

# Journal of Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

## **Feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral and LIDAR data**

Behnaz Bigdeli  
Farhad Samadzadegan  
Peter Reinartz

**SPIE.**

# Feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral and LIDAR data

Behnaz Bigdeli,<sup>a,\*</sup> Farhad Samadzadegan,<sup>a</sup> and Peter Reinartz<sup>b</sup>

<sup>a</sup>University of Tehran, Department of Geomatics Engineering, Faculty of Engineering,  
North Kargar Street, P.O. Box 11365-4563, Tehran, Iran

<sup>b</sup>German Aerospace Centre (DLR), Remote Sensing Technology Institute, Department of  
Photogrammetry and Image Analysis, P.O. Box 1116, D-82230 Weßling, Germany

**Abstract.** Interest in the joint use of different data from multiple sensors has been increased for classification applications. This is because the fusion of different information can produce a better understanding of the observed site. In this field of study, the fusion of light detection and ranging (LIDAR) and passive optical remote sensing data for classification of land cover has attracted much attention. This paper addressed the use of a combination of hyperspectral (HS) and LIDAR data for land cover classification. HS images provide a detailed description of the spectral signatures of classes, whereas LIDAR data give detailed information about the height but no information for the spectral signatures. This paper presents a multiple fuzzy classifier system for fusion of HS and LIDAR data. The system is based on the fuzzy K-nearest neighbor (KNN) classification of two data sets after application of feature grouping on them. Then a fuzzy decision fusion method is applied to fuse the results of fuzzy KNN classifiers. An experiment was carried out on the classification of HS and LIDAR data from Houston, USA. The proposed fuzzy classifier ensemble system for HS and LIDAR data provide interesting conclusions on the effectiveness and potentials of the joint use of these two data. Fuzzy classifier fusion on these two data sets improves the classification results when compared with independent single fuzzy classifiers on each data set. The fuzzy proposed method represented the best accuracy with a gain in overall accuracy of 93%. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JRS.8.083509](https://doi.org/10.1117/1.JRS.8.083509)]

**Keywords:** LIDAR data; hyperspectral data; feature grouping; classifier fusion; fuzzy classification.

Paper 14194 received Apr. 4, 2014; accepted for publication Nov. 11, 2014; published online Dec. 5, 2014.

## 1 Introduction

Data and sensor fusion methods appeared as a powerful technique for improving the classification performance. Based on the current various airborne and space borne remote sensing sensors, a wide range of data can be available for the same observed site. For many real applications, the information provided by individual sensors is incomplete, inconsistent, or imprecise. Fusion of the information of different sensors can provide a better understanding of the observed site, which is not possible with single sensor.<sup>1-3</sup>

Fusion of remote sensing data can be performed at the signal, pixel, feature, and decision levels. In signal level fusion, signals from different sensors are combined to make a new signal with a better signal-to-noise ratio than that of the primary signals. Pixel level fusion consists of merging information from different images on a pixel-by-pixel basis to improve the performance of image processing tasks such as segmentation or classification. Feature level fusion includes merging features obtained from different images. In feature level fusion, features are extracted from multiple sensor observations, then combined into a concatenated feature vector and classified using a classification strategy. Decision level fusion consists of merging information after

---

\*Address all correspondence to: Behnaz Bigdeli, E-mail: [bigdeli@ut.ac.ir](mailto:bigdeli@ut.ac.ir)

the decision making step. Based on the data from each single sensor, a preliminary classification is performed. Then a fusion method combined with the outputs from the preliminary classifiers is used. The ability to fuse different types of data from different sensors is cited as the benefit of decision level fusion methods rather than other level fusion methods.<sup>3-5</sup>

During the last decade and into the near future, the number of sensors and satellites has been increasing steadily, and the coverage of the Earth in space, time, and the electromagnetic spectrum is increasing correspondingly fast. Because of these advances in remote sensing sensors and the different abilities of each sensor, sensor fusion has become an important research topic in remote sensing and has been extensively studied and applied to many areas since it usually outperforms a single classifier.

Airborne light detection and ranging (LIDAR) provides accurate height information for objects on the Earth, which makes LIDAR more and more widespread in terrain and land surveying. On the other hand, HS imaging is a relatively new technique in remote sensing that gathers hundreds of images corresponding to different spectral channels. The rich spectral information of HS images increases the capability to distinguish different physical materials, leading to the potential of a more accurate image classification. As HS and LIDAR data provide complementary information (spectral reflectance and vertical structure, respectively), a promising and challenging issue is to fuse these data in the information extraction procedure.<sup>6,7</sup>

This paper proposes a classifier fusion system based on a fuzzy K-nearest neighbor (KNN) classifier for fusion of LIDAR and HS data. First, feature extraction strategies were applied on the two data sets to generate feature spaces on HS and LIDAR data. Then a feature grouping strategy was applied on HS data to split it into groups of features. This feature grouping methodology tries to overcome the Hughes phenomenon. After that, fuzzy KNN was applied as a fuzzy supervised classification strategy on LIDAR data and on each feature group of HS data. Finally, a fuzzy classifier fusion method fused all the resultant classifiers from LIDAR and HS data.

## 2 Background

LIDAR provides high resolution horizontal and vertical spatial point cloud data, and is increasingly being used in a number of applications such as classification, feature extraction, and change detection.<sup>8,9</sup> LIDAR has the advantage of being able to create elevation surfaces that are in three-dimensions (3-D). Because of these capabilities, the classification of LIDAR data into objects such as building, tree, and road in complex area is a challenging research topic in remote sensing studies.<sup>10,11</sup>

Some of these classification researchers try to benefit from 3-D information of LIDAR data to differentiate between ground and aboveground objects. Axelsson<sup>12</sup> and Ma<sup>13</sup> classified LIDAR data into features such as buildings. They try to separate ground and nonground points from LIDAR data to extract building objects. Some of the researchers try to extract trees or forests based on the definition of special features on LIDAR data.<sup>14,15</sup> In all of the proposed classification techniques on LIDAR data, shortages of spectral information reduce the accuracy of classification.

With the development of the remote sensing imaging technology, classification of HS image is becoming more and more widespread in different applications.<sup>16-18</sup> These data cover, in most cases, a wide spectral range resulting in hundreds of data channels. Due to this volume of information, it is feasible to deal with the applications that require a precise discrimination in the spectral domain. In this context, HS images have been successfully applied for supervised classification problems that require a very precise description in spectral feature space. Extensive literature is available on the classification of HS images. The maximum likelihood or Bayesian estimation methods,<sup>17</sup> decision trees,<sup>18</sup> neural networks,<sup>19</sup> genetic algorithms,<sup>20</sup> and kernel-based techniques<sup>21,22</sup> have been widely investigated in this direction. One of the most popular classification methods is support vector machines (SVM) defined by Vapnik, a large margin-based classifier with a good generalization capacity in a small-size training set problem with a high-dimensional input space. Recently, SVMs have been successfully applied in the classification of HS remote sensing data. Camps-Valls and Bruzzone demonstrated that SVMs perform equal to or better than other classifiers in terms of accuracy on HS data.<sup>22</sup>

At the same time, HS images are usually composed of tens or hundreds of close spectral bands, which result in high redundancy and a great amount of computation time for image classification. A large number of features can become a curse in terms of accuracy if enough training samples are not available, i.e., due to the Hughes phenomenon in most of the traditional classification techniques.<sup>14</sup> This implies that the required number of training samples for supervised classification increases as a function of dimensionality. Because of the Hughes problem and a shortage of height information, conventional classification strategies often cannot overcome the mentioned problem on HS data.

Recently, new researches have focused on data fusion strategies to overcome the weaknesses of single remote sensing sensors.<sup>1,23</sup> Multisensor image fusion combines information from different sensors to obtain more information than can be derived from a single sensor. Hsu and Bruke showed that fusion of HS imaging sensors with data from other sensors can enhance the overall detection and classification performance. They fuse HS data with synthetic aperture radar data and high-resolution imaging data at the feature level to fuse spectral and spatial information.<sup>23</sup>

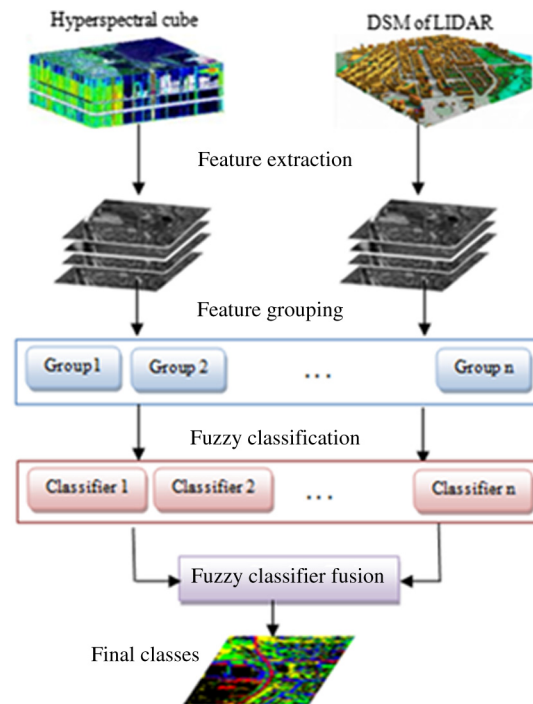
As HS images provide a detailed description of the spectral signatures of classes without any height information, fusion of this data with LIDAR data that give detailed information about the height but no information on the spectral signatures may improve classification results. Delpante investigated the potentialities of the joint use of HS and LIDAR data combined with advanced classification techniques based on SVM for forest classification.<sup>6</sup> He applied a feature selection strategy to automatically select the most effective features' subset of HS data. Then, LIDAR data were directly added to the selected HS bands for the classification. These types of research tried to select useful bands in dimension reduction techniques to overcome data redundancies. Nevertheless, the main drawback of dimension reduction techniques is related to the loss of information through the elimination of some bands.

Two recent researches, related to the classifier fusion of HS and LIDAR data, were published by Zhao et al.<sup>24</sup> and Uhlmann et al.<sup>25</sup> First, Zhao et al. applied four features: minimum noise fraction, principal component analysis, normalized difference vegetation index, and gray-level co-occurrence matrix (GLCM) on HS data. Then three classifiers, maximum likelihood classifier (MLC), SVM, and multinomial log regression, were applied on the features of HS data. On LIDAR data, they separated ground points and nonground points with an Axelsson filter and applied three-mentioned classifiers on LIDAR data. Finally, they fused all classifiers with majority voting.<sup>25</sup> Second, Uhlmann et al.<sup>24</sup> extracted some features from HS data. Then they combined each of the single features with the original HS bands and LIDAR data into five additional feature sets. In the classification step, they used SVM with a polynomial kernel to classify each feature set. Finally, they applied a majority voting to fuse the classification maps of the classifiers. Zhang et al.<sup>26</sup> tried to fuse HS and LIDAR data through a physical model for detecting objects under shadow. They developed a simple but efficient illumination correction method to remove the direct illumination component of the observed hyperspectral (HS) radiance data, and detected objects under shadows.<sup>26</sup> Liao et al. proposed a graph-based fusion method to fuse HS and LIDAR data. Their method first applies feature extraction on each individual data source, then concatenates all the features together into one stacked vector for classification. Finally, a graph-based fusion method is used to couple dimensionality reduction and data fusion of the spectral information (of original HS) and the features extracted by morphological features computed on both HS and LIDAR data together. Compared to the methods using only a single feature versus stacking all the features together, their proposed method has >10% and 5% improvements in overall classification accuracy, respectively.<sup>27</sup>

### 3 Proposed Multiple Fuzzy Classifier System on HS and LIDAR Data

A feature grouping-based multiple fuzzy classifier system for fusion of HS and LIDAR data is introduced in this paper. Figure 1 demonstrates the general structure of the proposed methodology.

First, a feature extraction strategy is applied on two data sets to generate feature spaces. Second, a feature grouping algorithm splits each data into a few feature groups. For HS data, we applied all spectral bands and extracted features in a cube. The major benefit of



**Fig. 1** A feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral (HS) and light detection and ranging (LIDAR) data.

this step is related to solving the high-dimensionality problem of HS data. Some previous researchers tried to select only useful bands in dimension reduction techniques to overcome data redundancies in HS data. Nevertheless, the main drawback of these methods is related to the loss of information through the elimination of some bands. By using a feature grouping approach, the proposed method tries to overcome this weakness with a system that enables the use of the entire high-dimensional HS image space. Third, the proposed methodology applies a fuzzy KNN classification to each feature group.

Finally after producing fuzzy classifiers, a fuzzy classifier fusion method [decision template (DT)] is applied to fuse the fuzzy classifiers. The main motivations of this proposed methodology for fusion of LIDAR and HS data are considered here:

- i. The proposed method tries to establish a system to fuse HS and LIDAR data. In this way, classification results benefit from the spectral information of HS data and height information of LIDAR data.
- ii. In the presence of mixed coverage pixels in remote sensing data, crisp classifiers may produce errors while fuzzy classifiers are not affected by such errors and, in principle, can produce a classification that is more accurate than any crisp classifier. The proposed method tries to apply the fuzzy classifier fusion concept for fusion of HS and LIDAR data.
- iii. Because of the higher performance of multiple classifiers rather than single classifiers, our proposed method benefits from this concept in order to fuse HS and LIDAR data.

### 3.1 Step 1: Feature Extraction on LIDAR and HS Data

The main step of the classification process on LIDAR and HS data is an extraction of the features from the data set. These features must contain useful information to discriminate between different regions of the surface. On LIDAR data, we have used different textural and topographical features. All types of these features on LIDAR data are represented in Table 1.

#### 3.1.1 Gray-level co-occurrence matrix

GLCMs are one of the earliest techniques used for image texture analysis. Let  $I$  be a given gray scale image. Let  $N$  be the total number of gray levels in the image. The GLCM defined by

**Table 1** Different features on light detection and ranging (LIDAR) data.

		Name	Formulation
Texture features	GLCM	Entropy	$Ent = \sum_{i,j=0}^{N-1} P_{ij} \times (-\ln P_{ij})$
		Correlation	$corr = \sum_{i,j=0}^{N-1} [(i - \mu_i)(j - \mu_j) / \sqrt{(\sigma_i^2)(\sigma_j^2)}]$
		Contrast	$Cont = \sum_{i,j=0}^N P_{ij}(i - j)^2$
		Mean	$Mean_i = \sum_{i,j=0}^{N-1} i \times P(i, j)$
		Standard deviation	$variance_i = \sum_{i,j=0}^{N-1} P(i, j) \times (i - Mean_i)^2$
		Homogeneity	$H = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1 + (i-j)^2}$
		Dissimilarity	$Diss = \sum_{i,j=0}^{N-1} P_{ij} i - j $
		Second moment	$M = \sum_{i,j=0}^{N-1} P_{ij}^2$
	Variogram	Semivariogram	$\gamma_k(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{DN_k(x_i) - DN_k(x_i + h)\}^2$
		Radogram	$\gamma_k(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \sqrt{ DN_k(x_i) - DN_k(x_i + h) }$
		Madogram	$\gamma_k(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)}  DN_k(x_i) - DN_k(x_i + h) $
Topography features	Slope		
	Aspect		
	nDSM		$nDSM = LidarDSM - M$
	Roughness		$Ra = 1/N \cdot \sum_{n=1}^N  h_n $

Haralick is a square matrix  $G$  of order  $N$ , where the  $(i, j)$  then try of  $G_d(i, j)$  represents the number of occasions that a pixel with intensity  $i$  is adjacent to a pixel with intensity  $j$ .<sup>28</sup>

The normalized co-occurrence matrix is obtained by dividing each element of  $G$  by the total number of co-occurrence pairs in  $G$ . The adjacency can be defined to take place in each of the four directions (horizontal, vertical, left, and right diagonals). The Haralick texture features are calculated for each of these directions of adjacency:<sup>26</sup>

$$GN_d(i, j) = \frac{1}{N} G_d(i, j). \quad (1)$$

### 3.1.2 Variogram

Variance operators are one of the useful statistical methods for texture analysis based on the differences between the pixels in a local neighboring window. A semivariogram as a geo-statistical operator is also in the family of variance operators. The operation of the semivariogram is based on the directional distances between each pixel and its neighboring pixels in a local window

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [I(x_i) - I(x_i + h)]^2. \quad (2)$$

Table 2 represents some of the spectral features on HS data.

## 3.2 Step 2: Feature Grouping on HS Data

HS images are usually composed of tens or hundreds of close spectral bands, which result in high redundancy (Hughes phenomenon) and a great amount of computation time for image classification. A feature grouping strategy tries to split the entire high-dimensional HS space into



**Table 2** Different spectral features on hyperspectral (HS) data.

Name	Formulation
Normalized difference vegetation index	$NDVI = (\rho_{NIR} - \rho_{RED}) / (\rho_{NIR} + \rho_{RED})$
Simple ratio	$SR = \rho_{NIR} / \rho_{RED}$
Enhanced vegetation index	$EVI = 2.5 \left( \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6\rho_{RED} - 7.5\rho_{BLUE} + 1} \right)$
Atmospherically resistant vegetation index	$ARVI = \frac{\rho_{NIR} - (2\rho_{RED} - \rho_{BLUE})}{\rho_{NIR} + (2\rho_{RED} - \rho_{BLUE})}$
Sum green index	Mean of 500 to 600 nm of spectrum
Red edge normalized difference vegetation index	$NDVI_{705} = (\rho_{750} - \rho_{705}) / (\rho_{750} + \rho_{705})$
Modified red edge simple ratio index	$mSR_{705} = (\rho_{750} - \rho_{445}) / (\rho_{705} - \rho_{445})$
Modified red edge normalized difference vegetation index	$mNDVI_{750} = (\rho_{750} - \rho_{705}) / (\rho_{750} + \rho_{705} - 2\rho_{445})$
Vogelmann red edge index 1	$VOG1 = \rho_{740} / \rho_{720}$
Vogelmann red edge index 2	$VOG2 = (\rho_{734} - \rho_{747}) / (\rho_{715} + \rho_{726})$
Red edge position index	
Photochemical reflectance index	$PRI = (\rho_{531} - \rho_{570}) / (\rho_{531} + \rho_{570})$
Structure insensitive pigment index	$SIPi = (\rho_{800} - \rho_{445}) / (\rho_{800} + \rho_{680})$
Red green ratio index	
Plant senescence reflectance index	$PSRI = (\rho_{680} - \rho_{500}) / \rho_{750}$
Carotenoid reflectance index 1	$CRI1 = (1/\rho_{510}) - (1/\rho_{550})$
Carotenoid reflectance index 2	$CRI2 = (1/\rho_{510}) - (1/\rho_{700})$
Anthocyanin reflectance index 1	$ARI1 = (1/\rho_{550}) - (1/\rho_{700})$
Anthocyanin reflectance index 2	$ARI2 = \rho_{800}[(1/\rho_{550}) - (1/\rho_{700})]$
Water band index	$WBI = \rho_{900} / \rho_{970}$

a few band/feature groups for classification by which it can overcome the Hughes phenomenon or the curse of dimensionality.

Feature grouping algorithms in some researches were applied as the premiere step of feature selection methods on HS data. This paper applies K-means as a feature grouping strategy and spectral angle mapping (SAM) as the distance measure. SAM is a widely used spectral similarity metric in remote sensing, especially for HS data. It measures the spectral similarity by finding the angle between the spectral signatures of two pixel vectors  $s_i$  and  $s_j$

$$SAM(s_i, s_j) = \cos^{-1}(s_i \cdot s_j / \|s_i\| \cdot \|s_j\|). \quad (3)$$

### 3.3 Step 3: Fuzzy Classification Based on Fuzzy KNN

The crisp nearest neighbor classification rule assigns an input sample vector  $y$ , which is of unknown classification, to the class of its nearest neighbor.<sup>29</sup> The most common distance function in KNN classification is a Euclidean distance:

$$d(x, y) = \left( \sum_{i=1}^b (x_i - y_i)^2 \right)^{1/2}. \quad (4)$$

The fuzzy KNN procedure is also a classification algorithm for which the results differ from the crisp version. The fuzzy KNN algorithm assigns class membership to a sample vector rather

than assigning the vector to a particular class. The basis of the algorithm is to assign membership as a function of the vector distance from its KNN and those neighbor memberships in the possible classes. Let  $w = \{x_1, x_2, \dots, x_n\}$  be the set on  $n$  labeled samples. Also let  $\mu_i(x)$  be the assigned membership of the vector  $x$  and  $\mu_{ij}$  be the membership in the  $i$ 'th class of the  $j$ 'th vector of the labeled sample set

$$\mu_i(x) = \frac{\sum_{j=1}^k \mu_{ij}(1/\|x - x_j\|^{2/(m-1)})}{\sum_{j=1}^k (1/\|x - x_j\|^{2/(m-1)})}. \quad (5)$$

As seen in Eq. (5), the assigned memberships of  $x$  are influenced by the inverse of the distances from the nearest neighbors and their class memberships.<sup>29</sup> The inverse distance serves to weight a vector's membership more if it is closer and less if it is farther from the vector under consideration. The variable  $m$  determines how heavily the distance is weighted when calculating each neighbor's contribution to the membership value.

### 3.4 Step 4: Fuzzy Classifier Fusion Based on DT

Multiple classifier systems (MCSs) are successfully applied on various types of data to improve single classifiers results. MCS can improve classification accuracy in comparison to a single classifier by combining different classification algorithms or variants of the same classifier.<sup>30</sup> In such systems, a set of classifiers is first produced and then combined by a specific fusion method. The possible ways of combining the outputs of the  $L$  classifiers in an MCS depend on what information can be obtained from the individual members. Kuncheva distinguishes between two types of classifier outputs which can be used in classifier combination methods. The first types are classifiers that produce crisp outputs. In this category, each classifier only outputs a unique class and a vector of classes is finally produced for each sample. The second type of classifier produces a fuzzy output which means that, in this case, the classifier associates a confidence measurement for each class and finally produces a vector for every classifier and a matrix for the ensemble of the classifier.<sup>28</sup>

The output of the fuzzy classifiers that stipulates which class an input pattern belongs to is arranged in a matrix form which is defined as a decision profile (DP) matrix. The DP matrix for  $L$  classifiers and  $C$  classes is shown in Fig. 2. Fuzzy classifier fusion methods applied different fusion strategies on DP to fuse the results of the fuzzy classifier ensemble. In fuzzy MCSs, some methods calculate the support for different classes using the corresponding column of the DP matrix regardless of the support for the other classes. This type of fusion methods that use the DP matrix in a class by class manner is called class-conscious (CC). In the CC group, various fuzzy aggregation methods such as maximum, minimum, product, sum, and mean can be applied. The alternative group is known as class-indifferent (CI). The former uses the context of the DP matrix, i.e., recognizes that a column corresponds to a class, but disregards part of the information with respect to rest of the classes. The CI methods use the whole DP matrix but disregard its context. In this paper, we applied DT as one of the CI fuzzy classifier fusion strategies. The idea of the DTs is to remember the most typical DP for each class  $w_j$ , called the decision template,  $DT_j$ , and then compare it with the current  $DP(x)$  using some similarity measure  $S$ . The closest match will be labelled  $x$ . The DT method is based on two steps:

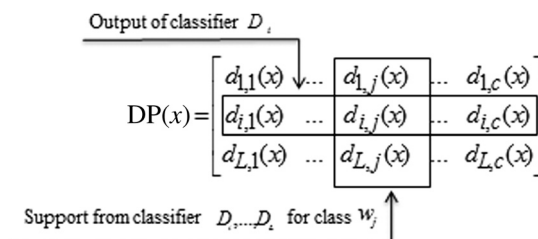


Fig. 2 Decision profile in fuzzy classifier fusion system.



Step 1: Training. For  $j = 1, \dots, C$  calculate the mean of the  $DP(x)$  of all members of  $w_j$  from the data set  $Z$ . Call the mean a  $DT_j$

$$DT_j = \frac{1}{N_j} \sum_{\substack{z_k \in w_j \\ z_k \in Z}} DP(z_k), \quad (6)$$

where  $N_j$  is the number of element of  $Z$  from  $w_j$ .

Step 2: Operation. Given the input  $x$ , calculate  $DP(x)$ , calculate the similarity measure  $S$  between  $DP(x)$  and  $DT_j$

$$\mu_j(x) = S(DP(x), DT_j), \quad j = 1, \dots, C. \quad (7)$$

The measure of similarity is based on the squared Euclidean distance  $[DT(E)]$ , so the ensemble support for class  $w_j$  is

$$\mu_j(x) = 1 - \frac{1}{L \times C} \sum_{i=1}^L \sum_{k=1}^C [DT_j(i, k) - d_{i,k}(x)]^2, \quad (8)$$

where  $DT_j(i, k)$  is the  $(i, k)$ 'th entry in  $DT_j$ .<sup>28</sup>

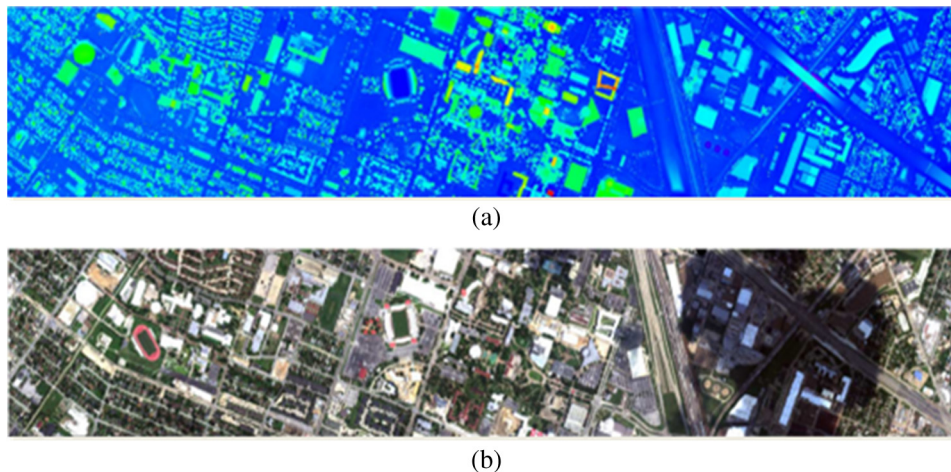
## 4 Experimental Results

### 4.1 Data Set

An HS image and an LIDAR-derived digital surface model (DSM), both with a spatial resolution of 2.5 m, have been utilized to evaluate the proposed method (Fig. 3).

The HS image has 144 spectral bands in the 380 to 1050 nm region. The data sets were captured over the University of Houston campus and the neighboring urban area and were acquired on June 22, 2012 by the NSF-funded Centre for Airborne Laser Mapping. The average heights of the sensors above ground were 2000 and 5500 ft for LIDAR and HS data, respectively.

We try to select 60% of each of 15 different land cover classes in Table 3 as training samples and the remaining as testing samples. These samples were selected from different areas of the images. They are spatially disjointed.



**Fig. 3** Data sets: (a) LIDAR-derived digital surface model (DSM) and (b) HS data over Houston campus.

**Table 3** Houston University lands cover classes and available reference samples.

ID	Class name	Reference number
1	Grass-healthy	198
2	Grass-stressed	190
3	Grass-synthetic	192
4	Tree	188
5	Soil	186
6	Water	182
7	Residential	196
8	Commercial	191
9	Road	193
10	Highway	191
11	Railway	181
12	Parking lot 1	192
13	Parking lot 2	184
14	Tennis court	181
15	Running track	187

## 4.2 Results

The first step of the presented methodology was designed to independently produce feature spaces on HS and LIDAR data. In the case of LIDAR data, all of the textural and topographical features in Table 1 were applied on data to generate the feature space. Figure 4 illustrates some of these features on the LIDAR data, and Fig. 5 demonstrated some of the spectral features on HS image.

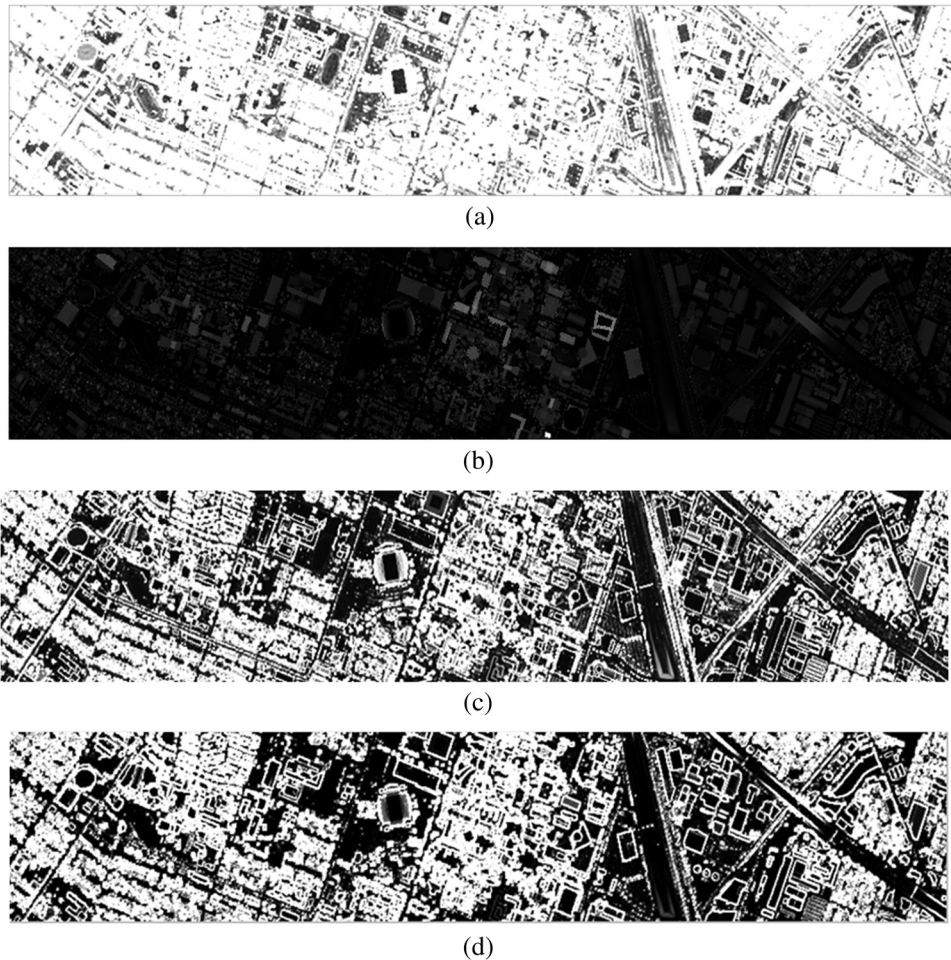
In order to compare between fuzzy and crisp classification, crisp SVM and fuzzy KNN were applied on each data. We have used error matrices and overall accuracies (OAs) of classification results as the main evaluation method for interpreting the quality of the classifiers.

Based on the further analysis in Table 4, comparison of OA values revealed that the use of the fuzzy KNN classifier is superior to the crisp SVM for both of data. For example, as can be seen in Table 4, the OAs of fuzzy KNN and crisp SVM classifiers on HS data are, respectively, 90.44% and 87.6%. The reasonable cause for this improvement is related to the ability of fuzzy classifier systems for the classification of mixed coverage pixels of HS data.

After definition of the feature spaces on LIDAR and HS data, K-means (SAM) is applied as a feature grouping strategy to group the features of HS data into some groups. (The group number is selected based on the spectral signature of the data set.) Table 5 demonstrates the results of feature grouping on HS data. In order to select the optimum features, a particle swarm optimization algorithm with OA as the objective function is applied on LIDAR data. From 16 features on LIDAR (15 features in Table 1 and original DSM), 10 optimum features are selected. Then fuzzy KNN was applied on each feature group of HS data as a fuzzy classifier. Table 6 shows the OAs of this ensemble of classifiers. Also a fuzzy KNN classifier was utilized on the optimum features of LIDAR data.

After producing an ensemble of classifiers, DT has been applied as a fuzzy classifier fusion approach for the fusion of 11 classifiers (10 for HS data and one for LIDAR data). Table 7 represents the OA results.

Based on these results, the proposed classifier fusion on HS and LIDAR data improves the results of the independent classifiers on each dataset. Furthermore, classifier fusion based on DT represented the best accuracy with a gain in OA of 93.2%.

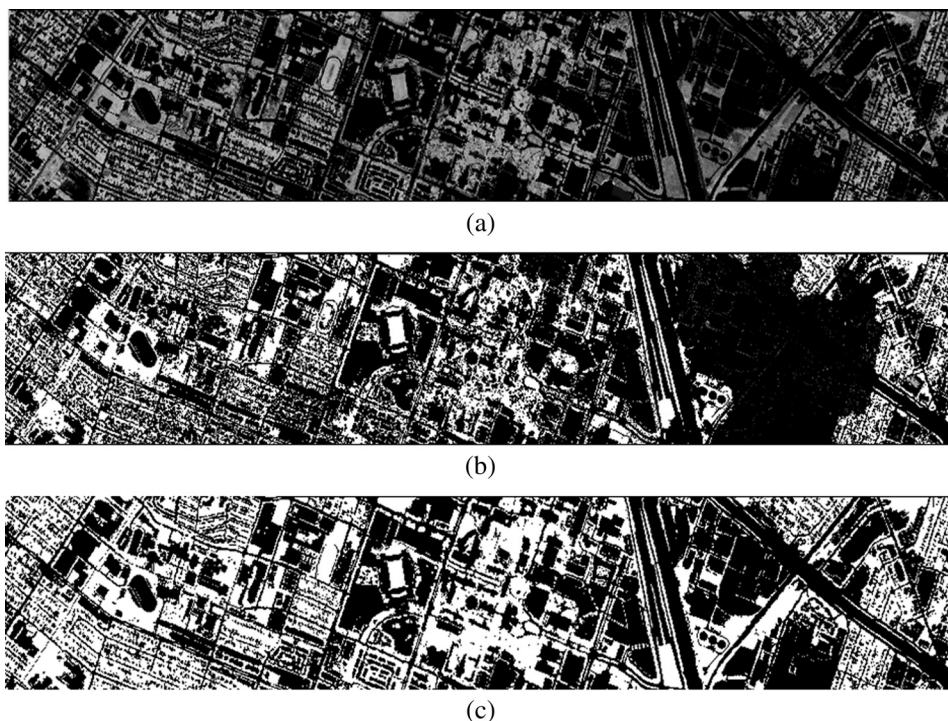


**Fig. 4** Some features of LIDAR data: (a) roughness, (b) nDSM, (c) gray-level co-occurrence matrix (std), and (d) variogram.

The modifications that resulted from application of the proposed methodology for all classes can be seen in Fig. 6. The analysis of this figure shows that for some classes such as grass-healthy, grass-stressed, and grass-synthetic, HS data outperforms LIDAR in terms of classification accuracy and for some classes such as commercial and residential, HS data outperform LIDAR data. These results confirm the superior performances of the fuzzy fusion method for most of the classes. A reasonable explanation of these results is that the height information of LIDAR data could highlight 3-D objects while the rich spectral information of HS data could be useful in extracting spectral classes such as grass, soil, and water. Figure 7 demonstrates the classification map of the fuzzy classifier fusion strategy on HS and LIDAR data.

### 4.3 Discussion

In this paper, the performance of a feature grouping based multiple fuzzy classifier system for fusion of HS and LIDAR data is assessed. The first objective of the proposed methodology concerns the effectiveness of definition of a data fusion methodology to fuse HS and LIDAR data. HS images provide a detailed description of the spectral signatures of classes but no information on the height of ground covers, whereas LIDAR data give detailed information about the height but no information on the spectral signatures. Consequently, the elevation information of LIDAR was very effective for the separation of species with similar spectral signatures but different mean heights. Also the spectral information of HS data was very effective for discrimination of similar elevation classes but different spectral information. The second objective of the proposed methodology concerns the effectiveness of feature



**Fig. 5** Some features of HS data: (a) normalized difference vegetation index, (b) red edge normalized difference vegetation index, and (c) water band index.

**Table 4** Accuracy assessment of crisp and fuzzy classifiers.

	SVM	FKNN
HS	87.69	90.446
LIDAR	64.4	70.22

**Table 5** Feature grouping results based on k-means (SAM) on HS data.

Groups	1	2	3	4	5
Features	1 to 2, 145 to 164	3 to 11, 120 to 127	12, 114 to 119, 128 to 131	13 to 18 57 to 72	19 to 34, 49
Groups	6	7	8	9	10
Features	35 to 56	73 to 77	81 to 83, 87 to 90	132 to 144 96 to 100 78, 85, 113	101 to 112 91 to 95 79 to 80 84, 86

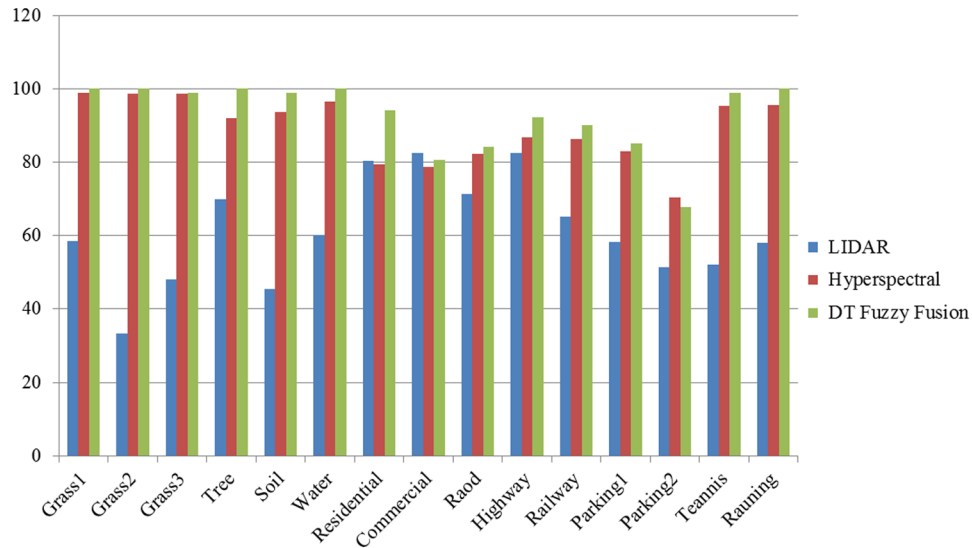
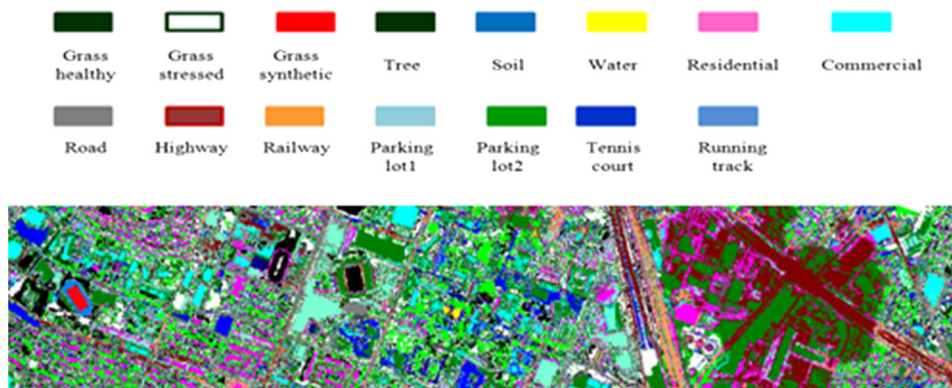
**Table 6** Classification results of different feature groups of HS data.

Groups	Feature group#1	Feature group#2	Feature group#3	Feature group#4	Feature group#5	Feature group#6	Feature group#7	Feature group#8	Feature group#9	Feature group#10
OA	55.16	85.08	82.53	87.06	88.52	87.49	73.03	73.76	82.78	80.66



**Table 7** Results of different classification strategies.

	LIDAR	HS	Fusion
OA	70.22	90.4	93.2

**Fig. 6** Comparison of class accuracies between different classification strategies.**Fig. 7** Classification map of proposed multiple classifier system on HS and LIDAR data.

grouping to solve the high-dimensionality problem (Hughes phenomenon) of HS data. Some previous researchers tried to select useful features based on dimension reduction techniques to overcome data redundancies. Nevertheless, the main drawback of these techniques is related to the loss of information through the elimination of some bands. By using feature grouping approaches, the proposed method tries to overcome this problem with a system that enables the use of the entire high-dimensional HS image space. The third objective of the proposed methodology concerns the high performance of the fuzzy classification of HS and LIDAR data. The reasonable cause for this ability is related to the mixed information in pixels and the complexities of remote sensing data. The fourth objective of the proposed method is related to the benefits of an ensemble of classifiers in comparison to single classifiers. The resulting fused classifier is generally more accurate than any of the individual classifiers that make up the ensemble. Furthermore, because of the natural differences between HS and LIDAR data, a decision level fusion based on the ensemble of classifiers could be more accurate than pixel level data fusion strategies.

The usability of the measurement setup was assessed with a coregistered HS and LIDAR-derived DSM. Based on the results, we conclude that fusion of classifiers for HS and LIDAR data could improve classification accuracy. Based on further analysis, comparison of the OA and kappa values revealed that the use of fuzzy classifier fusion based on feature grouping is superior to the single classification of each data set.

## 5 Conclusions

This paper presents a classifier fusion strategy for fusion of HS and LIDAR data based on a feature grouping strategy. We focus on three main improvements. First, the proposed method tries to establish a classifier fusion system to fuse HS and LIDAR data. In this way, classification results benefit from the spectral information of HS data and height information of LIDAR data. Second, the proposed method uses a feature grouping strategy to split HS data into some feature groups. By using a feature grouping approach, the proposed method tries to overcome the Hughes phenomenon (curse of dimensionality) with a system that enables the use of the entire high-dimensional HS image space rather than dimension reduction techniques. Third, the proposed method fused HS and LIDAR data at the decision level through an MCS. Decision level fusion is less sensitive to registration than pixel and feature level fusion. Further studies will focus on decision fusion methods and feature grouping strategies.

## Acknowledgments

The authors would like to thank the HS Image Analysis group and the NSF Funded Center for Airborne Laser Mapping at the University of Houston for providing the data sets used in this study, and the IEEE GRSS Data Fusion Technical Committee for organizing the 2013 Data Fusion Contest.

## References

1. G. Simone et al., "Image fusion techniques for remote sensing applications," *Inf. Fusion* **3**(1), 3–15 (2002).
2. C. Pohl and L. Van Genderen, "Multisensor image fusion in remote sensing: concepts, methods, and applications," *Int. J. Remote Sens.* **19**(5), 823–854 (1998).
3. P. Du et al., "Information fusion techniques for change detection from multi-temporal remote sensing images," *Inf. Fusion* **14**(1), 19–27 (2013).
4. J. Dong et al., "Advances in multi-sensor data fusion: algorithms and applications," *J. Sens.* **9**(10), 7771–7784 (2009).
5. Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens.* **6**(1), 657–661 (2004).
6. M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of HS and LIDAR remote sensing data for classification of complex forest areas," *IEEE Trans. Geosci. Remote Sens.* **46**(5), 1416–1427 (2008).
7. A. Swatantran et al., "Mapping biomass and stress in the Sierra Nevada using Lidar and HS data fusion," *Remote Sens. Environ.* **115**(11), 2917–2930 (2011).
8. R. Brenan and T. L. Webster, "Object-oriented land cover classification of Lidar-derived surfaces," *Can. J. Remote Sens.* **32**(2), 162–172 (2006).
9. L. C. Chen, "Detection of building changes from aerial images and light detection and ranging (LIDAR) data," *J. Appl. Remote Sens.* **4**(1), 041870 (2010).
10. M. Bartels and H. Wei, "Rule-based improvement of maximum likelihood classified LIDAR data fused with co-registered bands," in *Proc. of the Annual Conf. of the Remote Sensing and Photogrammetry Society*, 5–8 September, Cambridge, United Kingdom, pp. 1–9(2006).
11. S. Lodha et al., "Aerial Lidar data classification using support vector machines (SVM)," in *Third International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 567–574, Chapel Hill, North Carolina (2006).



12. P. Axelsson, "Processing of laser scanner data—algorithms and applications," *ISPRS J. Photogramm. Remote Sens.* **54**(2-3), 138–147 (1999).
13. R. Ma, "DEM generation and building detection from Lidar data," *Photogramm. Eng. Remote Sens.* **71**(7), 847–854 (2005).
14. J. Li, B. Hua, and T. Noland, "Classification of tree species based on structural features derived from high density Lidar data," *Agric. For. Meteorol.* **171**(1), 104–114 (2013).
15. H.O. Ørka, E. Næsset, and O. M. Bollands, "Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data," *Remote Sens. Environ.* **113**(6), 1163–1174 (2009).
16. S. Li et al., "An effective feature selection method for HS image classification based on genetic algorithm and support vector machine," *Knowledge Based Syst.* **24**(1), 40–48 (2011).
17. X. Jia, "Simplified maximum likelihood classification for HS data in cluster space," in *IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS '02)*, Toronto, Ontario, Canada, Vol. 5, pp. 2578–2580 (2002).
18. P. K. Goel et al., "Classification of HS data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agric.* **39**(2), 67–93 (2003).
19. F. Del Frate et al., "Use of neural networks for automatic classification from high-resolution images," *IEEE Trans. Geosci. Remote Sens.* **45**(4), 800–809 (2007).
20. C. Vaiphasa, "Innovative genetic algorithm for HS image classification," in *Proc. Int. Conf. Map Asia*, Kuala Lumpur, Malaysia (2003).
21. K. L. Müller et al., "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.* **12**(2), 181–202 (2001).
22. G. Camps-Valls and L. Bruzzone, "Kernel-based methods for HS image classification," *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1351–1362 (2005).
23. S. Hsu and H. Burke, "Multisensor fusion with HS imaging data detection and classification," *Lincoln Lab. J.* **14**(1), 145–159 (2003).
24. B. Zhao et al., "Outcome of the "Best Classification Challenge" of the 2013 IEEE GRSS data fusion contest: fusion of hyperspectral and LiDAR data," in *IEEE Int. Geosciences and Remote Sensing Symp.*, Melbourne, Australia (2013).
25. S. Uhlmann, S. Kiranyaz, and A. Yildirm, "Outcome of the "Best Classification Challenge" of the 2013 IEEE GRSS data fusion contest: fusion of hyperspectral and LiDAR data," in *IEEE Int. Geosciences and Remote Sensing Symp.*, Melbourne, Australia (2013).
26. A. Zhang et al., "Detecting objects under shadows by fusion of hyperspectral and Lidar data: a physical model approach," in *Proc. 5th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens. (WHISPERS)*, pp. 1–4 (2013).
27. W. Liao et al., "Feature fusion of hyperspectral and LiDAR data for classification of remote sensing data from urban area," in *5th Workshop of the EARSeL Special Interest Group on Land Use and Land Cover: Frontiers in Earth Observation for Land System Science*, Berlin, Germany (2014).
28. R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621 (1973).
29. J. Keller, M. Gray, and J. Givens, "A fuzzy k-nearest neighbour algorithm," *IEEE Trans. Syst. Man Cybern.* **SMC-15**, 580–585 (1985).
30. L. Kuncheva, *Combining Pattern Classifiers Methods and Algorithms*, John Wiley & Sons, Inc., Hoboken, New Jersey (2004).

**Behnaz Bigdeli** received her BSc degree in 2007 from KNT University at geomatics engineering. She received her MSc degree from Tehran University in 2010 at photogrammetry. She then started her PhD research in 2010 in fusion of hyperspectral (HS) and LIDAR data and got her PhD degree in 2014. Her main interests are in data fusion of space borne and airborne data, HS and LIDAR processing, decision fusion methods, support vector machine, and Bayesian theory.

**Farhad Samadzadegan** received his Bsc degree in 1990 from the University of Tehran in geomatics engineering. He received his Msc degree from KNT University in 1993 in photogrammetry. He then started his PhD research in 1997 in automatic 3-D building recognition and

reconstruction based on artificial intelligence techniques and got his PhD degree in 2001. He has been the head of Sensors and Platforms in Geomatics Research Center of Tehran University, Iran.

**Peter Reinartz** received his diploma (Dipl.-Phys.) in theoretical physics in 1983 from the University of Munich and his PhD degree (Dr.-Ing) in civil engineering from the University of Hannover in 1989. His dissertation is on optimization of classification methods for multi-spectral image data. He is a department head of the Department of Photogrammetry and Image Analysis at the German Aerospace Centre (DLR), Remote Sensing Technology Institute (IMF), and holds a professorship for geoinformatics at the University of Osnabrück.