

# List of Figures

1-1	Representation of a hyperspectral image. For each pixel in the scene, reflected light intensity is recorded in multiple spectral bands. Images acquired in one band appears gray-scale. . . . .	3
1-2	General framework for pixel-wise classification of Hyperspectral images. . . . .	5
1-3	Hughes Phenomenon: Y axis represents classification accuracy and X axis represents dimensionality. If the number of samples are constant, the classification accuracy increases till a few features and after that it actually decreases with the increase in dimensionality. .	7
1-4	The concept of principal component analysis. The data is transformed into a new dimension where the direction having maximum variance becomes first PC and the next high variance becomes second PC and so on. . . . .	12
1-5	A general framework for active learning. . . . .	15
1-6	Max-tree construction and attribute filtering considering area as attribute. . . . .	26
1-7	Attribute profile for a gray-scale image. . . . .	28
1-8	Extended attribute profile for an HSI. . . . .	29
1-9	Extended multi attribute profile for an HSI. . . . .	30
2-1	Block diagram of the proposed framework. . . . .	38
2-2	Average classification accuracy over ten runs versus the number of training samples provided by the different methods (KSC data set). .	47

## List of Figures

---

2-3	Classification maps provided by different approaches with 419 labelled samples on the KSC data set. . . . .	50
2-4	Average classification accuracy over ten runs versus the number of training samples provided by the different methods (University of Pavia data set). . . . .	51
2-5	Classification maps provided by different approaches with 407 labelled samples on the University of Pavia data set. . . . .	52
2-6	Average classification accuracy over ten runs versus the number of training samples provided by the different methods (Indian Pines data set). . . . .	53
2-7	Classification maps provided by different approaches with 948 labelled samples on the Indian Pines data set. . . . .	55
2-8	Average classification accuracy over ten runs versus the number of training samples provided by the different methods (University of Houston data set). . . . .	56
2-9	Classification maps provided by different approaches with 945 labelled samples on the University of Houston data set. . . . .	57
2-10	Average classification accuracy provided by the proposed technique varying the values of $k$ for the $k$ -means algorithm (KSC data set). . . . .	60
2-11	Average classification accuracy provided by the proposed technique by varying the values of $K$ for the $K$ -nearest neighbors algorithm (KSC data set). . . . .	61
3-1	Block diagram of the proposed architecture. . . . .	66
3-2	Classification maps of the best results obtained for KSC data set considering full feature space and the reduced feature space. . . . .	74
3-3	Classification maps of the best results obtained for University of Pavia data set considering full feature space and the reduced feature space. . . . .	78
3-4	Classification maps of the best results obtained for the Indian Pines data set considering full feature space and the reduced feature space. . . . .	81

3-5	Classification maps of the best results obtained for the University of Houston data set considering full feature space and the reduced feature space. . . . .	83
3-6	Average classification accuracy obtained using different number of features selected in the EEMAP for the Indian Pines data set. . . .	86
4-1	Steps of the proposed threshold-free attribute filtering technique considering a path from a leaf node to the root node obtained by applying depth first traversal. . . . .	92
4-2	(a) Leaf attribute functions (LAF); (b) gradient curves (GC); (c) attribute ratio curves (ARC); and (d) maximal suitability curves (MSC) obtained by analyzing three randomly selected paths from the max-tree. Each path is obtained from a randomly selected leaf node to the root node of the max-tree created by considering 1 <sup>st</sup> PC of University of Pavia data set. . . . .	93
4-3	An LAF representing the attribute values in the path from a randomly selected leaf node to the root. O is the starting position and R is the position on LAF having maximum gradient from O. . . . .	95
4-4	A synthetic tree and the filtered trees obtained by the proposed technique after detecting a node (represented with filled circle) from the path (b) A to D (c) A to E and (d) A to G.. . . .	97
4-5	Filtered images obtained by applying the state-of-the-art and the proposed filtering method to the 1 <sup>st</sup> principal component of the KSC data set by considering the <i>area of bounding box</i> attribute. The best filtered image obtained by the state-of-the-art method [36] considering (a) the 1 <sup>st</sup> threshold, (b) the 2 <sup>nd</sup> threshold and (c) the 3 <sup>rd</sup> threshold. The filtered image obtained by the proposed method after applying (d) the 1 <sup>st</sup> , (e) the 2 <sup>nd</sup> and (f) the 3 <sup>rd</sup> filtering operation. . . . .	101
4-6	Classification maps of the best results obtained by the State-of-the-art and the proposed method with 15 features for the KSC data set considering attribute Area, Perimeter, Abb, Dbb and Std. . . . .	103

4-7	Classification maps of the best results obtained by the State-of-the-art and the proposed method with 15 features for the University of Pavia data set considering attribute Area, Perimeter, Abb, Dbb and Std. . . . .	104
4-8	Filtered images obtained by applying the state-of-the-art and the proposed filtering method to the 1 <sup>st</sup> principal component of the University of Pavia data set by considering the <i>area of bounding box</i> attribute. The best filtered image obtained by the state-of-the-art method [36] considering (a) the 1 <sup>st</sup> threshold, (b) the 2 <sup>nd</sup> threshold and (c) the 3 <sup>rd</sup> threshold. The filtered image obtained by the proposed method after applying (d) the 1 <sup>st</sup> , (e) the 2 <sup>nd</sup> and (f) the 3 <sup>rd</sup> filtering operation. . . . .	105
4-9	Filtered images obtained by applying the state-of-the-art and the proposed filtering method to the 1 <sup>st</sup> principal component of the Indian Pines data set by considering the <i>area of bounding box</i> attribute. The best filtered image obtained by the state-of-the-art method [36] considering (a) the 1 <sup>st</sup> threshold, (b) the 2 <sup>nd</sup> threshold and (c) the 3 <sup>rd</sup> threshold. The filtered image obtained by the proposed method after applying (d) the 1 <sup>st</sup> , (e) the 2 <sup>nd</sup> and (f) the 3 <sup>rd</sup> filtering operation. . . . .	107
4-10	Classification maps of the best results obtained by the State-of-the-art and the proposed method with 15 features for the Indian Pines data set considering attribute Area, Perimeter, Abb, Dbb and Std. .	109
4-11	Filtered images obtained by applying the state-of-the-art and the proposed filtering method to the 1 <sup>st</sup> principal component of the University of Houston data set by considering the <i>area of bounding box</i> attribute. The best filtered image obtained by the state-of-the-art method [36] considering (a) the 1 <sup>st</sup> threshold, (b) the 2 <sup>nd</sup> threshold and (c) the 3 <sup>rd</sup> threshold. The filtered image obtained by the proposed method after applying (d) the 1 <sup>st</sup> , (e) the 2 <sup>nd</sup> and (f) the 3 <sup>rd</sup> filtering operation. . . . .	112
4-12	Classification maps of the best results obtained by the State-of-the-art and the proposed method with 15 features for the University of Houston data set considering attribute Area, Perimeter, Abb, Dbb and Std. . . . .	113

## List of Figures

---

A-1	Hyperspectral KSC image and its reference map. . . . .	121
A-2	Hyperspectral University of Pavia image and its reference map. . .	123
A-3	Hyperspectral Indian Pines image and its reference map. . . . .	124
A-4	Three band color image of the CASI University of Houston data set and its related map of available reference samples. . . . .	125

# List of Tables

1.1	Specification of some hyperspectral sensors. . . . .	3
2.1	Average overall classification accuracy ( $\overline{OA}$ ), its standard deviation ( $s$ ) and kappa accuracy obtained on ten runs for different training data sizes (KSC Data Set) . . . . .	48
2.2	Class wise average classification accuracies (%) obtained on ten runs (KSC Data Set) . . . . .	49
2.3	Average overall classification accuracy ( $\overline{OA}$ ), its standard deviation ( $s$ ) and kappa accuracy obtained on ten runs for different training data sizes (University of Pavia Data Set) . . . . .	50
2.4	Class wise average classification accuracies (%) obtained on ten runs ( University of Pavia Data Set) . . . . .	52
2.5	Average overall classification accuracy ( $\overline{OA}$ ), its standard deviation ( $s$ ) and kappa accuracy obtained on ten runs for different training data sizes (Indian Pines Data Set) . . . . .	53
2.6	Class wise average classification accuracies (%) obtained on ten runs (Indian Pines Data Set) . . . . .	54
2.7	Average overall classification accuracy ( $\overline{OA}$ ), its standard deviation ( $s$ ) and kappa accuracy obtained on ten runs for different training data sizes (University of Houston Data Set) . . . . .	56
2.8	Class wise average classification accuracies (%) obtained on ten runs (University of Houston Data Set) . . . . .	58

2.9	Obtained $Z$ -scores between the proposed and the state-of-the-art methods for all the considered data sets . . . . .	59
2.10	Computational time (in minutes) taken by the different AL methods on the considered data sets . . . . .	59
3.1	Average class-wise accuracy, average overall accuracy ( $\overline{OA}$ ), its standard deviation (std) and average kappa accuracy (kappa) obtained on ten runs by considering spectral and different spectral-spatial features (KSC data set) . . . . .	75
3.2	Average class-wise accuracy, average overall accuracy ( $\overline{OA}$ ), its standard deviation (std) and average kappa accuracy (kappa) obtained on ten runs by considering spectral and different spectral-spatial features (University of Pavia data set). . . . .	77
3.3	Average class-wise accuracy, average overall accuracy ( $\overline{OA}$ ), its standard deviation (std) and average kappa accuracy (kappa) obtained on ten runs by considering spectral and different spectral-spatial features (Indian Pines data set). . . . .	80
3.4	Average class-wise accuracy, average overall accuracy ( $\overline{OA}$ ), its standard deviation (std) and average kappa accuracy (kappa) obtained on ten runs by considering spectral and different spectral-spatial features (University of Houston data set) . . . . .	82
3.5	Details of average computational time (in seconds) required by the proposed and the Supervised-GAs techniques for different hyper-spectral data sets . . . . .	85
4.1	Classification results obtained for profiles constructed by the proposed and the state-of-the-art methods considering five different attributes (KSC). The best values are highlighted in bold face . . .	102
4.2	Classification results obtained for profiles constructed by the proposed and the state-of-the-art methods considering five different attributes (University of Pavia). The best values are highlighted in bold face . . . . .	106

4.3	Classification results obtained for profiles constructed by the proposed and the state-of-the-art methods considering five different attributes (Indian Pines). The best values are highlighted in bold face . . . . .	108
4.4	Classification results obtained for profiles constructed by the proposed and the state-of-the-art methods considering five different attributes (University of Houston). The best values are highlighted in bold face . . . . .	110
4.5	Overall accuracy (OA), average classwise-accuracy (AA) and kappa coefficient (kappa) provided by the proposed and several recent state-of-the-art spectral-spatial classification methods using standard training and test sets. (University of Houston data set) . . . .	110
4.6	Camparision of proposed method with EMAP-FS (the method proposed in Section 3). . . . .	114
4.7	Computational time in seconds required for constructing spectral-spatial profiles of size 35 using the state-of-the-art method and the proposed method . . . . .	115
A.1	Land-cover classes and the numbers of available labeled pixels on the KSC data set . . . . .	122
A.2	Land-cover classes and the numbers of available labeled pixels on the University of Pavia data set . . . . .	123
A.3	Land-cover classes and the numbers of available labeled pixels on the Indian Pines data set . . . . .	125
A.4	Land-cover classes and the numbers of available labeled pixels on the University of Houston data set . . . . .	126



# Glossary of Terms

ABD	Angle Based Diversity
AL	Active Learning
AMD	Adaptive Maximum Disagreement
AP	Attribute Profile
ARC	Attribute Ratio Curve
BT	Breaking Ties
CBD	Cluster Based Diversity
CC	Classification Confidence
CNN	Convolutional Neural Network
CP	Closing Profile
EAP	Extended Attribute Profile
ECBD	Extended Cluster Based Diversity
EEMAP	Entire Extended Multi-Attribute Profile
EMAP	Extended Multi-Attribute Profile
EMP	Extended Morphological Profile
GA	Genetic Algorithms
GC	Gradient Curve
HSI	Hyperspectral Image
JSRC	Joint SRC
KLD	Kullbach Leibler Divergence
LAF	Leaf Attribute Function
MASR	Multi-scale Adaptive SR
MCLU	Multi Class Level Uncertainty
MI	Mutual Information
MP	Morphological Profile
MRF	Marcov Random Field
MS	Margin Sampling
NMI	Normalized Mutual Information
OAA	One-Against-All
OAO	One-Against-One

OP	Opening Profile
PC	Principal Component
PCA	Principal Component Analysis
SAS	Shape adaptive sparse model
SBSDM	Superpixel based Discriminative Sparse Model
SE	Structuring Element
SR	Sparse Representation
SRC	SR-based Classification
SVM	Support Vector Machine
USRC	Unmixing + SRC

# Symbols and Notations

$a, b$	Spatial dimension
$d$	Spectral dimension
$G_i$	$i$ th spectral band
$C$	Sequence of class labels
$C_i$	$i$ th class label
$c$	Number of class label
$p_i$	$i$ th pixel
$P(C_j/p_i)$	Probability of $p_i$ to be of class $C_j$
$H$	A hyperspectral image
$X$	Data matrix without class labels
$n$	Number of pixels (patterns) ( $n = a \times b$ )
$Norm\_X$	Normalized data matrix $X$
$Cov(i, j)$	Covariance between the $i$ th and $j$ th feature
$V_i$	$i$ th eigenvector
$\lambda_i$	Eigenvalue corresponding to the eigenvector
$V$	Eigenvector matrix
$New_X$	Transformed data matrix
$U$	Set of unlabelled samples
$L$	Set of labelled samples
$EPT$	Expert for proving class label of a pixel
$\hat{x}^{KLD-max}$	Sample selected by maximizing Kullback-Leibler divergence between the distributions before and after adding the sample is maximized
$\hat{x}^{BT}$	Sample selected by maximizing Breaking Ties criterion
$H^{BAG}(x_i)$	Entropy for sample $x_i \in U$
$p^{BAG}(y_i^* = C_k x_i)$	The probability that the committee of $q$ classifier models will predict $C_k$ as the class level for sample $x_i$
$\hat{x}^{nEQB}$	Sample selected by using normalized entropy query-by-bagging criterion
$K(\cdot)$	Kernel function

$CC(x)$	Classification confidence of $x$
$Ang^{ABD}(x_i, x_j)$	The angle-based distance between $x_i$ and $x_j$ in feature space
$\delta_E(I)$	Dilation of image $I$ using structuring element $E$
$\epsilon_E(I)$	Erosion of image $I$ using starting element $E$
$E$	Structuring element
$\gamma_E(I)$	Opening of image $I$ using structuring element $E$
$\phi_E(I)$	Closing of image $I$ using structuring element $E$
$\gamma_R^E(I)$	Opening by reconstruction operation
$\phi_R^E(I)$	Closing by reconstruction operation
$R_\delta^E$	Reconstruction by dilation operation
$R_\epsilon^E$	Reconstruction by erosion operation
$MP(I)$	Morphological profile of image $I$
$t$	Number of filtering operation
$Tnp(I)$	Thinning profile for image $I$
$TkP(I)$	Thickening profile for image $I$
$AP(I)$	Attribute profile of image $I$
$r$	Number of attributes considered for constructing multi-attribute-profile
$Cl_i$	$i$ th cluster formed by $k$ -means clustering
$den(Cl_i)$	Density of the $i$ th cluster
$den(x)$	Density of a sample $x$
$Pen$	Penalty in objective function of GAs
$H_e(I)$	Entropy of image $I$ considering its gray-values
$I_i$	$i$ th filtered image in EEMAP
$\mathcal{G}_v$	Set of distinct gray-values in an image
$g$	$g \in \mathcal{G}_v$ , An specific gray-value
$P(g)$	Mass probability
$MI(I_i, I_j)$	Mutual information between images $I_i$ and $I_j$
$NMI(I_i, I_j)$	Dissimilarity based on normalized mutual information
$N_l$	Leaf node of a max-tree or min-tree
$Z$	Number of nodes between a leaf and root on the path