3D OBJECT CLASSIFICATION BASED ON THERMAL AND VISIBLE IMAGERY IN URBAN AREA

Hadiseh Hasani [a, *], Farhad Samadzadegan [a]

[a] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran - (hasani, samadz)@ut.ac.ir

Commission III, WG III/4

KEY WORDS: Fusion, Thermal Imagery, Visible Imagery, Classification, Urban Area

ABSTRACT:

The spatial distribution of land cover in the urban area especially 3D objects (buildings and trees) is a fundamental dataset for urban planning, ecological research, disaster management, *etc*. According to recent advances in sensor technologies, several types of remotely sensed data are available from the same area. Data fusion has been widely investigated for integrating different source of data in classification of urban area. Thermal infrared imagery (TIR) contains information on emitted radiation and has unique radiometric properties. However, due to coarse spatial resolution of thermal data, its application has been restricted in urban areas. On the other hand, visible image (VIS) has high spatial resolution and information in visible spectrum. Consequently, there is a complementary relation between thermal and visible imagery in classification of urban area. This paper evaluates the potential of aerial thermal hyperspectral and visible imagery fusion in classification of urban area. In the pre-processing step, thermal imagery is resampled to the spatial resolution of visible image. Then feature level fusion is applied to construct hybrid feature space include visible bands, thermal hyperspectral bands, spatial and texture features and moreover Principle Component Analysis (PCA) transformation is applied to extract PCs. Due to high dimensionality of feature space, dimension reduction method is performed. Finally, Support Vector Machines (SVMs) classify the reduced hybrid feature space. The obtained results show using thermal imagery along with visible imagery, improved the classification accuracy up to 8% respect to visible image classification.

1. INTRODUCTION

The spatial distribution of land cover in the urban area is a fundamental dataset for urban planning, ecological research, change detection, disaster management, *etc*. Over the past decades, different types of remotely sensed data widely applied for classification of urban areas, such as LiDAR (Chehata et al. 2009), aerial visible image (Myeong et al., 2001), satellite multispectral image (Moran, 2010), hyperspectral imagery (Samadzadegan et al. 2012), *etc*.

High spatial resolution images have been increasingly used for urban land use/cover classification, but the high spectral variation within the same land cover, the spectral confusion among different land covers, proximity of 3D objects and the shadow problem often encounter classification performance with some challenges (Moran, 2010). Therefore, the classification of urban area is ongoing topic in remote sensing community (Wentz et al. 2014; Weng, 2012).

By recent advances in sensor technologies, different types of remotely sensed data are available from the same area. Due to incompetence of just one dataset to model all characteristic of urban area and also availability of different source of data, much attention has been recently paid to multi-sensor data fusion. Multi-sensor data fusion seeks to integrate data from different sources to obtain more information than can be derived from a single sensor (Kumar et al. 2015). The assessment of the capacity of several combination of dataset are investigated in the literatures, such as hyperspectral and LiDAR data (Liao et al. 2015), optical image ad synthetic aperture radar (Zhu et al. 2012),

aerial image and LiDAR data (Huang et al. 2011), optical image, thermal image and LiDAR (Brook et al. 2012).

Due to high spatial resolution of the aerial imaging sensors, they play a very important role in the discrimination of land-cover classes in complex urban area. The visible camera observes a radiation reflected from Earth's surface over a visible wavelength range, whereas the thermal camera observes radiation emitted from Earth's surface over IR wavelength range. However, thermal imagery has coarse spatial resolution and its application has been restricted in urban areas. The brightness values of the visible and thermal images correspond to Earth's surface reflectance and temperature, respectively (Liao et al. 2014).

Fusion of the aerial visible and thermal images can enhance the spatial details of the thermal image and also add the temperature information to the visible image. Consequently, fused images of a target scene from visible and thermal imaging sensors provide additional information and make classification performance better than would be possible when the sensors are used individually.

Berni et al. generate quantitative remote sensing products for vegetation monitoring by means of a helicopter-based UAV equipped with inexpensive thermal and narrowband multispectral imaging sensors (2009). Moreover, fusion of thermal and visible imagery has application in real time people and vehicle tracking from unmanned aerial vehicle (Gaszczak et al. 2011). Xin et al. propose hierarchical classification framework to fuse thermal hyperspectral and visible imagery. Firstly, feature extraction is performed and then different classes were identified

---

* Corresponding author

successively using appropriate feature combinations. Finally, the pixel-based classification map was improved by post-processing steps (Lio et al. 2014).

This paper presents the method to fuse thermal hyperspectral imagery and aerial RGB image in feature level fusion for classification of urban area.

2. PROPOSED METHOD

This paper evaluate the capacity of aerial thermal hyperspectral and visible imagery fusion in classification of urban area. The proposed pipeline of multi-sensor data classification procedure composed of four steps: data registration, feature extraction, dimension reduction and classification.

2.1 Data Registration

The visible data were gereferenced and registered to the thermal data but they have different resolution. Thermal imagery was up-sampled to visible image by bicubic interpolation.

2.2 Feature Extraction

Feature extraction is the second step where spectral and spatial features are extracted from visible and thermal images to explain more about the scene. In this research, the following features were extracted:

- **PCA**: spatial analysis of thermal hyperspectral imagery is carried out on first two PCs.

- **Gray-Level Co-occurrence Matrices (GLCM):** 8 features are extracted includes: mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation. Table 1 presents the formula of the GLCM descriptors.

Table 1. GLCM descriptors formula

$\mu_x = \sum_{i=1}^{G}\sum_{i=1}^{G} P(i,j) \times i,$ $\mu_y = \sum_{i=1}^{G}\sum_{i=1}^{G} P(i,j) \times j$	$\sigma_i^2 = \sum_{i=0}^{G-1}\sum_{i=0}^{G-1} P(i,j) \times (i-\mu_i)^2$ $\sigma_j^2 = \sum_{i=0}^{G-1}\sum_{i=0}^{G-1} P(i,j) \times (j-\mu_j)^2$
$Homogeneity = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{P(i,j)}{1+|i-j|}$	$Contrast = \sum_{i=1}^{G}\sum_{j=1}^{G} P(i,j) \times (i-j)^2$
$Entropy = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \times \log(P(i,j))$	$Dissimilarity = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j)|i-j|$
$Second\ Moment = \sum_{i=0}^{G-1}\sum_{i=0}^{G-1} \{P(i,j)\}^2$	$Correlation$ $= -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{\{i,j\} \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y}$

where G is the number of grey level and the matrix element $P(i,j)$ is the relative frequency with which two pixels occur within a given neighbourhood, one with intensity i and the other with intensity j (Haralick et al. 1973).

All the GLCM features were extracted from each band of the visible image and 2 PCs of thermal hyperspectral imagery.

- **Morphological Profile (MP):** according to high-resolution of visible image, MPs are computed using the mean visible image bands to extract spatial features. MP is computed by applying opening and closing morphological operator with different structure element size. It is widely agreed that spatial features can significantly improve the classification

accuracy due to the consideration of spatial correlation between neighbouring pixels (Benediktsson et al. 2003).

- **Structural Feature Set (SFS):** this spatial feature is used to extract the statistical features of the direction-lines histogram. Direction lines can be defined as a series of a predetermined number of equally spaced lines through the central pixel. Six statistical features are considered, consist of length, width, pixel shape index (PSI), ω-mean, ratio, standard deviation (Table 2).

Table 2. SFS features formula

$Length = \max\{d_i(c)\}$	$Width = \min\{d_i(c)\}$
$PSI = \sum_{i=1}^{D} d_i(c)/D$	$\omega - mean = \sum_{i=1}^{D} \frac{a(k_i-1)}{st_i} d_i \Big/ D$
$Ratio = arctang\ \frac{\sum_{j=1}^{n} sort_{min}^{j}(H(c))}{\sum_{j=1}^{n} sort_{max}^{j}(H(c))}$	$SD = \frac{1}{D-1}\sqrt{\sum_{i=1}^{D}(d_i(c)-PSI)^2}$

where d_i is length of directional line, c denotes central pixel, $H(c)$ is directional line histogram of central pixel $H(c) = \{d_1(c), \dots, d_D(c)\}$, a is constant value, k_i is length of ith direction line, st_i is standard deviation of spectral value in the ith direction line and $sort_{min}^{i}$ and $sort_{max}^{i}$ are i^{th} minimum and maximum of all direction-line respectively (Huang et al. 2007).

- **Local statistics:** four statistic features include mean, variance, skewness and kurtosis are extracted from each visible image band.

- **Vegetation index:** Due to the biophysical characteristics of green plants, the normalized ratio between the mean thermal hyperspectral bands and the red band of visible image is used as vegetation index (Lio et al. 2014).

By merging all extracted features along with original images, the hybrid feature space is generated.

2.3 Dimension Reduction

High dimensionality of the hybrid feature space pose classification process some challenges, such as correlation between some features, time consuming and complexity of computation. Therefore, dimension reduction is applied to reduce the dimension of hybrid feature space, yet it preserve the main information. PCA is well-known dimension reduction method that is applied in this paper to reduce the dimension of hybrid feature space. The intrinsic dimension of hybrid feature space is determined based Eigen value and then corresponding PCs are chosen for classification.

2.4 Classification

Support Vector Machines (SVMs) are powerful and accurate classifiers, commonly used in remote sensing community. According to the stability of SVM, it is selected as classifier.

SVM is a learning technique derived from statistical learning theory. It is calculating an optimally separating hyperplane that maximizes the margin between two classes. If samples are not separable in the original space, kernel functions are used to map

data into a higher dimensional space with a linear decision function (Abe et al. 2010).

Given a dataset with n samples $\{(x_i, y_i) \mid i = 1,...,n\}$ where $x_i \in \Re^k$ is a feature vector with k components and $y_i \in \{-1,1\}$ denotes the label of x_i. The SVM looks for a hyperplane $w.\phi(x) + b = 0$ in a high dimensional space, able to separate the data from classes 1 and -1 with a maximum margin. w is a weight vector, orthogonal to the hyperplane, b is an offset term and ϕ is a mapping function which maps data into a high dimensional space to separate data linearity with a low training error. Maximizing the margin is equivalent to minimizing the norm of w. thus by solving the following minimization problem, SVM will be trained:

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i$$
$$\text{Subject to: } y_i(w.\phi(x) + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, for\ i = 1, ..., n \qquad (1)$$

where C is a regularization parameter that imposes a trade-off between the number of misclassification in the training data and the maximization of the margin and ξ_i are slack variables. The decision function obtained through the solution of the minimization problem in Equation (1) is given by:

$$f(x) = \sum_{x_i \in SV} y_i \alpha_i \phi(x_i).\phi(x) + b \qquad (2)$$

where the constants α_i are called Lagrange multipliers determined in the minimization process. SV corresponds to the set of support vectors, training samples for which the associated Lagrange multipliers are larger than zero. The kernel functions compute dot products between any pair of samples in the feature space. Gaussian Radial Basic Function (RBF) is a common kernel which is used in this paper and it is defined by (3).

$$K_{Gaussian}(x_i, x_j) = e^{\frac{-|x_i - x_j|}{2\sigma^2}} \qquad (3)$$

In the proposed method, the classification module plays an important role in evaluation of the fitness function where SVM is trained by training data and trained SVM is evaluated by testing (unseen) data.

It is notable that SVM parameters have strong effect on its performance. SVMs parameters are: (*i*) Regularization parameter (C) that determines the trade-off between minimizing the training error and minimizing the complexity of the model; (*ii*) Kernel parameters: bandwidth for Gaussian kernel functions (σ). In order to improve the classification performance, grid search is applied before classification process to determine the best value for these parameters.

3. EXPERIMENTAL RESULTS

3.1 Dataset

The potential of the proposed method, implemented on two airborne datasets acquired at different spectral ranges and spatial resolutions: 1) a coarser-resolution LWIR hyperspectral image and 2) fine-resolution visible imagery. The data were acquired and were provided for the 2014 IEEE Contest by Telops Inc., Québec, Canada. They cover an urban area, including buildings, roads and tree (Liao et al. 2014). The spectral variability of objects in the same class, spectral similarity of some different classes and proximity of 3D objects in urban area, make it complex scene for classification and appropriate for assessment of the proposed method.

The LWIR data were acquired using the Telops Hyper-Cam, an airborne long-wave infrared hyperspectral imager. It consists of 84 spectral bands in the 7.8−11.5 μm wavelength region with approximately 1m resolution. The VIS data consists of uncalibrated data with 0.2m spatial resolution.
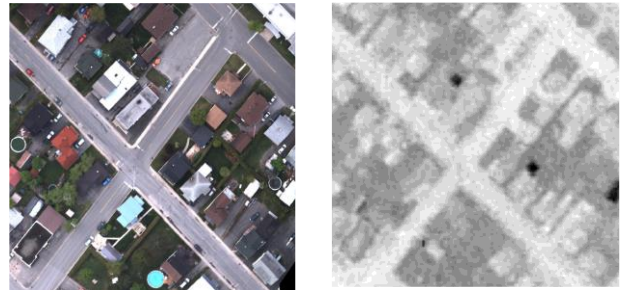

Figure 1. Dataset (a) VIS (b) TIR (second band)

3.2 Feature Space Generation

Generation of feature space is performed by processing both thermal hyperspectral imagery and visible image.

Visible feature space consists of original RGB bands (3 features), SFS features on each band of VIS (18 feature), MP on mean visible bands (12 feature), GLCM descriptors on each visible band (24 features), HSV transformation (3 features) and local statistics on each band of visible image (12). Therefore the visible feature space with 72 spectral and spatial features is generated.

Thermal hyperspectral image has 84 bands. PCA transformation is applied on thermal hyperspectral imagery and 2 first PCs are selected to use for texture analysis. GLCM descriptors apply on both PCs and 16 features are generated. Consequently, the thermal feature space compose of 102 descriptors.

Finally, feature stacking is performed to form feature space. Visible and thermal feature space in addition to vegetation index are merged to construct the hybrid feature space consists of 175 features.

3.3 Dimension Reduction

According to high dimensionality of the hybrid feature space with 175 features, dimension reduction is required to decrease the computational time and complexity. The intrinsic dimension of hybrid feature space is 9 based on the Eigen value. Therefore PCA transformation is applied on the feature space and corresponding PCs are chosen for the classification.

3.4 SVM Classifier Results

To evaluate the classification performance, several accuracy criteria are computed. Kappa coefficient and Overall Accuracy (OA) are common criterion in classification accuracy measurement which is applied in this paper to compare the classification accuracy of thermal image, visible image and fused image. Moreover K_{hat} is used to compare each class accuracy.

SVM classifier is applied to classify visible, thermal and hybrid images. Referring to effect of SVM parameters on classification performance, grid search is used to estimate the SVM parameters. The effect of regularization and kernel parameters is shown in Figure 2.
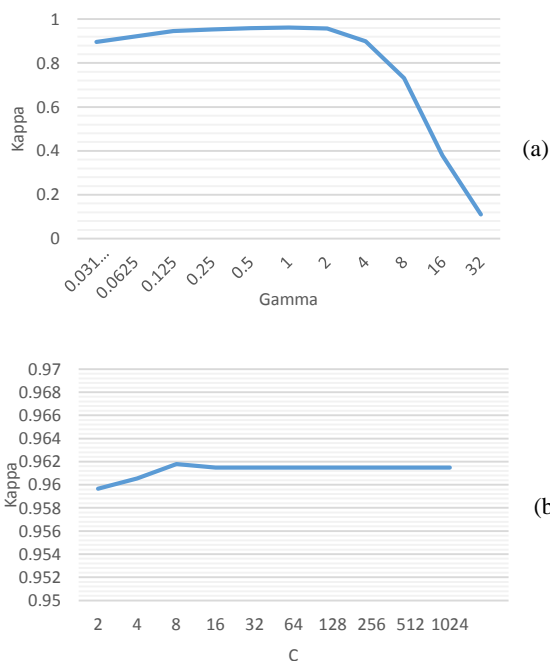
(a)

(b)

Figure 2. Effect of (a) Gamma, C =8 (b) C, Gamma =1 on classification accuracy

Figure 2 shows that SVM performance is strongly affected by the kernel parameter and less so by the regularization parameter. Table 3 presents the obtained results of grid search for visible image, thermal image and the proposed hybrid image.

Table 3. Results of grid search

Dataset	C	Gamma	Kappa	OA
VIS	512	16	0.8867	90.02%
TIR	32	2	0.8388	89.12%
Proposed	8	1	0.9618	97.96%

The obtained results clearly showed the superiority of the proposed algorithm in that the classification accuracy was reached. As Table 3 shows, the classification accuracy of the proposed method improve classification accuracy of visible and thermal imagery, 8% and 9% respectively. More detail about the each class accuracy is shown in Figure 3.
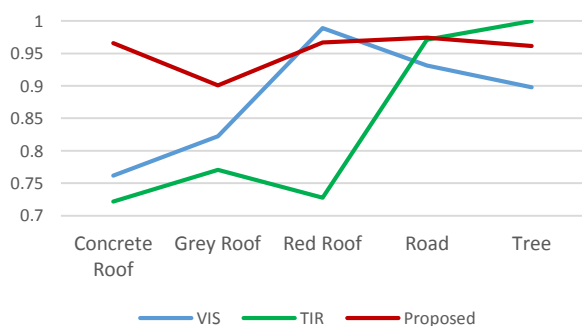

Figure 3. K$_{hat}$ measurement for each class

Per class analysis depicts the superior performance of the proposed method in almost all classes, and for red roof and tree, there is not meaningful difference with visible and thermal

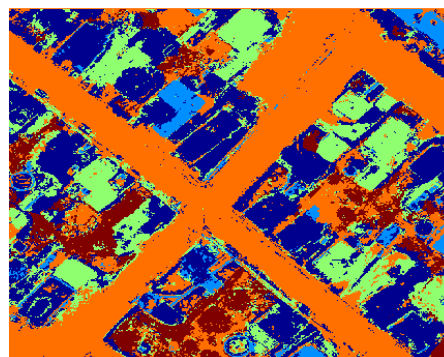image, respectively. The classification map based on proposed method is illustrated in Figure 4.

Figure 4. Classification map

4. CONCLUSION

The obtained results show that simultaneously using thermal and visible images base on feature level fusion, improves classification performance in complex urban area approximately 8% and 9% respect to visible and thermal image, respectively. Also the analysis of per class accuracy is performed for more details by K$_{hat}$ measurement. It proves thermal image has high potential to classify tree in urban area. Also in discrimination of road and building with same spectral reflectance, thermal imagery improves classification performance. Further investigations ought to be carried out for wrapper feature selection instead of applying PCA for dimension reduction. Moreover, object based analysis may be performed to improve the result of proposed pixel-based approach result.

REFERENCES

Benediktsson, J., Pesaresi, M. and Arnason, K., 2003. Classification and Feature Extraction for Remote Sensing Images From Urban Areas Based on Morphological Transformations. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1940-1949.

Berni, J. A., Zarco-Tejada, P. J., Suárez, L., and Fereres, E., 2009. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, No. 3, pp. 722-738.

Brook, A., Vandewal, M., and Ben-Dor, E., 2012. Fusion of optical and thermal imagery and lidar data for application to 3-d urban environment and structure monitoring. *Escalante-Ramirez, B. Remote sensing-Advanced techniques and platforms: China, InTech*, pp. 29-50.

Moran, E. F., 2010. Land cover classification in a complex urban-rural landscape with QuickBird imagery. *Photogrammetric engineering and remote sensing*, Vol. 76, No. 10, pp. 1159-1168.

Chehata, N., Guo, L., and Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. *International Archives of the Photogrammetry, Remote*

Sensing and Spatial Information Sciences, 39(Part 3/W8), pp. 207-212.

Gaszczak, A., Breckon, T. P., and Han, J., 2011. Real-time people and vehicle detection from UAV imagery. *In IS&T/SPIE Electronic Imaging*, pp. 78780B-78780B.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 6, pp. 610-621.

Huang, X., Zhang, L., and Gong, W., 2011. Information fusion of aerial images and LIDAR data in urban areas: vector-stacking, re-classification and post-processing approaches. *International Journal of Remote Sensing*, Vol. 32, No. 1, pp. 69-84.

Huang, X., Zhang, L. and Li, P., 2007. Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery. *IEEE Geoscience and Remote Sensing Letters*, Vol. 4, No. 2, pp. 260-264.

Kumar, U., Milesi, C., Nemani, R. R., and Basu, S., 2015. Multi-sensor multi-resolution image fusion for improved vegetation and urban area classification. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, Volume XL-7/W4, Kona, Hawaii, USA.

Liao, W., Huang, X., Van Coillie, F., Gautama, S., Pižurica, A., Philips, W., Liu, H., Zhu, T., Shimoni, M., Moser, G. and Tuia, D., 2014. Processing of Multiresolution Thermal Hyperspectral and Digital Color Data: Outcome of the 2014 IEEE GRSS Data Fusion Contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 8, No. 6, pp. 2984-2996.

Liao, W., Pizurica, A., Bellens, R., Gautama, S., and Philips, W., 2015. Generalized Graph-Based Fusion of Hyperspectral and LiDAR Data Using Morphological Features. *IEEE Geoscience and Remote Sensing Letters*, Vol. 12, No. 3, pp. 552-556.

Myeong, S., Nowak, D. J., Hopkins, P. F., and Brock, R. H., 2001. Urban cover mapping using digital, high-spatial resolution aerial imagery. *Urban Ecosystems*, Vol. 5, No. 4, pp. 243-256.

Samadzadegan, F., Hasani, H., and Schenk, T., 2012. Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using Ant Colony Optimization. *Canadian Journal of Remote Sensing*, Vol. 38, No. 2, pp. 139-156.

Wentz, E. A., Anderson, S., Fragkias, M., Netzband, M., Mesev, V., Myint, S. W., Quattrochi, S., Rahman, A., and Seto, K. C., 2014. Supporting global environmental change research: a review of trends and knowledge gaps in urban remote sensing. *Remote Sensing*, Vol. 6, No. 5, pp. 3879-3905.

Zhu, Z., Woodcock, C. E., Rogan, J., and Kellndorfer, J., 2012. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sensing of Environment*, Vol. 117, pp. 72-82.

Weng, Q. 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, Vol. 117, pp. 34-49.