

3D SEMANTIC LABELING OF ALS DATA BASED ON DOMAIN ADAPTION BY TRANSFERRING AND FUSING RANDOM FOREST MODELS

Jianwei Wu^{1,2}, Wei Yao^{2,*}, Jiyong Zhang³, Yunfan Li⁴

¹School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China - jianwei_wu@whu.edu.cn

²Dept. of Land Surveying & Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong - wei.hn.yao@polyu.edu.hk

³State Grid Economic and Technological Research Institute CO.LTD, Beijing, China - zhangjiyong@chinasperi.sgcc.com.cn

⁴Changjiang River Scientific Research Institute of Changjiang Water Resources Commission, Wuhan, China - 179727719@qq.com

Commission III, WG III /6

KEY WORDS: 3D semantic labelling, ALS data, random forest, domain adaption, decision fusion

ABSTRACT:

Labeling 3D point cloud data with traditional supervised learning methods requires considerable labelled samples, the collection of which is cost and time expensive. This work focuses on adopting domain adaption concept to transfer existing trained random forest classifiers (based on source domain) to new data scenes (target domain), which aims at reducing the dependence of accurate 3D semantic labeling in point clouds on training samples from the new data scene. Firstly, two random forest classifiers were firstly trained with existing samples previously collected for other data. They were different from each other by using two different decision tree construction algorithms: C4.5 with information gain ratio and CART with Gini index. Secondly, four random forest classifiers adapted to the target domain are derived through transferring each tree in the source random forest models with two types of operations: structure expansion and reduction-SER and structure transfer-STRUT. Finally, points in target domain are labelled by fusing the four newly derived random forest classifiers using weights of evidence based fusion model. To validate our method, experimental analysis was conducted using 3 datasets: one is used as the source domain data (Vaihingen data for 3D Semantic Labelling); another two are used as the target domain data from two cities in China (Jinmen city and Dunhuang city). Overall accuracies of 85.5% and 83.3% for 3D labelling were achieved for Jinmen city and Dunhuang city data respectively, with only 1/3 newly labelled samples compared to the cases without domain adaption.

1. INTRODUCTION

Assigning each airborne laser scanning (ALS) point with correct object class - 3D semantic labelling of ALS data, is still a challenging and complicated task in both computer vision and remote sensing community. It is the basic step of much application processing such as highly accurate mapping, object extraction, building modelling and so on. Sufficient and large volume labelled samples are always essential to accurately classifying 3D point cloud with supervised learning methods. While for traditional semantic labelling tasks of ALS point clouds, the labelled samples are data and scene dependent. It needs to spend a lot of efforts and time to re-collect training samples for new data. Therefore, how to exploit existing labelled LiDAR points and the derived models to mitigate the needs of large volume of samples has been attracting more and more concerns from geospatial computer vision community.

Transfer learning is an important sub-branches of machine learning, which aims to improve the learning of the target prediction function using the knowledge in source domain \mathcal{D}_s and source learning task \mathcal{T}_s (Pan and Yang, 2010). Knowledge in source domain \mathcal{D}_s can be labelled samples or derived models. Thus, in both computer vision (Gong et al., 2014) and remote sensing community (Tuia et al., 2016), transfer learning methods have been researched to reuse collected samples and existing models to mitigate the needs of large volume of samples for supervised classification.

There are three kinds of transfer learning methods in terms of detailed implementation techniques: instances transfer, feature transformation and model adaption (Pan and Yang, 2010). While according to the different settings, transfer learning includes inductive transfer learning, transductive transfer learning and unsupervised transfer learning. Domain adaption (DA) has been attracting lots of attention from remote sensing community (Persello and Bruzzone, 2016; Tuia et al., 2016), which belongs to transductive transfer learning. In general, DA aims to adapt models trained to solve a specific task to a new yet related task, for which the knowledge of the initial model is sufficient, although not perfect (Tuia et al., 2016).

Existing methods are mainly applied on image classification; thus it is very interesting to investigate whether domain adaption methods apply to 3D point cloud labelling. This paper studied a model transfer method by (Segev et al., 2015) with domain adaption concept, which adapts the learned random forest model in source domain to target domain with fewer samples for semantic labelling of ALS data. As non-linear models, decision trees (DTs) can excel in learning non-linear decision rules and their hierarchical structure enables detection and accommodation of non-linear transformations from source to target. On the other hand, local adjustment of the tree structure can solve the domain shift to some extent. A single classifier can describe two identical domains. Therefore, as one domain drifts, the changes can be captured via small

* Corresponding author

modifications to the tree structure. Thus, through transferring all the tree models to target domain, newly adapted random forest models are derived. To enhance final labelling accuracy, the weights of evidence based fusion model is adopted and discussed. The remainder of this paper is organized as follows. Section 2 describes the method. Experiment and results are shown in section 3. Finally, it concludes and proposes the future work in section 4.

2. METHOD

2.1 Main workflow

The main steps of the proposed method are as following:

1. Training two initial random forest models (Breiman, 2001) using labelled samples in source domain based on Gini index (Breiman, 1984) and information gain ratio (Quinlan, 1993), respectively;
2. Adapting learned random forest models of source domain to target domain using some labelled samples from the target domain:
 - a. updating each classification tree in the initial models through *structure expansion and reduction*-SER (Segev et al., 2015);
 - b. adapting the structure of each tree in the initial model through modifying the splitting node's threshold based on maximizing Gini index and information gain ratio, which is equal to tree *structure transfer*-STRUT (Segev et al., 2015);
3. Fusing the four adapted random forest models by weight-based evidence fusion to label each point in the target domain accurately.

2.2 Structure Expansion and Reduction (SER) for Tree Model Adaption

In the construction process of decision tree, node splitting and leaf pruning are the two main important operations. Thus, tree structure expansion and reduction are naturally an approach to adapt the source domain model to target domain using labelled samples from target domain. The detailed steps of SER for each tree in the random forest classifier are as following (Segev et al., 2015):

1. For each node V , calculating the set D_v^T of all labelled points in the target data D^T that reaches V ;
2. Tree structure expanding: for every leaf node V_l in the tree, expanding its structure to a full tree based on its corresponding reached sample set D_v^T ;
3. Tree structure reduction: this step is similar to leaf pruning working in the bottom-up pattern. Here, for each node V , it relates to two types of error: one is defined as SubTreeError (V)- corresponding to the error when node V is not pruned and taken as the root of a non-empty subtree; the other is called as LeafError (V), which is the error when V is pruned and together with all its child nodes become as the direct child of V 's parent node. When the error LeafError(V) is smaller than SubTreeError(V), the internal node V is pruned to be a leaf node. It should be noted that structure expanding operation should be performed before structure reduction. Following this operation

order, the original structure can be retained better when adapting the decision tree's structure, which equals to transfer the source domain information better. After SER operation, the new decision rule at each leaf node in modified tree will depend more on the target domain distribution.

2.3 Structure Transfer (STRUT) Method for Tree Model Adaption

For classification of two similar scenes, their data classification trees should also be with similar structures. Motivated by such observation, the decision tree can be modified just through tree structure transfer operation: adjusting each node's decision value (threshold of the selected numeric feature) to be adapted to target domain (Segev et al., 2015). The correctness of decision value determines decision tree's accuracy. When using target domain samples to optimize the numeric feature's threshold of the source domain's tree, it tends to cause the model to be negative or over-transferred without constraints. STRUT works in top-down pattern. For a decision tree trained with source domain data, its structure is modified by STRUT through modifying the node's decision value meanwhile with the samples from target domain when traversing the decision tree from root to each node:

1. For the node with reached target domain samples, its decision value is optimized as following:

$$\begin{aligned} & \text{Maximize InformationGainRatio}(D_v^T, \theta, x) \text{ or Giniindex}(D_v^T, \theta, x) \\ & \text{s.t.} \\ & \max(\text{differ}(Q_v^L, Q_v^L), \text{differ}(Q_v^R, Q_v^R)) < \beta \cdot \min(|Q_v^L|, |Q_v^R|) \end{aligned} \quad (1)$$

where InformationGainRatio(D_v^T, θ, x) or Giniindex(D_v^T, θ, x) calculate the information gain ratio or Gini index when adjust the threshold(x) of decision feature θ at current reached node V ; Q_v^L, Q_v^L is the target sample distribution of V 's left subtree before and after adjusting the decision value; similar to $Q_v^L, Q_v^L, Q_v^R, Q_v^R$ is for V 's right subtree; $\text{differ}(Q_v^L, Q_v^L)$ counts the number of different samples between Q_v^L and Q_v^L ; $|Q_v^L|$ is the sample size in Q_v^L , β is a parameter to control how much the decision value can be adjusted in fact.

2. While for those nodes without reached target domain samples, they should be pruned.

2.4 Decision Level Fusion of Adapted Random Forest Classifiers based on Weights of Evidence Model

WofE model adopts the prior and conditional probability to generate a logit posterior odds, which is used to examine the support for a given hypothesis (Good, 1985). Given N independently trained classifiers f_1, f_2, \dots, f_N of data \mathbf{D} with C classes, for each predicted sample in \mathbf{D} , the weight of evidence $W(w_j : f_i)$ for the j th class (w_j) with the i th classifier f_i can be expressed as follows (Song and Li, 2014):

$$W(w_j : f_i) = \log_e \frac{P(w_j | f_i)}{1 - P(w_j | f_i)} \quad (2)$$

where, the conditional probability that one sample is classified as w_j or not with classifier f_i is denoted as $P(w_j | f_i)$ and $1 - P(w_j | f_i)$ respectively.

Given $P(w_j)$ as the prior probability for class w_j , then the log posterior odds for classifier f_i can be calculated as:

$$\log_e O(w_j | f_i) = \log_e O(w_j) + W(w_j : f_i) \quad (3)$$

In equation (3), $O(w_j) = P(w_j) / (1 - P(w_j))$ is an odds operation, $P(w_j)$ is determined based on the average result of N classifiers' results (in this paper, $N=4$ and the classification model is random forest). Considering the different contributions of the N independent classifiers, the log posterior odds for classification fusion can be calculated as:

$$\log_e O(w_j | f) = \log_e O(w_j) + \sum_{i=1}^N \alpha_i W(w_j : f_i) \quad (4)$$

where, the weight α_i represents the classification reliability. Lastly, the final label for a sample can be determined as class w_j , which is obtained with the maximal logit posterior odds:

$$\text{Class}(j) = \arg \max_{j=1,2,C} \log_e O(w_j | f) \quad (5)$$

Through using different classification rule (Gini index and information gain ratio) and tree model adaption methods (SER and STRUT), we obtain four independent classifiers based on standard random forest. Given the fact that the probability output of the classifiers is used as input, final decision level fusion is performed with WofE model. The main workflow for combining the four different random forest classifiers to label point cloud data in target domain is depicted as following (Figure 1):

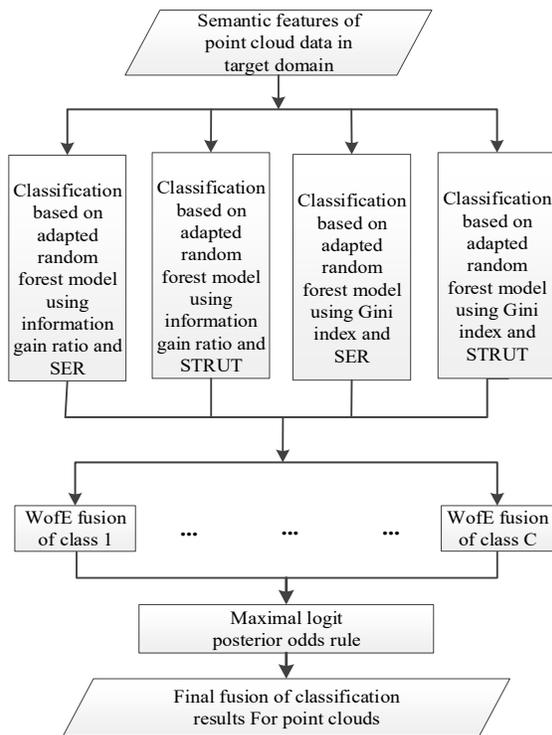


Figure 1. Workflow for point cloud data classification based on WofE fusion

3. EXPERIMENT AND RESULTS

3.1 Data material

Three ALS point cloud data sets were used in this experiment: one was exploited as source domain data to train the initial random forest models; the other two data sets were used as target domain data to validate the proposed complete approach based on model transfer and decision fusion for point cloud data labelling. The source data is from one of the 3D semantic labelling benchmark training data-the labelled Vaihingen data set (Niemeyer et al., 2014), which consists of 9 classes: powerline, low vegetation, impervious surfaces, car, fence, roof, facade, shrub and tree. Two test data of target domain are acquired from Jinmen city (China) and Dunhuang city (China), collected by Leica ALS 50-II and ALS 60 respectively. The point density of about 2~3 points/m² (area size: 1600m*500m) and 1~2 points/m² (area size: 1400m*600m) can be obtained for the Jinmen and Dunhuang data, respectively. The main object classes in the target domain data comprise buildings, bare grounds, vegetation and some vehicles.

3.2 Experiments and Results

In this experiment, only 4 classes were labeled considering the point density: ground, low vegetation, tree and building. To make source domain has same classes with the target domain, in the initial random forest models' training in source domain, the labelled impervious surfaces points were used as the training samples for ground and roof points as building points. For the Jinmen data, the classes' ratio is about: ground (44.1%), low vegetation(8.41%), tree(13.3%), building(25.69%); while for the Dunhuang data, the classes' ratio is about: ground (32.8%), low vegetation(6.41%), tree(10.9%), building(44.74%).

3.2.1 Features used. To label the point cloud, features including some semantic features were extracted and used, which included (Wei et al., 2012):

- I: Intensity, which is provided by the LiDAR system for each point;
- ΔI: Intensity difference between points having the highest and lowest intensities within the cuboid neighbourhood;
- σI: Standard deviation of intensity of points within the cuboid neighbourhood;
- ΔZ: Height difference between the highest and lowest points within the cuboid neighbourhood;
- σZ: height standard deviation of points within the cuboid neighbourhood;
- Two eigenvalue-based features: the planarity P_λ and omnivariance O_λ .

To derive the above features, a 3D cuboid neighbourhood is defined with help of a 2D square with radius of 2m in horizontal dimension. For the calculation of eigenvalue-based features, the 3D covariance matrix $C \in \mathbb{R}^{3 \times 3}$ is calculated for a given point and its neighbours. Since C is a symmetric, its three eigenvalues exist. The three eigenvalues $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ represent the extent of a 3D covariance ellipsoid along its main axes and are thus suitable for describing

the local 3D structure. Then, the planarity P_λ can be calculated as equation (6):

$$P_\lambda = (\lambda_2 - \lambda_3) / \lambda_1 \quad (6)$$

omnivariance O_λ can be calculated as equation (7):

$$O_\lambda = \sqrt[3]{\prod_{i=1}^3 \lambda_i} \quad (7)$$

3.2.2 Experiment process. Using the 7 features defined in section 3.2.1 and labelled samples of the source domain data, two initial classifiers based on random forest were generated based on the Gini index and information gain ratio respectively. Both random forest models were trained with 200 trees and four classes' samples as described in section 3.1: ground, low vegetation, tree and building. Then, each tree in the two source models were adapted with some labelled samples in the target domain using the two methods: SER and STRUT. Thus, four new random forest models adapted with target domain samples were derived. To avoid negative transfer or over-transfer to some extent, in the experiment of STRUT, we constrained that the adjusted threshold should cause not more than 20% samples' label results changed and achieved maximal gain in information gain ratio or gini index at the same time. Finally, each point in the target domain was labelled through fusing outputs of the four new random forest models with WoFE model.

3.2.3 Results and discussion. The final classification results with our approach are shown in Figure 2(b) and Figure 3(b) in the following:

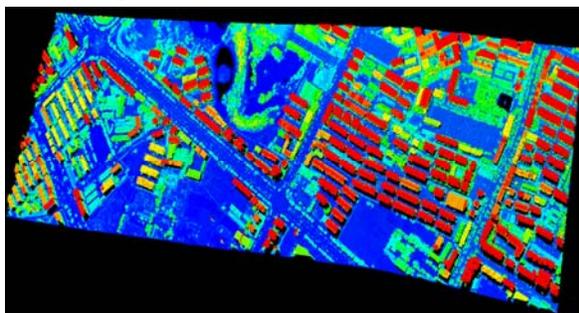


Figure 2(a). Jinmen data coloured by height

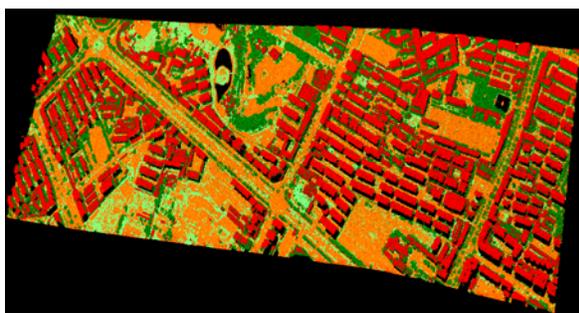


Figure 2(b). Jinmen data classification result, with: ■ ground, ■ low vegetation, ■ tree, ■ building

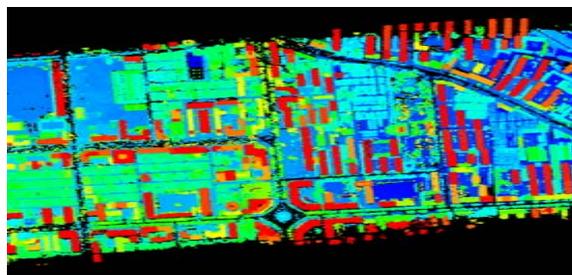


Figure 3(a). Dunhuang data coloured by height

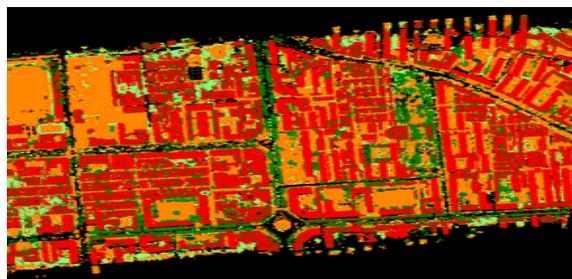


Figure 3(b). Dunhuang data classification result, with: ■ ground, ■ low vegetation, ■ tree, ■ building

The overall accuracies for labelling Jinmen city and Dunhuang city data sets are 85.5% and 83.3%, respectively. As with more similar point density with the source domain training data and lower building ratio in the data, the Jinmen city data achieved better results. As class facade, class car and class shrub were not trained in the source domain, the facade points were always mislabelled as low vegetation or tree points, while the car and shrub points were usually mislabelled as points of low vegetation. Since only 7 types of ALS features were used and the neighbourhood size was fixed, the classification could be improved further if more diverse semantic features were exploited and the local neighbourhood size for each point was optimized adaptively (Weinmann, 2015).

If only target domain labels were used without the trained models from the source domain, more labelled samples are needed. In this experiment, only 1/3 samples from target domain were essential to achieve similar accuracy compared to the cases without adopting domain adaption processing.

4. CONCLUSION AND FUTURE WORK

This paper developed and studied a method to transfer source domain information to target domain for the semantic labelling of 3D point cloud data, which was implemented through knowledge transfer modelling by adapting decision tree structure of source domain into target domain of labelled samples. Using some labelled samples in target domain, the SER and STRUT methods were adopted to adjust decision tree structure by expanding or pruning tree and re-optimizing new feature threshold for each tree internal node. Besides, a decision fusion approach based on WoFE model was exploited to enhance the final classification accuracy. To validate the complete approach, an initial experiment was conducted. In the

experiment, overall accuracies of 85.5% and 83.3% for point labelling were achieved in two urban scenes. More importantly, it reduced 2/3 samples needed to be collected in the target domain, which demonstrated that the method could exploit the source domain information to some extent and reduce the number of samples needed to be labeled in target domain. However, there are needs for more experiments to test the method in following aspects:

1. In this paper, only 4 classes objects were labeled, test work on point labelling with more classes should be conducted in further;
2. More diverse semantic features can be derived and local neighbourhood optimization for feature extraction should be adopted to enhance the feature space representation.

For the future work, the ability of model transfer for this method with regard to classification scenes, data characteristics (point density, point attributes like intensity, echo and colour information) and object types, and how much the fusion processing can improve point labelling accuracy in the context of domain adaption will be evaluated and researched.

ACKNOWLEDGEMENTS

This work is supported by the National Key Research and Development Program of China (No. 2016YFF0103503) and "The Application Research of the Remote Sensing Technology on Global Energy Internet" (No.JYYKJXM (2017)011) in part. The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010): <http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html>.

REFERENCES

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L. I., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees (cart). *Encyclopedia of Ecology*, 40(3), 582-588.
- Gong B., K. Grauman, and F. Sha, (2014). Learning kernels for unsupervised domain adaptation with applications to visual object recognition, *Int. J. Comput. Vis.*, vol. 109, no. 1–2, pp. 3–27.
- Good, I J. (1985). Weight of Evidence: A Brief Survey. *Bayesian Statistics*, 1(3):249-269.
- Niemeyer, J.; Rottensteiner, F.; Sörgel, U. (2014). Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 87, January 2014, pp 152-165.
- Pan, S J. and Yang, Q. (2010). A Survey on Transfer Learning[J]. *IEEE Transactions on Knowledge & Data Engineering*, 22(10):1345-1359. <http://doi.ieeecomputersociety.org/10.1109/TKDE.2009.191>

Persello, C. and Bruzzone, L. (2016). Kernel-based domain invariant feature selection in hyperspectral images for transfer learning, *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2615–2626

Quinlan, J. R. (1993). C4.5: programs for machine learning. *Morgan Kaufmann*, vol. 1.

Segev, N., Harel, M., Mannor, S., Crammer, K., & Ran, E. Y. (2015). Learn on source, refine on target: a model transfer learning framework with random forests. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, PP(99), 1-14

Song B, Li, P. (2014). A novel decision fusion method based on weights of evidence model. *International Journal of Image & Data Fusion*, 5(2):123-137.

Tuia, D., Persello, C., & Bruzzone, L. (2016). Domain adaptation for the classification of remote sensing data: an overview of recent advances. *IEEE Geoscience & Remote Sensing Magazine*, 4(2), 41-57.

Wei, Y., Yao, W., Wu, J., Schmitt, M., & Stilla, U. (2012). AdaBoost-based feature relevance assessment in fusing airborne LiDAR and image data for classification of trees and vehicles in urban scenes, *XXII ISPRS General Congress, Proceedings. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume 1-7, 2012, Melbourne, Australia.

Weinmann, M., Jutzi, B., Hinz, S., & Mallet, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry & Remote Sensing*, 105, 286-304.