

TENSOR MODELING BASED FOR AIRBORNE LiDAR DATA CLASSIFICATION

N. Li^{a,b}, C. Liu^{a,*}, N. Pfeifer^b, J.F., Yin^c, Z.Y. Liao^c, Y. Zhou^a

^a College of Survey and Geoinformation, Tongji University, 200092, Shanghai, China

^b Department of Geodesy and Geoinformation, Technische Universität Wien, 1040 Vienna, Austria

^c Department of Mathmetrics, Tongji University, 200092, Shanghai, China

Commission III, WG III/2

KEY WORDS: Feature Selection, Tensor Processing, KNN classification

ABSTRACT:

Feature selection and description is a key factor in classification of Earth observation data. In this paper a classification method based on tensor decomposition is proposed. First, multiple features are extracted from raw LiDAR point cloud, and raster LiDAR images are derived by accumulating features or the “raw” data attributes. Then, the feature rasters of LiDAR data are stored as a tensor, and tensor decomposition is used to select component features. This tensor representation could keep the initial spatial structure and insure the consideration of the neighborhood. Based on a small number of component features a k nearest neighborhood classification is applied.

1. INTRODUCTION

Land Cover Classification in urban scenes is an important application in Airborne LiDAR point cloud processing. Urban scene classification based on aerial LiDAR points can guide surface reconstruction techniques in urban modeling, piecewise planar surfaces are used for precise building modeling, while vegetation is best represented based on height information provided by LiDAR (Carlberg et al, 2008). An urban scene is usually composed of a complex combination of artificial ground, natural ground, roads, railway, buildings, high vegetation, low vegetation or other objects such as fences and vehicles. Depending on the aim of classification, various classes are defined by researchers. To distinguish objects efficiently, an amount of features derived from LiDAR data are explored, meanwhile automatic classification methods are proposed for efficient classification in urban scene.

For automatic classification of objects, machine learning classifiers were widely used recently. Common machine learning machine learning method includes support vector machine (SVM) algorithm, adaboost, decision trees, random forest and other classifiers. SVM seeks out the optimal hyperplane that efficiently separates the classes, Secord (Secord and Zakhor, 2007) uses a Gaussian kernel function to map non-linear decision boundaries to higher dimensions where they are linear. Adaboost is a binary algorithm, but several extensions are explored for multiclass categorization, a hypothesis generation routines are used to classify terrain and non-terrain area (Lodha et al., 2007). A C4.5 decision tree is used to carry out the classification (Garcia-Gutierrez et al, 2009), by training data and make a hierarchical binary tree model, new objects can be classified based on previous knowledge. Random Forest is an ensemble learning method that used a group of decision trees, provide measures of feature importance for each class (Guo et al., 2011), and runs efficiently on large datasets. All the machine learning need data set for training, so ideal features and thresholds are needed to obtain a good classification result.

The commonly used features can be summarized into spatial information features, amplitude features and multiple-return feature and texture features. The capability of acquiring 3D data increases awareness of LiDAR for land cover classification and

object recognition (Yan et al., 2015). An advanced type of LiDAR spatial information features, i.e. normalized height are used for ground and non-ground segmentation (Scoonl, 2006), point distribution frequency criteria of different type of land cover is considered (Antonarakis et al., 2008), 3D geometry factor are calculated to identify vegetation (Brodu and Lague, 2012). Using airborne LiDAR amplitude data features for urban landcover classification have been discussed in many researches. The amplitude of LiDAR response varies among different object materials, thus amplitude is often used as an input feature for road, buildings and trees classification (Charaniya et al, 2004), and Xu (2013) calibrate the amplitude to improve the land cover classification accuracy. Multiple-return features such as first return LiDAR points are used to detect building edges distorted by multi-path errors in the last return LiDAR data, then threshold of first and last return LiDAR height values are employed to classify roof and edge surfaces (Weed et al, 2001). To achieve texture feature information, multi-spectral bands and low-sampled airborne laser scanning data are combined to distinguish buildings from vegetation (Sohn and Dowman, 2007). An overview on features is given in (Otepka et al., 2013).

Present method are focussed on the processing or feature extraction, but relevance of features is less often studied. Tensor decomposition was used in hyperspectral remote sensing to select relevant features from the spectral bands (Renard and Bourennane, 2009). We transfer this method to LiDAR data. To consider relevance between spatial, amplitude and echo features, a feature tensor is generated and decomposed to get principal features. Then a KNN classification method is applied for urban scene classification. An airborne LiDAR point cloud of Vienna city is used and classified into 4 classes: ground, buildings, vegetation, other point

2. TENSOR BASED CLASSIFICATION METHOD

The principal approach to classify the LiDAR data is to first compute a (large) number of features, some of which may be redundant w.r.t. each other. In the second step, based on the tensor representation and decomposition, the relevant features are extracted. In the final step the “k Nearest Neighbors” method is used to classify the data.

2.1 Feature description

Multiple features are extracted from raw LiDAR points to generate a high-dimensional vector at each point. The input LiDAR features are summarized into 5 groups: height-based features, amplitude-based features, echo-based features, roughness features, and geometry features. The feature vector is composed of 18 components, which will be described below.

The features are first computed, either at each individual point and then aggregated into raster cells, or directly computed from the echo and its inherent features (e.g. the coordinates) for a cell. In the first case the high resolution and 3D content of the LiDAR point cloud is maintained on echo basis, which is then, however, aggregated by taking the mean as representative value in the raster cell. The results are 18 raster feature images.

Some of the features, especially those that are close to the raw measurement, need no further computation. Others, such as “Heightrange”, “HeightRMS” are calculated by points falling into the cell, and “slope” are calculated in a defined neighborhood, which would be larger than cell size.

2.1.1 Height-based Features: Height is defined as z coordinate that is directly recorded by the sensor, which could easily reflect various objects, such as ground, building and high vegetation. Height-based features include:

Heightdiffground: The height from ground is the height difference between individual points and the ground, which indicates ground and non-ground points. If a precise terrain model is already available, this certainly simplifies the classification task considerably. However, in the suggested approach a very simple ground extraction method, using a block minimum filter (Pfeifer and Mandlbauer, 2008) is used in a preprocessing step to obtain a rough approximation of the terrain model. This terrain model is used for computing the height above ground for each point of a cell, which is then averaged to obtain the raster cell value.

Heightrange: Height range is the difference between the highest and lowest echo for the area of each cell. This feature could help discriminating plane targets and trees.

Heightdiffecho: height difference between the average of the first and of the last echos per cell. In this work the definition is used that in the case of single echoes they are simultaneously first and last echo. The vegetation or building edges could be extracted from this feature

2.1.2 Amplitude Features: The amplitude is related to the object reflectance, and same targets should have similar amplitude values. It is, however, output of the detector and not a calibrated geophysical quantity, but depends on the sensor as well as mission parameters. However, given small height variations in elevation and flight path, the analysis of the LiDAR equation (Wagner et al., 2010) shows that the variations should be limited. Intensity normalization (Lin, 2015) would be another option next to calibration in order to reduce the variability. All features are aggregated by taking the mean value per cell.

FirstAmplitude, LastAmplitude, and AmplitudeDiff are the average of the first echo amplitudes, the last echo amplitudes, and their difference, respectively.

Ground and building have a single return, so their value should be (close to) zero. Vegetation, on the other hand, should also be detected based on this feature.

2.1.3 Echo-based Features: NrofEcho: this feature depends on the number of echoes per emitted pulse. It is the average number of echoes per cell. This feature is high for vegetation and building facades and relatively low for ground and building roofs

2.1.4 Roughness Features: The surface roughness is a significant feature for vegetation and building façade identification. Multiple roughness features are used in this paper.

HeightRMS: the root mean square of all height value in the raster cell.

NormalPlaneOffset: mean offset from the current points to the estimated local plane in the neighborhood.

NormalZ RMS: The root mean square of the normal vector (see below) z-component.

Slope: The steepest slope (maximum value) in the cell, given in percent.

ER (Echo ratio): the echo ratio is a measure for local transparency and roughness. It is defined as follows (Höfle et al., 2012):

$$ER = n_{3D}/n_{2D} \times 100$$

With $n_{3D} \leq n_{2D}$, n_{3D} is the number of neighbors found in a certain search distance measured in 3D and n_{2D} is the number of neighbors found in same distance measured in 2D. The ER is nearly 100% for flat surface, whereas the ER decreases for penetrable surface parts since there are more points in a vertical search cylinder than there are points in a sphere with the same radius.

2.1.5 Geometry Features: The spatial distribution of points in a fixed neighborhood can be reflected according to the geometry features. It could help discriminating buildings from vegetation. NormalX, NormalY, NormalZ are the normal vectors of local planes, which are estimated by points in a small neighborhood. The covariance matrix for the normal vectors is computed to find the eigenvalues $\lambda_1, \lambda_2, \lambda_3$.

NormalEigenvalue3: stands for λ_3 and used as features. λ_3 has low values for planar object and higher values for voluminous point clouds. Two structure features derived from that are anisotropy and sphericity and introduced to describe the spatial local points' distribution (West, 2004).

$$\text{Anisotropy} = (\lambda_1 - \lambda_3) / \lambda_1$$

$$\text{Sphericity} = \lambda_3 / \lambda_1$$

2.2 Tensor representation for LiDAR data

In each raster cell the feature vector with 18 components is computed as described above (Heightdiffground, Heightdiffecho, Heightrange, FirstAmplitude, LastAmplitude, Amplitudediff, NrofEcho, HeightRMS, Slope, NormalPlaneOffset, ER, NormalZ RMS, NormalX, NormalY, NormalZ, NormalEigenvalue3, Anisotropy, Sphericity).

The high-dimensional feature vectors of LiDAR data are considered as a third-order tensor, the entries of which are accessed via three indexes. It is denoted by $R \in R^{I \times J \times K}$, $R \in R^{I \times J \times K}$, with element arranged as r_{ijk} , where $i = 1, \dots, I$; $j = 1, \dots, J$; $k = 1, \dots, K$; and R is the real manifold. Each index is called mode: two spatial modes and one feature mode characterize the LiDAR feature tensor (see also Figure 1). The tensor representation is explored to process the whole data from spatial and feature perspectives. The Tucker decomposition applied here is a form of higher-order principal component analysis (Renard and Bourennane, 2009). It decomposes a tensor into a core tensor multiplied by a matrix along each mode as shown in Figure 2. Tucker decomposition is expressed as:

$$R \approx C \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$$

Here, $U^{(1)} \in \mathbb{R}^{I \times P}$, $U^{(2)} \in \mathbb{R}^{J \times R}$, $U^{(3)} \in \mathbb{R}^{K \times Q}$ are the factor matrices and could be considered as the principal components in each mode. $C \in \mathbb{R}^{P \times J \times Q} \times_n$ is the n -mode product, $C \in \mathbb{R}^{P \times J \times Q}$ is the core tensor, and its entries show the level of interaction between the different components (Tamara G Kolda and Brett W Bader, 2007). If $P, R, Q < I, J, K$, $P, R, Q < I, J, K$, the core tensor C can be considered as a compressed version of the raw tensor. Thus, the principal features are achieved by projecting raw features into lower dimensional subspace, the projection is based on following equation:

$$R_{pc} = R \times_3 U^{(3)}$$

R_{pc} is a reduced three-order tensor $\mathbb{R}^{I \times J \times P}$, holding the p components, generalizing the product between a tensor and a matrix along an n -mode. And $U^{(3)}U^{(3)}$ is the eigenvectors of raw tensor along 3-mode. Finally, R_{pc} holds the principal features after projection.

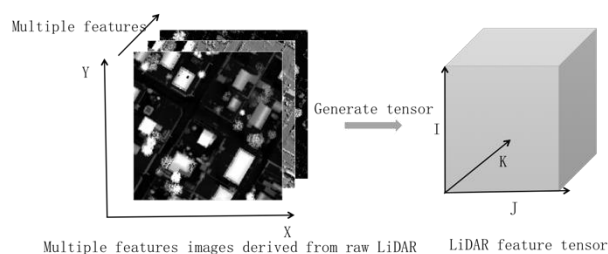


Figure 1. Feature Tensor generation

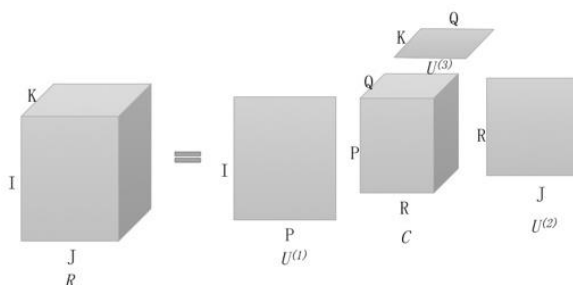


Figure 2. Tucker decomposition

2.3 LiDAR Classification based on KNN

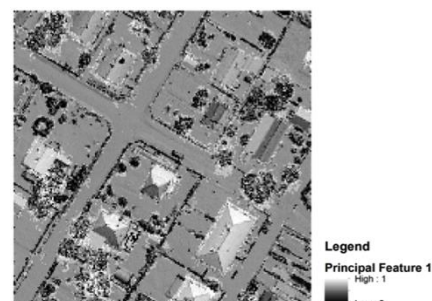
KNN (k nearest neighbours) is an algorithm that stores all available cases and classifies new cases based on a similarity measurement, such as distance function. The case is classified by a majority vote of its neighbours, with the case being assigned to the class most common amongst its k nearest neighbours measured by a distance function. This paper defined Classification in this paper is based on single raster pixel, principal features after Tucker decomposition are considered as input data, and distance function is defined as Euclidean distance between each features vector.

3. CLASSIFICATION RESULT AND EVALUATION

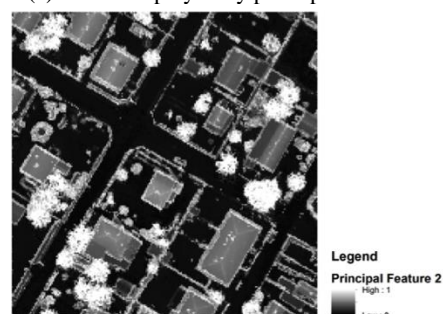
A section of airborne LiDAR points of Vienna city is used and classified into 4 classes: ground, buildings, vegetation and other points. The area is $100m \times 100m$, and cell size of the raster image is defined as $0.5m$. 18 features images are extracted from raw data based on spatial, amplitude and echo attributes, then a 3-order $201 \times 201 \times 18$ feature tensor is generated. By Tucker decomposition 5 principal features are selected, and normalized principal features are taken as input data for classification based on KNN. Figure3 is the points cloud displayed by 2 principal feature(due to the limitation of pages, only 2 principal feature

displayed), object differences are enhanced by the principal features.

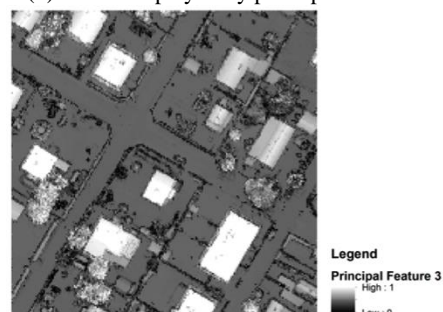
This paper takes 30% data as training data, and 70% as test data to evaluate the classification result. Figure4(a) is the reference classification result and Figure4(b) is the classification result achieved by this paper.



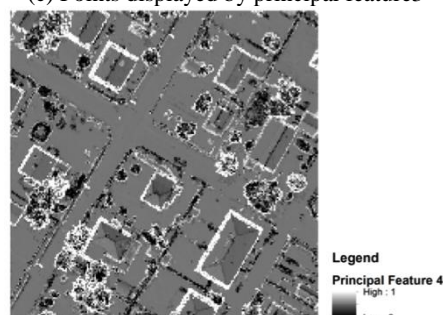
(a) Points displayed by principal feature1



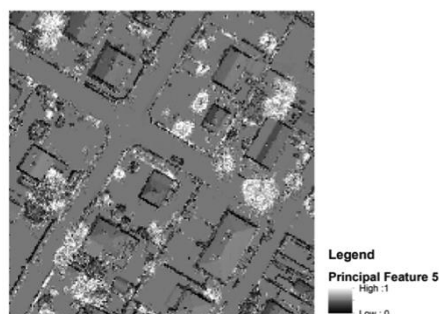
(b) Points displayed by principal feature2



(c) Points displayed by principal feature3



(d) Points displayed by principal feature4

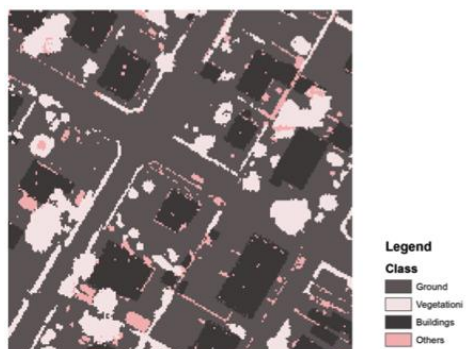


(e) Points displayed by principal feature5

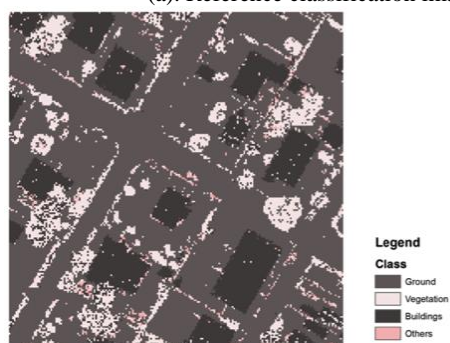
Figure 3. Point cloud displayed by 5 principal features

This paper takes 30% data as training data, and 70% as test data to evaluate the classification result. Figure4(a) is the reference classification result and Figure4(b) is the classification result achieved by this paper.

Table1 indicates overall accuracy and misclassified rate in each class. The ground is well classified, the overall accuracy reaches 97%. Building classification overall accuracy is 88.3%, a few building points are misclassified into ground. However the algorithm has difficulty in classifying vegetation and ground, and a amount of vegetation are classified into buildings. Other points such as construction and cars can hardly be detected.



(a). Reference classification image



(b) Classification in this paper

Figure 4. Classification reference and result for comparison

| | Classes provided by this paper | | | |
|------------|--------------------------------|------------|----------|--------|
| Class | Ground | Vegetation | Building | Others |
| Ground | 97.0% | 2.1% | 0.6% | 0.3% |
| Vegetation | 26.3% | 61.0% | 11.6% | 1.1% |
| Building | 7.0% | 4.6% | 88.3% | 0.1% |
| Others | 50.0% | 30.3% | 13.6% | 6.3% |

Table 1. Classification accuracy

4. CONCLUSION

In this paper the Tucker decomposition of tensors was demonstrated for LiDAR data classification. In this first work a set of features was selected and rasterized features were input for the tensor decomposition. The five most important components (principal features), in comparison to 18 features, were selected for a KNN classification. The results show that ground and buildings could be well detected. Other classifiers (neural network, SVM) can be considered as alternatives for the classification task.

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China (Project No.41371333)

REFERENCES

- Antonarakis, A.S., Richards, K.S., Brasington, J., 2008. Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment* 112, 2988–2998.
- Brodu, N., Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68, 121–134.
- Carlberg, M.A., Gao, P., Chen, G., Zakhor, A., 2008. Urban Landscape Classification System Using Airborne LiDAR. in *IEEE International Conference on Image Processing*, 2009.
- Garcia-Gutierrez, J., Goncalves-Secob, L., and Riquelme-Santosa, J. C., 2009. Decision trees on LiDAR to classify land uses and covers, *IAPRS Conference on Laserscanning09* (pp. 323–328), Paris: ISPRS.
- Charaniya, A.P., Manduchi, R., and Lodha, S.K., 2004. Supervised Parametric Classification of Aerial LiDAR Data. *IEEE Workshop on Real Time 3D Sensor and their Use*, Washington DC.
- Guo, L., Chehata, N., Mallet, C., Boukir, S., 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 56–66.
- Höfle, B., Hollaus, M., Hagenauer, J., 2012. Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 67, 134–147.
- Lin, Y.-C., 2015. Normalization of Echo Features Derived from Full-Waveform Airborne Laser Scanning Data. *Remote Sensing* 7(3), 2731–2751.
- Lodha, S.K., Fitzpatrick, D.M., Helmbold, D.P., 2007. Aerial Lidar Data Classification Using AdaBoost *Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling*, p.435–442, August 21–23, 2007
- Otepka, J., Guhffar, S., Waldhauser, C., Hochreiter, R., Pfeifer, N., 2013. Georeferenced Point Clouds: A Survey of Features and Point Cloud Management. *ISPRS Journal of Geo-Information* 2(4), 1038–1065;

Pfeifer, N., Mandlbauer, G., 2008. Filtering and DTM Generation. In: "Topographic Laser Ranging and Scanning: Principles and Processing", Shan, Toth (eds.). CRC Press, 307-333.

Renard, N., Bourennane, S., 2009. Dimensionality Reduction Based on Tensor Modeling for Classification Methods. IEEE Trans. Geosci. Remote Sensing 47, 1123–1131.

Scoonl, 2006. Object-oriented land cover classification of lidar-derived surfaces 1–11.

Secord, J., Zakhor, A., 2007. Tree Detection in Urban Regions Using Aerial Lidar and Image Data. IEEE Geosci. Remote Sensing Lett. 4, 196–200.

Sohn, G., Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 62, 43–63.

Tamara G Kolda and Brett W Bader, S.N.L., 2007. Tensor Decompositions and Applications 1–71.

Wagner, W., 2010. Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. ISPRS Journal of Photogrammetry and Remote Sensing, 65(6), 505-513.

Weed, C.A, Crawford, M.M., Neuenschwander, A.L., Gutierrez, R., 2001. Classification of LIDAR data using a lower envelope follower and gradient-based operator - Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International 1–3.

Xu, G.C., Pang, Y., Li, Z.Y., Zhao, D., and Li, D., 2013. Classifying land cover based on calibrated full-waveform airborne light detection and ranging data. Chin. Opt. Lett. 11, 082801.

Yan, W.Y., Shaker, A., El-Ashmawy, N., 2015. Urban land cover classification using airborne LiDAR data: A review. Remote Sensing of Environment 158, 295–310