# Contents

Abstract / Resumen / Resum iii

Acknowledgments v

Contents vii

List of Figures xi

List of Tables xiii

Notation xv

Abbreviations and Acronyms xvii

## 1 Introduction

1.1 Goals of the Thesis ............................. 3
1.2 Organization of the Thesis .......................... 4

## 2 Classification Problems and Dimensionality Reduction

2.1 Statistical Pattern Classification ..................... 5
2.1.1 Review of Related Work ...................... 6
2.2 Dimensionality Reduction in Classification ................ 8
2.2.1 Review of Related Work ...................... 9

## 3 Learning Projections and Prototypes for Classification

3.1 Estimation of the 1-NN Error Probability ................. 12
3.2 Optimization Approximations .......................... 14
3.3 The LDPP Algorithm ................................ 17
3.4 Discussion ........................................... 18
3.5 Using LDPP only for Dimensionality Reduction .......... 19
3.6 Orthonormality Constraint .......................... 20
3.7 Algorithm Convergence and the Sigmoid Slope .......... 21
3.8 Distances Analyzed .................................. 21
3.8.1 Euclidean Distance .............................. 21
3.8.2 Cosine Distance .................................. 23
CONTENTS

3.9 Normalization and Learning Factors ........................................ 24
3.10 Experiments  ................................................................................ 26
  3.10.1 Data Visualization ................................................................. 27
  3.10.2 UCI and Statlog Corpora ......................................................... 30
  3.10.3 High-Dimensional Data Sets .................................................... 33
3.11 Conclusions  .................................................................................. 38

4 Modeling Variability Using Tangent Vectors  41
  4.1 Overview of the Tangent Vectors .................................................... 41
    4.1.1 Tangent Distance .................................................................. 43
    4.1.2 Estimation of Tangent Vectors ................................................. 45
  4.2 Principal Component Analysis ......................................................... 47
    4.2.1 Tangent Vectors in PCA ........................................................ 48
  4.3 Linear Discriminant Analysis ......................................................... 49
    4.3.1 Tangent Vectors in LDA .......................................................... 50
  4.4 Spectral Regression Discriminant Analysis .................................. 51
    4.4.1 Tangent Vectors in SRDA ...................................................... 53
  4.5 LDPP Using the Tangent Distances ............................................... 54
  4.6 Experiments  ................................................................................... 56
    4.6.1 Gender Recognition .............................................................. 57
    4.6.2 Emotion Recognition ............................................................ 58
    4.6.3 Face Identification ............................................................... 59
    4.6.4 LDPP Using the Tangent Distances ......................................... 61
  4.7 Conclusions  .................................................................................. 64

5 Regression Problems and Dimensionality Reduction  65
  5.1 Regression Analysis ...................................................................... 65
    5.1.1 Review of Related Work ........................................................ 66
  5.2 Dimensionality Reduction in Regression ..................................... 67
    5.2.1 Review of Related Work ........................................................ 67
  5.3 Learning Projections and Prototypes for Regression .................. 68
    5.3.1 Normalization and the LDPPR Parameters ............................ 72
  5.4 Experiments  ................................................................................... 73
    5.4.1 StatLib and UCI Data Sets ..................................................... 74
    5.4.2 High-Dimensional Data Sets ................................................... 76
  5.5 Conclusions  .................................................................................. 78

6 Ranking Problems and Score Fusion  81
  6.1 Review of Related Work ............................................................... 82
  6.2 Score Fusion by Maximizing the AUC ....................................... 83
    6.2.1 Score Normalization ............................................................. 83
    6.2.2 Score Fusion Model ............................................................... 84
    6.2.3 AUC Maximization ............................................................... 84
    6.2.4 Notes on the Implementation of the Algorithm ...................... 86
    6.2.5 Extensions of the Algorithm .................................................. 87
  6.3 Biometric Score Fusion ............................................................... 87
6.4 Estimation of Quality by Fusion ................................. 90
  6.4.1 Proposed Quality Features ................................. 91
  6.4.2 Quality Fusion Methods for Frame Selection ............. 91
  6.4.3 Experimental Results ................................. 92
6.5 Conclusions ........................................... 95

7 General Conclusions ........................................... 97
  7.1 Directions for Future Research ............................... 99
  7.2 Scientific Publications ................................... 100

A Mathematical Derivations .................................... 103
  A.1 Chapter 3 ........................................... 103
    A.1.1 Gradients of the Goal Function in LDPP ................. 103
    A.1.2 Gradients of the Euclidean Distance in LDPP ........... 103
    A.1.3 Gradients of the Cosine Distance in LDPP ............. 104
    A.1.4 Dependence of the LDPP Parameters on the Distance .... 105
    A.1.5 Normalization Compensation .......................... 106
  A.2 Chapter 4 ........................................... 106
    A.2.1 The Single Sided Tangent Distance ..................... 106
    A.2.2 Principal Component Analysis .......................... 107
    A.2.3 Linear Discriminant Analysis .......................... 108
    A.2.4 Between Scatter Matrix Accounting for Tangent Vectors 108
    A.2.5 Covariance Matrix Accounting for Tangent Vectors .... 109
    A.2.6 Gradients of the Tangent Distance in LDPP ........... 110
  A.3 Chapter 5 ........................................... 111
    A.3.1 Gradients of the Goal Function in LDPPR ............... 111
  A.4 Chapter 6 ........................................... 112
    A.4.1 Gradients of the Goal Function in SFMA ............... 112
    A.4.2 Constraints in SFMA ................................ 113

Bibliography ........................................... 115