Data-Driven Agricultural Machinery Activity Anomaly Detection and Classification

Yang Wang[1], Andrew Balmos[1], James Krogmeier[1], Dennis Buckmaster[2]

[1] Electrical and Computer Engineering, Purdue University, West Lafayette, IN
[2] Agricultural and Biological Engineering, Purdue University, West Lafayette, IN

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract. *In modern agriculture, machinery has become the one of the necessities in providing safe, effective and economical farming operations and logistics. In a typical farming operation, different machines perform different tasks, and sometimes are used together for collaborative work. In such cases, different machines are associated with representative activity patterns, for example, in a harvest scenario, combines move through a field following regular swaths while grain carts follow irregular paths as they ferry grain from combines to trucks. Sometimes unusual conditions due to field conditions, machine status, weather, or human factors may cause anomalous activity patterns. Detecting and classifying anomalous patterns can be used for planning and efficiency improvements.*

Zhang performed rudimentary work (Zhang et al. 2017) utilizing GPS paths for classifying different machines' activity patterns using a rule-based algorithm to understand what has happened in the field for farm logistics improvements. In this paper, we use machine data in addition to GPS tracks for anomaly detection and classification. The algorithm uses Kalman filtering and the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al. 1996) to cluster a point cloud that consists of a combine's engine load, speed data, and the Kalman filter residual.

Keywords. *ISOBUS, GPS, Machine Data, Anomaly Detection, Classification, Kalman Filter, DBSCAN.*

Introduction

The motivation for this paper summary is to explore the feasibility of an algorithm for classifying activities for combine operations using data filtering and clustering technique with the aid of GPS, vehicle speed and engine load data. In normal combine operation (either in harvesting or moving between fields), the combine is usually traveling in near constant speed in a straight line. On the other hand, changes in combine speed typically indicate in-field or on-road maneuvers like turning and accelerating. Hence, in the scope of this paper, each spatial point is represented by a three-vector consisting of engine load, speed, and the residual from the Kalman filter applied to the GPS coordinates. These points are classified by a clustering algorithm discussed below. The classes represent different patterns of the combine. Classes with very few points relative to the other classes are considered as anomalous.

The GPS data were collected in a CASE IH 6130 combine using an ISOBlue 2.0 (Wang et al., 2017) during two wheat harvest sessions. Each GPS data sample contains epoch timestamp, vehicle speed, latitude and longitude. The latitude and longitude are also converted to easting and northing for computational convenience.

The engine load data were also acquired using an ISOBlue 2.0 in the same harvesting sessions. The samples appear more frequently than the GPS data. A mapping between each GPS data sample and an engine load value is created by computing the average engine load values between two consecutive GPS timestamps and assigning the computed value to the smaller GPS timestamp.

Recent works (Celik et al., 2011; Ranjith et al., 2015; Tran and Kim, 2011) have shown that by choosing the appropriate input parameters, the DBSCAN algorithm can be used to cluster point clouds for anomaly detection, pattern recognition, and performance analysis on agricultural big data sets (Mehta et al., 2015). In our paper, a novel algorithm is presented. A Kalman filter is first employed to filter the GPS data and compute the residual. The computed residual is combined with the vehicle speed and the engine load data to form a point cloud. Finally, DBSCAN is used to cluster the point cloud. The detailed description of the algorithm as well as the results are described in the following sections.

Algorithm

The algorithm can be visualized as in Figure 1 .

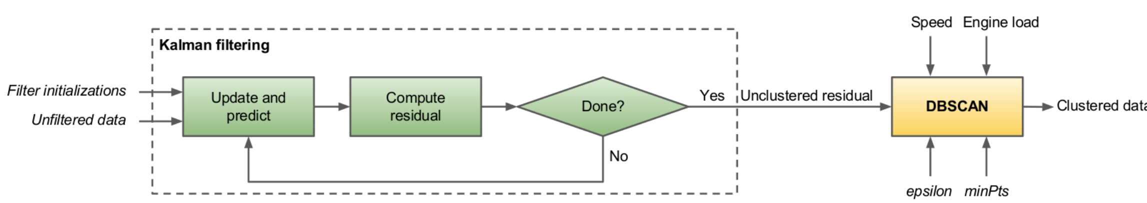

Figure 1: Proposed algorithm consists of a Kalman filter as well as the DBSCAN algorithm.

The algorithm uses a Kalman filter that is based on a constant velocity (CV) dynamical model for filtering the easting and northing coordinates recursively for the entire length of the unfiltered data. Noise covariances, initial error covariances and initial state estimates are initialized before the start of the filter. Within each Kalman filter iteration, the filtered easting and northing values are used to compute the Kalman filter residual. The residual is a value that indicates how much the filter estimates deviate from the actual measurements. In this context of this algorithm, given that the Kalman filter uses a CV model, the residual (in meters) serves as an indicator to show the smoothness of the combine motion. In other words, if the combine undergoes any sudden change in motion, it would result in a higher residual value.

The second part of the algorithm utilizes the DBSCAN algorithm. It requires two input parameters: ε and *minPts*. ε is the minimum Euclidean distance between two points for expanding clusters and *minPts* is the minimum number of points in a cluster. If ε is too small, the DBSCAN will result redundant clusters that carry similar traits. On the other hand, an exaggerated ε value means that the algorithm would falsely merge points together and output very few clusters. It is worthwhile to note that the DBSCAN is an unsupervised algorithm; it does not require an initial guess on the number of clusters and the total number of clusters are not known. Once the clustered data are obtained, one would typically assign classification labels to different clusters based on practical knowledge in the area.

Within the scope of this paper, appropriate ε and *minPts* values were chosen as 2.2 and 120 to group the three-dimensional point cloud (engine load, vehicle speed, and computed Kalman filter residual) so that the number of clusters were reasonable and optimal.

Results

The speed, engine load and residual data were clustered into a total of four clusters. By visualizing the GPS coordinates according to the clusters, one could give labels to the four clusters:

- Cluster #1: uniform motion in field.
- Cluster #2: uniform motion on road.
- Cluster #3: stationary.
- Cluster #4: nonuniform motions (in-field turns, on-road turns, decelerating, accelerating).

Conclusion

The proposed algorithm successfully clustered the speed, engine load and residual point cloud into clusters that reflect the different operational modes of the 6130 combine. However, there are some false positives as well as some anomalies that have not been detected or falsely merged onto another cluster by the algorithm. Future work includes developing a systematic way to estimate the input parameters to both the Kalman filter and the DBSCAN algorithm as well as employing machine learning techniques to automate the clustering and labeling process.

References

Celik, Mete, Filiz Dadaser-Celik, and Ahmet Sakir Dokuz. 2011. “Anomaly Detection in Temperature Data Using DBSCAN Algorithm.” In , 91–95. IEEE. https://doi.org/10.1109/INISTA.2011.5946052.

Ester, Martin, Hans-Peter Kriegel, Jiirg Sander, and Xiaowei Xu. 1996. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” 6.

Mehta, Parth, Hetasha Shah, Vineet Kori, Vivek Vikani, Soumya Shukla, and Mihir Shenoy. 2015. “Survey of Unsupervised Machine Learning Algorithms on Precision Agricultural Data.” In , 1–8. IEEE. https://doi.org/10.1109/ICIIECS.2015.7193070.

Ranjith, R., J. Joshan Athanesious, and V. Vaidehi. 2015. “Anomaly Detection Using DBSCAN Clustering Technique for Traffic Video Surveillance.” In , 1–6. IEEE. https://doi.org/10.1109/ICoAC.2015.7562795.

Tran Manh Thang, and Juntae Kim. 2011. “The Anomaly Detection by Using DBSCAN Clustering with Multiple Parameters.” In , 1–5. IEEE. https://doi.org/10.1109/ICISA.2011.5772437.

Wang, Yang, Andrew D Balmos, Alexander W Layton, Samuel Noel, Aaron Ault, James V Krogmeier, and Dennis R Buckmaster. 2017. “An Open-Source Infrastructure for Real-Time Automatic Agricultural Machine Data Processing.” In . American Society of Agricultural and Biological Engineers. https://doi.org/10.13031/aim.201701022.

Zhang, Yaguang, Aaron Ault, James V. Krogmeier, and Dennis Buckmaster. 2017. “Activity Recognition for Harvesting via GPS Tracks.” In . American Society of Agricultural and Biological Engineers. https://doi.org/10.13031/aim.201700813.