

ÍNDICE GENERAL

Índice General	I
1. Introducción	1
1.1. Clasificación de los sistemas de RATM.	2
1.1.1. Con respecto al módulo de adquisición	2
1.1.2. Con respecto al módulo de parametrización	3
1.1.3. Con respecto al módulo de clasificación/reconocimiento	4
1.1.4. Con respecto al módulo de entrenamiento	5
1.2. Estado del arte	8
1.2.1. Clasificación de caracteres manuscritos aislados	9
1.2.2. Reconocimiento de palabras aisladas	10
1.2.3. Reconocimiento general de texto manuscrito	11
1.3. Sistemas comerciales	11
1.3.1. Reconocimiento de cantidades numéricas en cheques	11
1.3.2. Reconocimiento de direcciones postales	12
1.4. Objetivos de la tesis	12
2. Fundamentos Teóricos	15
2.1. Modelos ocultos de Markov	16
2.1.1. Definición de HMM continuo	17
2.1.2. Algoritmos básicos para HMMs	18
2.2. Modelos de lenguaje	23
2.2.1. n-gramas	24
2.2.2. Modelos de estados finitos	26
2.3. Métrica para la evaluación del sistema de RATM	28
2.3.1. Tipos de experimentos	29
2.3.2. SER	29
2.3.3. WER	29
3. Corpus	31
3.1. ODEC	31
3.2. IAMDB	32
4. Preproceso	37
4.1. Normalización a nivel de página	39
4.1.1. Umbralización	39
4.1.2. Reducción del ruido	44

4.1.3.	Corrección del <i>skew</i> o desencuadre	47
4.1.4.	Segmentación en líneas	54
4.2.	Normalización a nivel de texto	54
4.2.1.	Corrección del <i>slope</i>	54
4.2.2.	Corrección del <i>slant</i>	60
4.2.3.	Normalización del tamaño	74
4.3.	Resumen	79
5.	Extracción de Características	81
5.1.	Taxonomía de las características	82
5.2.	Características usadas en esta tesis	86
5.2.1.	Nivel de gris normalizado	87
5.2.2.	Derivada vertical y horizontal	87
6.	Adaptación al Escritor	91
6.1.	Técnicas de adaptación al escritor	92
6.2.	<i>Maximum Likelihood Linear Regression</i> : MLLR	93
6.2.1.	Árboles binarios de regresión	94
6.2.2.	Estimación de las matrices de transformación	97
6.3.	Experimentación	99
6.4.	Resumen	101
7.	Verificación de hipótesis	103
7.1.	Estimación de medidas de confianza	103
7.1.1.	Estimación basada en grafos de palabras	104
7.1.2.	Estimación basada en un modelo probabilístico <i>Naïve Bayes</i>	107
7.2.	Verificación de la hipótesis	109
7.3.	Medidas de evaluación	109
7.3.1.	Curvas ROC	110
7.3.2.	AROC	110
7.3.3.	CER: Tasa de error de clasificación	111
7.4.	Experimentación	112
7.5.	Resumen	118
8.	Conclusiones y Trabajos Futuros	119
8.1.	Trabajos Futuros	119
8.2.	Publicaciones relacionadas con este trabajo	120
A.	Ejemplos de salida del verificador	123
A.1.	Ejemplos de salida del verificador: corpus ODEC	123
A.2.	Ejemplos de salida del verificador: corpus IAMDB	127
	Índice General de Tablas	131
	Índice General de Figuras	133