

ON A KNOWLEDGE-BASED APPROACH TO THE CLASSIFICATION OF MOBILE LASER SCANNING POINT CLOUDS

Mathias Lemmens^{1,*}

¹ GIS Technology Section, Department OTB, Faculty of Architecture and the Built Environment, Delft University of Technology, 2628 BL Delft, The Netherlands; M.J.P.M.Lemmens@tudelft.nl

Commission IV, WG IV/8

KEY WORDS: Point clouds, mobile laser scanning, knowledge-based system, feature extraction, classification

ABSTRACT:

A knowledge-based system exploits the knowledge, which a human expert uses for completing a complex task, through a database containing decision rules, and an inference engine. Already in the early nineties knowledge-based systems have been proposed for automated image classification. Lack of success faded out initial interest and enthusiasm, the same fate neural networks struck at that time. Today the latter enjoy a steady revival. This paper aims at demonstrating that a knowledge-based approach to automated classification of mobile laser scanning point clouds has promising prospects. An initial experiment exploiting only two features, height and reflectance value, resulted in an overall accuracy of 79% for the Paris-rue-Madame point cloud bench mark data set.

1. INTRODUCTION

Around 2003 mobile laser scanning (MLS) systems became operational for 3D mapping of outdoor scenes. Today MLS systems are used for surveying a variety of scenes with a point density reaching a thousand or more points per cubic meter. Within photogrammetry the exploration of MLS point clouds is directed towards 3D mapping: the outlining of objects along road scenes which are of interest for a particular task at hand, that means use case.

Interpretation of images and point clouds concerns the identification and outlining of objects. For more than a century, the interpretation of aerial images is a profession of which methodologies and technologies are underpinned by sound concepts. Since manual image interpretation is a labour intensive and tedious task, automation has always been at the forefront of ambitions of photogrammetric professionals. With the emergence of computers in the fifties and sixties, it seemed that a completely automated pipeline from image to map would become a reality. Or simply stated: scan an aerial photo, store the pixels in the computer and extract a topographic map out of it without human intervention. At that time computer vision systems came into practice, which could automatically detect and outline objects lying on a conveyor belt in a factory or other restricted scenes. Inspired by the successes of computer vision - why would the same not be possible for aerial images? - own research on automatically detecting road networks started in the mid-eighties (Lemmens, et al., 1988).

Indeed, exploiting the achievements of Artificial Intelligence (AI) and computer vision seemed to be the Holy Grail for fully automation. Later on the knowledge-based systems came in view (de Gunst et al. 1991; de Gunst and Lemmens, 1992; Vosselman and de Gunst, 1997; Zhang and Baltsavias, 2000).

However, after a dozen years of toil, disappointment came – fully automation of mapping from aerial and satellite images appeared to be an illusion; only parts of the photogrammetric pipeline could be fully automated, including aero-triangulation, generation of Digital Surface Models (DSM), and creating orthoimages and digital landscapes. The latter consist of the superposition of orthoimages on DSMs from which a surveyor can measure objects of interest by roaming and clicking a mouse. However, the generation of Digital Elevation Models (DEM) from DSMs through ground filtering, which is a binary classification process, still requires manual editing step.

Today, computer vision has developed new powerful tools which spark the hope on automated mapping, while passive image sensors got accompanied by active sensors, particularly Lidar, the last two decades. Mounted on manned or unmanned aircraft, cars, vans or boats, and also on human back or on hand-held sticks Lidar is able to produce very dense point clouds consisting of billions of points in only a view hours of surveying.

Today's Holy Grail is called Deep Learning. The amount of papers published within the photogrammetric and remote sensing community shows that many do have a deep belief Deep Learning is the ultimate solution. Particularly, Deep Learning approaches based on convolution neural networks (CNN) are extensively investigated. A CNN is not a magic tool box, but a software programme, made by humans, at the basis consisting of a concatenation of 2D differential filters, such as the Laplace operator, and 2D integrating filters to construct a hierarchy of image pyramids by aggregating small neighbourhoods, e.g. windows of 3x3 or 5x5 pixels, called pooling. Usually the minimum or maximum value within the window is taken in the creation of the hierarchy of images, which makes the approach sensitive to the presence of noise and

* Corresponding author

texture. High success rates could be achieved for applications such as identifying whether the pet in an image is a cat or a dog. CNN research also focusses on making the dream of self-driving cars a reality.

There is nothing new about the use of neural networks for mapping purposes and the same is true for the knowledge-based approach. Already in the eighties and nineties of the previous century both approaches enjoyed intensive research efforts. A criticism on neural networks is that it is a black box approach which impedes a clear view on how the classifier obtained the solution; the rules learned by the neural network from feeding the system with training data are concealed in the weights of the hidden neurons of the hidden layers (Jensen et al., 2001).

In the wake of AI and neural networks going through a revival in recent years, this paper aims at advocating studying the knowledge-based approach for creating 3D maps from MLS point clouds. We start with analysing the complexity of outdoor scenes to arrive at insight in the complexity of the automation of image interpretation and to derive four generic rules to transfer images and point clouds into 3D maps. Next the elements of visual image interpretation are in focus, followed by an analysis how to explore human craftsmanship in a knowledge-based approach. This paper ends-up with an experiment, which demonstrates the feasibility of the knowledge-based approach.

2. CLASSIFICATION OF OUTDOOR SCENES

2.1 Basics of Classification

Classification aims at assigning the most likely class to individual objects (Duda and Hart, 1973). In a general classification problem the set of classes is known in advance and depends on the use case, i.e. the task at hand. To assign a label to each object information has to be available in the form of measurements. From the measurements features, which give an indication about the possible class label, can be derived. A good decision rule should optimally explore all available information. The state of information in a labelling problem may thus consist of:

- A number of objects waiting to receive a label representing a class
- A set of classes which we want to assign to individual objects, which set is defined in advance and depends on the use case or task at hand
- A number of measurements joined in a measurement vector. From these measurements features can be computed which provide the information on which the classification is based. Also the measurements themselves may be used as features. The size of the measurement vector and the feature vector are not necessarily the same
- An appropriate classifier.

Till so-far the concepts. How does it work out on complex scenes?

2.2 Complex Scenes

Many people react astonished when they hear that the automatic classification of images and point clouds of urban scenes is an elusive problem, yet unsolved. Google can do that, will be their convinced answer. Why should something a human being can do so easily not be done by an advanced computer? Researchers

often state that the results are below expectation because of the complexity of the scene. What is meant by complexity? One of the biggest issues causing the complexity of scenes is the presence of real objects which are not relevant for the use case. Examples of such semantic noise are bicycles, parked cars, flags, balconies or shrubs when the use case consists of 3D mapping of an urban area in which buildings, streets, lampposts, and traffic signs have to be captured.

Occlusion is another issue. From the viewpoint of the sensor a façade may be partly hidden by a truck, a donkey or other dynamic or static object. The human visual system is so well developed that it can fill in the missing parts. Objects may appear in different sizes, shapes and orientations with respect to the sensor. This issue arises when feeding a machine learning system with examples of traffic signs, which have various shapes, including triangle, circle, diamond and so on. So, the question arises: to be able to map the generic class of traffic signs, should subclasses be defined and the system trained by prototypes of these subclasses?

The photometric characteristics of passive sensors depend on sun light conditions, which may considerably vary during a survey of several hours. Shadow may cut one object into pieces with different photometric properties. Machine learning systems explore the same type of features for all pixels, cells or points, without distinction. The range of feature values per class are determined in a training stage and next classes are assigned to unknown objects in bulk – each pixel, cell or point undergoes the same treatment. No refinement takes place by identifying differences in the local structure or arrangement of objects. The local situation is embedded in the features themselves, e.g. in the form of eigenvalues or normal vectors.

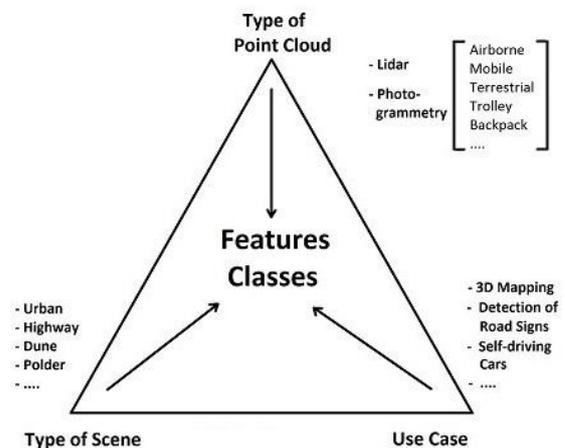


Figure 1. Dependence of features and classes on data source, type of scene and use case

2.3 Rules

The above considerations result in the formulation of four generic rules.

Generic Rule 1: In 3D mapping one is not interested in the semantics of individual pixels or points but in what is present on-top of the surface of the Earth, or bare ground, at specified locations.

Generic Rule 2: A point cloud is a blind sampling of the scene, meaning that we have to deal with arbitrary points being part of

the surfaces of objects; the determination of object types requires an interpretation stage.

Generic Rule 3: Everything is connected to something else and ultimately to the surface of the Earth.

Generic Rule 4: The types of object (classes) which can be recognized in an image or point cloud depend on: (1) type of data source; (2) type of scene and (3) use case.

Figure 1 illustrates Generic Rule 4. The data source is key for defining classes, e.g. a satellite image is not suited for mapping traffic signs. An aerial image provides a different view on a road scene than an MLS point cloud and it is obvious that one cannot extract classes from a geo-data set which are not implicitly or explicitly present in the data set. Also the type of scene restricts class definition. It may be a good idea to map banana trees in the city of Abuja, the capital of Nigeria, but not in the capital of Iceland. Also the task at hand or use case determines the classes to be mapped. A municipality may want to map other object types than a maintenance service of highways.

To get a further grip on the complexity of classification of outdoor scenes it is appropriate to briefly look at the sound concepts of visual image interpretation developed within the century old craftsmanship of manual interpretation of aerial photographs. In other words: how does a human operator accomplish the task?

3. ELEMENTS OF VISUAL INTERPRETATION

Mapping from images and point clouds is a profession which requires expert knowledge. Every human can recognize objects and give these objects a name and even outline them. But when given ten people the same aerial image one will end-up with ten completely different maps. Also people need training to map the objects which are relevant for a use case. It is well-known that a professional mapping operator uses a number of visual cues, in particular: tone, colour, texture, shape, size, height, shadow, site and association (Figure 2). See Estes et al. (1983) for an excellent discussion of these so-called elements of image interpretation and the exploitation of each.

The cues may be subdivided into geometric and photometric (or radiometric) elements. The statistical pattern recognition techniques used in remote sensing primarily explore the photometric features: tone, colour (multispectral bands) and texture. Texture can be quantified using the texture measures of Haralick et al. (1973). Tone and shape are the primary cues in image analysis from which the other basic elements can be derived using the prescriptions: extension, combination and repetition (Figure 3). This figure demonstrates why it is possible to develop a machine vision system to detect and outline objects in well-conditioned scenes. The background (e.g. surface of a conveyor belt) contrasts maximal with the objects. After histogram thresholding, edge detection or region growing objects can be fully outlined, followed by the calculation of shape measures and size. The light conditions are optimized so that no shadow is present. Other objects are absent and when present they do not occlude or clutter with the object under analysis. The complex cues of site and association have no meaning here.

Everyone who has made an attempt to apply edge detectors or region growing techniques on aerial images, knows that the contours seldom correspond to boundaries of relevant objects.

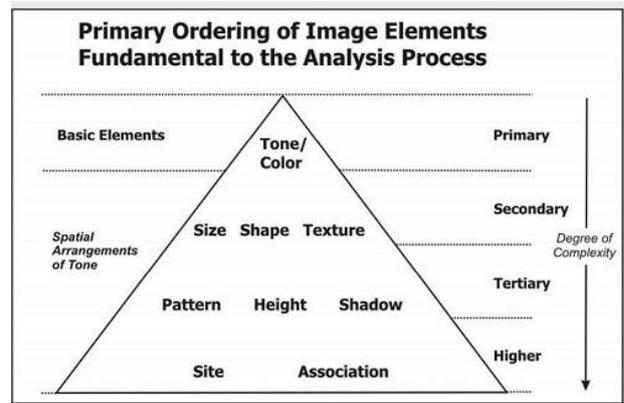


Figure 2. Basic elements of image interpretation arranged by increasing complexity (source: Estes et al. (1983))

So, cues such as shape and size are difficult to exploit in images of outdoor scenes, resulting in the exploration of tone, colour and texture only in remote sensing classification tasks.

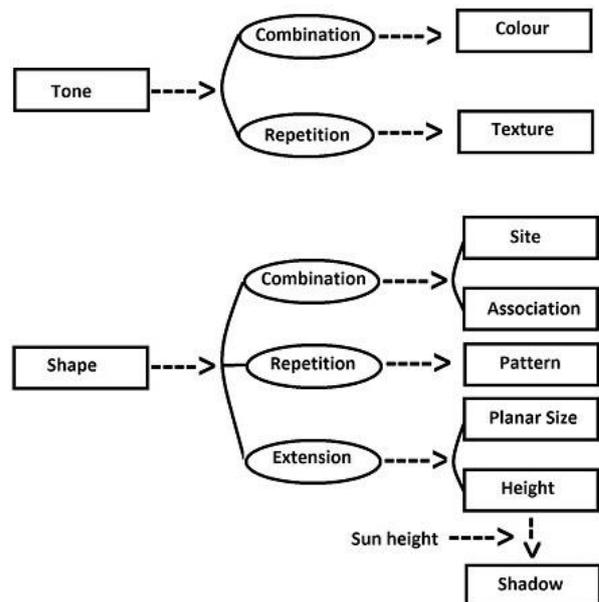


Figure 3. From tone and shape the other basic elements can be derived using prescriptions (source: Lemmens (1987))

Based on the above considerations an approach can be developed in which the knowledge of the human expert is implemented in a knowledge-based system. This idea has been intensively investigated in the early nineties (see de Gunst, 1996 and the references sited in there) but because of lack of success, research has been faded out. But as with neural networks, the time has come to reconsider the feasibility of a knowledge-based approach. At least three developments underpin this proposal:

- The availability of Lidar sensors provide the height cue, which is an important basic element of visual interpretation, in an easy way; Lidar sensors become progressively cheaper, smaller, lighter and produce high point densities,
- Dense image matching which also provides the height cue in the form of point clouds derived from overlapping images
- Powerful computer combined with huge storage facilities.

Instead of a classifier trained by getting fed by examples, a knowledge-based system consists of an inference machine which iteratively exploits knowledge stored in a knowledge base often supported by additional geo-data (Figure 4).

4. KNOWLEDGE EXPLORATION

4.1 Objects

Classification requires the definition of objects in advance. In multispectral classification the individual pixels are considered to be the objects. This approach is prone to error and as a result extensive post-processing is required to ‘clean’ the data set in an attempt to improve the classification result. To tackle the high sensitivity to errors of the *pixel-is-the-object* approach, methods have been developed in which adjacent pixels are aggregated based on similarities of distinguishable features, such as tone or texture, using region growing techniques. This so-called object based image analysis (OBIA) is now implemented in commercial image analysis software, e.g. eCognition.

Similarly, in point-wise classification approaches the point is considered to be the object. Point-wise classification exploits the intensity of the return and geometric properties derived from the point itself and surrounding points. For point clouds also OBIA types of approach have been developed in which planes, spheres, cylinders or other geometric primitives are fitted through an ensemble of neighbouring points (see Grilli et al. (2017) for a review). The descriptive parameters of these segments are used as features for further grouping, classification and mapping.

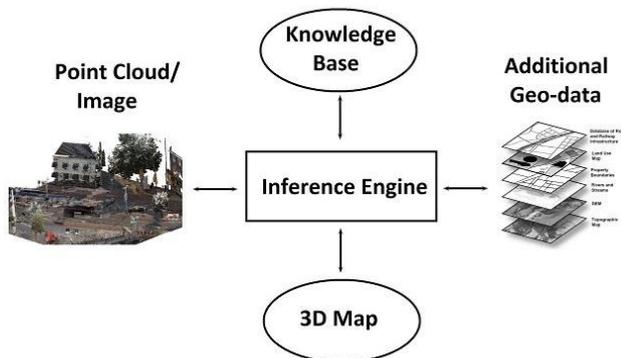


Figure 4. Schematic overview of a knowledge-based system for 3D mapping from point clouds

One of the characteristics of objects, which extend in the vertical direction, is that they are characterized by differing heights. For example, a building façade varies in height which may start at seven metres, or higher, depending on the urban area, while the height of a traffic sign mounted on a pole from ground level upwards usually does not exceed three metres. Many points reflected on traffic signs, façades, lamp posts, cars, pedestrians and trees all may have the same height. So, height above ground level weakly discriminates among the different classes and thus is not well-suited for point-wise classification.



Figure 5. MLS point cloud of an urban scene enriched with RGB (source: Cyclomedia, the Netherlands)

Generic Rule 1 states that 3D mapping aims at finding objects present at a certain location at the surface of the Earth. A location on the Earth’s surface can be defined in many ways, e.g. as a point, a circle, or as a square. The extension of a point on the bare ground in vertical direction is a line, which will only intersect with a few points of the point cloud. So, a bare ground point would not be a proper option for being introduced as the object. A circle extended in vertical directions will become a cylinder. Circles will lead to gaps, i.e. uncovered parts of the surface, and/or overlaps. Therefore, square tiles seem to be the best choice to act as objects in a knowledge-based approach. Of course, the boundary of a tile can cut through a lamppost, traffic sign or other object. For example, when the height difference within a tile indicates a lamppost, but the fingerprint (see Subsection 4.3) violates this assumption, it may happen that the lamppost points are distributed over two or more adjacent tiles. The knowledge-based approach allows a piecewise, iterative refinement and to define new tile which covers the entire object. Next, the fingerprints of the column of this tile can be re-examined.

4.2 Features

Figure 5 shows a part of an MLS point cloud enriched with RGB data simultaneously acquired with a fish-eye camera using Cyclomedia’s Mobile Mapping technology. In general, MLS point clouds are attribute poor. In addition to the 3D coordinates in a local, national or regional reference system, usually only the reflectance value of each point – often represented as a digital number in the range from 0 to 255 – is available in a point cloud (Lemmens, 2017). As a result, many classification approaches rely on enriching the attribute set with RGB values from imagery, which may not always be available, and on examining the local geometric structure of a set of neighbouring points. The suitability of the local geometric structure is based on the observation that many objects differ in shape; e.g. buildings do have a planar shape, pole-like objects a cylindrical shape, while foliage is characterized by normal vectors which point in arbitrary directions. Geometric features have been extensively explored by Weinmann et al. (2015). Becker et al. (2017) have shown that point clouds enriched with RGB results in a significant increase in classification accuracy.

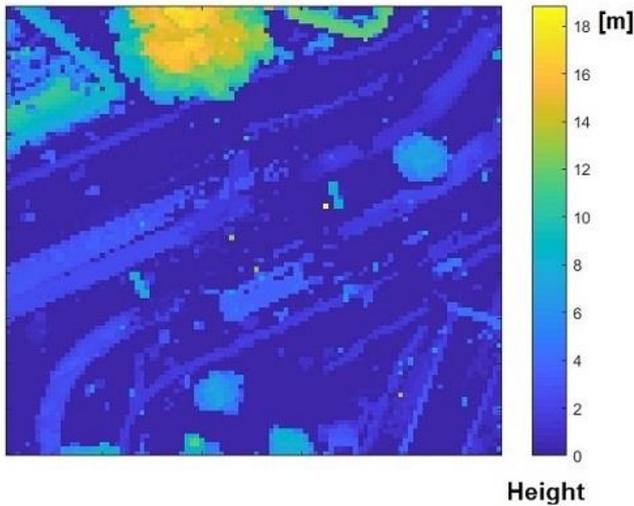


Figure 6. Maximum height above ground level in 50x50cm tiles

The above features are assigned to individual points resulting in a point-wise classification, i.e. the individual point is considered as the object. When introducing tiles as objects, other features may be derived, including number of points present in the column (N) and difference between maximum and minimum height values (Δh). Figure 6 shows Δh for tiles of 50x50cm of the point cloud shown in Figure 5.

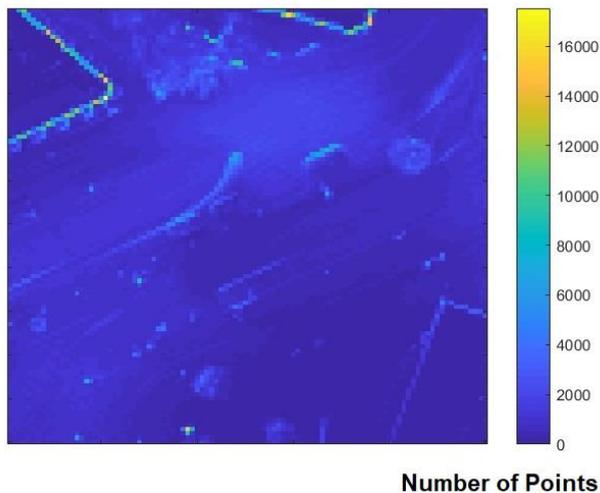


Figure 7. Number of points present within the 50x50cm tiles

One would expect a large correlation between Δh and N – the higher the object in the column above a tile the more laser pulses will be reflected. Figure 7, which represents the number of points per 50x50cm tile, shows that this assumption is only partly true. Since vegetation is semi-permeable for laser pulses, the trees show high Δh but only a modest number of points. Furthermore, the façades of buildings are clearly visible as lines in Figure 7 and it would probably be possible to trace them by a line-following technique. Many points passed the windows of buildings and next reflected on indoor walls, furniture and so on (see upper-left corner of Figures 6, 7). The height values in Figure 6 are rather large, however the number of points is relatively modest enabling to distinguish building indoors from façades. The windows appear also in the fingerprints of façades (Figure 8).

4.3 Fingerprints

The diverse fingerprints of each tile may be compared with prototype fingerprints stored in the knowledge base and based on the matching result a class or several classes, in case more than one object is present within the column, may be assigned to the column or part of the column. In case no unique matches can be found the points in the columns may be scrutinized more in depth by computing eigenvalues, normal vectors or other geometric features of a subset of the points within the column. Also other feature types may be computed.

The differences found in these additional features may give the clue for proper class assignment. This approach includes that the exploitable set of features is not assigned at forehand to the objects but may be extended depending on the progress of the classification process and the local structure. In other words, the selection of features depends on a decision tree with at the top a generic set of features (FS-1). If FS-1 is sufficient to obtain a proper classification result, the task for that tile is finished. If not, a second set of features (FS-2) can be brought into position and so on.

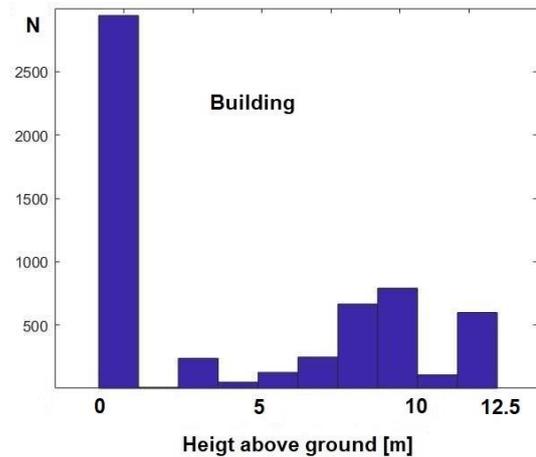


Figure 8. Histogram of the number of points present within a façade tile of 50x50cm (see black arrow in Figure 5)

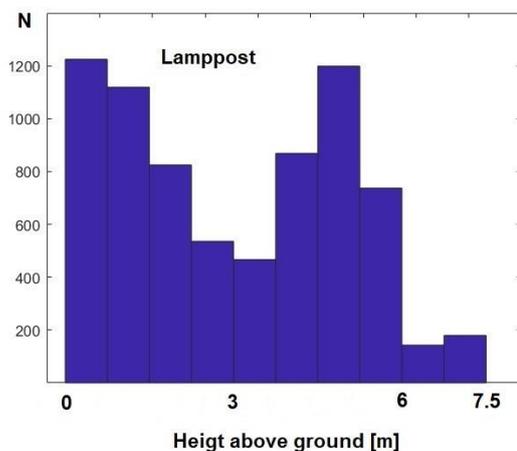


Figure 9. Histogram of the number of points present within a lamppost tile of 50x50cm (see red arrow in Figure 5).

4.4 Scene Knowledge

Different object types may have similar feature values. So, when only using the limited amount of features of point clouds the classification may be prone to error. To avoid that the classification is only depending on the use of features, a priori scene knowledge can be brought in.

Roads and their vicinity are man-made. So, the placing of objects, their shape, size and orientation have to obey official regulations. The specifications may differ in various countries, but will be usually consistent in the same jurisdiction. As a result lampposts, traffic signs, and other road objects appear in zones parallel to the main road direction, while the distance to the road edge stays within certain limits. Added to this, the orientation of pole-like traffic signs is usually perpendicular to the road direction. Rule 3 can be specified as: all objects on and near roads are situated either at surface level or expand in the vertical direction, while their heights are often defined by regulations. In a first stage ground points can be separated from non-ground points, for which algorithms are available although editing is usually required (Meng et al., 2010).

Added to regularities in the zonal distribution, there may also be regularities available about the along-track arrangement of objects at road sides. For example, lampposts are usually placed at regular distances of each other. This knowledge may be explored to remove erroneous lamppost assignments to a tree or a flagpole, or guide the searching of missing lampposts. Yang et al. (2017) demonstrated that exploration of scene knowledge improves classification accuracy.

5. INITIAL EXPERIMENT

To demonstrate the feasibility of a knowledge-based approach an experiment has been conducted using the benchmark dataset of the Robotics laboratory (CAOR) at MINES ParisTech, France. The point cloud covers a 160m-traject in Paris acquired February 8, 2013 and contains 20 million points of which the X,Y,Z coordinates, the reflectance value per point and the object class are given (Serna et al., 2014).

In this experiment we use a point-wise classification scheme exploring two features: height and reflectance value. Façades, cars, pedestrians, motorcycles and traffic signs are the classes selected. When combining knowledge about the scene with a thorough inspection of height values (H) and reflection values (R) a rule pops-up which states that the heights of objects in classes other than façade are less than 2.25m (Figure 10).

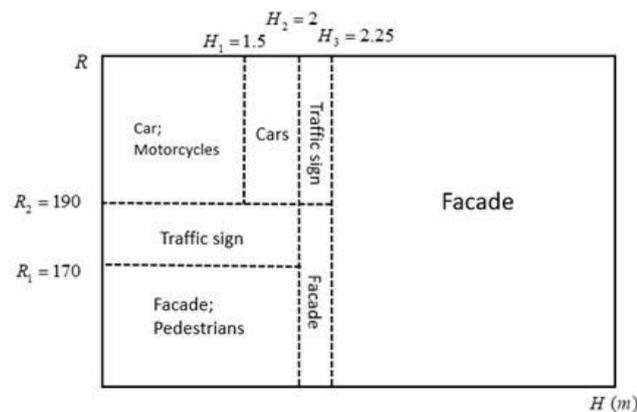


Figure 10. Knowledge on Height (H) and Reflectance (R) shown in a 2D feature space (Adopted from Zheng et al., 2018)

Points with heights smaller than 2.25m may be reflected on other object classes but on façades as well. The 2.00 – 2.25m height interval may contain points reflected on façades and traffic signs points. Scrutinizing the reflectance values reveals a threshold of $R = 190$: points with $R > 190$ are likely traffic sign and when R is below this value the point will likely be part of a façade. Points within the 1.70 – 2.00m height interval may belong to façade or pedestrian. Since the reflectance value gives no clue about the type of object, the assignment of façade or pedestrian to the point is done randomly. The points with $H < 2.00m$ lying within the 170 – 190 reflectance interval are likely traffic sign. The points with $1.5m < H < 2.0m$ and $R > 190$ can be identified as cars. The points with $H < 1.5m$ and $R > 190$ belong either to cars or to motorcycles. The assignment either the class car or motorcycle is done in a random way.

This knowledge-based approach has been used as a baseline reference in Zheng et al. (2018). With this simple knowledge-based approach, implemented as what is known in remote sensing as box classification, an overall accuracy could be achieved of 79%.

6. CONCLUSIONS

There is nothing new about the use of neural networks for the classification of point clouds and the same is true for the knowledge-based approach. Neural networks in particular in the form of Convolutional Neural Networks have enjoyed a remarkable revival in recent years. This paper has argued and demonstrated that a knowledge-based approach is likewise feasible and worthwhile to scrutinize in future research. With a relatively simple knowledge-based rule schema an overall accuracy of 79% could be achieved for the Paris-rue-Madame point cloud bench mark data set. Conventional classification schemes may be looked at as a bulk approach, one size fits all, while a knowledge-based approach focusses on a stepwise, iterative refinement depending on the fingerprints of the diverse features in the vertical direction.

ACKNOWLEDGEMENTS

Thanks are due to Mingxue Zheng for assisting in conducting the initial experiment.

REFERENCES

- Becker, C., Häni, N., Rosinskaya, E., d'Angelo, E., Strecha, C.I., 2017. Classification of Aerial Photogrammetric 3D Point Clouds, Presented at ISPRS 2017 workshop, Hannover
- Duda, R. O., Hart, P. E., 1973. Pattern classification and scene analysis, Wiley, New York. Friedman
- Estes, J. E., Hajic, E. J., Tinney, L. R., 1983. Fundamentals of Image Analysis: Analysis of Visible and Thermal Infrared Data, Chapter 24 Manual of Remote Sensing. Second Edition (R. N. Colwell, ed.), American Society for Photogrammetry & Remote Sensing, Bethesda, pp. 987-1124.
- de Gunst, M.E., Han, C.S.L.A., Lemmens, M.J.P.M., van Munster, R.J., 1991. Automatic extraction of roads from SPOT images. *Proceedings of ISPRS and OEEPE joint workshop on updating digital data by photogrammetric methods*, OEEPE Official Publication, 27, pp. 131-140.

- de Gunst, M., Lemmens, M.J.P.M., 1992. Automatized updating of road databases from scanned aerial photographs. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 29, Part B4, pp. 477-484
- de Gunst, M. 1996. Knowledge-based interpretation of aerial images for updating of road maps, PhD thesis, Delft University of Technology, The Netherlands, Netherlands Geodetic Commission, New Series, Vol. 44, 184 pp.
- Grilli, E., Menna, F., Remondino, F., 2017. A review of point clouds segmentation and classification algorithms. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W3, pp. 339-344
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-3, pp. 610-621
- Jensen, J.R., Qiu, F., Patterson, K., 2001. A Neural Network Image Interpretation System to Extract Rural and Urban Land Use and Land Cover Information from Remote Sensor Data. *Geocarto International*, Vol. 16, No. 1, pp. 19-28.
- Lemmens, M.J.P.M., 1987, Remote Sensing Techniques, Earth Observation from Air and Space, Delft University of Technology, Faculty of Geodesy, 270 pp. (in Dutch).
- Lemmens, M., 2017. Mobile Laser Scanning Point Clouds – Status and Prospects of Automatic 3D Mapping of Road Objects, *GIM International*, 31(8), pp. 18-21.
- Lemmens, M.J.P.M., Bruel, E.W., Fennis, F. 1988. Linear feature extraction and road recognition from large scale digital aerial images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 27 (B8) pp. 20-31, Comm. VII
- Meng, X., Currit, N., Zhao, K., 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues, *Remote Sensing*, no. 2, pp. 833-860
- Serna, A., Marcotegui, B., Goulette, F., Deschaud, J.-E., 2014. Paris-rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods, *4th International Conference on Pattern Recognition, Applications and Methods ICPRAM 2014*
- Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers External Link. *ISPRS journal of photogrammetry and remote sensing*, 105, pp. 286–304. doi:10.1016/j.isprsjprs.2015.01.016
- Vosselman, G., de Gunst, M., 1997. Updating Road Maps by Contextual Reasoning. Automatic, In: *Extraction of Man-Made Objects from Aerial and Space Images (II)*, A. Gruen, E.P. Baltsavias, O. Henricsson (Editors), Birkhäuser Verlag, Basel, pp. 267-276
- Yang, B., Dong, Z., Liu, Y., Liang, F., Wang, Y., 2017. Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 126, pp. 180–194
- Zhang, C., Baltsavias, E., 2000. Knowledge-based Image Analysis for 3D Edge Extraction and Road Reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 33, PartB3/2, pp. 1008-1015
- Zheng, M., Lemmens, M., van Oosterom, P., 2018. Classification of Mobile Laser Scanning Point Clouds of Urban Scenes Exploiting Cylindrical Neighbourhoods, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, pp. 1225-1228, <https://doi.org/10.5194/isprs-archives-XLII-2-1225-2018>