests (voir dans ./lst pour la repartition train/test)

aide : utilisez les commandes suivantes

ComputeTest.exe --config ./cfg/target_seg_female.cfg --ndxFilename ./ndx/tests_female.ndx --worldModelFilename world --outputFilename female.res --debug false --verbose true

ComputeTest.exe --config ./cfg/target_seg_male.cfg --ndxFilename ./ndx/tests_male.ndx --worldModelFilename world --outputFilename male.res --debug false --verbose true

Travail à effectuer :

-donner le taux d'erreur de ce premier système

-re-apprendre d'autres modèles avec des variantes (sans normalisation des paramètres, sans

détection de silence, sans passer par un modèle du monde, avec un nombre différents de parametres et de gaussiennes...).

attention a la ligne featureServerMask dans les fichiers Train.cfg !!

3) Reconnaissance du locuteur (semaines 3 et 4)

En utilisant les meme sequences de commandes, réaliser un premier système d'identification du locuteur

Evaluer le taux d'erreur de ce système

OU BIEN enregistrez votre propre voix, ajoutez la dans la base, et faites des tests d'identification avec de nouveaux échantillons de votre voix. OU BIEN faites une interface de type serrure vocale