Fusion of Hyperspectral and LiDAR Data for Landscape Visual Quality Assessment

Naoto Yokoya, *Member, IEEE*, Shinji Nakazawa, Tomohiro Matsuki, and Akira Iwasaki

Abstract—Landscape visual quality is an important factor associated with daily experiences and influences our quality of life. In this work, the authors present a method of fusing airborne hyperspectral and mapping light detection and ranging (LiDAR) data for landscape visual quality assessment. From the fused hyperspectral and LiDAR data, classification and depth images at any location can be obtained, enabling physical features such as land-cover properties and openness to be quantified. The relationship between physical features and human landscape preferences is learned using least absolute shrinkage and selection operator (LASSO) regression. The proposed method is applied to the hyperspectral and LiDAR datasets provided for the 2013 IEEE GRSS Data Fusion Contest. The results showed that the proposed method successfully learned a human perception model that enables the prediction of landscape visual quality at any viewpoint for a given demographic used for training. This work is expected to contribute to automatic landscape assessment and optimal spatial planning using remote sensing data.

Index Terms—Hyperspectral data, landscape visual quality, least absolute shrinkage and selection operator (LASSO) regression, light detection and ranging (LiDAR) data, multisensor classification, openness.

I. INTRODUCTION

LANDSCAPE quality assessment is becoming a major component in spatial planning owing to increasing interest in the concept of sustainability including aspects related to the quality of life (QOL) and of landscapes [1]. Previous studies and theories [2] suggest an influence of the landscape visual quality in urban areas on criminal [3], [4] and health behaviors [5], [6], which are significant factors determining the QOL. Therefore, evaluative maps of landscape visual quality based on human perception or, in the case of this study, machine-modeled human perception, are helpful for policy makers, space planners, and architects to plan and design the appearance of urban areas that are attractive to inhabitants [7].

The assessment of landscape visual quality is a challenging issue because there are many objective factors that affect subjective human perceptions. Psychophysical preference modeling is one of the quantitative holistic techniques of landscape evaluation that mix subjective and objective methods [8], [9].

This approach relates psychological human preference data to physical landscape component data in a mathematical manner, which mainly consists of three steps: 1) collection of human perception data for a number of different landscapes; 2) physical feature extraction from these landscapes; and 3) learning the relationship between human perception and physical feature datasets. Quantitative human perception data of landscape quality have commonly been collected by surveys that use ground-based color photographs and scoring methods, such as paired comparisons, Likert scales, and ranking scales [4], [9], [10]. Physical characteristics of landscapes can be extracted from color photographs by manual segmentation [9]. Multivariable linear regression analysis have been widely used to find psychophysical predictive models that relate physical features to landscape quality [1], [11], [12]. Even when a predictive model is obtained, the creation of evaluative maps of landscape visual quality is time-consuming, and it is not feasible to take photographs everywhere to extract physical features. Geographic information systems (GISs) and computer vision are useful for calculating key visual features from any viewpoint over a large area to assess landscape quality [1], [13]. Wu *et al.* reported that biological and physical characteristics defined by classification have major influences on landscape visual quality [1]. Weitkamp *et al.* proposed a GIS-based procedure for measuring landscape openness [13], [14] using depth information, which is an important part of our understanding and appreciation of landscapes [15]. Therefore, remotely sensed data, computer vision, and machine learning techniques have major potential for extracting the physical features of landscapes at a low cost.

Hyperspectral imaging is a promising remote sensing technology for generating classification maps. A continuum spectrum enables the accurate identification and classification of land-cover classes that are spectrally similar. Many researchers have investigated hyperspectral classification and shown its validity in mineral mapping, tree species discrimination, and urban classification [16]–[18]. Land-use and land-cover classification maps are basic information of GISs. Light detection and ranging (LiDAR) is an active remote sensing technique that uses electromagnetic energy in the optical range to determine the distance between the instrument and a target surface and deduce physical properties of the target based on the interaction of the radiation with the target [19]. LiDAR has a wide range of applications, such as atmospheric monitoring and canopy height analysis [20], [21]. Mapping LiDAR data is an accurate tool for obtaining a digital surface model (DSM) and a digital terrain model (DTM). Such data are rasterized 2.5D data and useful for visible area analysis. Openness can be quantified using depth information of sight [13], [14], which is calculated from a DSM.
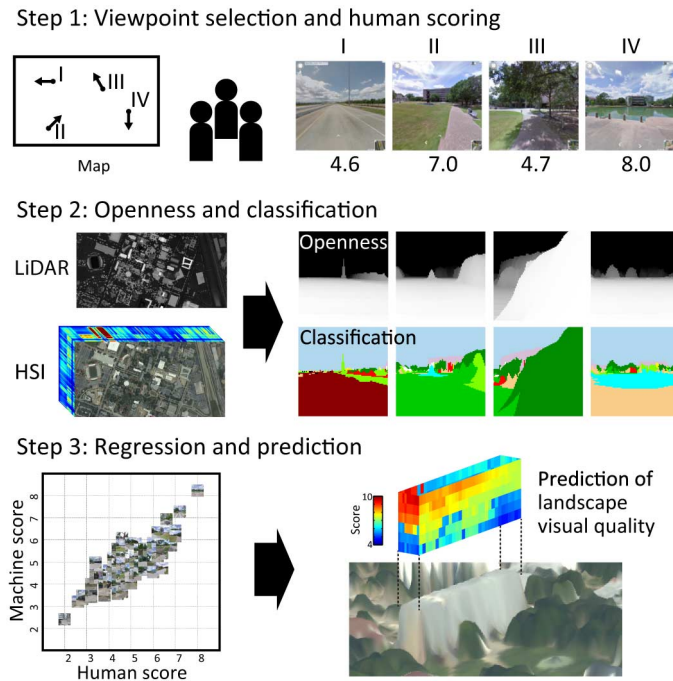
1939-1404 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Fig. 1. Overview of the proposed method.

Fig. 2. (a) Terrain and building datasets. (b) Visible space in 2-D with contour lines. (c) Depth map from a viewpoint.

Moreover, the joint use of optical and LiDAR data is effective for accurate classification, e.g., tree species classification [22]–[25]. Therefore, the fusion of hyperspectral and LiDAR data is promising for extracting important physical features related to landscape quality.

In this work, we present a new methodology that fuses an airborne hyperspectral image and a LiDAR-derived DSM for landscape visual quality assessment. The main objective of this work is to propose a framework that relates landscape physical features to the human perception of such features, not to develop a general psychophysical relationship. Landscape physical features, such as openness and proportions of land-cover classes for a given view, are extracted from the hyperspectral image and a LiDAR-derived DSM. The relationship between physical features and human landscape quality assessment based on ground-based images is learned by least absolute shrinkage and selection operator (LASSO) regression [26]. This framework enables the prediction of landscape quality from any viewpoint using large-scale remote sensing observation.

This paper is organized as follows. Section II describes the methodology for predicting landscape visual quality by fusing an airborne hyperspectral image and a LiDAR-derived DSM. An experimental study using the datasets provided for the 2013 IEEE GRSS Data Fusion Contest and images collected from Google Street View is presented in Section III. The conclusion is given in Section IV.

II. METHODOLOGY

The methodology presented in this paper is divided into three main steps, as shown in Fig. 1. Each step is summarized as follows.

1) *Viewpoint selection and human scoring:* We prepare human-perception data by collecting landscape photographs taken a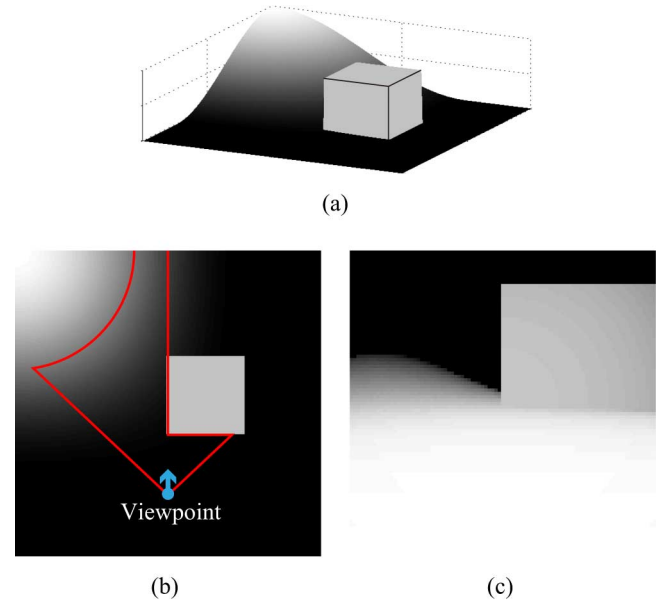t various locations and view angles and asking people to rate their visual quality, which is a common technique for measuring landscape preferences [7], [9], [10]. Ground-based landscape photographs are acquired from Google Street View [27].

2) *Openness and classification:* We calculate physical features, such as openness, proportions, and average depths of land-cover classes in a field of view, at the collected viewpoints in the first step using the DSM and classification map derived from hyperspectral and LiDAR data.

3) *Regression and prediction:* LASSO regression is applied to the datasets using physical features and human preference scores as input and target variables, respectively [26]. Once the predictive model is learned by LASSO regression, the landscape visual quality can be predicted at any viewpoint.

The three novel points of our method are summarized as follows: 1) ground-based color photographs are collected using Google Street View; 2) physical features are automatically extracted using remote sensing data and computer vision; and 3) LASSO regression is adopted to learn the psychophysical model that relates physical features to human-perception-based landscape quality data. More details about the calculation of openness and classification using the hyperspectral image and LiDAR-derived DSM and as well as LASSO regression are given in Sections II-A, II-B, and II-C respectively.

A. Openness

Weitkamp *et al.* measured the visible space using the average line of sight in a two-dimensional (2-D) space to assess landscape openness [13], [14]. Terrain and building datasets [Fig. 2(a)] are merged to create contour lines, and a visible space is identified using viewing limitations as shown in Fig. 2(b). The average length of radials is highly correlated with perceived openness [14]. We extend this approach to a three-dimensional (3-D) space and define openness as the average value of a depth image

obtained at a viewpoint [Fig. 2(c)]. Note that depth values are assigned to nonsky objects to avoid including the infinite depth of the sky in the calculation of openness. When the location and direction of an observer are given, the depth of the direction can be calculated by computing the intersection of the DSM and the line-of-sight vector. Depth values are estimated on radials with $1°$ spacing in the vertical and horizontal directions. The maximum horizontal and vertical ranges of the viewing angle were both set to $90°$ considering the approximate field of view of the Google Street View images. This field of view can be an approximation of human binocular vision that covers $120°$ horizontally and $60°$ vertically [28].

B. Classification

In this work, we begin with the use of airborne data taken over urban areas with a small number of training samples. Shadows contained in observed data have a major influence on classification. Therefore, as a preprocessing step, shadows in the hyperspectral data need to be modified. Shadows caused by buildings and clouds are modified using line-of-sight analysis and by thresholding illumination distributions, respectively [29], [30]. When the DSM is directly used for classification, various heights of the terrain may result in classification errors. Therefore, the DTM is estimated from the DSM by extracting buildings and trees by thresholding local height information [31] and interpolating extracted segmentations. The normalized DSM (NDSM), which is the height of features relative to the ground, is generated by subtracting the DTM from the DSM.

After the preprocessing, we apply the nearest neighbor (NN) algorithm to a fused feature, i.e., the cosine similarity of spectra and the NDSM difference. The NDSM difference is normalized and fused with the cosine similarity using a weighted summation. A class label $f(\mathbf{x})$ of a test point $\mathbf{x} = (x, y, z)$ is estimated by

$$\hat{f}(\mathbf{x}) = \arg\max_c \frac{\mathbf{s}(\mathbf{x}) \cdot \mathbf{s}(\mathbf{x}_c)}{\|\mathbf{s}(\mathbf{x})\|\|\mathbf{s}(\mathbf{x}_c)\|} + \omega \mathcal{N}(\tilde{z} - \tilde{z}_c | 0, \sigma^2) \quad (1)$$

where $\mathbf{x}_c = (x_x, y_x, z_c)$ is a reference point belonging to class $c \in \{1, \ldots, C\}$, with C being the number of classes. $\mathbf{s}(\mathbf{x})$ denotes the spectrum of point $\mathbf{x}$, and z and $\tilde{z}$ indicate the DSM and NDSM values, respectively. σ^2 is the variance of a Gaussian function defined by statistical analysis of the NDSM data and ω is a weight coefficient determined by cross-validation. The final classification result is obtained by applying iterative bilateral filtering [32] to the NN classification map to reduce noise. We extend the normal bilateral filter to higher dimensions, i.e., a 3-D location for geometric closeness and a spectrum for photometric similarity. Finally, the test point $\mathbf{x}$ is classified as

$$\hat{f}(\mathbf{x}) = \arg\max_c \ k^{-1}(\mathbf{x}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta_c(\xi) \mathcal{N}\left(\|\xi - \mathbf{x}\| | 0, \sigma_g^2\right) \mathcal{N}$$
$$\left(\mathbf{s}(\xi) - \mathbf{s}(\mathbf{x}) | 0, \sigma_p^2 \mathbf{I}\right) d\xi \quad (2)$$

where $\delta_c(\xi) = 1$ when ξ belongs to class c and $\delta_c(\xi) = 0$ otherwise. $\|\cdot\|$ denotes the Euclidean distance and σ_g and σ_p are the geometric and photometric spread parameters, respectively.

$k(\mathbf{x})$ is a normalization term defined by $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{N}(\|\xi - \mathbf{x}\| | 0, \sigma_g^2) \mathcal{N}(\mathbf{s}(\xi) - \mathbf{s}(\mathbf{x}) | 0, \sigma_p^2 \mathbf{I}) d\xi$. An iterative bilateral filter with a small window acts as an edge-preserving and noise-reducing smoothing filter for the classification map. By combining the classification map with the 3-D virtual view, proportions and average depths of land-cover classes occupied in a given field of view can be calculated, which are key physical features influencing perceived landscape quality [1], [15]. In addition to land-cover classes, we also consider the proportion of sky as a physical feature of landscapes. Here, the proportion of a class in a field of view indicates how much of the class exists in a 2-D view image. The average depth of a class indicates how far pixels assigned to the class in a 2-D view image are located from the viewpoint on average, which can be interpreted as class-labeled openness.

C. LASSO Regression

The objective of the learning phase is to reveal the relationship between physical features of a landscape and human landscape perception and to construct a prediction model for a new input. Multivariable linear regression is useful owing to its mathematical simplicity and physical interpretability. Openness, the proportion of sky, and the proportions and average depths of land-cover classes in a field of view are selected as features and used as input variables ($\mathbf{X} \in \mathbb{R}^{N \times M}$). N and M denote the numbers of samples and variables, respectively. Human scores for landscape visual quality are used as target values ($\mathbf{y} \in \mathbb{R}^{N \times 1}$). All input variables are standardized with zero mean and unit variance to investigate their impacts on landscape visual quality. The least-squares method is the simplest method of regression; however, it suffers from overfitting. Regularized least-squares methods that add penalty terms of regression parameters to the objective function are useful to avoid this problem and produce robust results. We adopt LASSO regression [26], which is one of the regularized least-squares methods and uses the L^1-norm of the parameter vector as a penalty term. We can assume without loss of generality that $\bar{\mathbf{y}} = 0$. The LASSO estimate $\hat{\mathbf{w}}$ is formulated by

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \|\mathbf{y} - \mathbf{X}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_1 \quad (3)$$

where $\mathbf{w} \in \mathbb{R}^{M \times 1}$ is the parameter vector and λ is the tuning parameter. This problem may be solved by convex optimization. In LASSO regression, more of the parameters become zero as the penalty is increased, whereas in ridge regression, which uses the L^2-norm for the penalty, the parameters are reduced but remain nonzero. The former property results in the straightforward physical interpretation of regression parameters, which is the reason why we choose LASSO regression as the regression method. The difference between least-squares and LASSO regression is discussed in Section III. The tuning parameter λ can be determined by leave-one-out cross-validation in which the tradeoff between the sparsity of parameters and the fitting of the regression is examined.

III. EXPERIMENTAL STUDY

A. Datasets

In this work, we used the hyperspectral image and LiDAR-derived DSM distributed by the 2013 IEEE GRSS Data Fusion Contest. The two images have the same size (349×1905 pixels)

(a)

Healthy grass　Stressed grass　Trees　Soil　Water　Residential
Multistory buildings　Roadways　Highways　Railways　Parking lots　Walkways
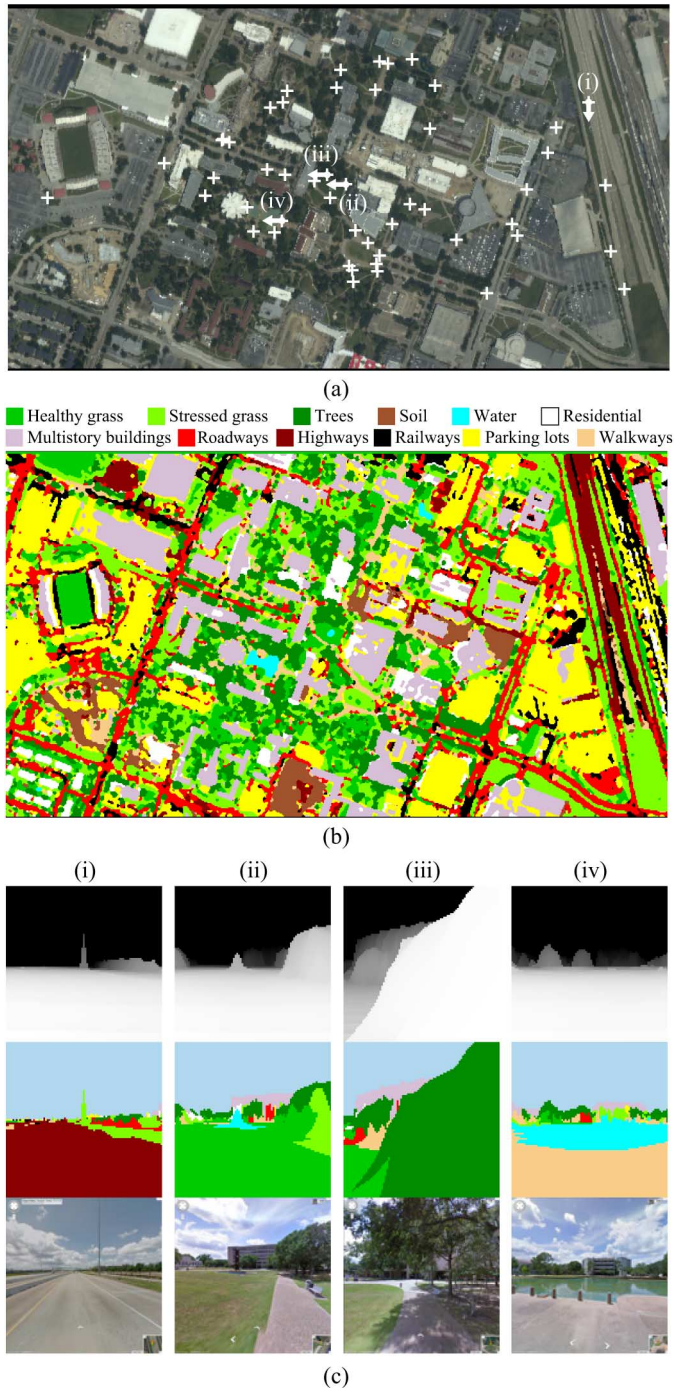
(b)

(i)　　(ii)　　(iii)　　(iv)

(c)

Fig. 3. Physical features obtained from hyperspectral and LiDAR data. (a) RGB. (b) Classification. (c) Depth and classification images with views from top to bottom.

and ground sampling distance (GSD) (2.5 m). The hyperspectral data have 144 spectral bands in the 380–1050 nm region. The dataset was acquired over the University of Houston campus and the neighboring urban area. The RGB image of the study area is shown in Fig. 3(a). The ground truth of 15 classes is provided by the Best Classification Challenge of the 2013 IEEE GRSS Data Fusion Contest. Each class of the ground truth consists of approximately 200 pixels.

Parking lots and cars were merged, and synthetic grass, tennis courts, and running tracks, which do not exist in the study area,

TABLE I
R-SQUARED AND PEARSON'S CORRELATION VALUES OF LEAST SQUARES AND LASSO

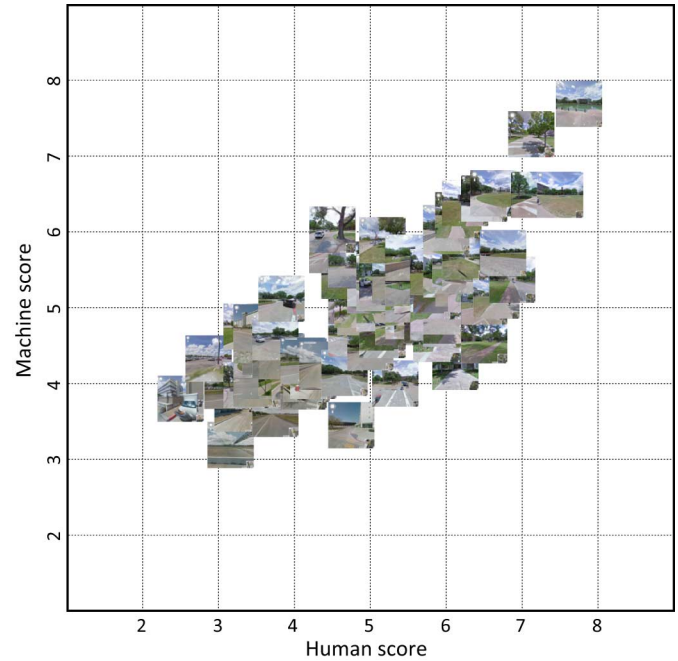|	R^2		Pearson's correlation	
	Training	Test	Training	Test
Least squares	**0.702**	0.470	**0.838**	0.715
LASSO	0.664	**0.527**	0.822	**0.740**

Fig. 4. Scatter plot showing human-perceived and machine-predicted scores for the test data with the corresponding views.

were excluded. Furthermore, we manually assigned the "walkway" class to 255 pixels from pavements on the campus into the ground truth, so that they can be distinguished from spectrally similar classes, such as "roadways." The final ground truth consisted of 12 classes, i.e., healthy grass, stressed grass, trees, soil, water, residential, multistory buildings, roadways, highways, railways, parking lots, and walkways. Since we use openness, the proportion of sky, and the proportions and average depths of 12 classes as inputs, the number of variables is 26.

We collected 200 view images taken on the ground at 50 locations from four directions using Google Street View in this area. The 50 locations are shown in Fig. 3(a) as white crosses and the four directions are north, south, east, and west. In this experiment, we fixed the training data to 100 randomly chosen views and used the remaining 100 views as test data. Subjective human landscape preferences can be influenced by many factors, such as culture, nationality, generation, and personality. The main objective of this work is to present a framework that relates physical features to human perception to construct a predictive model of landscape quality using remote sensing data, not to reveal a general psychophysical relationship. Therefore, we chose a specific group of people, i.e., eight students from the University of Tokyo, Japan, and asked them to give preference scores between 1 and 10 assuming that they have similar backgrounds owing to their common culture and generation. Their average scores were used as target values.
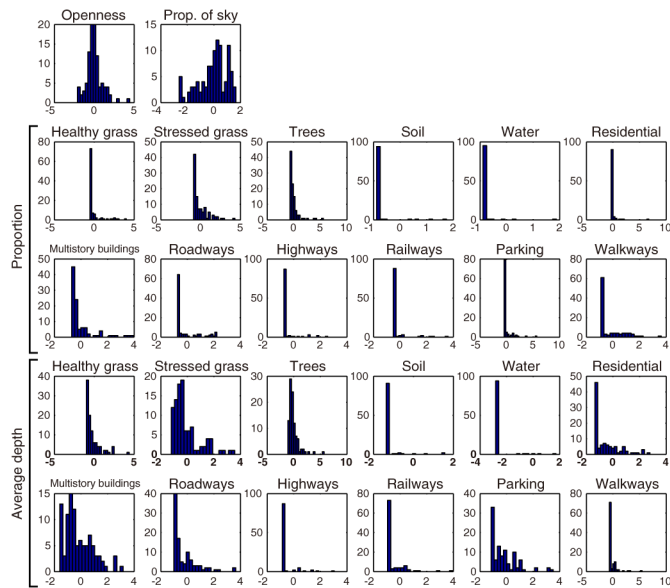
Fig. 5. Histograms of training samples for all input variables.

Fig. 6. Regression coefficients (blue: least squares and red: LASSO).

B. Results and Discussion

Fig. 3(b) shows the classification map estimated by our method. The presented method achieved 90.65% classification accuracy in the Best Classification Challenge of the 2013 IEEE GRSS Data Fusion Contest [33]. Fig. 3(c) shows the depth, classification, and view images for the locations and directions shown by four arrows in Fig. 3(a). Although the camera model of the depth images and classification maps is different from that of the subimages of Google Street View, Fig. 3(c)(i), (ii), and (iv) show the consistent appearance of the major-class distribution between classification images and ground-based photographs, which indicates that the physical features at a viewpoint can be approximately extracted by the proposed method. The advantage of the use of remote sensing data is that physical features can be computed at any viewpoint of a large area. Some differences between the photographs and the classification images of views are caused by the limitation of remote sensing images. In the classification image of Fig. 3(c)(ii), the walkway is misclassified as grass owing to the mixed pixel spectra. A smaller GSD may allow the more accurate extraction of features from a landscape. Most attempts to create regression models of landscape quality have been applied to rural landscapes [9], [10]. There are more buildings in urban areas, resulting in a significant difference in the landscape at a smaller geometric scale depending on the location and viewing direction. Therefore, spatially detailed information is more important for urban areas, whereas spatially larger datasets may be required for rural areas owing to there being fewer obstacles and better viewing conditions. In Fig. 3(c)(iii), the visibility of the classification view image under trees is different from that of the photograph owing to the complicated 3-D structures of trees and the occlusion of the walkway. The LiDAR-derived DTM and DSM have only 2.5D information, meaning that a single value of elevation is provided for an entire cell, which is the elevation of the perceived ground or the elevation of the first detected return, respectively. If the full 3-D information contained in the LiDAR point cloud is used,
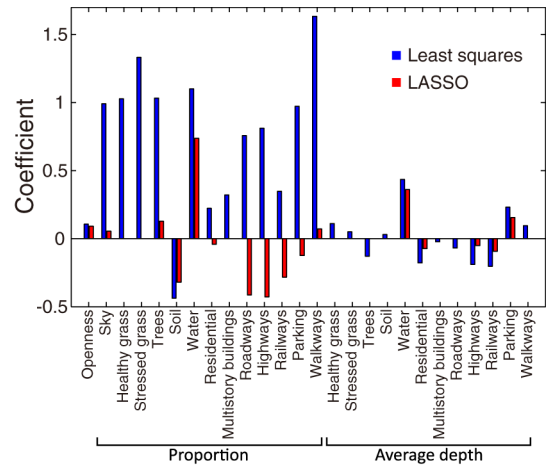
better-matching synthetic images could have been obtained with respect to the ground-based photographs used for the regression.

R-squared and Pearson's correlation values were used to evaluate the agreement between the machine-predicted score, which is the output of the regression model, and the human-preference score. Table I shows a comparison between LASSO regression and the least-squares method for both the training and the test data. LASSO regression exhibits a more accurate and robust performance than the least-squares regression for the test data. Fig. 4 shows a scatter plot of the human-preference and machine-predicted scores obtained using the test data along with the corresponding views. This plot intuitively visualizes the similarity between machine predictions and human perceptions and reveals both the correlation and the misalignment obtained by the proposed method. Fig. 5 shows histograms that represent the distributions of the 100 training samples for all input variables. A wide range in the relative occurrence of different variables can be observed. Fig. 6 shows the coefficients of the least-squares and LASSO regression, which were used to investigate physical factors that influence landscape visual quality. We can find differences between the two methods in terms of the scale and sparsity of the coefficients as well as their signs for class proportions of human-made structures. In the least-squares regression, it is more difficult to interpret the coefficients because all the coefficients of class proportions except for soil have positive values. In LASSO regression, several coefficients are zero owing to the L^1 regularization, which results in a straight-forward interpretation of the trained model. Openness and the proportions of sky, trees, and water have positive impacts on landscape quality, whereas those of most human-made structures have negative impacts. This result is consistent with other landscape studies [1], [7], [9], [12]. In particular, the proportion of water has the largest positive impact; water exists in the form of fountains and a small natural lake on this campus. It is interesting that there are clear differences between the landscape-quality impact of walkways and other human-made structures and also between grass and trees, which are spectrally similar classes. Detailed and accurate classification by the joint use of a hyperspectral image and a LiDAR-derived DSM enables a meaningful model of landscape quality prediction to be learned.
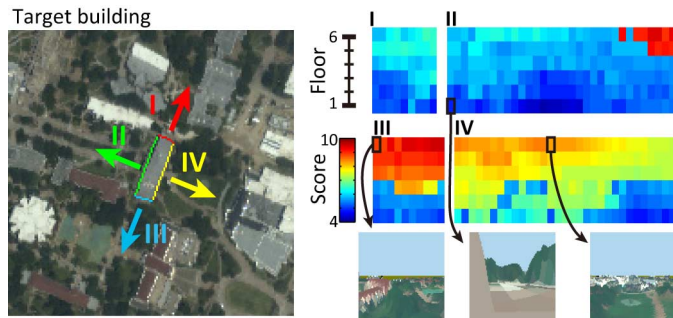
Fig. 7. Landscape assessment from the target building.

Finally, we show an application of the proposed method of landscape assessment. Since the model is already learned and all physical features can be calculated from the hyperspectral image and LiDAR-derived DSM, we can automatically predict the landscape visual quality from any 3-D location and direction. We focused on one building, as shown on the left of Fig. 7, and predicted the landscape visual quality scores for the views from all floors. The upper-right images of Fig. 7 show the predicted score maps for the views from the four sides of the building. At this location, the landscape quality is high mainly because of the high openness and proportions of sky, trees, and water. Consequently, views from the higher floors, where a fountain can be seen and the openness and the proportion of sky are greater, have higher landscape scores. The three images in the bottom-right of Fig. 7 show virtual RGB views from three representative points. As described earlier, the proposed system can predict the landscape visual quality at any viewpoint in the study area. This predictive method will effectively predict the landscape visual quality for the given demographic that it was trained for.

IV. Conclusion

In this work, we presented a novel method of fusing a hyperspectral image and a LiDAR-derived DSM for landscape visual quality assessment. The physical properties of objects and the openness in a view are assumed to be important factors influencing landscape quality. The datasets of the 2013 IEEE GRSS Data Fusion Contest were used in the experimental study, and an accurate classification map was obtained by fusing spectral and 3-D spatial information. Using a LiDAR-derived DSM and the classification map, depth and classification maps from any viewpoint could be estimated. LASSO regression was adopted to learn a psychophysical predictive model that relates physical features to human preference scores with meaningful regression coefficients. Physical feature extraction by the joint use of the hyperspectral image and LiDAR-derived DSM reveals the differences in the landscape-quality impact between spectrally similar land-cover classes and enables the prediction of landscape quality from any 3-D viewpoint over a large area. This work is expected to contribute to automatic landscape assessment and optimal spatial planning using remote sensing data.

References

[1] Y. Wu, I. Bishop, H. Hossain, and V. Sposito, "Using GIS in landscape visual quality assessment," *Appl. GIS*, vol. 2, no. 3, pp. 18.1–18.20, 2006.

[2] J. Q. Wilson and G. L. Kelling, "Broken windows," *Atl. Mon.*, vol. 249, no. 3, pp. 29–38, 1982.

[3] K. Keizer, S. Lindenberg, and L. Steg, "The spreading of disorder," *Science*, vol. 322, no. 5908, pp. 1681–1685, 2008.

[4] P. Salesses, K. Schechtner, and C. A. Hidalgo, "The collaborative image of the city: Mapping the inequality of urban perception," *PLoS One*, vol. 8, no. 7, pp. 1–12, 2013.

[5] D. Cohen, S. Spear, R. Scribner, P. Kissinger, K. Mason, and J. Wildgen, "'Broken windows' and the risk of gonorrhea," *Amer. J. Public Health*, vol. 90, no. 2, pp. 230–236, 2000.

[6] C. E. Ross and J. Mirowsky, "Neighborhood disadvantage, disorder, and health," *J. Health Social Behav.*, vol. 42, no. 2, pp. 258–276, 2001.

[7] J. K. Nasar, *The Evaluative Image of the City.* Newbury, CA, USA: Sage, 1997.

[8] E. L. Shafer, "Perception of natural environments," *Environ. Behav.*, vol. 1, pp. 71–82, 1969.

[9] J. R. Wherrett, "Creating landscape preference models using internet survey techniques," *Landscape Res.*, vol. 25, no. 1, pp. 76–96, 2000.

[10] H. Schroeder and T. C. Daniel, "Progress in predicting the perceived scenic beauty of forest landscapes," *Forest Sci.*, vol. 27, no. 1, pp. 71–80, 1981.

[11] G. J. Buhyoff, P. A. Miller, J. W. Roach, D. Zhou, and L. G. Fuller, "An AI methodology for landscape visual assessments," *AI Appl.*, vol. 8, no. 1, pp. 1–13, 1994.

[12] M. Arriaza, J. F. Cañas-Ortega, J. A. Cañas-Madueño, and P. Ruiz-Aviles, "Assessing the visual quality of rural landscapes," *Landsc. Urban Plann.*, vol. 69, pp. 115–125, 2004.

[13] G. Weitkamp, A. Bregt, and R. V. Lammeren, "Measuring visible space to assess landscape openness," *Landsc. Res.*, vol. 36, no. 2, pp. 127–150, 2011.

[14] G. Weitkamp, "Mapping landscape openness with isovists," *Res. Urban. Ser.*, vol. 2, no. 1, pp. 205–223, 2011.

[15] I. D. Bishop, J. R. Wherrett, and D. R. Miller, "Using image depth variables as predictors of visual quality," *Environ. Plann. B: Plann. Des.*, vol. 27, pp. 865–875, 2000.

[16] F. A. Kruse, J. W. Boardman, and J. F. Huntington, "Comparison of airborne hyperspectral data and EO-1 hyperion for mineral mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1388–1400, Jun. 2003.

[17] M. L. Clark, D. A. Roberts, and D. B. Clark, "Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales," *Remote Sens. Environ.*, vol. 96, pp. 375–398, 2005.

[18] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.

[19] J. C. F. Diaz, W. E. Carter, R. L. Shrestha, and C. L. Glennie, "Lidar remote sensing," in *Handbook of Satellite Applications,* 2013, pp. 757–808.

[20] F. Molero and F. Jaque, "The laser as a tool in environmental problems," *Opt. Mater.*, vol. 13, pp. 167–173, 1999.

[21] M. A. Lefsky, W. B. Cohen, G. G. Parker, and D. J. Harding, "Lidar remote sensing for ecosystem studies," *Bioscience*, vol. 52, no. 1, pp. 19–30, 2002.

[22] M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1416–1427, May 2008.

[23] M. Pedergnana, P. R. Marpu, M. D. Mura, J. A. Benediktsson, and L. Bruzzone, "Classification of remote sensing optical and LiDAR data using extended attribute proles," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 7, pp. 856–865, Nov. 2012.

[24] R. Dinuls, G. Erins, A. Lorencs, I. Mednieks, and J. Sinica-Sinavskis, "Tree species identification in mixed Baltic forest using LiDAR and multispectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 594–603, Apr. 2012.

[25] J. Vauhkonen, T. Hakala, J. Suomalainen, S. Kaasalainen, O. Nevalainen, M. Vastaranta *et al.*, "Classification of spruce and pine trees using active hyperspectral LiDAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1138–1141, Sep. 2013.

[26] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Roy. Stat. Soc. Ser. B (Stat. Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996.

[27] Google Maps, *About Street View* [Online]. Available: http://www.google.com/maps/about/behind-the-scenes/streetview/ accessed on Jan. 14, 2014.

[28] D. B. Henson, *Visual Fields.* London, U.K.: Oxford Univ. Press, 1993.

[29] O. Friman, G. Tolt, and J. Ahlberg, "Illumination and shadow compensation of hyperspectral images using a digital surface model and non-linear least squares estimation," *Proc. SPIE*, vol. 8180, pp. 1–8, 2011.

[30] B. D. Wemett, J. K. Riek, and R. A. Leathers, "Dynamic thresholding for hyperspectral shadow detection using Levenberg-Marquardt minimization on multiple Gaussian illumination distributions," *Proc. SPIE*, vol. 7334, pp. 1–9, 2009.

[31] A. Brunn and U. Weidner, "Extracting buildings from digital surface models," in *Proc. IAPRS,* Stuttgart, Germany, 1997, vol. 32, part 3–4w2.

[32] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.* (ICCV'98), 1998, pp. 839–846.

[33] *Outcome of the "Best Classification Challenge" of the 2013 IEEE GRSS Data Fusion Contest* [Online]. Available: http://hyperspectral.ee.uh.edu/?page_id=695, accessed on Jan. 14, 2014.

Naoto Yokoya (S'10–M'13) received the M.Sc. and Ph.D. degrees in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2010 and 2013, respectively.

He was a Research Fellow with the Japan Society for the Promotion of Science from 2012 to 2013. He is currently an Assistant Professor with the University of Tokyo. His research interests include image analysis and data fusion in remote sensing.

Shinji Nakazawa received the M.Sc. degree in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2014.

His research interests include image processing and optical system.

Tomohiro Matsuki received the M.Sc. degree in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2014.

His research interests include hyperspectral data analysis.

Akira Iwasaki received the M.Sc. degree in aerospace engineering and the Doctoral degree in engineering from the University of Tokyo, Tokyo, Japan, in 1987 and 1996, respectively.

In 1987, he joined the Electrotechnical Laboratory, where he was engaged in research on space technology and remote sensing system. He is currently a Professor with the University of Tokyo.