

Índice

Introducción	i
Capítulo I El reconocimiento de voz	5
1.1 Antecedentes	6
1.2 Características acústicas	7
1.2.1 Realización fonética	7
1.2.2 Clasificación de los fonemas	8
1.2.2.1 Fonemas consonánticos	8
Lugar de articulación	8
Manera de articulación	9
Articulaciones sonoras o sordas	10
1.2.2.2 Fonemas vocálicos	10
1.3 Componentes básicos de un reconocedor	13
1.4 Tipos de reconocedores	13
1.5 Aplicaciones	16
1.6 Representaciones de la señal de voz	17
Capítulo II El CSLU <i>toolkit</i>	20
2.1 Arquitectura del CSLU <i>toolkit</i>	20
2.2 El proceso de reconocimiento basado en <i>frames</i>	21
2.2.1 Extracción de características	23
2.2.1.1 Vectores MFCC de características	24
2.2.1.2 Ventanas de contexto	24
2.2.2 Clasificación de <i>frames</i>	25
2.2.3 La búsqueda Viterbi y la red de pronunciaciones	25
2.3 Modelado dependiente del contexto	25
2.4 División del fonema en partes	26
2.5 Redes neuronales y el CSLU <i>toolkit</i>	27
2.6 Entrenamiento de una red neuronal	27

Capítulo III	El corpus de voz	29
3.1	<i>Tlatoa common questions corpus</i>	29
3.2	Condiciones de grabación	30
3.3	División del corpus	31
3.4	Protocolo de etiquetado	32
3.4.1	Transcripciones ortográficas	32
3.4.2	Transcripciones a nivel palabras	33
3.4.3	Transcripciones fonéticas	33
3.4.4	Conjunto de símbolos	34
3.5	Formato y estructura de los archivos	36
Capítulo IV	Experimento base	37
4.1	Conjunto de fonemas del vocabulario	37
4.2	Distribución de los datos	39
4.3	Otras investigaciones similares	40
4.4	Tcl y Tk	41
4.5	Fase de entrenamiento	42
4.5.1	Crear el archivo de corpora y el archivo de files	42
4.5.2	Crear el archivo de parts y de vocab	44
4.5.3	Correr los scripts de entrenamiento	45
4.6	Fase de desarrollo	46
4.6.1	Script de conversión	47
4.7	Resultados con CSLU toolkit	48
4.8	Nist	48
4.8.1	Alineación	49
4.8.2	Scoring	50
4.9	Resultados finales (con Nist)	50

Capítulo V	Experimento base	53
5.1	Experimentos independientes del contexto	53
5.1.1	Modelos de duraciones	54
5.1.2	Duraciones 4% en experimento base	55
5.1.3	Duraciones 6% en experimento base	56
5.1.4	Duraciones 8% en experimento base	56
5.2	Categorías dependientes del contexto	58
5.3	Número de partes	58
5.4	Experimentos dependientes del contexto	59
5.4.1	Dependiente del contexto 10,000 muestras	61
5.4.2	Dependiente del contexto 10,000 muestras con duraciones 6%	62
5.4.3	Dependiente del contexto 50,000 muestras	63
5.4.4	Dependiente del contexto 50,000 muestras con duraciones 6%	64
5.5	Fase de prueba final	66
5.6	Prueba a nivel de palabras	69
	Conclusiones y trabajo a futuro	71
	Apéndice A	74
	Apéndice B	86
	Referencias	97