Image-Based "Campus Positioning System" With Data Mining Techniques

Jichuan Jiang **Pengfei Weng** **Hebo Yang** **Yue Zhao** **Frank Wood**
Columbia University, New York, NY 10027, USA
{fwood}@stat.,{jj2560, pw2336, hy2326, yz2416}@columbia.edu

Abstract

One of the major data mining applications is mining image data. We want to implement a data mining model to specifically identifying photos taken on Columbia main campus and thus determining the location it is taken as well as its view direction. This project will likely involve a hidden Markov model and achieve high accuracy and quick reaction.

1 Introduction

Every semester, there are freshmen admitted to Columbia and they have to walk around the campus searching for the building they wanted to go to, sometimes have no sense of where they are at all. Moreover, there are numerous visitors coming to Columbia every day and it is quite easy for them to get confused on their current location on campus. Motivated by the idea of identifying buildings on Columbias main campus, we wanted to move one step further to build an application which could help the user identify his current position on campus and view direction by implementing data mining techniques.

2 Problem Description

By taking a short video on users front view with his phone and sending it to the server, the user could use our app to get his current position on campus map as well the direction he is facing.

A set of pictures will be obtained by taking a few frames from the video and then compared with the pictures in the servers database. We want to implement a data mining model to accurately and quickly determine a series of matches between the set and another set of pictures in the database so that further predictions can be made by maximizing the probability of such matches.

Two major challenges we have to face are: prediction accuracy and system reaction time. Since there are difficulties such as exposure, shadow, distance, obstruction, blur, weather and day/ night, high-accuracy is not easy to achieve. System reaction time is another challenge which highly depends on the model complexity, since our goal is to create a cellphone app to give user feed-back in quick response. Therefore, our model and its parameters have to be carefully chosen and determined, in order to handle these difficulties.

Our model will involve conditional probability measures, possibly a Hidden Markov Model [1] and other techniques we will learn on class. The hidden Markov model could be represented as a Bayesian net graph in Fig. 1.

We will divide the walking area of main campus into a large number of grids. Upon determining the matches, we can return which grid the users video is taken from and which view direction the user is facing. Currently we are considering an application with a capacity of one hundred partitions and eight directions. Actually, since there are lots of buildings at campus, the places where users need location services are limited. Several hundred is a reasonable sampling frequency. Our application will represent the users location by a pin point on a campus map shown on cellphone and an arrow indicating his view direction. Our preliminary goal is to build a standalone version of the application for the purpose of this project first. Then, if time permits, we shall develop an android or iOS app on cellphones.

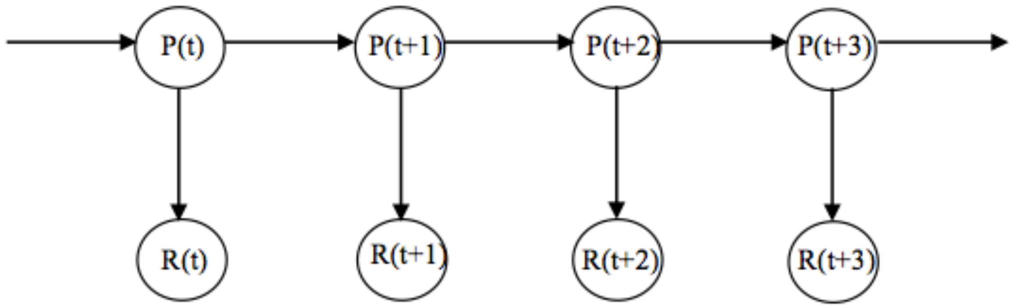

Figure 1: Hidden Markov Model

3 Data Source

We will gather the data by walking on all the routes on campus and taking pictures at center point of each grid. A set of pictures will be taken on each of the eight directions. Therefore, the picture database will consist of one hundred positions, with every position having a set of pictures on each of the eight directions. Then we will extract features from the picture database and build the database using just the features. Features to be used might include *Histogram of Oriented Gradients (HOG)*[2], *Scale-invariant feature transform (SIFT)* [4], *Bag of Words (BOW)*[3] and etc.

4 Expected Conclusion

Upon finishing the project, we want to take a tour around the campus, taking pictures and testing the performance of our application. The expected conclusion is that our application will show relatively high accuracy (like 70%) and short reaction time (like less than 5 seconds).

References

[1] L.E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, 41(1):164–171, 1970. URL `http://www.jstor.org/stable/2239727`.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886 –893 vol. 1, june 2005. doi: 10.1109/CVPR.2005.177.

[3] David D. Lewis. Naive (bayes) at forty: The independence assumption in information retrieval. pages 4–15. Springer Verlag, 1998.

[4] D.G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150 –1157 vol.2, 1999. doi: 10.1109/ICCV.1999.790410.