

Escuela de Informática – Universidad Nacional de Trujillo
Tópicos Especiales en Ciencia de la Computación
Reconocimiento Automático del Habla
Jorge Luis Guevara Diaz

LABORATORIO NUMERO 1

Presentación : Lunes 22 de octubre

1 Recoger muestras de habla

digitos 1-10 (5 veces por persona)

vocales a-e-i-o-u (5 veces por persona)

comandos

abrir-cerrar-ven-adios-caminar-parar-hola-saltar-correr-coger (5 veces por persona)

texto (esta grabacion lo haremos mas adelante)

""

Descripción

Con estas muestras haremos nuestros experimentos la grabación lo haremos en el laboratorio, llevare el micro, Juan llevara la laptop, por favor instalar el Adobe Audition para realizar las grabaciones

2. Construir un programa para visualizar una señal de habla

No utilizar Matlab en esta sección (opciones C++ o Java), piense son 16000 datos por un segundo, no necesariamente se tiene que graficar todos los puntos pues resultaría muy lento.

Graficar cada cierto número de puntos, luego construir la curva (una recta, interpolación, etc)

3. Hacer un programa para eliminar segmentos inútiles de una señal de habla (palabras aisladas)

Descripción

Para nuestro primer proyecto de construir un reconocedor de palabras aisladas es necesario eliminar las partes de información importante de la señal de habla, ejemplo : ruido al inicio y al final causado por el micro, silencios no necesarios , etc,

Algoritmo Eliminación de Segmentos Inútiles

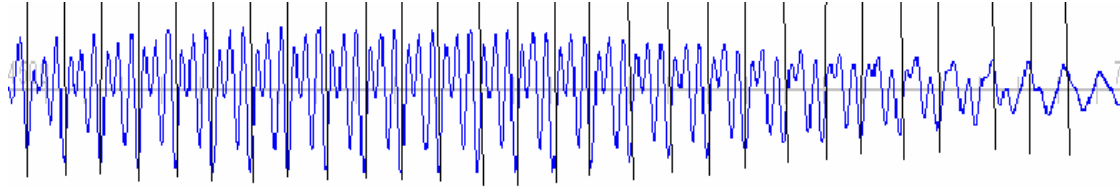
INPUT señal de habla con segmentos inútiles

SALIDA señal con solo la información relevante

4. Hacer un programa para determinar el periodo de "pitch" de una palabra (probar con los archivos dados en la página web del curso: habla.rar)

Descripción

Tal y como lo hemos visto en clase existe un comportamiento periodico de las ondas en una señal de habla, ejemplo diapositiva 8 , diapositiva 11, de las diapositivas vistas en clase, el objetivo es determinar el periodo de “pitch” determinar el periodo de pitch es importante para realizar síntesis de habla, y para realizar reconocimiento automático del habla utilizando el seguimiento y cambio del periodo de pitch a travez de una señal de habla, el periodo del pitch esta relacionado con la frecuencia fundamental generada por las cuerdas vocales



Los periodos de pitch en la imagen superior están marcados con líneas verticales (aproximadamente), se debera diseñar un algoritmo el cual marque seguramente los periodos de pitch (marcar los picos ”mas grandes” de cada periodo)

Los resultados deberan ser mostrados en un archivo de texto como sigue:

Marcas	Tipo	Muestras	Periodo	Frecuencia
17	S	121	1/f	$f = F_s * \text{marcas} / \text{muestras}$

F_s = periodo de muestreo ejem: 16000 muestras por segundo es decir el archivo se grabo a una tasa de 16000 Hz

Marcas = numero de secciones marcadas tanto para sonidos sonoros como no sonoros

Tipo = S=sonoro, NS= no sonoro ¿Cómo cree usted de que se diferenciaran?

Muestras = numero de muestras para el presente periodo

Algoritmo Determinación de pitch

INPUT señal de habla

SALIDA archivo de texto con la información + opcional (visualizacion de la señal con las marcas del periodo de pitch)

5 Presentar reporte conteniendo:

Descripción

Su reporte deberá contener lo siguiente

Una descripción detallada de sus algoritmo propuestos

Una sección de conclusiones la cual valore críticamente su algoritmo y indique que tipos de palabras habladas de las que fueron exitosamente eliminados los segmentos inútiles, si investigo algún algoritmo referenciarlo y comentarlo

Una sección de conclusiones la cual valore críticamente su algoritmo y indique que tipos de palabras habladas fueron marcadas con el periodo de pitch en forma

satisfactoria y cuales no ¿cuan seguro es su algoritmo?. Si investigo algún algoritmo referenciarlo y comentarlo

Nota

Para las preguntas 3 y 4 puede utilizar cualquier LP, (java, C++ matlab, etc)

Para la pregunta 5 puede idear su propio algoritmo o investigar algun algoritmo que haga determinación del pitch en la señal de habla, (en este caso referenciarlo),

Enlaces

<http://cmusphinx.sourceforge.net/html/cmusphinx.php>

<http://htk.eng.cam.ac.uk/>

determinación del pitch

<http://www.owl.net.rice.edu/~elec431/projects96/digitalbb/pitch.html>

<http://www.cnmat.berkeley.edu/~tristan/Report/node4.html>

papers sobre determinación del pitch

ver pagina web del curso

<http://www.inf.unitru.edu.pe/~jlgd/topicosiii-2007ii/index.htm>