1 Introduction

Dense point clouds may contain billions of points and hence fast classification methods are required to extract meaningful information within a reasonable amount of time. Point cloud classification can be decomposed into (1) feature extraction, (2) point representation, and (3) class assignment. Feature extraction aims at finding suitable features for each point derived from the original data. Here we use the altitude above street level. In this study point representation aims at transforming the features derived in step one into a vector expression, one vector per point. This paper focuses on the latter step, i.e. step two. Much research on improving image classification performance has been conducted in the field of multispectral image classification. The promising results encourage us to apply the classification methods of multispectral images on point clouds. In particular, we consider histogram encoding (VQ) (Sivic et al., 2003) and Kernel codebook encoding (KC) (Philbin et al., 2008) because they have proven to be not computationally demanding.

The classification flowchart is shown in Figure 1.

2 Basics

Before treating both methods, we briefly consider K-means.

Given a set \( x_1, \ldots, x_n \in \mathbb{R}^p \), K-means seeks \( k \) vectors \( \mu_1, \ldots, \mu_k \in \mathbb{R}^p \) and a data-to-means assignments \( q_1, \ldots, q_n \in \{1, \ldots, k\} \) such that the cumulative approximation error \( \sum_{i=1}^{n} \| x_i - \mu_{q_i} \|^2 \) is minimized, through alternating between seeking the best means given the assignments \( \mu_k = \text{avg}(x_i; q_i = k) \), and seeking the best assignments given the means \( q_{ki} = \arg\min_k \| x_i - u_k \|^2 \).

The classification accuracy is below 50% but the computational performances encourage further investigation using more features.

Keywords: Classification, point clouds, feature encoding.
encoding starts by learning k-means average vectors $\mu_1, \ldots, \mu_k$. Given a set $x_1, \ldots, x_n$, let $q_i$ be the assignments of each point sample $x_i$ as given by (1). The histogram encoding is the non-negative vector $f_{\text{hist}} \in R^k$ such that $[f_{\text{hist}}]_k = [i; q_i = k]$.

2.2 Kernel codebook encoding

Kernel codebook encoding is a variant in which descriptors are assigned to $[f_{\text{kcb}}(x_i)]_k = K(x_i, \mu_k)/\sum_{j=1}^{n} K(x_j, \mu_k)$. Especially $K(x, \mu) = \exp(-\frac{1}{2} \|x - \mu\|^2)$ is a common kernel function. $\gamma$ is a smoothing parameter. A set of $n$ descriptors is extracted from an image as $f_{\text{kcb}} = \frac{1}{n} \sum_{i=1}^{n} f_{\text{kcb}}(x_i)$.

3 Experiments

3.1 Data

We use a benchmark dataset created by Serna et al. (2011), which consists of two PLY files with 10 million points each. To each point (X, Y, Z) coordinates, reflectance value, label and class have been assigned. There are 26 classes. We use one PLY file, i.e. 10 million points, and choose five classes: pedestrians, motorcycles, traffic signs, trash cans and fast pedestrians. The selection is based on the similar amount of points (around 10,000) reflected on each object surface. Figure 2 shows an orthophoto of the test site. In our tentative test we assume that street level is everywhere the same, i.e. the points at street level do have the same height. So, we use the original height values (Z) as feature.

Figure 2: Rue Madame, Paris (France). Orthophoto from IGN-Google Maps.

3.2 Results

To conduct our experiments we use the public Library for SVMs (LIBSVM) package (Chang et al., 2011). After applying both the VQ and KC classifiers the overall accuracy, kappa coefficient, and computation times are computed (Table 1).

The accuracies of both classes are lower than 60%, but VQ has a better accuracy than KC. In addition, the more accurate a method is (i.e. VQ) the more computation time is required. This demonstrates that accuracy comes at the cost of increasing computational efforts. The Recall and Precision of two methods are shown in Figure 3. The class Trash cans acquires the best Recall and Precision values with VQ and KC. The average values of Recall and Precision are nearly 50% with VQ and nearly 40% with KC.

Figure 3: the Recall and Precision of KC (left) and VQ (right).

4 Conclusions

We tested two classifiers, developed for use on multispectral images, on their feasibility for classifying massive amounts of points rapidly. Using one feature (height above street level) VQ and KC demonstrated to be fast classifiers but the classification accuracy is low. Nevertheless the results of our tentative tests are promising, especially with respect to computation time. So, we continue to carry out refinements by using more features, including reflectance values. In addition, we will store data in a Database Management System (DBMS) to manage the massive amount of points efficiently and to incorporate classification functionality into the DBMS to reduce further computation time.

References


<table>
<thead>
<tr>
<th>Method</th>
<th>VQ</th>
<th>KC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy</td>
<td>52%</td>
<td>42%</td>
</tr>
<tr>
<td>Kappa coefficient</td>
<td>43%</td>
<td>35%</td>
</tr>
<tr>
<td>Computation time (sec.)</td>
<td>2.9</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Table 1: Performance