International Journal of Innovative Computing, Information and Control
Volume **8**, Number **8**, August **2012**

ICIC International ©2012 ISSN 1349-4198
pp. 5455–5472

FUSION OF LIDAR HEIGHT DATA FOR URBAN FEATURE CLASSIFICATION USING A HYBRID METHOD

Shiahn-Wern Shyue[1], Ming-Jer Huang[2,*], Liang-Hwei Lee[3] and Chih-Chung Kao[4]

[1]Department of Marine Environment and Engineering
National Sun Yat-Sen University
No. 70, Lien-hai Rd., Kaohsiung 804, Taiwan
swshyue@faculty.nsysu.edu.tw

[2]Department of Applied Science
Naval Academy
No. 669, Junxiao Rd., Kaohsiung 813, Taiwan
*Corresponding author: mjhuang@cna.edu.tw

[3]Department of Civil Engineering
National Kao-Hsiung University of Applied Sciences
No. 415, Chien Kung Rd., Kaohsiung 80778, Taiwan
lhlee@cc.mail.kuas.edu.tw

[4]Department of Information Management
Fortune Institute of Technology
14FL, No. 62, Chien-An St., San-Min Dist., Kaohsiung 831, Taiwan
kcc7879@center.fotech.edu.tw

Received April 2011; revised August 2011

Abstract. *There is growing interest in three-dimensional (3D) modeling of urban areas in various fields, and in response many efficient light detection and ranging (LIDAR) systems have been developed and integrated with new types of digital aerial cameras. Intensive research has also been undertaken on the classification of natural and human-made structures from LIDAR data and high-resolution imagery, with the aim of constructing detailed 3D urban models. Here, a new, simplified hybrid classification algorithm is presented for efficient urban feature extraction. It can handle topographic and image data without sacrificing accuracy compared with artificial neural network (ANN) and support vector machine (SVM) methods. Numerical and graphical analyses are conducted to assess the importance of LIDAR height information, and this is then used to resolve classification confusion between high and low objects by applying two-height-level knowledge-based rules. A maximum likelihood classification method is used to sort input features into low/high categories, and these results are then merged. The proposed hybrid classification method revealed the utility of considering LIDAR height and demonstrated higher efficiency and equivalent or improved accuracy compared with traditional ANN/SVM methods.*

Keywords: Artificial neural network (ANN), Hybrid classification method, Maximum likelihood classification (MLC), Support vector machine (SVM), Two-level height rule-based scheme

1. **Introduction.** There is growing interest in obtaining three-dimensional (3D) information on urban areas for various applications, such as urban planning, cyber city construction, and urban public works. This demand has spurred research into digital surface models (DSMs) that can automatically classify and extract natural and human-made structures to construct detailed 3D models [1-3]. At the same time, many manufacturers

are now producing highly efficient light detection and ranging (LIDAR) systems that integrate new types of digital aerial cameras, such as the digital mapping camera (DMC) or ADS40 camera, and provide both red-green-blue (RGB) and near-infrared (NIR) images [1]. While large number of data can be acquired efficiently, the ultimate goal is to provide highly efficient urban classification in combination with LIDAR height data to fulfill 3D modeling demands.

Classifying LIDAR data and aerial imagery in an urban area is a complex, multi-source problem. Many studies have applied artificial neural network (ANN) or support vector machine (SVM) approaches to multi-source data classification and have reported significantly improved classification accuracy [4-8]. These methods may be useful when data types differ in their statistical distribution in one stacked dataset. The ANN and SVM techniques have been used to overcome the limitations of traditional parametric algorithms for multi-source data classification, resulting in higher accuracies. However, these techniques also have some drawbacks. For example, an ANN has a high computational cost, and it does not have a straightforward approach for identifying the optimal number of network layers, training rate, and training momentum. Using an SVM, the kernel function and penalty parameters can be identified by a grid search involving various parameters, which are tested to find the combination that yields the highest accuracy based on cross-validation [9,10]. In both methods, deriving these optimal parameters and adjusting the classifier for individual cases are time-consuming. These solutions merely meet the accuracy requirement for urban classification.

Expert systems or knowledge-based classification methods may also be used to improve classification accuracy and better resolve heterogeneous multi-source data [11-13]. Knowledge-based classification is derived from expert rules based on features that can be classified from the data themselves. For urban feature extraction, many studies have used LIDAR-derived normalized DSMs (nDSMs) as additional channels in combination with multispectral images, in an integrated classification or Dempster-Shafer fusion method. This approach enables the extraction of urban features such as buildings, roads, bare soil, and grass-covered areas. Classification methods have also been used as pre-processing steps in urban building extraction and construction [2,14]. For example, in a previous study, maximum likelihood classification (MLC) was used to quantitatively and qualitatively evaluate how adding LIDAR nDSMs and intensity to color or multispectral aerial imagery data may effect urban classification; the results demonstrated that nDSMs significantly improved classification accuracy [15]. The results of [15] and other studies suggest that the integration of nDSMs and aerial images may have great potential for urban classification.

In the present study, we propose a hybrid algorithm for urban classification that performs efficient classification with the same accuracy as that of the ANN and SVM approaches. First, we applied hybrid classification approaches to topographic and image datasets using a two-level-height rule-based scheme. Then we performed MLC to manipulate the image data at each level. We evaluated the feasibility of the proposed hybrid algorithm in terms of computational efficiency and overall accuracy. The details of the methodology are introduced in the following section. The experimental results are presented and compared with the ANN and SVM approaches in Section 3. The conclusions are presented in Section 4.

2. **Hybrid Classification Method.**

Most previous classifiers are based on single-classification methods, even when handling different types of data. We developed a multi-source classifier based on different algorithms to manipulate different data types, such as topographic (LIDAR height) and digital aerial image data. Here, we present the basic process,

which uses a hybrid algorithm for hierarchical classification. First, it was necessary to identify which features would contribute to the classification by applying numerical (for separability) and graphical (for feature-space plots) feature selection to LIDAR data and aerial imagery. Next, we applied rules to specific features to separate them and to form several levels beneath the top level. According to this design, the selected feature was not used as a texture in the classification system but rather was used to divide other features in another dimension into several vertical stratification levels. Because the categories in each level were reduced by this approach, the image characteristics of the objects in each level had similar statistical distributions. Therefore, the ground features in each level were much easier to discriminate using traditional statistical parametric classification. Hence, we expected more accurate classification results using this hybrid approach compared to the conventional single-step process. Interestingly, the hierarchical hybrid classifier inherits the advantages of both knowledge-based and traditional statistical parametric classification algorithms.

2.1. **Feature selection.** We considered feature selection a problem of selecting the best subset of features from a hyperspectral dataset to be used in a specific classification. This process was also applied to the selection of key features from multi-source datasets. The Jeffries-Matusita (JM) distance is widely used as a selection criterion [16]. The JM distance between a pair of distributions of spectral classes c and d is defined as

$$J_{cd} = 2(1 - e^B), \tag{1}$$

where

$$B = \frac{1}{8}(M_c - M_d)\left\{\frac{\sum_c + \sum_d}{2}\right\}^{-1}(M_c - M_d) + \frac{1}{2}\ln\left\{\frac{|(\sum_c + \sum_d)/2|}{|\sum_c|^{1/2}\,|\sum_d|^{1/2}}\right\} \tag{2}$$

is the Bhattacharyya distance, M_c and M_d are the mean vectors for classes c and d, respectively, and $\sum_c$ and $\sum_d$ are the summations of classes c and d, respectively.

The JM separability measure is calculated for four different band combinations, as shown in Table 1. Most of the JM distances increased when adding the LIDAR nDSM to the RGB band combination, compared to the distances obtained by adding the LIDAR intensity. There was only one exception out of the 12 pair separation values, indicated by the gray background in Table 1. This finding indicated that the LIDAR nDSM was the key factor in different combinations.

TABLE 1. Jeffries-Matusita distances between categories for different feature combinations

	Jeffries-Matusita Distance					
Bands	TR	BR	BT	BG	TG	RG
R, G, B	1.0692	1.1799	1.2547	1.6402	1.7013	1.7255
R, G, B, I	1.5681	1.6225	1.7325	1.8876	1.7530	1.9999
R, G, B, D	1.9795	1.9497	1.7841	1.9891	1.9825	**1.6720**
R, G, B, I, D	1.9944	1.9634	1.8218	1.9992	1.9907	1.9999
Notes: Band information: R: red, G: green, B: blue, I: LIDAR intensity, D: LIDAR nDSM Class information: T: trees, B: buildings, R: roads, G: grass						

2.2. **Feature space inspection for urban objects.** Aerial imagery with multispectral bands is widely used to investigate natural resources and classify land use and cover. The red edge between the red and near-infrared (NIR) bands can effectively discriminate between vegetation and other land cover. In this study, only RGB imagery acquired simultaneously with the LIDAR data was used. The LIDAR intensity data were used as an alternative band, because the characteristics of the LIDAR laser beam were very similar to those of the NIR. Geographically registered RGB and LIDAR intensity data could be incorporated into the classification procedure using stacked vectors [16]. Training samples of traditional urban feature types, such as buildings, trees, roads, and grass, could then be picked from the stacked vectors.

After constructing the stacked vectors, the next step was to investigate the image characteristics of different categories. Many visual tools are available for this purpose, such as spectral plots, histograms, and feature space plots. Feature space plots can provide great insight into the real information content of aerial imagery and LIDAR intensity images through training samples, as well as the degree of correlation between bands [17]. After inspecting different combinations of feature space plots, we found that most plots correlated well. The plot of the red band and LIDAR intensity showed exceptional dispersion throughout the image band space (Figure 1). The upper box in Figure 1 indicates that even though the LIDAR intensity data (similar to the NIR effect) helped separate vegetation from non-vegetation, there was some confusion between trees and grass. Similarly, non-vegetation objects such as buildings and roads may also be confused, as shown in the middle box of Figure 1. Applying a non-parametric classification algorithm on its own may resolve the confusion but would incur computational costs and require case-by-case parameter searches.

However, the addition of LIDAR height information, identified by applying feature selection (as described in the previous section), can help resolve confusion regarding urban

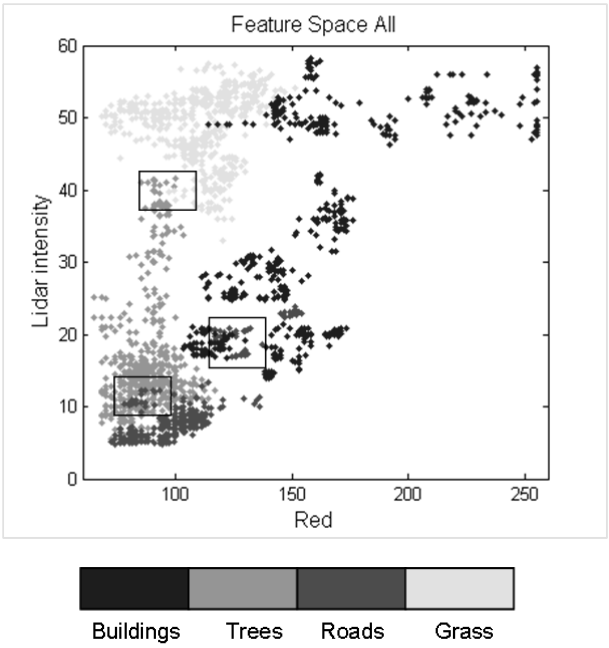

FIGURE 1. Feature space plots for four categories. The upper box indicates that some confusion exists between trees and grass. Similarly, as shown in the middle box, non-vegetation objects such as buildings and roads may also be confused.

ground features. In urban areas, objects at low height levels are mainly roads and grass. Conversely, the two main categories of high-height-level objects are buildings and trees. Therefore, we divided Figure 1 into two subplots according to object height, as shown in Figure 2. The subplot on the left side of the figure is the feature space plot of tall objects (buildings and trees) while the right subplot shows the other objects (roads and grass). It is easy to distinguish between the two categories. Thus, the parametric classification could handle individual classifications at each height level. Nevertheless, for more complex cases, non-parametric classification algorithms can be applied to further increase classification accuracy.

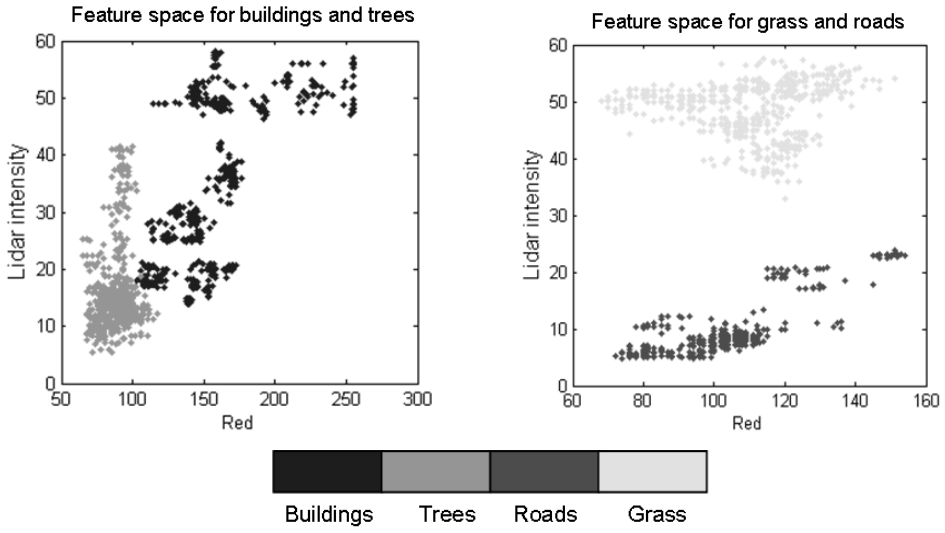

FIGURE 2. Feature space plots for objects at high (left) and low (right) height levels. It is easy to distinguish between the two categories using parametric classification.

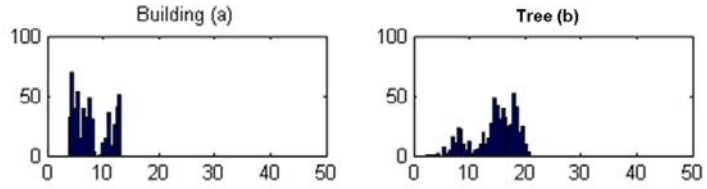

FIGURE 3. Histograms of buildings and trees. The statistical distributions of buildings and trees were bimodal, i.e., below and above 10 m. In the data fusion case, height information could not be handled by the same parametric method as was used for the images.

Buildings and trees in an urban area will naturally have diverse heights. Figure 3 shows that the trees and buildings sampled in the study area fell into two major height groups divided by the 10-m mark. Therefore, the statistical distribution of buildings and trees may be bimodal. Moreover, in each group, the distribution is not normal. For this reason, it was not appropriate to use parametric algorithms to handle height information. Instead, the height information required further processing using knowledge-based rules. The major benefit of using knowledge-based rules to separate objects according to height is that only one simple threshold is required. After separation, the feature space plots

in Figure 2 indicated that a traditional parametric method such as MLC could classify buildings and trees accurately according to the other image bands.

2.3. **Two-level selection and definition.** Height information is critical for resolving confusion between short and tall objects in urban feature classification. However, it must be applied properly to derive accurate results. Generally, height information is used as an additional feature that can increase the accuracy of urban classification. The effect of adding LIDAR nDSMs to color or multispectral aerial imagery data has been studied, and MLC has been used to demonstrate that nDSMs significantly improve classification accuracy [15]. Some studies have used height texture to increase feature dimensions and thus improve classification accuracy [18]. For object-based classification (OBC), height information may be used as the horizontal segmentation factor to reduce internal inconsistencies in high-resolution remote-sensing data.

A knowledge-based approach offers another way to resolve height data. In our previous study, we proposed a three-height-level, rule-based classification scheme to clarify the importance and functionality of height in a classification system [15]. For this scheme, LIDAR height was not only used for features or texture in the classification system, but also to divide the vertical dimension into three stratification levels and then extract ground-feature segments from each level. This had two major advantages. First, the ground objects in each level were separated before classification, so the number of categories at each level was reduced. Second, the ambiguity between low and high objects was easier to resolve than by conventional statistical parametric classifiers. This hybrid approach is different than the OBC method, which only uses LIDAR height data to represent objects horizontally, without considering the vertical dimension.

2.3.1. *Two-level-height rules.* It is not easy to model mid-height segments when integrating three-height-level knowledge-based rules with MLC. We implemented a two-height-level, rule-based classification scheme to model objects at low and high height levels. Two height levels were easily established for the LIDAR nDSM by selecting one height threshold (ThresholdH). The threshold was defined according to the lowest building height (2.5 m) in each study area from in situ knowledge. The rules defining the two levels were as follows:

$$\textbf{If } nDSM < Threshold_H \textbf{ then } low\_height = true \tag{3}$$

$$\textbf{If } nDSM >= Threshold_H \textbf{ then } high\_height = true \tag{4}$$

The segments derived from the above rules were the kernels of the two-height-level scheme. This unique scheme is not found in traditional pixel-based classifications, which motivated us to use segments derived from the knowledge-based rules in MLC as a mask layer to filter out ambiguity within each level. The classification results for each level were merged for the final classification output.

2.3.2. *Narrow lane considerations.* In old parts of Asian cities, there are many small, narrow lanes between buildings. Such features require special treatment for successful urban classification. For example, classification results may show some small lanes as blocked by buildings because the buildings are too close together in the generated DSM. This problem also influences the correctness of random samples that are very close to the edges of buildings. Adjusting the weighting function of the DSM interpolation method to improve the quality of the DSMs can solve this problem. For example, the weighting function for inverse distance weighting (IDW) is a simple method for multivariate interpolation. It consists of a procedure for assigning values to interpolated points using those

from reference points. The height of an interpolated point can be interpolated using IDW as follows:

$$Z_{int} = \frac{\sum_{i=1}^{N} W_i Z_i}{\sum_{i=1}^{N} W_i} \tag{5}$$

where Z_{int} is the height of the interpolated point, Z_i is a reference value, N is the total number of reference points used in the interpolation, and W_i is the weighting function used for the interpolation. A simple weighting function for the IDW can be defined as follows [19]:

$$w(d) = \frac{1}{d^p}, \tag{6}$$

where $w(d)$ is the weighting factor applied to a reference point, d is the distance from the reference point to the interpolated point, and p is a power parameter that is assigned different values for different applications. Normally, the value of p is 2. Changing it to 4 resolves this problem.

2.4. **Hybrid classification implementation.** A major advantage of the proposed hybrid classification method is its simplicity. The method can be used with common commercial remote-sensing software tools, such as ERDAS Imagine, Exelis ENVI, or PCI Geomatics, without any programming. Figure 4 shows the workflow of the proposed method. Initially, we used the lowest building height in the study area to define the height threshold. Then we applied the threshold value and the two-level-height rules to the LIDAR nDSM layer to derive two mask layers: low and high. Second, MLC was performed twice: once for roads and grass in the low level, and again for buildings and trees in the high level. This step classified stacked vectors including RGB aerial imagery and LIDAR intensity data using training samples selected from the low and high levels, respectively. Third, we masked out pixels that did not belong to individual levels using the mask layers, as shown in Figure 4. Finally, the MLC classification results for the two layers were combined into four major urban categories. ENVI was used to implement this procedure. The ENVI MLC classifier was embedded with mask functionality, as described in the third step. The classification output was stored in ENVI native image format, and the classification accuracy was assessed in terms of the overall accuracy and Kappa index.

3. Experiments and Discussion.

3.1. **Study area and datasets.** The study site, Kaohsiung City, is a harbor city in southwestern Taiwan with a long history of development. This area contains a mix of apartment buildings, small lanes, industrial facilities, and a world-class harbor. Two study areas with two different types of urban features were selected. Study Area 1 was an urban area with a variety of small to very large buildings, a big parking lot, and a grassy field. Study Area 2 was on an island across the harbor from Study Area 1. Area 2 was not level and sloped down toward the harbor. A large park with dense tree cover is located on the seaward side of the island, which also includes a mixture of variously sized buildings.

The LIDAR data were acquired in February 2005 using an Optech ALTM 3070c, operating at a sampling rate of 46 Hz. The flying height was approximately 1200 m above ground level (AGL), with a scanning angle of 13 degrees, swath width of approximately 554 m, and scan interval of approximately 0.7 point/m. An nDSM was derived by subtracting the DEM (Digital Elevation Model) derived by a LIDAR data-filtering procedure [20] from the DSM. The nDSM was subsequently resampled to the same grid size as the

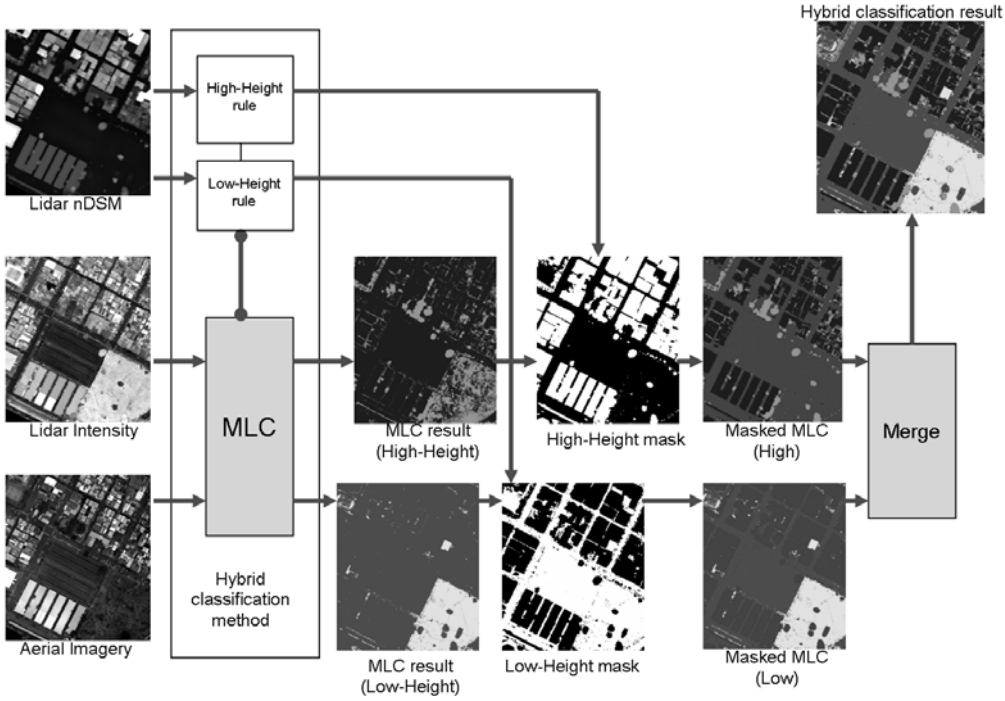

FIGURE 4. Workflow of the hybrid classification method

TABLE 2. Specifications of the digital camera used in this study

Camera focus	55.073 mm
Principal point shift (X):	-0.202 mm $\pm$ 0.0036 mm
Principal point shift (Y):	0.188 mm $\pm$ 0.0036 mm
Exposure rate	2.5-2.7 s
CCD array size	4077 (cross flight) $\times$ 4092 (along flight) pixels
Pixel size	9 μm
Image size	36.693 $\times$ 36.828 mm

spatial resolution of the aerial imagery. A true-color images were collected simultaneously with the LIDAR data by an Optech ALTM4k02 digital camera with direct exterior orientation parameters and rectified to 0.2 m $\times$ 0.2 m grid resolution. Table 2 lists the specifications of the camera used in this study. To allow for comparison, the study areas were the same as those of our previous knowledge-based classification system study [15].

For all experiments, training samples from four ground categories (i.e., trees, grass, buildings, and roads) were selected from the aerial images and LIDAR intensity data using ENVI software. The training sample size needed to be large enough to represent the characteristics of ground objects. For an N-dimensional feature space, 100 N points for each ground type were sampled as the training data to ensure the trained MLC classification capability [21].

Next, we overlapped randomly sampled points over the aerial imagery and the LIDAR height data to assign the feature type of each test point. If the feature types from the images were not certain, we confirmed them through field investigations. The test data were also used for accuracy evaluations of the SVM, ANN, and the hybrid classifications. For accuracy assessment, at least 50 samples were randomly sampled as test data for each category, following previous studies [22,23]. However, simple random sampling tended to

under-sample some small but important areas. To ensure that the sampled data were statistically valid, we iteratively applied a random sampling method in which each category accumulated more than 50 samples. Table 3 lists the number of training and test samples for the hybrid classification experiments of the two study areas.

TABLE 3. The training data and tests used for study Areas 1 and 2

	Categories	Training Samples	Test Samples
Area 1	Buildings	637	485
	Trees	556	73
	Roads	733	476
	Grass	542	117
Area 2	Buildings	602	189
	Trees	536	226
	Roads	561	55
	Grass	602	189

Finally, we needed to determine the height threshold for the two layer-height rule. We used the LIDAR nDSM data and a 3-m threshold to generate a binary image to identify the potential lowest buildings in the experimental areas. The field survey indicated that the lowest building heights were 2.3 m and 1.8 m for Areas 1 and 2, respectively.

3.2. **Hybrid classification experiments.** The hybrid classification method based on knowledge-based rules and traditional statistical MLC was used to investigate the performance and accuracy of urban classification while integrating LIDAR height data with aerial images. Aerial imagery, LIDAR intensity data, and LIDAR nDSM were selected to perform the MLC as a standard case (MLC1) for comparison, because the LIDAR height was used as an additional layer here. Next, the hybrid method (knowledge-based rules + MLC) was applied (case HMLC2). In this case, we applied the two-level-height rules proposed in this study. Third, the three-level-height rules were applied for the HMLC3 case to test whether the three-level-height rules proposed in our previous work [15] were superior to the two-level scheme for the hybrid classification method.

Table 4 lists the performance results for the Area 1 experiments, as Area 1 was larger than Area 2. The computational time for each MLC was only 5 seconds for Area 1, making this a very efficient tool for hybrid classification. The computational time for setting up the two-level height rule by segmentation using ENVI was only 3 seconds. The total computational time for each classification using the proposed hybrid approach was under 30 seconds. Table 5 lists the accuracy assessment results for all experiments. The overall accuracy of the standard MLC1 experiments was about 79% for Area 1 and 85% for Area 2. The lower accuracy for Area 1 was due to the low accuracy for trees and roads in the standard MLC1 compared to that of other categories. The user accuracy, an accuracy assessment index corresponding to commission error, for trees was even worse (only about 44% correct). This finding agreed with the visual inspection mentioned above. However, the overall accuracy improved approximately 12% for Area 1 and 5% for Area 2 when the two-level-height hybrid method (case HMLC2) was used.

Figures 5 and 6 show the resulting MLC images for Areas 1 and 2, respectively. Figures 5(a) and 6(a) illustrate the classification results for standard cases using all available features in the stacked vectors. Some shadows in the narrow lanes between buildings were misclassified as trees (Figure 5(a)) or buildings (Figure 6(a)). In addition, as shown in Figure 5(a), some parking lot signs on the ground were misclassified as buildings. This confusion was due to the similar image characteristics of these ground objects, as

indicated in the lower box (tree vs. roads) and middle box (building vs. roads) in the feature space plots (Figure 1). These were misclassifications, and some small lanes can be clearly seen in results derived from the hybrid method (Figures 5(b), 5(c), 6(b), and 6(c)). The misclassifications and confusion were resolved using LIDAR height information before applying the MLC. The classification results derived from the three-level-height hybrid method were not better than those obtained using the two-level height, but the former method required more effort when modeling training samples for the mid-height level.

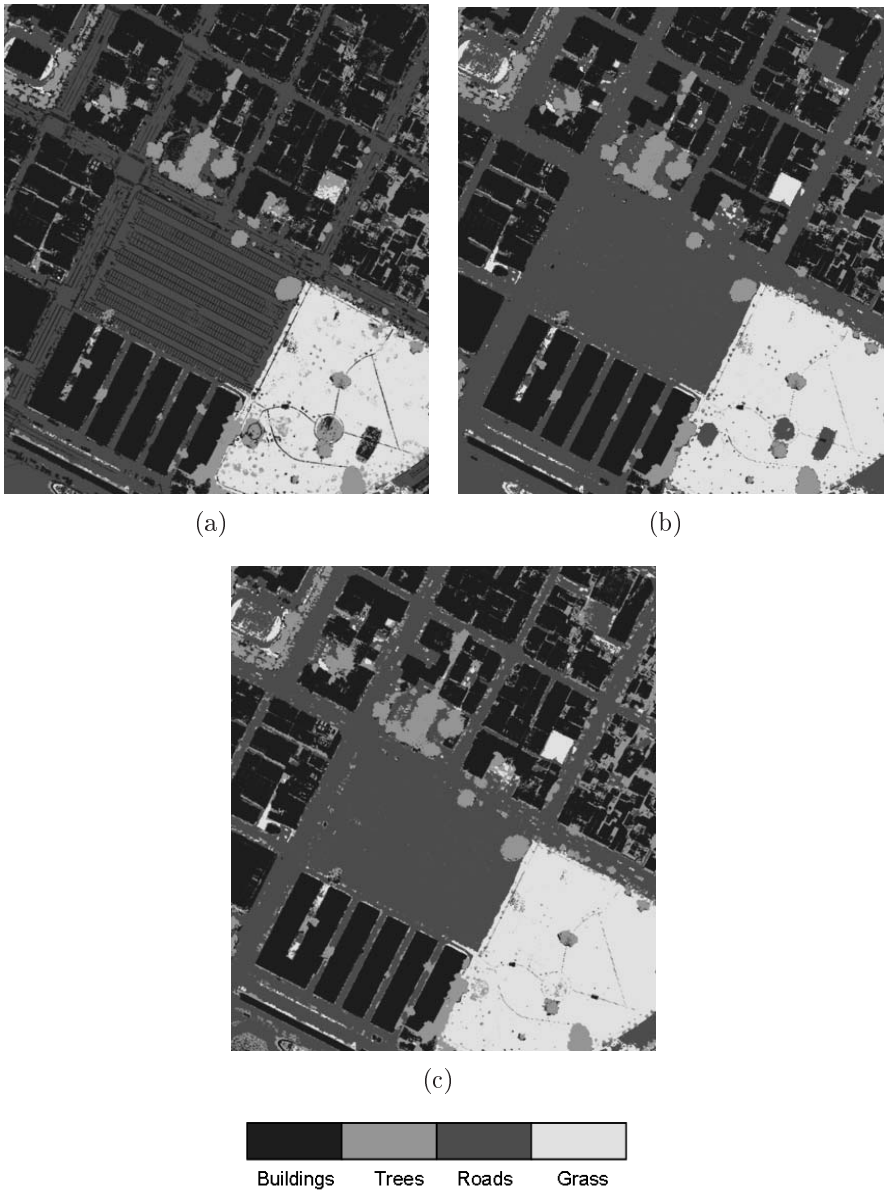

FIGURE 5. Results of (a) MLC, (b) HMLC2, and (c) HMLC3 for Area 1

In Area 1, there were major improvements in user accuracy for buildings, which increased from 74% to 96%, while producer accuracy, an accuracy index corresponding to error of omission, remained about the same. At 96% accuracy, almost all of the classified building pixels were correct in the two-level-height scheme. A similar result was found in the tree category. Even though producer accuracy decreased from 75% to 66%, user accuracy increased from 44% to 59%, resulting in better overall accuracy for trees. Nevertheless, the classification performance was still not accurate enough because trees

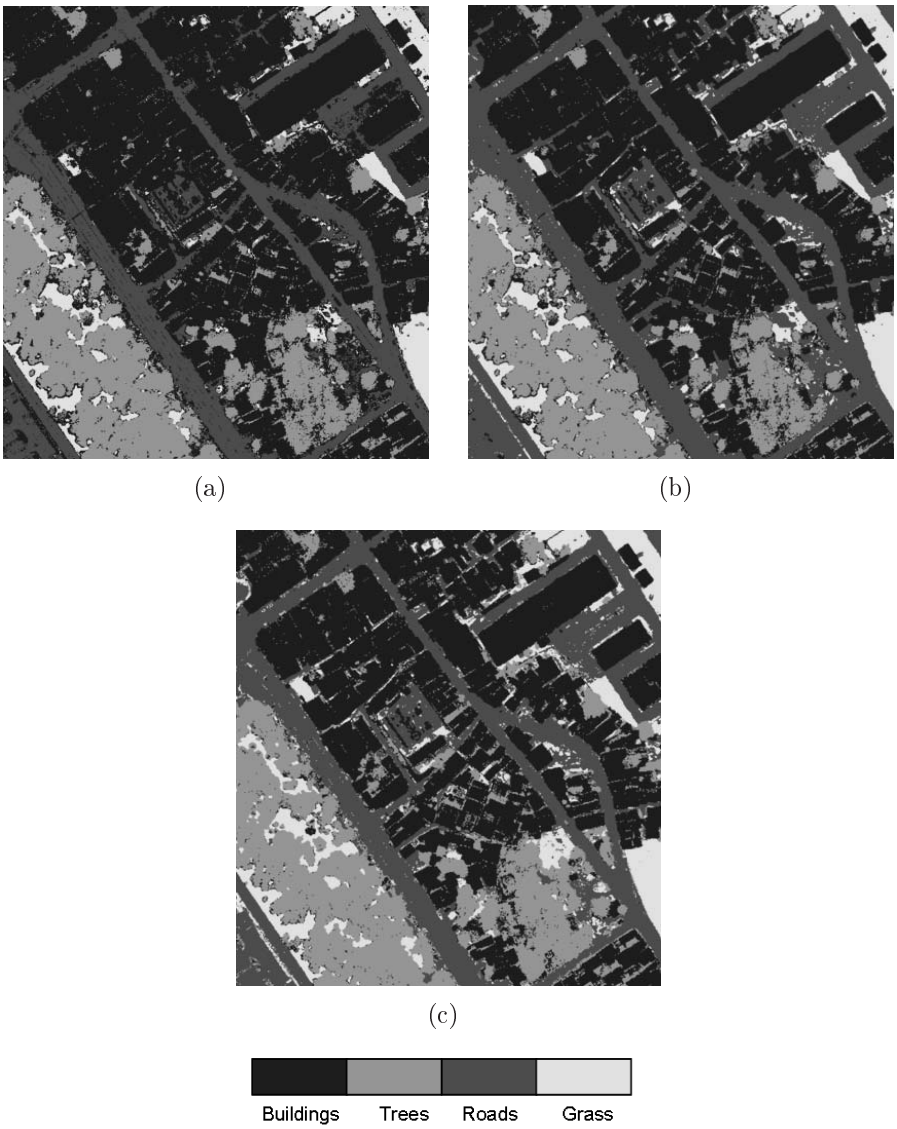

FIGURE 6. Results of (a) MLC, (b) HMLC2, and (c) HMLC3 for Area 2

were rare. Visual checks of the classification results indicated that too many building pixels were misclassified as trees, as shown in Figure 5(b). In Area 2, user accuracy for trees increased from 65% to 93%, while other categories showed little significant improvement or equivalent results. In contrast, the three-level-height scheme did not significantly improve the results compared to the two-level-height scheme, as indicated not only by visual checks of the classified images but also by the accuracy assessment. Therefore, our results indicate that the two-level-height scheme was superior for the proposed hybrid classification method in this study.

3.3. **The ANN and SVM experiments.** ANN and SVM experiments were performed for the two study areas to investigate the performance of our hybrid classification method. The feature vectors used for this experiment were RGB aerial imagery, LIDAR intensity, and nDSM. Because each of these non-parametric classifiers required numerous parameters, the classifier with the set of parameters resulting in the highest accuracy is reported here. To identify parameters for the SVM effectively, we adopted the libSVM and imageSVM [9,10] tools to obtain optimal penalty parameters and the gamma value of the

TABLE 4. Process time of the MLC, ANN, and SVM experiments for Area 1

MLC Experiment			
Training/Classification	5 seconds		
SVM Experiment			
Grid search for optimal penalty parameters and the gamma value	52 seconds		
Training/Classification	96 * 2 $\sim$ 3 seconds		
Total	244 $\sim$ 340 seconds		
Note: Grid search using libSVM and imageSVM. Training/classification using ENVI. After obtaining the range of penalty parameters and the gamma value of the radial basis kernel function, it takes 2 or 3 calculations to get optimal results. Total process time: 4 $\sim$ 5.6 minutes.			
ANN Experiment (all layer values expressed in seconds)			
Training iteration	Hidden Layer (1)	Hidden Layer (2)*	Hidden Layer (3)*
1000	121	177	230
2000	246	360	468
3000	369	540	701
4000	492	720	935
5000	619	905	1177
10000	1258	1840	2391
Total training	3105	4542	5902
Classification	60	78	100
Total	3165	4520	6002
Note: Process time for different training iterations (moment = 0.3, training rate = 0.2). The process time for each moment and training rate was 14977 seconds (4.16 hours). Optimal parameter search: Moment: 0 $\sim$ 0.9 with interval 0.1 (10 times). Training rate: 0 $\sim$ 0.5 with interval 0.2 (6 times). Total process time estimate: $4.16 * 60 = 248.9$ hours			

radial basis kernel function. The ANN was run using a back-propagation neural network classifier to compare the accuracy to that of the SVM and the hybrid classification method. Table 6 lists the parameters with the highest accuracy for each classifier in each area. The ANN and SVM outputs were stored in ENVI native image format. The same accuracy evaluation was then performed with the same test datasets used for the MLC and hybrid classification experiments, as shown in Table 2.

The process time for ANN and SVM are also shown in Table 4. All of the experiments were run using MLC, ANN, and SVM classification tools of ENVI software on the same computer. However, the ENVI software does not provide grid search tools for SVM. The optimal penalty parameters and the gamma value of the radial basis kernel function obtained by libSVM and imageSVM tools were inputted into ENVI SVM to calculate the process time for the SVM. Our experiments showed that the process time for the MLC, SVM, and ANN methods were 5 seconds, 4 $\sim$ 5.6 minutes, and about 249 hours, respectively, for Area 1. The ratios between these values were approximately the same as those for Area 2 among the three methods. Applying the two-layer scheme to the MLC, SVM, and ANN methods would double the process time. In this study, Areas 1 and 2

TABLE 5. Accuracy assessments of the MLC, hybrid classification, ANN, and SVM experiments

Experiment				Producer Accuracy (%)				User Accuracy (%)			
		OA	KI	Bd	Tr	Rd	Gs	Bd	Tr	Rd	Gs
Area 1	MLC1	79.93	0.7002	93.20	73.97	65.97	83.62	73.86	43.90	99.05	93.08
	HMLC2	91.91	0.8780	92.58	65.75	94.75	93.22	95.53	59.26	94.15	91.16
	HMLC3	89.27	0.8426	90.31	76.71	87.39	96.61	96.48	39.44	98.11	89.53
	ANN1	85.80	0.7928	83.92	76.71	84.66	97.74	96.90	44.80	92.22	75.55
	HANN2	85.55	0.7920	81.44	78.08	85.71	99.44	98.01	38.51	98.08	72.13
	SVM1	86.62	0.8026	84.33	73.97	91.18	85.88	96.69	39.13	91.18	87.36
	HSVM2	90.42	0.8577	89.90	64.38	92.02	98.31	95.61	49.47	97.33	82.86
Area 2	MLC1	84.19	0.7735	83.53	89.42	80.09	87.27	94.14	65.25	96.28	77.42
	HMLC2	89.58	0.8478	93.93	80.42	90.71	89.09	90.03	93.25	91.52	72.06
	HMLC3	89.09	0.8419	92.77	86.77	85.40	89.09	92.51	85.42	94.61	67.12
	ANN1	90.20	0.8584	89.31	90.48	90.71	92.73	94.21	89.06	90.71	72.86
	HANN2	86.76	0.8118	79.48	93.65	90.27	94.55	96.83	73.75	92.31	73.24
	SVM1	86.27	0.8040	82.37	84.66	91.59	94.55	94.37	85.11	84.84	63.41
	HSVM2	88.84	0.8403	85.84	92.06	89.82	92.73	97.06	79.82	91.86	71.83

Notes:
MLC1, ANN1, SVM1
Features: RGB + LIDAR intensity + LIDAR nDSM
HMLC2, HANN2, HSVM2:
Two-height-level hybrid classification method
MLC3: Three-height-level hybrid classification method

OA: Overall accuracy (%)
KI: Kappa index
Bd: Buildings
Tr: Trees
Rd: Roads
Gs: Grass

TABLE 6. Parameters resulting in the highest accuracy in the ANN and SVM experiments

	ANN parameters					SVM parameters		
	Case	Hidden layer	Training iterations	Training momentum	Training rate	Case	Gamma	Penalty
Area 1	ANN1	2	10000	0.3	0.2	SVM1	0.2	100
	HANN2	1	10000	0.3	0.2	HSVM2	0.25	100
Area 2	ANN1	3	10000	0.4	0.2	SVM1	0.2	100
	HANN2	1	10000	0.5	0.2	HSVM2	0.25	100

were only a very small portion of Kao-Hsiung City. The assessment results shown in Table 4 demonstrate the practical superiority of the proposed hybrid method over the ANN and SVM approaches. The accuracy assessment results for the ANN and SVM experiments are shown in Table 5. The overall accuracy and Kappa index values of the hybrid method were 5-6% higher than those of the ANN and SVM in Area 1. They were equivalent to (0.7% lower than) those of the ANN but 4% higher than those of the SVM in Area 2.

Figures 7 and 8 show the resulting classified images from the ANN (Figures 7(a) and 8(a)) and SVM (Figures 7(b) and 8(b)) experiments for Areas 1 and 2, respectively. As expected, the ANN and SVM images (cases ANN1 and SVM1) have superior visual interpretability compared to the MLC image (case MLC1 in Figures 5(a) and 6(a)). They are even equivalent to the hybrid classification cases (case HMLC2 in Figures 5(b) and

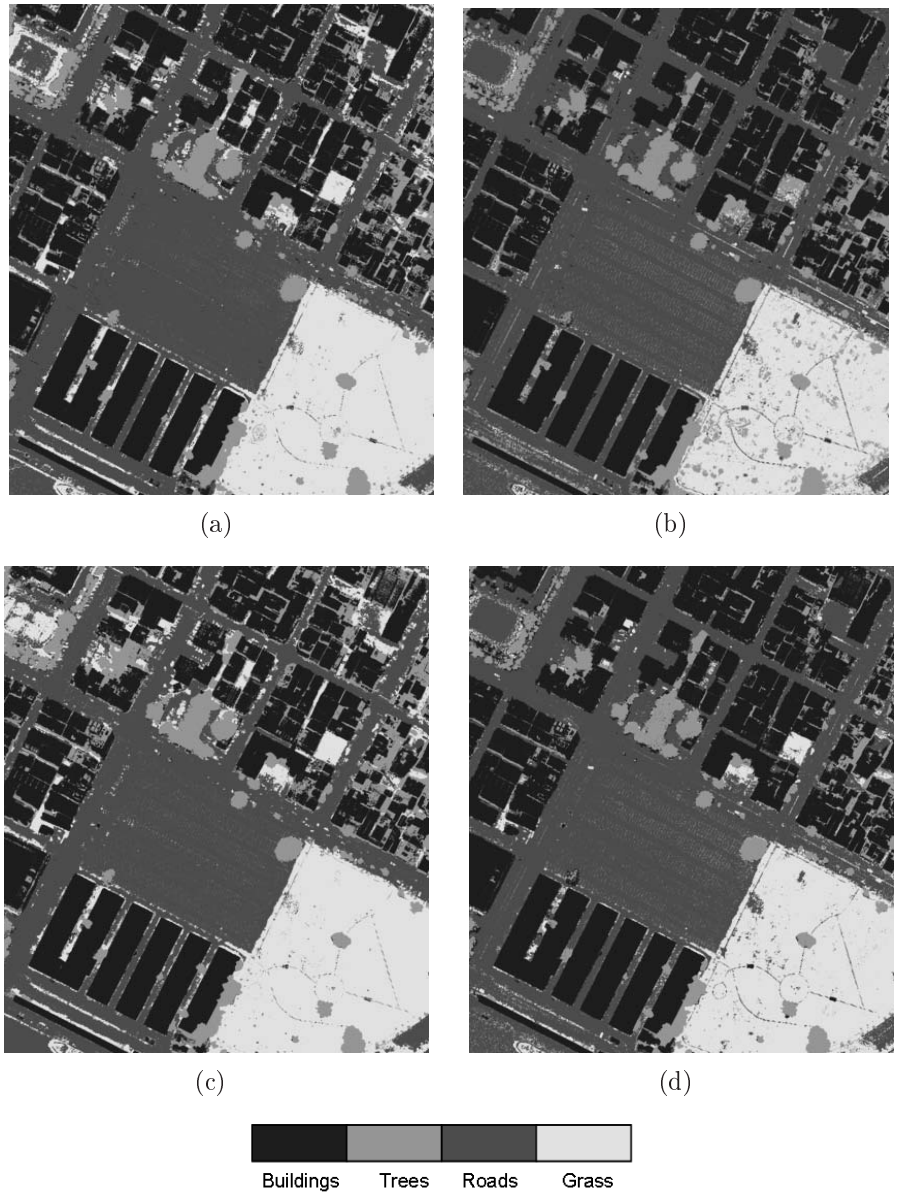

FIGURE 7. Results of (a) ANN, (b) SVM, (c) HANN2, and (d) HSVM2 for Area 1

6(b)). Thus, the computational and time costs of parameter identification are worthwhile. However, there were some differences among the ANN, SVM, and hybrid classifications. In the SVM experiment, many misclassifications of small trees occurred inside a large, grassy field in Area 1 (case SVM1 in Figure 7(b)). The vegetation discrimination between trees and grass was better by ANN than by SVM. Conversely, ANN misidentified many areas of road and small pavement patches inside the grassy field as grassy spots, as shown in Figure 7(a), more so than did SVM. However, the SVM misclassified buildings and trees in an area beside the park with dense tree cover and on the roof of a large building in Area 2 (case SVM1 in Figure 8(b)).

To further analyze the hybrid classification method, additional ANN and SVM experiments were performed to evaluate the improvement when the two-level-height scheme was applied to the ANN (case HANN2) and to the SVM (case HSVM2). As shown in Table 5, the overall accuracy of HANN2 did not improve for Area 1, while the accuracy of HANN2

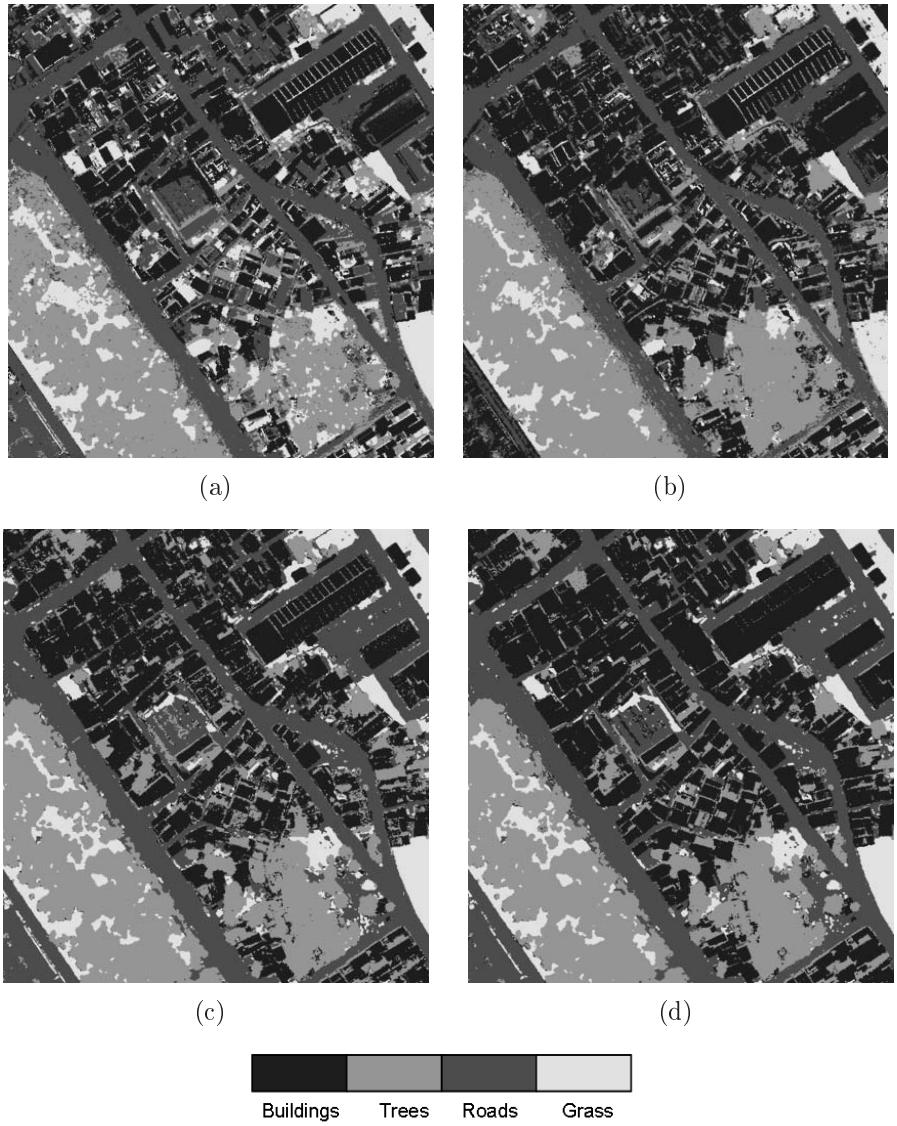

FIGURE 8. Results of (a) ANN, (b) SVM, (c) HANN2, and (d) HSVM2 for Area 2

decreased 3% for Area 2. This indicated that the hybrid approach did not improve ANN performance. However, the overall accuracy of SVM2 improved by about 4% for Area 1 and 3% for Area 2, indicating that the hybrid approach did improve the performance of SVM. However, neither the hybrid ANN nor the hybrid SVM (case HANN2 and HSVM2) had higher classification accuracy than the hybrid MLC (case HMLC2). Furthermore, even with a 3-4% improvement in accuracy, the HSVM2 method still required twice the computational cost and more time to identify parameters compared to the hybrid case. Instead, the 3-4% improvement could be achieved by simply using the hybrid MLC.

The above findings were confirmed by the HANN2 and HSVM2 classification result images for Areas 1 (Figure 7(c) and 7(d)) and 2 (Figure 8(c) and 8(d)). Figures 7(c) and 8(c) show many buildings surrounded by false trees, which indicates more misclassification of trees and buildings compared to the ANN1 case (Figures 7(a) and 8(a)). On the other hand, many misclassifications of small trees inside a large, grassy field in Area 1 in the SVM1 case were improved in the HSVM1 case, as shown in Figure 7(c). In addition, the misclassification between buildings and trees in an area beside the park with dense

tree cover and on the roof of a large building in Area 2 (case SVM1 in Figure 8(b)) was reduced in the hybrid SVM (case HSVM1 in Figure 8(d)).

Generally, a non-parametric approach (e.g., ANN or SVM) may be able to overcome the ambiguity of multi-source data for urban classification. However, some misclassifications between high and low objects may still occur in the ANN and SVM results. The above experiments confirmed the role of height, represented by the unique two-level-height scheme embedded in the hybrid classification method, as a key factor in urban classification. The ANN or SVM approach cost about 60 times or 197724 times more, respectively, in terms of computing power compared to MLC, as shown in Table 4. These methods may not be worth their additional costs.

There are several advantages to using the proposed hybrid classification method to separate ground features into categories. The major advantages of the new two-height-level classification framework not only demonstrate the computational efficiency for practical urban classification applications but also reveal the importance of LIDAR height in urban classification algorithms. The method offers a mechanism for reducing the number of categories at each level to overcome the ambiguity between low and high objects. In the hybrid method, height information plays a key role in feature selection and is a starting point for classifying multi-source data for urban classification. The proposed hybrid framework could be extended using other key features in other classification applications.

4. **Conclusions.** For the simultaneous acquisition of LIDAR data and aerial imagery in urban classification applications, it is critical to use spatially heterogeneous multi-source data effectively. LIDAR height information can be used to overcome difficulties in discrimination between trees and grass, as well as misclassification of buildings due to diverse roof compositions and shadow effects. The hybrid classification method proposed in this study integrates a two-height-level scheme into the traditional pixel-based MLC. This scheme demonstrates the importance and functionality of LIDAR height in a classification system from a new viewpoint. LIDAR height is not only used for texture in the classification system but also divides the vertical dimension into two stratification levels and then extracts ground-feature segments from each level. In this way, high efficiency can be achieved in practical urban classification applications. This scheme is not used in traditional pixel-based classification. The OBC approach uses LIDAR height data to represent objects horizontally only, without considering the vertical dimension.

Furthermore, as our results demonstrate, the proposed method has two major advantages over ANN or SVM non-parametric classification methods in terms of efficiency and accuracy. First, the computing efficiency of the hybrid classification method is higher than the SVM or ANN methods by a factor of about 60 and 197724, respectively, due to the two-layer approach based on a simple MLC method, which can be easily applied with general image-processing tools. In addition, the classification accuracy of the hybrid method was 4-7% higher than that of the ANN or SVM. Its accuracy may not always be superior to these other methods, but when used properly, it is equivalent in value to complex computational algorithms.

The proposed hybrid classification method provides a way to integrate knowledge-based rules into traditional pixel-based MLC classification. In addition, it can be incorporated into the non-parametric SVM method to handle more complex problems. Finally, this method is straightforward and understandable and can be replicated by applying the MLC function rules in commercial classification software such as ENVI, ERDAS Imagine, and PCI Geomatics, without any programming. We have outlined the fundamentals of the hierarchical hybrid classification method for classifying multi-source data for urban areas. The results provide a solid base for further study and research. In the future,

more subcategories could be added to the hybrid classification method according to user requirements. More critical discrimination of ground-feature rules should also be explored by the feature selection method if more subcategories are required.

Acknowledgment. This work was supported by National Science Council, Taiwan with grant NSC94-2611-E-110-007. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] P. E. Baltsavias and A. Gruen, A comparison of aerial photos, LIDAR and IKONOS for monitoring cities, in *Remotely Sensed Cities*, V. Mesev (ed.), New York, Taylor & Francis, 2003.

[2] F. Rottensteiner, J. Trinder, S. Clode and K. Kubik, Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.62, no.2, pp.135-149, 2007.

[3] M. Awrangjeb, M. Ravanbakhsh and C. S. Fraser, Automatic detection of residential buildings using LIDAR data and multispectral imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.65, no.5, pp.457-467, 2010.

[4] G. M. Foody and A. Mathur, A relative evaluation of multiclass image classification by support vector machines, *IEEE trans. on Geoscience and Remote Sensing*, vol.42, no.6, pp.1335-1343, 2004.

[5] J. T. Morris, D. Porter, M. Neet, P. A. Noble, L. Schmidt, L. A. Lapine and J. R. Jensen, Integrating LIDAR elevation data, multi-spectral imagery and neural network modelling for marsh characterization, *International Journal of Remote Sensing*, vol.26, no.23, pp.5221-5234, 2005.

[6] C.-J. Lin, C.-H. Chen and C.-Y. Lee, Classification and medical diagnosis using wavelet-based fuzzy neural networks, *International Journal of Innovative Computing, Information and Control*, vol.4, no.3, pp.735-748, 2008.

[7] X. Song, W. Chen and B. Jiang, Sample reducing method in support vector machine based on K-closet sub-clusters, *International Journal of Innovative Computing, Information and Control*, vol.4, no.7, pp.1751-1760, 2008.

[8] L. Bruzone and C. Persello, Approaches based on support vector machine to classification of remote sensing data, in *Handbook of Pattern Recognition And Computer Vision*, 4th Edition, C. H. Chen (editor), Singapore, World Scientific Publishing Co., 2009.

[9] C. W. Hsu, C. C. Chang and C. J. Lin, *A Practical Guide to Support Vector Classification*, Department of Computer Science, National Taiwan University, http://www.csie.ntu.edu.tw/∼cjlin/papers/guide/guide.pdf, 2009.

[10] S. Van der Linden, A. Rabe, A. Okujeni and P. Hostert, ImageSVM classification, *Application Manual: ImageSVM Version 2.0*, Humboldt-Universität zu Berlin, Germany, 2009.

[11] W. Zhou, G. Huang, A. Troy and M. L. Cadenasso, Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study, *Remote Sensing of Environment*, vol.113, no.8, pp.1769-1777, 2009.

[12] Y. Ke, L. J. Quackenbush and J. Im, Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification, *Remote Sensing of Environment*, vol.114, no.6, pp.1141-1154, 2010.

[13] X. Meng, *Determining Urban Land Uses Through Building-Associated Element Attributes Derived from LIDAR and Aerial Photographs*, Ph.D. Thesis, Texas State University-San Marcos, 2010.

[14] N. Haala and C. Brenner, Extraction of buildings and trees in urban environment, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.54, no.2-3, pp.130-137, 1999.

[15] M.-J. Huang, S.-W. Shyue, L.-H. Lee and C.-C. Kao, A knowledge-based approach to urban feature classification using aerial imagery with LIDAR data, *Photogrammetric Engineering and Remote Sensing*, vol.74, no.12, pp.1473-1485, 2008.

[16] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*, 4th Edition, Springer-Verlag, Berlin, 2006.

[17] J. R. Jensen, *Introductory Digital image Processing: A Remote Sensing Perspective*, Prentice Hall, Upper Saddle River, New Jersey, 2005.

[18] A. Charaniya, R. Manduchi and S. Lodha, Supervised parametric classification of aerial Lidar data, *Proc. of Conference on Computer Vision and Pattern Recognition Workshop*, Washington DC, vol.3, 2004.

[19] D. Shepard, A two-dimensional interpolation function for irregularly-spaced data, *Proc. of the ACM National Conference*, pp.517-524, 1968.
[20] L. H. Lee, M. J. Huang, S. W. Shyue and C. Y. Lin, An adaptive filtering and terrain recovery approach for airborne LiDAR data, *International Journal of Innovative Computing, Information and Control*, vol.4, no.7, pp.1783-1796, 2008.
[21] P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*, McGraw-Hill, New York, 1978.
[22] R. G. Congalton, A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, vol.37, no.1, pp.35-46, 1991.
[23] K. Fitzpatrick-Lins, Comparison of sampling procedures and data analysis for a land-use and land-cover map, *Photogrammetric Engineering and Remote Sensing*, vol.47, no.3, pp.343-351, 1981.