

Contents

Abstract	page
Acknowledgements	v
Contents	vii
List of Figures	ix
List of Tables	xiii
List of Symbols	xvii
List of Abbreviations	xix
Chapter 1: Introduction	1
1.1 Challenges/Problems Addressed in this Thesis	1
1.2 Our Contributions	5
1.3 Organization of this Thesis	5
1.3.1 State of the Art: Printed and Handwritten Content Separation (Chapter 2)	5
1.3.2 State of the Art: Writer Identification (Chapter 3)	6
1.3.3 Printed and Handwritten Annotation Segmentation using Spectral Partitioning (Chapter 4)	6
1.3.4 Annotation Localization using a Supervised Model for Top-down Visual Saliency (Chapter 5)	7
1.3.5 Annotation Localization using a Weakly Supervised Model for Top-down Visual Saliency (Chapter 6)	7
1.3.6 Core Region Extraction for Off-line Unconstrained Handwritten Words (Chapter 7)	8
1.3.7 Writer Identification for Handwritten Words (Chapter 8)	8
1.3.8 Conclusion and Future Scope (Chapter 9)	8
Chapter 2: State-of-the-Art : Printed vs Handwritten	9
2.1 Introduction	9
2.2 Classification at Character Level	10
2.3 Classification at Block Level	12
2.4 Classification at Word Level	14
2.5 Classification at Connected Component Level	17
2.6 Classification at Pixel Level	22
2.7 Classification at Line Level	23
2.8 Improving the Classification Rate: Post Processing	25
2.8.1 Delaunay Triangulation	26
2.8.2 Contextual Re-labeling	26
2.9 Conclusion	27
Chapter 3: State-of-the-Art : Off-line Writer Identification	29
3.1 Introduction	29
3.2 Character/Grapheme Based Writer Identification	31
3.2.1 Feature Based Writer Identification	32
3.2.1.1 Character Feature Based Writer Identification	32
3.2.1.2 Grapheme/Connected Component Feature Based Writer Identification	33
3.2.2 Codebook Based Writer Identification	34
3.2.2.1 Character Codebook Based Writer Identification	35
3.2.2.2 Grapheme/Connected Component Codebook Based Writer Identification	36
3.2.3 Hybrid Approach for Writer Identification	37
3.3 Document Based Writer Identification	40
3.4 Word Based Writer Identification	46
3.5 Line Based Writer Identification	49
3.6 Conclusion	52
Chapter 4: Printed and Handwritten Annotation Segmentation using Spectral Partitioning	55
4.1 Introduction	55
4.2 Related Work	56
4.3 Dataset Used	58
4.4 Unsupervised Clustering	58
4.5 Our Work	60
4.5.1 Preprocessing	61
4.5.2 Feature Extraction	61
4.5.3 Segmentation	63

4.5.4	Observation and Result Analysis	64
4.6	Conclusion	65
Chapter 5: Annotation Localization using a Supervised Model for Top-down Visual Saliency		71
5.1	Introduction	71
5.2	Visual Saliency	72
	5.2.1 Bottom Up Visual Saliency	73
	5.2.2 Top Down Visual Saliency	74
	5.2.3 Hybrid Visual Saliency	74
5.3	Dataset	75
5.4	Feature Extraction	76
5.5	Experimental Protocol	77
	5.5.1 Image Representation	77
	5.5.2 Prediction Model	77
	5.5.3 Joint CRF and Dictionary Learning	79
	5.5.3.1 Computing the gradients of the loss function	80
	5.5.4 Saliency Computation	80
	5.5.5 Proposed Model Testing	81
	5.5.6 Implementation Details	81
	5.5.7 Parameter Settings	81
5.6	Experiments and Results	81
	5.6.1 Accuracy metrics for performance evaluation	82
	5.6.2 All Annotation vs Printed Text	82
	5.6.3 Category-wise Annotation vs Printed Text	83
	5.6.4 Results on Standard Datasets	84
5.7	Comparison with SVM	86
5.8	Conclusion	87
Chapter 6: Annotation Localization using a Weakly Supervised Model for Top-down Visual Saliency		
6.1	Introduction	95
	6.1.1 Computational principles for top-down saliency	96
	6.1.2 Optimal feature selection for discriminant saliency	97
	6.1.2.1 Implementation of discriminant saliency	100
	6.1.2.2 Estimating the models	101
6.2	Dataset	103
6.3	Feature Extraction	103
6.4	Saliency Threshold Details	103
6.5	Experiments and Results	103
	6.5.1 Accuracy metrics for performance evaluation	104
	6.5.2 All Annotation vs Printed Text	104
	6.5.3 Category-wise Annotation vs Printed Text	104
	6.5.4 Results on Standard Datasets	106
6.6	Comparison with SVM	106
6.7	Conclusion	107
Chapter 7: Core Region Extraction for Off-line Unconstrained Handwritten Words		117
7.1	Introduction	117
7.2	Related Work	118
7.3	Methodology	121
7.4	Experimental Details and Results	124
7.5	Conclusion	126
Chapter 8: Writer Identification for Handwritten Words		129
8.1	Introduction	129
8.2	Related Work	130
	8.2.1 Allographs features	130
	8.2.1.1 Grapheme level:	130
	8.2.1.2 Character level:	131
	8.2.1.3 Bi-gram level:	131
	8.2.2 Word-level writer identification	132
	8.2.3 Our Contribution	133
8.3	Methodology	133
	8.3.1 Feature Extraction	133
	8.3.2 Classification of segment features	134
8.4	Experiments	135
	8.4.1 Experiments at Grapheme Level:	135
	8.4.2 Experiments at Character Level:	135
8.5	Conclusions	138

Chapter 9: Conclusion and Future Scope	139
9.1 Summary of the proposed methods	139
9.2 Summary of the Experimental Results	142
9.3 Future Scope	144
References	147

