

UNIVERSIDAD POLITECNICA DE CATALUÑA

Departamento de Teoria de la señal y comunicaciones

**TECNICAS DE PROCESADO Y
REPRESENTACION DE LA SEÑAL
DE VOZ PARA EL
RECONOCIMIENTO DEL HABLA
EN AMBIENTES RUIDOSOS**

Autor: Francisco Javier Hernando Pericas

Director: Climent Nadeu i Camprubi

Barcelona, mayo 1993

A Patricia

AGRADECIMIENTOS

Esta tesis es fruto de cuatro años de trabajo en el área del procesado de señal aplicado al reconocimiento automático del habla realizado en el Grupo de Procesado de Señal del Departamento de Teoría de la Señal y Comunicaciones.

Quiero expresar mi más sincero agradecimiento a todas aquellas personas que directa o indirectamente han hecho posible que este trabajo vea la luz. Sería difícil pretender citar todos los nombres. Todos aquellos con los que me creo en deuda, ya lo saben.

Especialmente importante ha sido la tutoría, no sólo en esta tesis, sino a lo largo de mi aprendizaje como investigador, del profesor Climent Nadeu. Su contribución en muchas de las ideas incluidas en esta tesis, sus valiosos consejos y sugerencias y su apoyo moral en las ocasiones difíciles han resultado fundamentales en la culminación del presente trabajo de investigación.

Asimismo expreso mi gratitud a los profesores José B. Mariño y Eduardo Lleida por sus contribuciones en diferentes aspectos de esta tesis y de mi formación investigadora.

Vaya también mi agradecimiento a mis compañeros del Grupo de Procesado de Señal por la inestimable ayuda profesional y humana que me han brindado con la aportación de medios materiales y con sus sugerentes discusiones y ayuda personal. Agradezco de manera especial la ayuda constante y desinteresada del profesor Francesc Vallverdú en tareas docentes, que me ha permitido dedicar a esta tesis el tiempo necesario.

Por último, agradezco la colaboración de los Sres G. Babini y J. Gómez Mena, que hizo posible la utilización en esta tesis de una base de datos con ruido real.

INDICE

1. INTRODUCCION	1
1.1. OBJETIVOS GENERALES.....	1
1.2. ESTRUCTURA DE LA TESIS	4
2. RECONOCIMIENTO DEL HABLA EN ENTORNOS ADVERSOS.....	7
2.1. SISTEMAS DE RECONOCIMIENTO AUTOMATICO DEL HABLA	9
2.1.1. PROBLEMAS ASOCIADOS AL RECONOCIMIENTO DEL HABLA	10
2.1.2. RESTRICCIONES DE LOS SISTEMAS DE RECONOCIMIENTO	12
2.1.3. DIFERENTES APROXIMACIONES AL RECONOCIMIENTO	16
2.1.3.1. COMPARACION DE PATRONES.....	17
2.1.3.2. MODELOS OCULTOS DE MARKOV	21
2.1.3.3. APROXIMACION CONEXIONISTA.....	28
2.1.3.4. METODOS BASADOS EN EL CONOCIMIENTO.....	32
2.2. ESTADO DEL ARTE DEL RECONOCIMIENTO DEL HABLA EN ENTORNOS ADVERSOS	33
2.2.1. ENTORNOS ADVERSOS	36
2.2.1.1. RUIDO	36
2.2.1.2. DISTORSION.....	37
2.2.1.3. EFECTOS ARTICULATORIOS.....	38
2.2.2. TECNICAS DE RECONOCIMIENTO DEL HABLA EN ENTORNOS ADVERSOS	38
2.2.2.1. TRANSDUCTORES ESPECIALES	39
2.2.2.2. NUEVAS REPRESENTACIONES DE LA SEÑAL DE VOZ	43
2.2.2.3. PREPROCESADO DE MEJORA DE LA SEÑAL DE VOZ	44
2.2.2.4. ENMASCARAMIENTO DEL RUIDO Y MODELOS ADAPTATIVOS	47
2.2.2.5. MEDIDAS DE DISTORSION ROBUSTAS.....	49

2.2.2.6. COMPENSACION DE ESTRES.....	52
--------------------------------------	----

3. TECNICAS ROBUSTAS DE REPRESENTACION DE LA SEÑAL DE

VOZ	55
3.1. MODELADO DIGITAL DE PRODUCCION DE LA SEÑAL DE VOZ	57
3.1.1. PRINCIPIOS FISIOLÓGICOS BÁSICOS	57
3.1.2. LA SEÑAL DE VOZ	59
3.1.3. MODELO LINEAL DE PRODUCCION DEL HABLA.....	61
3.2. PREDICCIÓN LINEAL CLÁSICA DE LA SEÑAL DE VOZ	65
3.2.1. ESTIMACION DE MÍNIMOS CUADRADOS	68
3.2.1.1. MÉTODO DE AUTOCORRELACION. ECUACIONES DE YULE-WALKER (YWE).....	70
3.2.1.2. MÉTODO DE COVARIANZA.....	72
3.2.2. MODELADO ESPECTRAL.....	73
3.3. PREDICCIÓN LINEAL EN PRESENCIA DE RUIDO	77
3.3.1. MODELADO AR, MA Y ARMA	81
3.3.2. ECUACIONES DE YULE-WALKER DE ORDEN SUPERIOR (HOYWE)	84
3.3.3. ECUACIONES SOBREDETERMINADAS DE YULE-WALKER DE ORDEN SUPERIOR (OHOYWE)	86
3.3.4. ECUACIONES SOBREDETERMINADAS DE YULE-WALKER (OYWE).....	88
3.4. INTERPRETACION COMO PREDICCIÓN LINEAL DE LA SECUENCIA DE AUTOCORRELACION	89
3.5. PREDICCIÓN LINEAL DE LA PARTE CAUSAL DE LA AUTOCORRELACION	92
3.5.1. LA PARTE CAUSAL DE LA AUTOCORRELACION. ESPECTRO ANALÍTICO Y ENVOLVENTE ESPECTRAL.....	93
3.5.2. MODELADO INVERSO DE LA AUTOCORRELACION CAUSAL (MIAC).....	98
3.5.3. PREDICCIÓN LINEAL DE LA PARTE CAUSAL DE LA AUTOCORRELACION (OSALPC).....	103
3.5.4. RELACION CON LA COHERENCIA MODIFICADA LOCALIZADA (SMC).....	110
3.6. TRANSFORMACION DE LA ESCALA DE FRECUENCIAS.....	116

4. MEDIDAS DE DISTANCIA ROBUSTAS.....	1 2 3
4.1. CEPSTRUM	125
4.2. DISTANCIA DE MAHALANOBIS	128
4.3. DISTANCIAS EUCLIDEAS CEPSTRALES PONDERADAS.....	133
4.4. DISTANCIAS DE PROYECCION CEPSTRAL.....	141
4.5. INCORPORACION DE INFORMACION DINAMICA Y ENERGIA.....	145
5. RECONOCIMIENTO DEL HABLA MEDIANTE MODELOS OCULTOS DE MARKOV.....	1 5 1
5.1. MODELOS OCULTOS DE MARKOV. DEFINICION Y TIPOS.....	152
5.2. MODELOS OCULTOS DE MARKOV DISCRETOS.....	156
5.2.1. SOLUCION A LOS TRES PROBLEMAS BASICOS.....	157
5.2.1.1. EVALUACION.....	157
5.2.1.2. DECODIFICACION.....	162
5.2.1.3. ENTRENAMIENTO.....	164
5.2.2. IMPLEMENTACION DE LOS MODELOS.....	168
5.2.2.1. ESCALADO DINAMICO Y COMPRESION	
LOGARITMICA	168
5.2.2.2. INICIALIZACION DE LOS PARAMETROS.....	169
5.2.2.3. MULTIPLES SECUENCIAS DE OBSERVACIONES	
.....	169
5.2.2.4. INSUFICIENCIA DE DATOS DE ENTRENAMIENTO.	
SUAVIZADO	170
5.2.3. APLICACION AL RECONOCIMIENTO DEL HABLA.....	176
5.2.3.1. DESCRIPCION GENERAL DEL SISTEMA DE	
RECONOCIMIENTO.....	176
5.2.3.2. DISCRETIZACION DEL ESPACIO DE	
CARACTERISTICAS	181
5.2.3.3. UTILIZACION DE VARIAS INFORMACIONES	185
5.3. MODELOS OCULTOS DE MARKOV CONTINUOS	187
5.4. MODELOS OCULTOS DE MARKOV SEMICONTINUOS.....	191
5.5. MODELOS OCULTOS DE MARKOV CON MULTIPLE ETIQUETADO	198
6. RESULTADOS EXPERIMENTALES.....	2 0 5
6.1. BASES DE DATOS Y TIPOS DE PRUEBAS.....	206
6.1.1. RUIDO BLANCO	206
6.1.2. RUIDO DE COCHE.....	209

6.2. SISTEMA BASICO DE RECONOCIMIENTO	2 1 0
6.3. RESULTADOS OBTENIDOS CON RUIDO BLANCO.....	2 1 2
6.3.1. OPTIMIZACION DEL SISTEMA BASICO	2 1 2
6.3.1.1. PREENFASIS	2 1 4
6.3.1.2. ESTRUCTURA DEL MODELO DE MARKOV.....	2 1 6
6.3.1.3. TAMAÑO DEL DICCIONARIO DEL CUANTIFICADOR VECTORIAL.....	2 1 7
6.3.1.4. ORDEN DE PREDICCION LINEAL Y PONDERACION CEPSTRAL.....	2 1 9
6.3.2. PARAMETRIZACIONES ALTERNATIVAS.....	2 2 3
6.3.2.1. HOYWE, OHOYWE Y OYWE.....	2 2 4
6.3.2.2. MIAC	2 2 7
6.3.2.3. OSALPC Y SMC	2 3 0
6.3.2.4. TRANSFORMACION BILINEAL.....	2 4 2
6.3.3. DISTANCIAS ALTERNATIVAS.....	2 4 3
6.3.3.1. DISTANCIA DE MAHALANOBIS.....	2 4 3
6.3.3.2. DISTANCIAS DE PROYECCION.....	2 4 4
6.3.4. INCORPORACION DE ENERGIA Y PARAMETROS DINAMICOS.....	2 4 8
6.3.4.1. PARAMETROS DINAMICOS CEPSTRALES.....	2 4 8
6.3.4.2. ENERGIA Y PARAMETROS DINAMICOS	2 5 6
6.3.5. SUAVIZADO DE LOS MODELOS	2 6 1
6.3.6. MODELOS SEMICONTINUOS Y MULTIPLE ETIQUETADO	2 6 2
6.3.7. COMPARACION Y COMBINACION DE TECNICAS.....	2 6 6
6.4. RESULTADOS OBTENIDOS CON RUIDO DE COCHE.....	2 7 3
6.4.1. PARAMETRIZACION LPC.....	2 7 4
6.4.2. PARAMETRIZACION OSALPC2	2 7 9
6.4.3. COMPARACION.....	2 8 2
7. CONCLUSIONES.....	2 8 5
REFERENCIAS.....	2 9 1