2016

Real-time Traffic Flow Prediction using Augmented Reality

Minxuan Zhang
University of Windsor

Real-time Traffic Flow Prediction Using Augmented Reality

By

Minxuan Zhang

A Thesis
Submitted to the Faculty of Graduate Studies
through the School of Computer Science
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2016

Real-time Traffic Flow Prediction Using Augmented Reality

by

Minxuan Zhang

APPROVED BY:

Dr. Kevin Li
Odette School of Business

Dr. Imran Ahmad
School of Computer Science

Dr. Xiaobu Yuan
School of Computer Science

March 8, 2016

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Traffic congestion is one of the most difficult problems in the 21st century. Different approaches have been developed to deal with traffic congestion and manage traffic flow. In comparison with prediction based upon historical datasets only, real-time methods take vehicle operators and travelers into consideration, and develop new algorithms/models to improve accuracy for efficient traffic management.

This thesis is going to first highlight current research in traffic flow prediction, and then use chaotic dynamic complexity to discuss the scale-free characteristics of traffic flow. As sharp variation points provide rich information to analyze the fluctuation and sharp variations of traffic flow, it is used in a new method developed in this thesis to guide the classification of historical datasets and to combine real-time datasets from multiple sources of traffic-relevant information. In addition, an augmented reality system is constructed to visualize traffic flow under the influence of different factors.

Keywords: real-time traffic flow prediction, sharp variation points, augmented reality

AKNOWLEDGEMENTS

I would like to present my appreciation to my supervisor Dr. Xiaobu Yuan for his priceless assistance and support during the past two years.

I would also like to express my gratitude to all the committee members, Dr. Imran Ahmad and Dr. Kevin Li. Thank you all for your valuable comments and suggestions to this thesis.

Finally, I want to thank my parents, my girlfriend and all my friends who gave me consistent support during the past two years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

The number of the vehicles increases with the booming increase of the population density around the world. The amount of cars has been increasing in a high speed, and the car has transferred from luxury to necessary for common people. Vehicle is the most usual tools in transportation, and driving cars is now an indispensable part of people's life. Because of this, humans have to face an unavoidable problem, i.e., how to solve the traffic congestion.

Traffic congestion has become one of the most difficult problems in 21st century. Every day, many people suffer from traffic congestions, and lose time everywhere. Similar to air pollution, the problem of traffic congestions cannot be solved completely, unless there are much less cars in the world. In order to enjoy a better life, people have been working hard to find ways to reduce the impact of traffic congestion in the last few decades[18].

Real-time traffic flow prediction has been required by authorities in an increasing number[21]. In the previous decade, the collection of real-time traffic data was a foremost goal. Real-time traffic data is very important, because the travelers need to know what is happening now and what will happen in the near future. Though historic data is useful for prediction, it cannot provide high accuracy in most of the time. The goal of real-time prediction is to achieve higher accuracy by using the latest traffic data.

There are still some big problems to be solved in traffic flow prediction. Researchers are trying to create a model which is effective under all kinds of traffic conditions with higher accuracy. Some other factors also have effects on prediction, like weather changes. Therefore, the prediction should be under the same condition, and the results are expected to be acceptable. As people are finding ways to make prediction visible for travelers, the technology of simulation is also required in this area.

In order to get more accurate results for prediction, many methods and models have been created in the past few decades. In this thesis, a new idea based on sharp variation points will be discussed in the following chapter, and the result will be shown in a simulated virtual environment by augmented reality. Wavelet transform is a good way to analyze the fluctuation from original data. With this technology, researchers can make a deeper understanding that cannot get from the original data. Sharp variation point is defined as zero-crossing point of the wavelet transform, and it indicates a beginning or ending for a varied trend. For real-time prediction, a length of two sharp variation points is reasonable, because it keeps the integrity of variation trend. To make prediction visible, augmented reality is the best tool. This technology can combine multiple sources together, like weather conditions and accident issues. In Unity 3D, by using augmented reality, the simulated virtual world can make prediction more realistic.

In the remaining of this thesis, Chapter II shows the brief preview of the traffic flow prediction for the past 30 years. Chapter III shows the related work in traffic flow prediction in different ways. Chapter IV gives an introduction to the augmented reality. Chapter V discusses the proposed method based on sharp variation points. Chapter VI covers the experiments and analysis. Chapter VII is the system design and its implementation using augmented reality in Unity 3D. Chapter VIII gives the conclusion and future work.

CHAPTER II

Traffic Flow Prediction

After the Industrial Revolution, human civilization has been developing in an increasingly speed. However, people also suffer from some problems caused by the rapid development of civilization. One of them is the heavy traffic congestions.

Traffic congestions occur everywhere in people's daily life. Whenever it happens, it causes huge impact for human beings, including wasting time and money. There are three main reasons for traffic congestions[3]. The first one is the increasing amount of vehicles. The second one is the low capacity of road. The third one is the large numbers of intersections. For these three reasons, none of them can be solved completely. Therefore there is a huge need for traffic management to reduce the traffic congestions.

In the last three decades, many policies and methods were implemented in traffic managements. In some countries, governments force to reduce the flow of family cars and make extensions for main roads and highways. As the results have not been satisfying, researchers are working hard to find better ways to solve the problem.

Like diseases and disasters prevention, scientists and researchers tried to make prediction of traffic flow, hoping to reduce the bad effects of traffic congestion to the lowest level. The research in traffic flow prediction has been developing for more than 30 years. Here are some representative publications for different eras in this area.

In 1984, I Okutani and YJ Stephanedes published the paper about dynamic traffic prediction. It was almost the first generation of methods in this area. Two models based on Kalman filtering theory have been proposed to predict short-term traffic volume[15]. In their method, they used the most recent error as parameters to predict traffic, and better volume for prediction was achieved[15].

In 1994, BL Smith and MJ Demetsky published a paper using neural network approach for short-term prediction. In their paper, they first analyzed previous methods and grouped them into three different categories. They also pointed out the possibility of using neural network models in real-time applications[17].

A paper combining Kohonen maps with ARIMA model for short-term prediction was authored by M Van Der Voort and S Watson in 1996. In the paper, the initial classifier was generated by a Kohonen self-organizing map with an individual ARIMA model[19]. They pointed out the algorithm could be easily retrained for small number of classes[19].

In 1999, BM Williams and LA Hoel established the seasonal time series methods, especially using ARIMA modeling for traffic flow. In addition, the research contributed a specific application using this modeling theory[23].

B Abdulhai, H Porwal and W Recker improved a new method for short-term prediction in 2002. Their new method was based on a Time Delay Neural Network(TDNN) model, and its structure was synthesized by a Genetic Algorithm(GA)[1]. The results tested by both simulation and real data were all acceptable.

In 2005, an optimized genetic approach in neural network was published by EI Vlahogianni and JC Golias. Genetic algorithms seemed to be a new direction in this area. This paper pointed out that some traffic parameters played important roles in modern Intelligent Transportation System(ITS) research and practice[20]. It also suggested that neural networks should be one of the best choices for selecting parameters for modeling and predicting traffic.

In 2009, Support Vector Regression(SVR) came to people's sight in this area. M Castro-Neto and LD Han improved this method on online-SVR under typical and atypical conditions. In the paper, they explained the need of accurate prediction in short-term traffic flow under atypical conditions. The results showed a high accuracy in pattern recognitions for short-term traffic with SVR[8].

In 2013, a new method based on Kalman filtering was published by LL Ojeda and AY Kibangou. In their paper, they claimed the importance of continuous traffic flow prediction in most of ITS researches[14]. The main goal of this study was to make a multi-step ahead prediction for traffic flow that can meet various requests, including high accuracy and low memory capacity. In order to achieve the goal, two new approaches were proposed.

Big data is a new concept in recent years, and Y Lv and Z Li used it in their prediction for a deep learning approach in traffic flow prediction. Over the last few years, data has been exploding in all areas including the traffic. Existing prediction methods mainly use models with some limitations and cannot satisfy many real-world requests. This problem has stimulated people to reconsider the architecture of models with big traffic data[12].

Table 1 shows those publications highlighted in this chapter. These publications represent the main methods in the area of traffic flow prediction. The first direction of methods focused on some basic theory, like Kalman filtering theory. The second direction focused on prediction models combined with genetic theory. The third direction focused on prediction in the new direction of big data.

Title	Authors	Year	Cited by
Dynamic prediction of traffic volume through Kalman filtering theory	I Okutani, YJ Stephanedes	1984	468
Short-term traffic flow prediction: neural network approach	BL Smith, MJ Demetsky	1994	261
Combining Kohonen maps with ARIMA time series models to forecast traffic flow	M Van Der, Voort M, Dougherty, S Watson	1996	257
Modeling and forecasting vehicular traffic flow as a seasonal stochastic time series process	BM Williams, LA Hoel	1999	97
Short-term traffic flow prediction usingneuro-genetic algorithms	B Abdulhai, H Porwal, W Recker	2002	71
Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach	EI Vlahogianni, MG Karlaftis, JC Golias	2005	260
Online-SVR for short-term traffic flow prediction under,typical and atypical traffic conditions	M Castro-Neto, YS Jeong, MK Jeong, LD Han	2009	51
Adaptive Kalman filtering for multi-step ahead traffic flow prediction	LL Ojeda, AY Kibangou	2013	12
Traffic flow prediction with big data: a deep learning approach	Y Lv, Y Duan, W Kang, Z Li	2015	5

TABLE 1: Review of prediction methods

CHAPTER III

Related Work

In this chapter, two different research directions are discussed in details with examples. The first category is the data-driven methods, which means it depends on data collection. This kind of methods requires chaotic data for prediction. The second category is the model-driven methods, which means the key factor for this category is various model for prediction. The third part of this chapter covers the theory of the proposed method using wavelet transform of sharp variation points.

1 Data-driven Methods

Data-driven methods require chaotic data for prediction, because the traffic flow can be described as a dynamic system under the chaotic characteristic[9]. Lyapunov exponent is one of the methods to define the chaotic data. Phase-space reconstruction is an effective way to analyze the non-linear time series for a dynamic system, and it will be discussed with an example in this part. In addition, some algorithms help to obtain the parameters during the process of reconstructions, such as C-C Algorithm[11], Wolf Algorithm[7] and Jocobian Algorithm[5].

The following discussion shows an example of prediction with Phase-space reconstruction. The first step is to define a univariate time series $\{x(t)|t = 1, 2, ..., N\}$. In the second step the phase-space is reconstructed via delay coordinates to get the delay vectors and track matrices[9]:

$$X = [X_1 X_2 ... X_M]^T = \begin{bmatrix} x_1 & x_2 & ... & x_M \\ x_{1+\tau} & x_{2+\tau} & ... & x_{M+\tau} \\ ... & ... & ... & ... \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & ... & x_{M+(m-1)\tau} \end{bmatrix} \tag{1}$$

In the matrix, $M = N-(m-1)\tau$ is the number of phase points; m is the embedding dimension; and τ is the time delay. When the value of m and τ is appropriate, it can restore the original dynamic characteristics of the system in the topologically equivalence via Tokens Theory. In the process of phase-space reconstruction, the value of time delay and embedding dimension will directly affect the quality.

Embedding dimension and time delay are dependent between each other and need to be calculated one by one. While recent research showed that there is a tight relationship between embedding dimension and time delay, the selection of time delay should not depend on embedding dimension, and should combine the time window to define the parameter via the equation $\tau_w = (m-1)\tau$[9].

C-C Algorithm[11] creates the statistics by using the correlation integral sequence, where the statistics represents the relevance of nonlinear time series. It calculates the τ_w and τ by the relation between statistics and time delay, from which the embedding dimension m can be obtained.

With the calculated parameters, predictions can be made in the following steps.

Suppose the state vector in T is

$$X(T) = (x(T), x(T-\tau), ..., x(T-(m-1)\tau)) \tag{2}$$

The nearest neighborhoods of $X(T)$: $X(T_1), ..., X(T_K)$ can be used to predict the state vector of $X(T+1)$:

$$\hat{X}(T+1) = \frac{1}{K}\sum_{K=1}^{K} X(T_K+1) \tag{3}$$

After reformulating the above equation in the following format:

$$\hat{x}(T+1) = \frac{1}{K}\sum_{K=1}^{K} x(T_K+1) \tag{4}$$

The prediction equation can be expressed as:

$$\hat{x}(T+1) = c_{0,0} + c_{1,0}x(T) + c_{2,0}x(T-\tau) + ... + c_{m,m}x(T-(m-1)\tau)^2 + e \tag{5}$$

Here, e is the random error, while $c_{i,j}$ are parameters to be defined.

To calculate $c_{i,j}$, two matrices should be defined first:

$$y = [x(T_1+1), x(T_2+1), ...x(T_K+1)]^T \tag{6}$$

$$A = \begin{bmatrix} 1 & x(T_1) & x(T_1-\tau) & ... & x(T_1+1)^2 & ... & x(T_1-(m-1)\tau)^2 \\ 1 & x(T_2) & x(T_2-\tau) & ... & x(T_2+1)^2 & ... & x(T_2-(m-1)\tau)^2 \\ ... & ... & ... & ... & ... & ... & ... \\ 1 & x(T_K) & x(T_K-\tau) & ... & x(T_K+1)^2 & ... & x(T_K-(m-1)\tau)^2 \end{bmatrix} \tag{7}$$

So, $c_{i,j}$ can be calculated as

$$C = [c_{0,0}, c_{1,0}, ..., c_{m,m}]^T = (A^T A^{-1})A^T y \tag{8}$$

Suppose $T = 1000, m = 10, \tau = 10, K = 50$, the final equation for prediction is

$$\hat{x}(1001) = c_{0,0} + c_{1,0}x(1000) + c_{2,0}x(990) + ... + c_{10,10}x(910)x(910) + e \tag{9}$$

2 Model-driven Methods

The key factors of model-driven methods are the completeness and calculability of the model for short time intervals. In order to obtain better results for prediction, two aspects should be considered. One is the parameters for the prediction model, and the other is the minimum number of those parameters[2].

The method from [13] is one of the model-driven methods. The general model is defined below[13]:

$$X_t - \sum_{t=1}^{p}\sum_{t=1}^{I}[\Phi_{lir} \bigotimes S^{ri}]X_{t-i,r} = \alpha_t + \sum_{j=1}^{q}\sum_{i=1}^{I}[\Theta_{jir} \bigotimes S^{ri}]\alpha_{t-1,r} \tag{10}$$

In this equation, the key factor is every single matrix of S^{ri}. Here, i refers to the number of correlation matrices, and r is the template. The value of S^{ri} is either 0 or 1. The symbol $\bigotimes$ refers to the Hadamard matrices.

Every single matrix of S^{ri} shows the result of traffic flow, using 1 to indicate that the flow can reach a certain section in the time interval, and 0 to indicate that it cannot do so[13]. Given the volume of cars and the matrix S^{ri}, the model can make the prediction. In order to define all the matrices of S^{ri}, the parameter r and i should be defined first. Because r refers to the templates of the traffic flow, it is usually defined as $r = 2$ in order to describe the peak and off-peak traffic flow. Parameter i is associated with the number of intervals.

Here is an example with $r = 2$ and $i = 2$. The following figure shows the connections of the roads. In order to simplify the calculation, the time interval is ten minutes for $i = 1$, and all the roads are assumed to have the same length and width. There will be 16 sections of the roads.

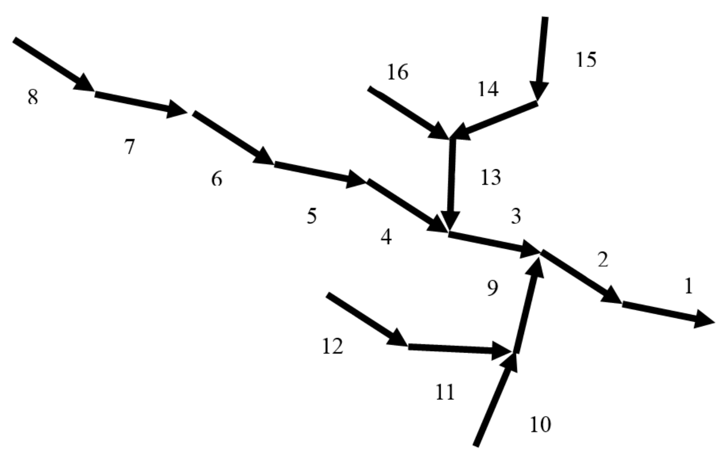

FIGURE 1: Road connections

Based on the figure, it is easy to obtain the matrices of S^{ri}.

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	1
4	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
5	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

TABLE 2: Values of S^{11}

In Table 2, 1 represents that traffic flows can shift from location i to j, while 0 represents that the flows cannot finish shifting or there is no route from i to j. Table 2 describes the traffic flow during off-peak hours, while Table 3 shows the flow during peak hours. Therefore, with these values, it is easy to determine the trends of traffic flow. With the specific volume number, the model can make predictions.

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	1	1	1	0	0	0	0	1	1	1	0	1	0	0	0
2	0	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1
3	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1
4	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

TABLE 3: Values of S^{21}

The accuracy of the model is related to the values of matrices. Therefore, an improved method for calculation of distance and speed will make improvements to the whole prediction.

3 Sharp Variation Points by Wavelet Transform

Traffic flow time series show different fluctuation characteristics at different time scales[24]. The wavelet transform is a good way to divide the sequence into multiple scales. Those scales are divided based on the number of the zero-crossing points from wavelet transform, which are called sharp variation points. Researchers have proven that on a certain scale, the number of sharp variation points in traffic flow shows the characteristic of self-similarity[24].

According to changes in signal, the definition of the wavelet transform[26] is

$$W_f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x)\varphi(\frac{x-b}{a})dx \tag{11}$$

Here, parameters a and b represent factor of scale and time shift respectively. By changing these two parameters, the information of $f(x)$ can be obtained at any location b under different scale a. Function $\varphi(x)$ is defined as below[25]:

$$\varphi(x) = (1-x^2)e^{-\frac{x^2}{2}} \tag{12}$$

This is a continuous smooth even function. Before using it in traffic flow calculation, it should be made discrete. The new formulation is changed[25]:

$$W_f(a,b) = \frac{1}{\sqrt{a}} \sum_{n=1}^{N} f(i\triangle x)\varphi(\frac{n\triangle x - b}{a}) \tag{13}$$

Then put the time series $q(n)(n = 1, 2, ..., N)$ into the function, the new form will be[24]:

$$W_q(a,t) = \frac{1}{\sqrt{a}} \sum_{n=1}^{N} q(n)\varphi(\frac{n-t}{a}) \tag{14}$$

Figure 2 shows an example after the wavelet transform.

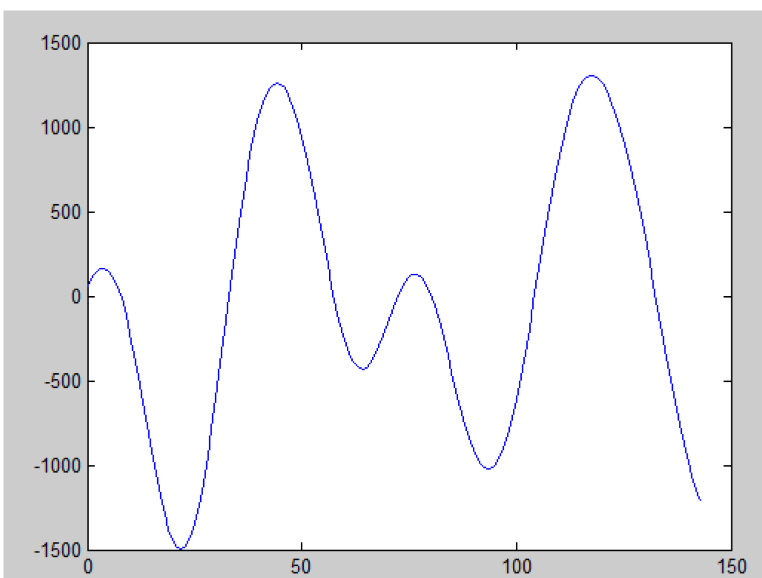

FIGURE 2: Example figure of wavelet tranform

The key word of this section is sharp variation points. A sharp variation point is a point to describe the trend's beginning or end. With this important characteristic, the detection work is paramount. Sharp varation points are defined as the zero-crossing points of the wavelet transform. Based on the figure, it is easy to see that there are seven sharp variation points. Because the data includes 144 points, it may not get zero in an accurate time position. There is another way to define the sharp variation points[9]. While $q(n) * \theta(\frac{t-n}{a})$ is the Gaussian smoothing function of series $q(n)$ under the scale a in time t, it keeps the properties in $t = n$, and $q(n) * \theta(\frac{t-n}{a}) = q(n)$. Therefore the zero-crossing points of $W_q(a,t)$ can be regarded as the sharp variation points of $q(n)$[24]. For any scale a, there is a way to detect the sharp variation points:

$$W_q(a,t)W_q(a,t+1) < 0 \tag{15}$$

Using this function, the positions of sharp variation points are detected:
$t = 7, 33, 57, 72, 80, 104, 132$.

CHAPTER IV

Augmented Reality

Augmented reality is a brand new word for most people. It refers to a new way to interact with the virtual world. Actually, it is changing our world in all aspects. Although it is not popular at the moment, the fast speed of development in this area will certainly make it expand worldwide.

1 Introduction

Augmented reality, based on virtual reality, is a variation of virtual environment[4]. However, there is some difference between augmented reality and virtual reality. Augmented reality is more concerntrated and interactive while virtual reality focuses more on experiencing than augmented reality does.

In the final decade of the 20th century, with huge development in the technology of cameras and computation, augmented reality came to people's sight. At the beginning, the technology was just an imagination. Then, along the development of advanced imaging technologies, people were able to experience a newly created virtual world using heavy equipment. During the past 20 years, the equipment has become lighter, and the quality of imaging has become better than before, with higher resolution.

The most common tool to realize augmented reality is a monitor. Unlike druing the 1990s, every monitor can now be an interface for augmented reality. Smartphones, computers, GPS devices, iPads can all provide a scene for augmented reality[16]. However, some monitors can only represent the virtual world to certain scales. If people want to experience the virtual world in any location without limitations, augmented reality glasses are the best choice. These glasses can provide a 360-degree virtual world.

Besides hardware, software is also significant for augmented reality. Software is related to the speed and quality of imaging. Additionally, it provides a development environment for the virtual world. Unity 3D is a good engine that can be used to construct an interface between the user and the virtual environment[22].

2 Applications

Augmented reality is going to change people's lives in different fields. This section shows some applications that have been realized or are being developed[16].

- Medical: Medical research is always important because it is related to people's lives. Many surgeries are very dangerous and required very experienced doctors. Therefore, the problem is how to cultivate experienced doctors. Because there are not enough cases for new doctors and surgery itself is dangerous, it is unlikely that new doctors will perform the operation. Augmented reality is an effective way to solve this problem. It can create a virtual environment for the surgery, and it allows anyone to perform operation using augemented reality(AR) equipment. After several training sessions in the virtual world, surgeons become experienced. Addtionally, AR can help to demonstrate some new progress in medicine that works in molecular levels in human's bodies.

- Military: The cost of running the military is extremely high all the time because all military experiments should have high accuracy. Doing military experiments also cause some environmental problems, such as air pollution and nuclear pollution. AR is an effective way to conduct simulated experiments, instead of doing them in reality. For example, experiments involving missiles, it can be simulated on screens. When changing the parameters, the track of the missile can be observed on a screen. This saves much time and money. AR can also simulate battle fields in order to allow new soldiers to practice their shooting or driving skills. That will help greatly in terms of survival.

- Entertainment: AR technology provides a new way for entertainment, especially the way to play games[6]. Some video game companies like Sony and Microsoft are devoloping new AR technologies. When people put on AR equipment, a scene in a game will be created immediately from a first person's perspective. When the user moves, the scene moves as well. That is a revolution in entertainment that provides interactive functions between users and virtual environments.

3 Future Development

Augmented reality has the potential to be one of the most popular technologies in the future. The technology is not perfect now. It still has much to improve:

- Hardware: This is now the greatest obstacle for the AR technology. All of the advanced equipment is very large and cannot meet people's demands for portability. Addtionally, more sensors need to be added to the equipment in order to make the simulation more accurate and realistic.

- Software: More software that provides an environment for building virtual worlds needs to be developed. Software must improve in terms of calculation and imaging technology in order to help create the virtual world quickly with high image resolutions.

- Applications: More applications should be developed as well. This is the key for people to experience the power of augmented reality. Applications should be easy to get at low cost. Applications in more areas are needed for augmented reality.

CHAPTER V

Analysis and Prediction based on Sharp Variation Points

1 Statement of Problems

Before doing research about traffic flow prediction, the problems that need to be solved must be identified. This thesis tries to make some improvements in the following directions.

- Model: This research is going to develop a model that is easy to operate and applicable to all conditions. The model should get deep information and can analyse the fluctuations of traffic.
- Accuracy: Higher accuracy is required under the same conditions than previous models. The accuracy should be relatively stable as time increases.
- Visualization: Another important part of this thesis is to combine the results of the new idea to be presented in the following section with augmented reality in order to make a simulated word. The virtual world can simulate how the traffic flows in general.

2 New idea of Prediction

As shown in the previous section, the time series of traffic flow is scale-free and self-similar, and the sharp variation points record what happened at sharp variation time. They show how the traffic flow fluctuates in a day. Every sharp variation point means something happens or ends at that time, which is useful for the prediction. The model can have the prediction up to 2 sharp variation points.

Because every figure of wavelet transform contains much information about the traffic flow, the fluctuation can be classified according to different situations. In other words, it can put every section of figures from the wavelet transform into certain categories. In order to obtain accurate classification results, some typical parameters are needed for all kinds of conditions. As sharp variation points mean a lot for the prediction, $n = n/t$ is defined to determine whether the time period is the one with issues or not. For weather conditions, $\bar{X} = \sum_{i=1}^{N} X_i/N$ can be used to evaluate the weather and $M = X_{max} - X_{min}$ to evaluate the level of weather. In addition, $\sigma = \sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2/N}$ is a good choice to consider situations of accidents. Simulaition results to be reported in the next chapter also points out that adding σ will increase accuracy and gain more important information. In the selection of parameters, N can be a boolean parameter. The other three parameters $\bar{X}$, M and σ are all in some floating ranges.

3 Algorithms

3.1 Historical Data Collection

As all of the predictions will use historical data, the more data becomes available, the more accurate the analysis will be. All the data should be collected from the same section of road, or it will be useless. Figure 3 is the algorithm for data classification.

The aim of this process is to classify collected historical data. The first step is to make transformation using wavelet transform, and to detect the positions of sharp variation points.The n in Figure 3 is the parameter to determine whether there are sharp variation points in a day. The next step is to define the weather conditions using parameters $\bar{x}$ and M. After that, the situations of accidents can be defined using σ. The last step is to classify a certain day's data into a certain Category N by

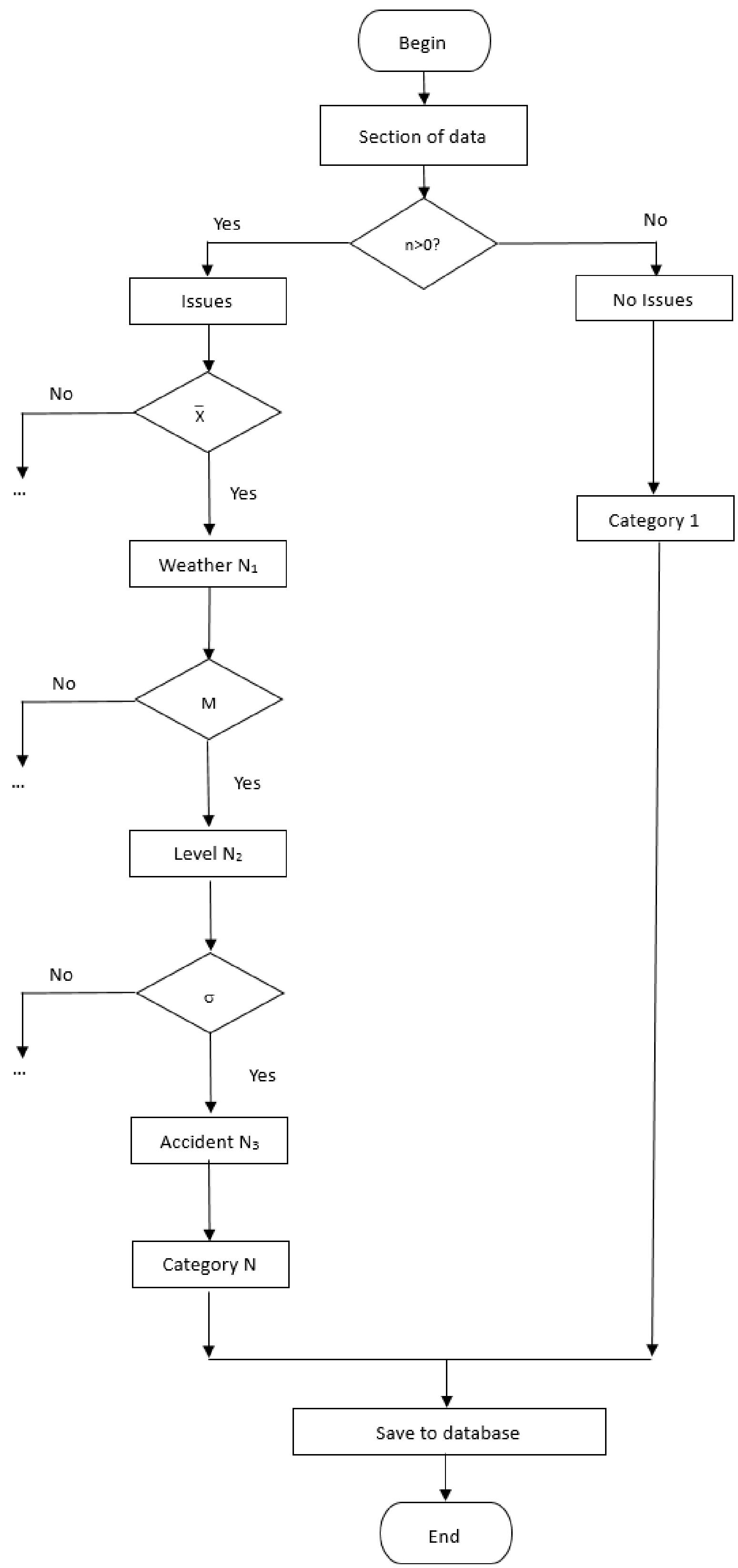

FIGURE 3: Diagram for algorithm of historical data collection

combining these three parameters.

3.2 Similarity Comparison and Prediction

This section discusses how to use real-time traffic data for comparison and prediction based on sharp variation points.

Figure 4 shows the algorithm for similarity comparison. After getting real-time data, the first step is to follow the algorithm in the previous section to determine what category it belongs to. If the real-time data comes with information of weather conditions and accidents' situations, for example, the parameters of weather and accidents will be used to find the category for the real-time data. The next step is to search all the data from database under the same category. The aim of this step is to find out the best sample for prediction, and comparison will be made in the same time periods among these data. The function used for comparison is $\mu_{j_i} = \sqrt{\sum_{T=1}^{T}(X_i - X_j)^2/T}$, where X_j is the time series of real-time data and X_i is the time series of historical data. The index of i refers to the best sample when μ_{j_i} obtains its minimum value.

After finding the best sample, the next step is to make the prediction. Before that, it needs to find out the positions of sharp variation points from the best sample. If the last time point of real-time data is T_0, the work is to find the next two positions of sharp variation points T_{X_i+1} and T_{X_i+2}. Prediction is then achieved by using the corresponding data segment from the best sample from T_0 to T_{X_i+2}. Figure 5 shows the algorithm for prediction.

The accuracy can be calculated as following formula:

$$Accuracy = 1 - \frac{1}{T}\sum_{i=1}^{T}[\frac{forcast.vol(i) - vol(i)}{vol(i)}]$$

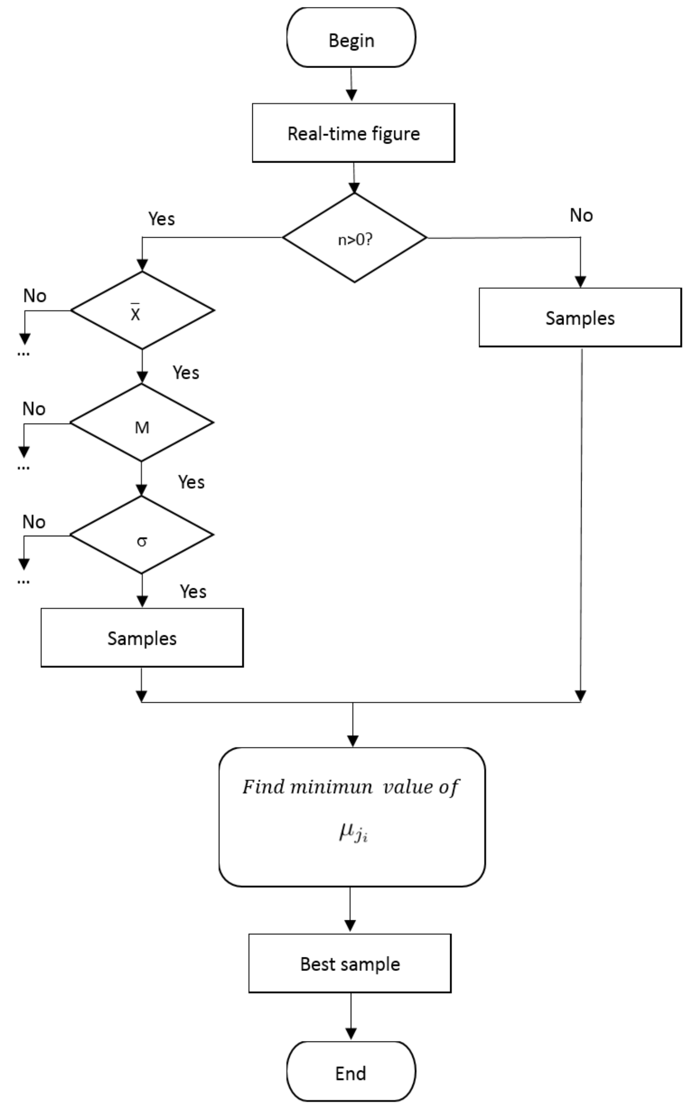

FIGURE 4: Diagram for algorithm of comparison

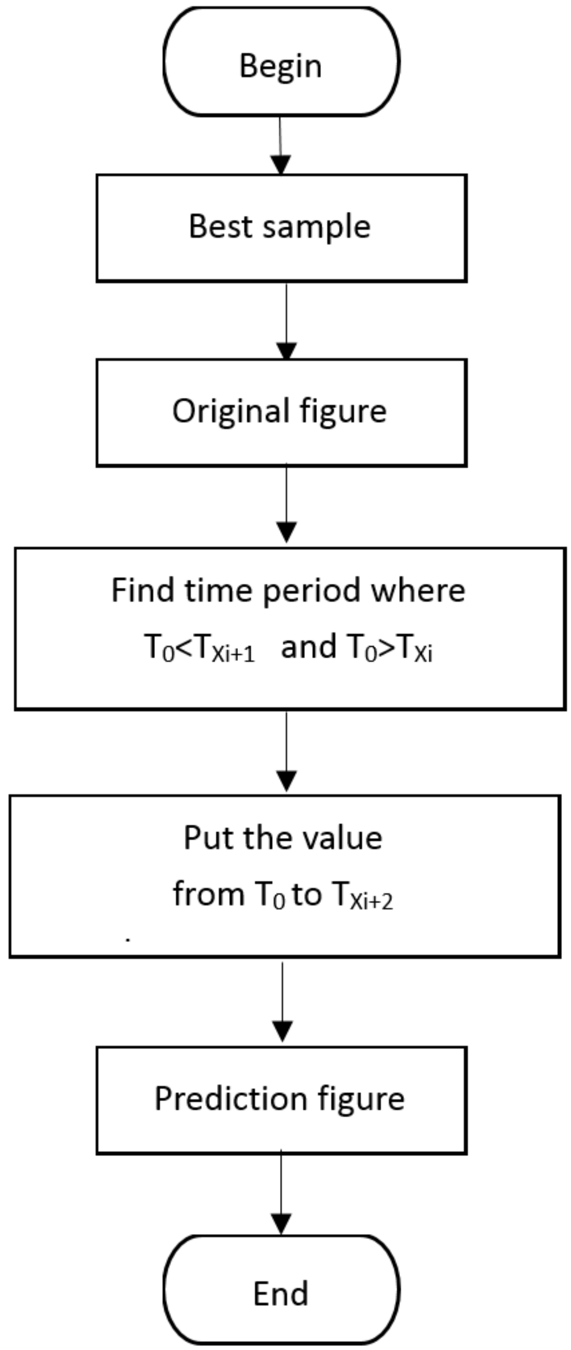

FIGURE 5: Diagram for algorithm of prediction

4 Time Complexity

The most time is spent on the preparation for the prediction. If n is the number of the data, the time complexity will be:

- Wavelet transform: $O(n^2)$.
- Sharp variation points detection: $O(n)$.
- Classification: $O(n)$.
- Similarity comparison: $O(n)$.
- Prediction: $O(n)$

If the prediction repeats for $m(m < n/2)$ times, the total time complexity will be: $O(n^2) + O(n) + O(n) + m(O(n) + O(n)) = O(n^2)$.

CHAPTER VI

Experiments with the Proposed Method

1 Historical Data Collection

The simulation data used in this thesis is created based on the data presented in a published paper[24]. The original data was collected by Beijing Jiaotong University. The simulation data recorded the volume of the traffic flow for ten days. The first eight days is selected for trainings in this experiment, and the rest for testing. Shown in Figure 6 is an example of one day's stream of traffic, where the time interval is ten minutes. That means there will be 144 points to record the trend of a day. As the time interval is ten minutes, when $t = 0$ refers to 12:00 a.m., $t = 1$ refers to 12:10 a.m. and so on. In this hypothetical environment, weather conditions were assumed to be the same, and no accidents happened on these days($\bar{X}$ is around -32, M is about 2795 and σ is around 65). The method works no matter the day is a workday or during weekend. Though there is some difference between workdays and weekends, it will not change the way to calculate the positions of sharp variation points. Some little fluctuations in wavelet transform can be detected in the process of similarity comparisons, and the best samples must have some conditions.

Figure 6 shows how traffic flows in a day. It is clear that, there are three high peaks in a day. That means there are rush hours in the morning, afternoon and evening. The estimated time of those three peaks occurs at about eight o'clock in the morning, twelve o'clock in the noon, and six o'clock in the evening. During these

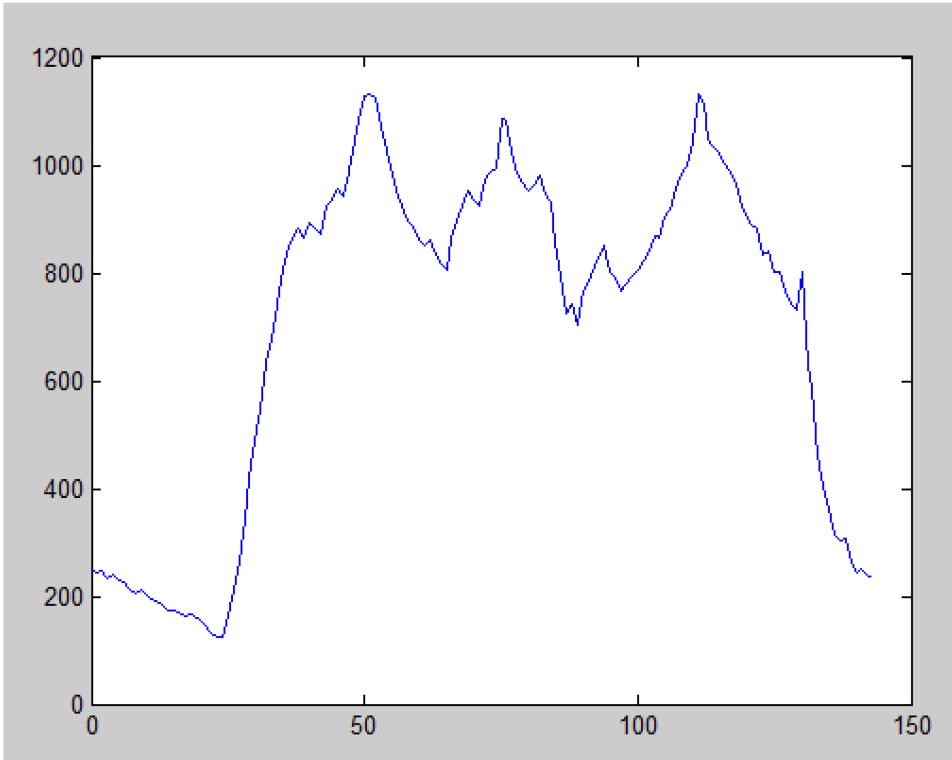

FIGURE 6: Traffic flow of Day1

time periods, most people go to or from work. Therefore, the flows in the figure seem reasonable. There are also some minor trends after the second high peak, which may indicate that some people go back home earlier or some people go out for dinner or shopping in the afternoon. The last high traffic trend is in the late evening. That means people who have finished dinner go back home.

Figure 7 is a figure that shows all of the trends over ten days. It is obvious to see that trends are similar for each day because people's lifestyles are almost the same. For the experiments in this section, 80 percents of the information is used for training sets, while the rest is used as testing sets.

In order to gain a deep understanding of the traffic flow patterns, wavelet transform can be used to obtain additional information. Because wavelet transform can describe the fluctuation of traffic, it is an effective way to analyse the traffic flow.

The formula for the wavelet transform is[24]: $W_q(a,t) = \frac{1}{\sqrt{a}} \sum_{n=1}^{N} q(n)\varphi(\frac{n-t}{a})$, and $\varphi(x) = (1-x^2)e^{-\frac{x^2}{2}}$. In the formula, $q(n)$ is associated the number of the vehicles

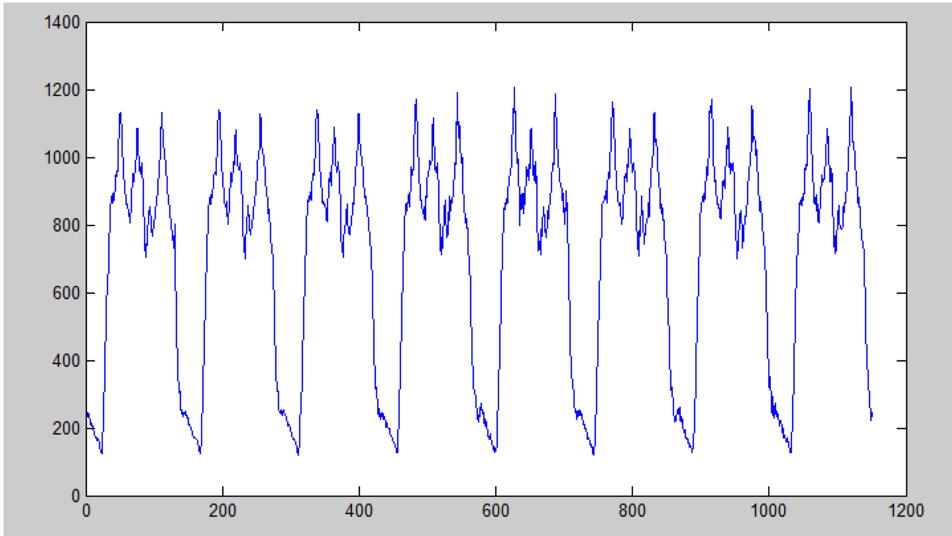

FIGURE 7: Traffic flow of training sets

at time t and a is the factor of scale, which is ten in this dataset because of the ten-minute time interval. After applying the wavelet transform to the original data, the figure is shown below

Based on Figure 8, the fluctuation of the traffic can be easily observed. In order to obtain more detailed information, some calculations need to be done based on the figure. The most important parameter for the wavelet transform is the sharp variation point. Sharp variation points are defined as zero-crossing points from wavelet transform, which can be detected using another formula[24]: $W_q(a,t)W_q(a,t+1) < 0$.

In experiments, sharp variation points were detected at t=7,33,57,72,80,104,132. It means that there are sharp variations at 1:10 a.m., 5:30 a.m., 12:00 p.m., 1:20 p.m., 5:20 p.m. and 10:00 p.m. At these time points, some sharp changes occur. Additionally, with some other parameters, the whole traffic flow can be described in a certain category: $\bar{X} = \sum_{i=1}^{N} X_i/N = -32.25$, $M = Max - Min = 2795.82$, $\sigma = \sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2/N} = 65.93$.

Figure 9 shows data from the traning sets after wavelet transform.

Obviously, with similar values of the original data, the trend of wavelet transforms also shows in high similarity.

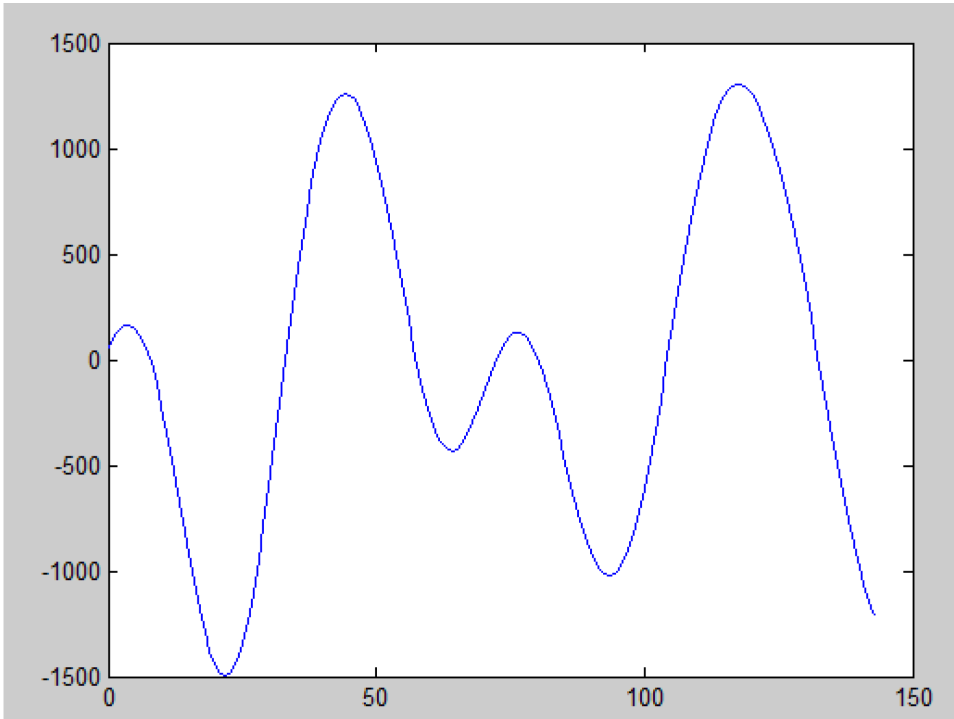

FIGURE 8: Wavelet transform for Day1

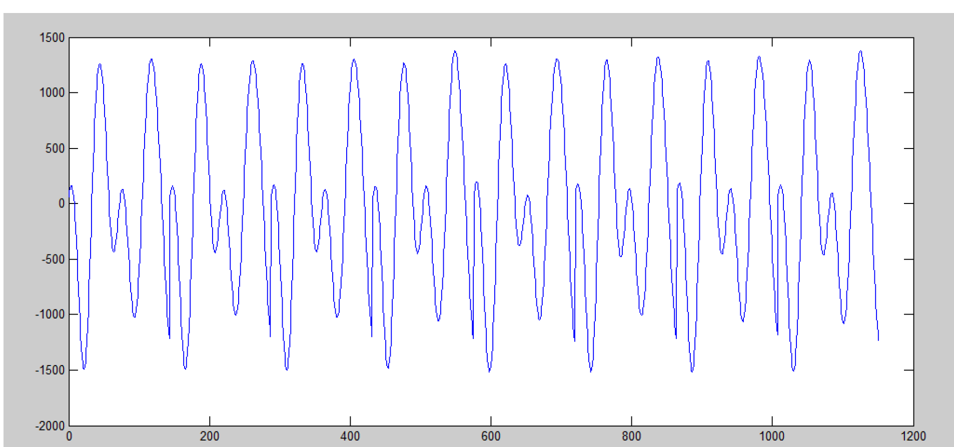

FIGURE 9: Wavelet transform for training sets

2 Prediction

After trainning, the experiment will take the data from Day 9 and Day 10 for testing, whose original data are shown in Figure 10 and Figure 11.

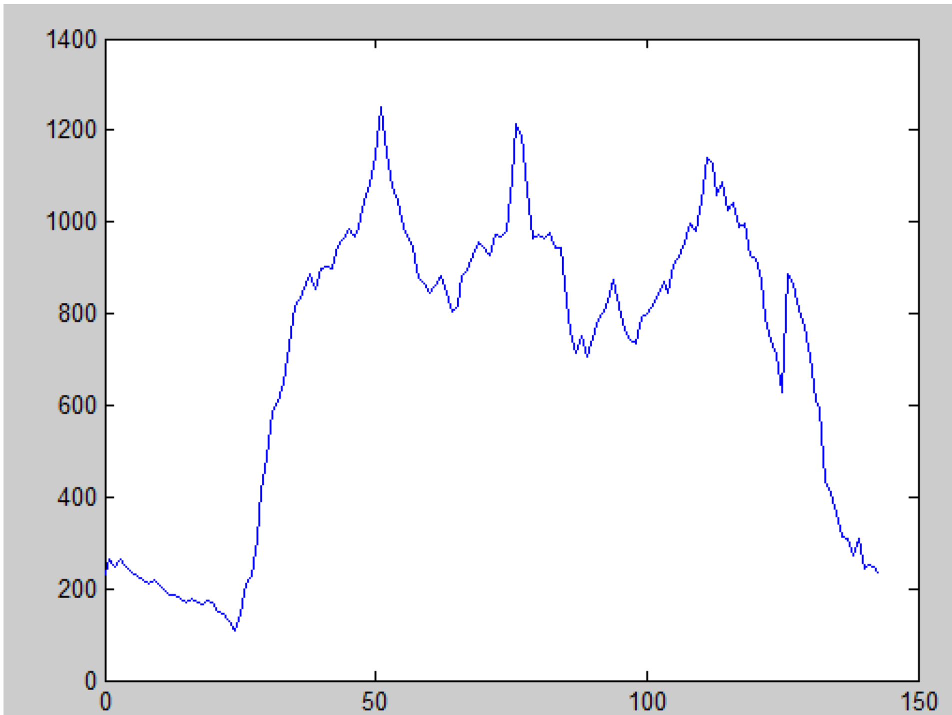

FIGURE 10: Original traffic of Day 9

The prediction will use the basic data from 12:00 a.m. to 8:00 a.m.. The first step is to apply wavelet transform to this data, and obtain the information about every value of these points. In order to calculate the similarity, the data must be in the same period. The formula used for this part is defined as: $\mu_{9_i} = \sqrt{\sum_{i=1}^{T}(X_i - X_j)^2/T}$.

Here i is the index of the day from the training sets, and X_j is the value of testing sets. Figures 12 and 13 show the data before prediction for Day 9. The first one is the original traffic flow, and the second one is the figure after applying the wavelet transform.

After applying the wavelet transform to Day 9, the first sets of μ_{9_i} show as: $\mu_{9_1} = 50.82, \mu_{9_2} = 50.45, \mu_{9_3} = 48.20, \mu_{9_4} = 51.52, \mu_{9_5} = 35.21, \mu_{9_6} = 28.61, \mu_{9_7} = 30.32, \mu_{9_8} = 36.78$. The minimum of these values can be easily identified as: $\mu_{9_6} =$

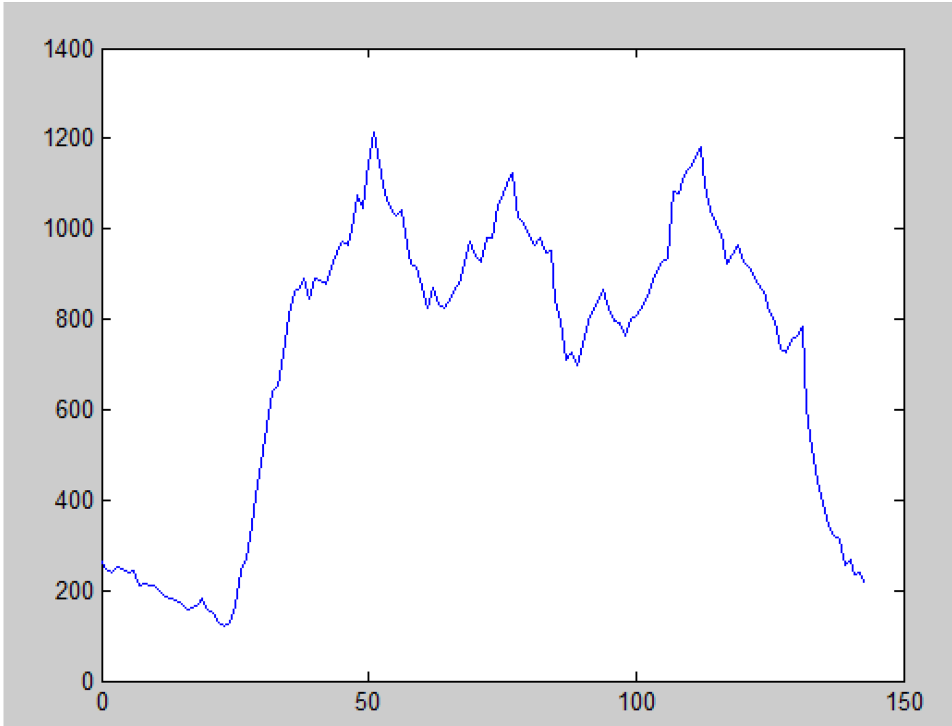

FIGURE 11: Original traffic of Day 10

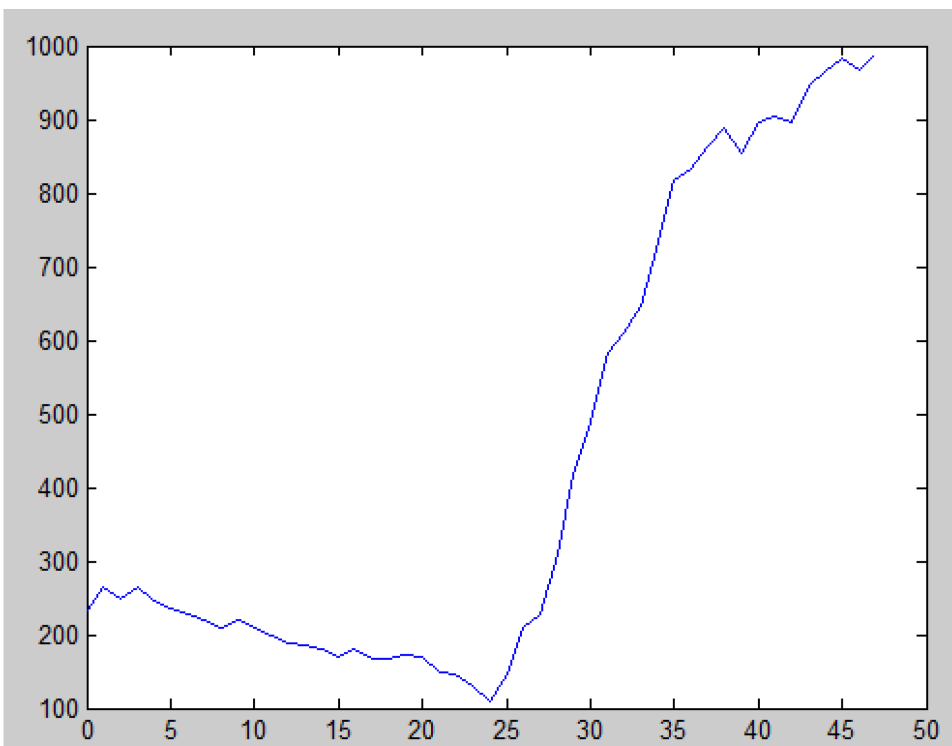

FIGURE 12: Known data of Day 9

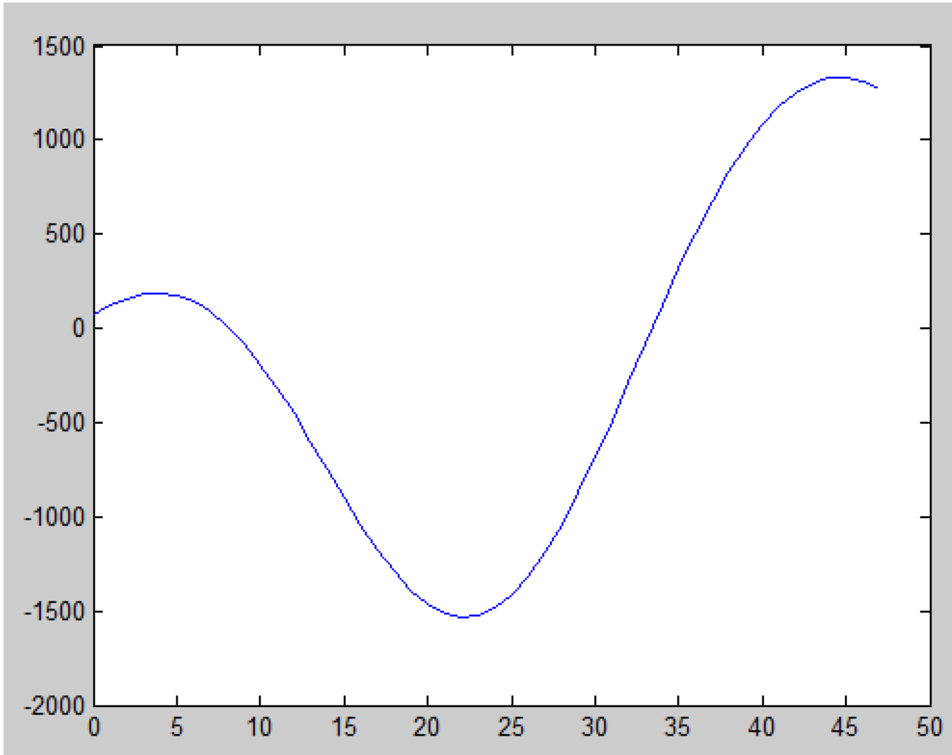

FIGURE 13: Known wavelet transform of Day 9

28.61.

The next step is to check the position of sharp variation points of Day 6 from its wavelet transforms. All the positions of sharp varation points have been recorded in the files before predictions. The next two positions of sharp variation points of Day 6 are $t = 57$ (9:30 a.m.) and $t = 72$ (12:00 p.m.). Therefore, the data from $t = 48$ (8:00 a.m.) to $t = 72$ (12:00 p.m.) can be used for prediction.

Figure 14 shows the first prediction for Day 9. In the figure, the red line with stars is the prediction results while the blue line is the original data. Obviously, these two lines are very similar.

Then, the steps above are repeated for the second prediction. Because this is a real-time traffic flow prediction, it must use treal-time traffic data for prediction. The first step now is to calculate the data from $t = 0$ (12:00 a.m.) to $t = 72$ (12:00 p.m.) using wavelet transform, and find the minimum μ_{9_i}. The values are $\mu_{9_i} : \mu_{9_1} = 50.74, \mu_{9_2} = 49.03, \mu_{9_3} = 48.34, \mu_{9_4} = 48.27, \mu_{9_5} = 54.53, \mu_{9_6} = 30.54, \mu_{9_7} = 32.41, \mu_{9_8} = 38.71.$

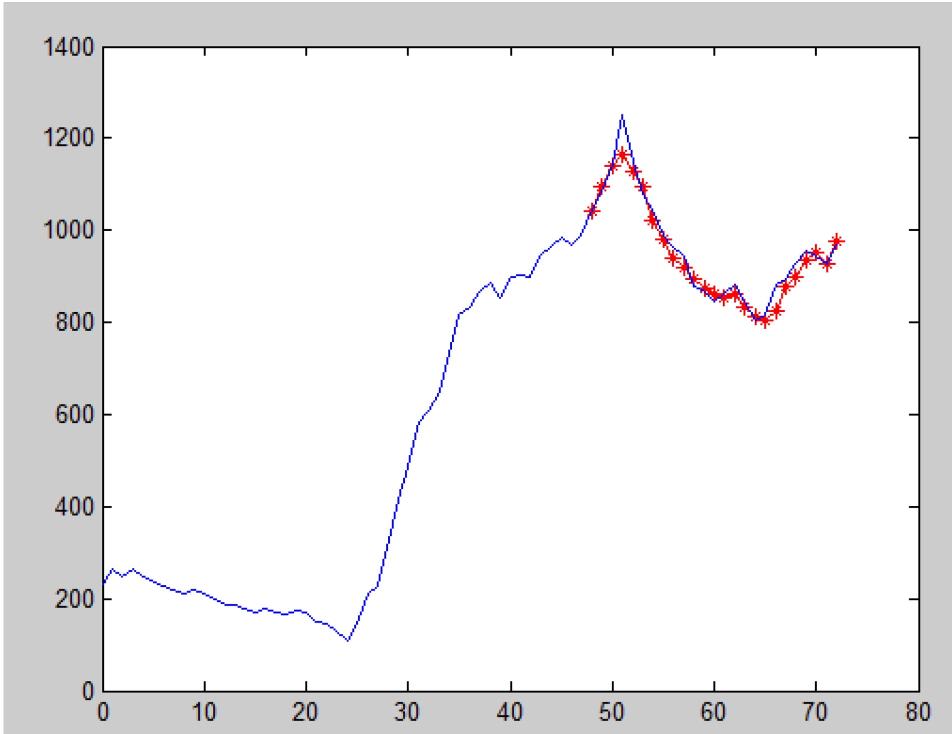

FIGURE 14: The first prediction of Day 9

The minimum of this set is $\mu_{9_6} = 30.54$. So, the data from Day 6 will be used for the second prediction.

The next step is to check the position of the sharp variation points of Day 6. The next two sharp variation points are $t = 80$ (1:20 p.m.) and $t = 104$ (5:20 p.m.). Then, use the data from $t = 72$ (12:00 p.m.) to $t = 104$ (5:20 p.m.) from Day 6 is used to make the second prediction.

Figure 15 shows the results of the second prediction. The prediction figure has a high similarity with the original figure from $t = 48$ (8:00 a.m.) to $t = 104$ (5:20 p.m.).

The third prediction uses the data from $t = 0$ (12:00 a.m.) to $t = 104$ (5:20 p.m.) to obtain the third sets of μ_{9_i} : $\mu_{9_1} = 58.11, \mu_{9_2} = 61.75, \mu_{9_3} = 56.66, \mu_{9_4} = 51.15, \mu_{9_5} = 74.05, \mu_{9_6} = 51.07, \mu_{9_7} = 42.36, \mu_{9_8} = 57.49$. The minimum value in this set is $\mu_{9_7} = 42.36$. So, the data from Day 7 will be used for the third prediction.

On Day 7, the only position of sharp variation point left is $t = 132$ (10:00 p.m.). Therefore, the prediction uses the data from $t = 104$(5:20 p.m.) to $t = 144$ (12:00 a.m.) of Day 7.

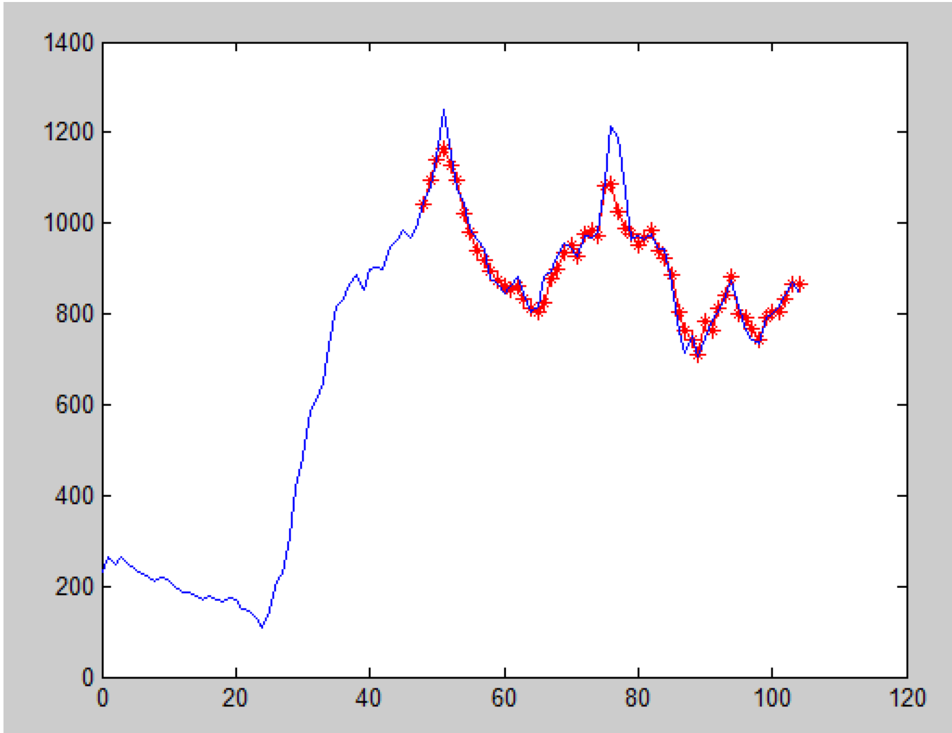

FIGURE 15: The second prediction of Day 9

Figure 16 shows the final prediction for Day 9

After all the data for prediction were determined, the accuracy can be calculated. The result is 96.34%.

Similarly, the prediction for Day 10 repeats the procedure for Day 9. The first step is to obtain the basic data from $t = 0$(12:00 a.m.) to $t = 48$ (8:00 a.m.). Figure 17 shows the basic data of Day 10 before prediction.

The second step is to calculate its form in wavelet transform and obtain the same set of μ_{10_i} for further prediction.

Figure 18 is the results for Day 10 after applying wavelet transform from $t = 0$ (12:00 a.m.) to $t = 48$ (8:00 a.m.).

The first sets of μ_{10_i} show as: $\mu_{10_1} = 32.16, \mu_{10_2} = 31.48, \mu_{10_3} = 31.62, \mu_{10_4} = 31.56, \mu_{10_5} = 13.72, \mu_{10_6} = 23.23, \mu_{10_7} = 25.85, \mu_{10_8} = 23.44$. The minimum value in this set is $\mu_{10_5} = 13.72$, In order to obtain the next two sharp variation points for prediction, the information of Day 5 needs to be checked. The positions for next two sharp variation points are $t = 58$ (9:40 a.m.) and $t = 72$ (12:00 p.m.). Then, use the

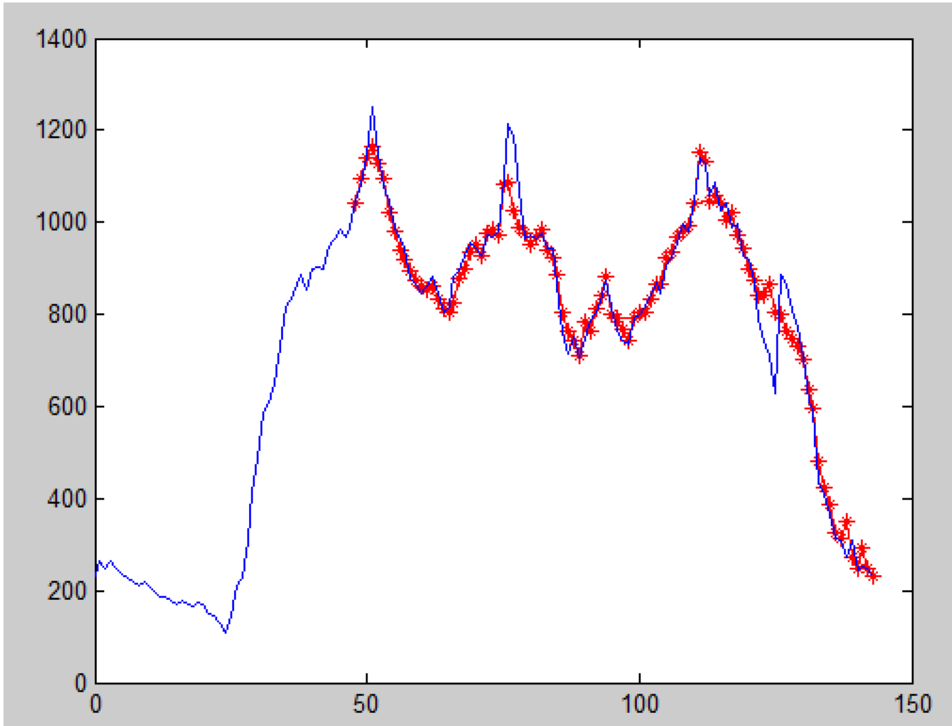

FIGURE 16: Final prediction of Day 9

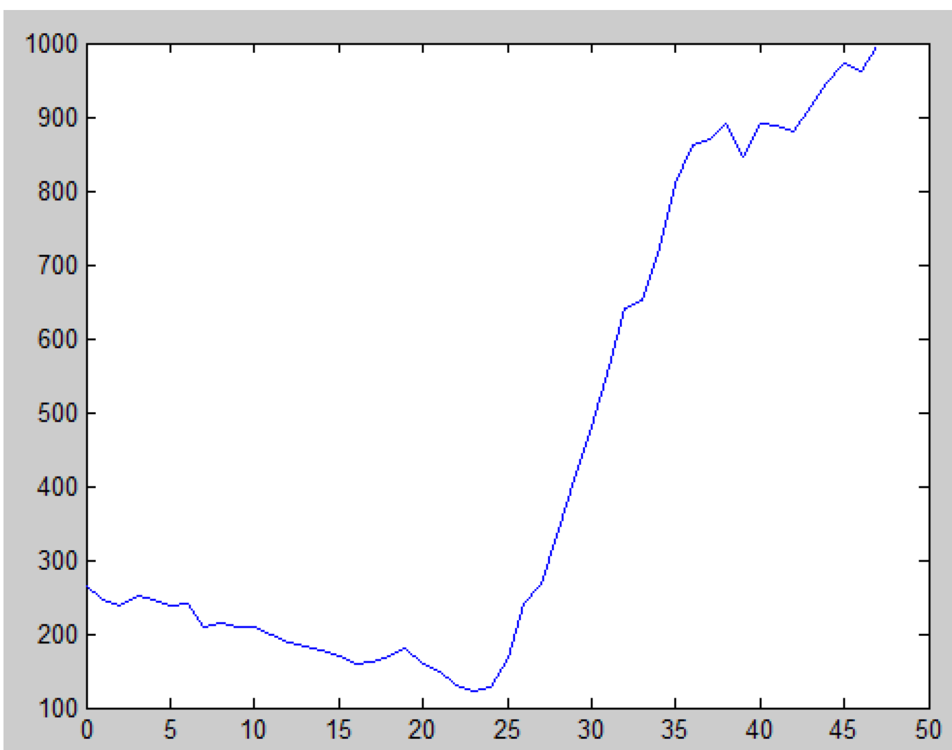

FIGURE 17: Known data of Day 10

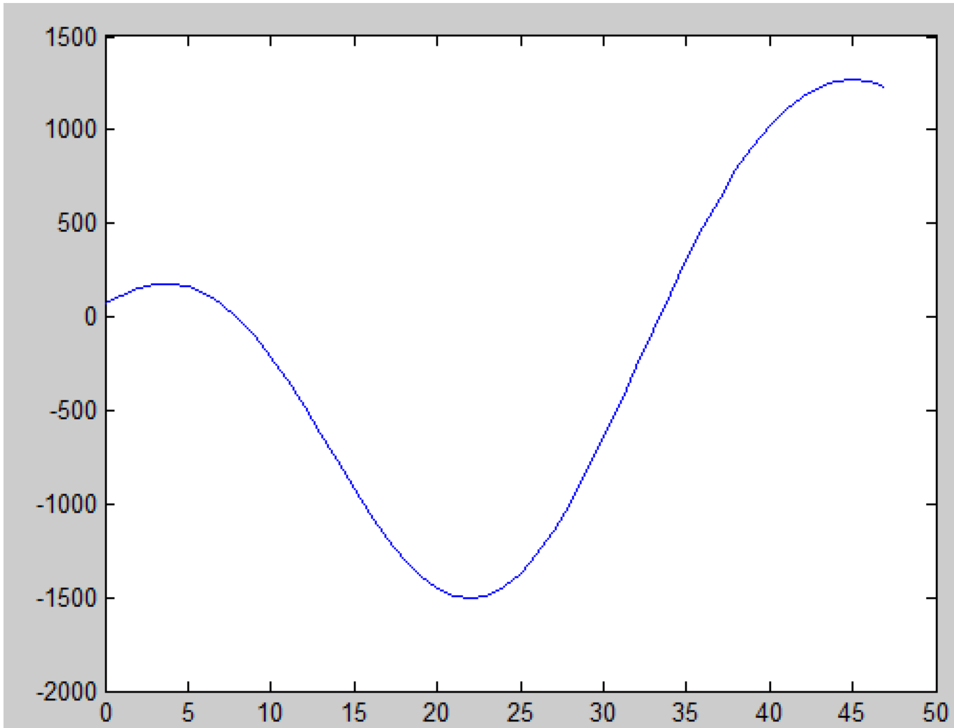

FIGURE 18: Known wavelet transform of Day 10

data from $t = 58$ (9:40 a.m.) to $t = 72$ (12:00 p.m.) in Day 5 is used to make the first prediction, whose result is shown in Figure 19.

Then, the beginning is updated to $t = 72$ (12:00 p.m.). The next step is to obtain the second sets of μ_{10_i} using the wavelet transform from $t = 0$ (12:00 p.m.) to $t = 72$ (12:00 p.m.). Here is the information about the second sets of μ_{10_i} : $\mu_{10_1} = 41.21, \mu_{10_2} = 38.44, \mu_{10_3} = 40.89, \mu_{10_4} = 45.47, \mu_{10_5} = 24.87, \mu_{10_6} = 31.07, \mu_{10_7} = 31.97, \mu_{10_8} = 24.44$. The minimum of this set is $\mu_{10_8} = 24.44$. In order to obtain the next two sharp variation points for prediction, the information of Day 8 needs to be checked. Based on the file, the next two positions are $t = 79$ (1:10 p.m.) and $t = 103$ (5:10 p.m.).

Figure 19 is the second prediction using the data of Day 8 from $t = 79$ (1:10 p.m.) to $t = 103$ (5:10 p.m.)

As shown in Figure 20, there is still unprocessed time periods for Day 10. Therefore, the third sets of μ_{10_i} needs to be obtained: $\mu_{10_1} = 50.68, \mu_{10_2} = 53.78, \mu_{10_3} = 50.26, \mu_{10_4} = 46.05, \mu_{10_5} = 38.55, \mu_{10_6} = 59.98, \mu_{10_7} = 45.30, \mu_{10_8} = 30.13$. Obviously,

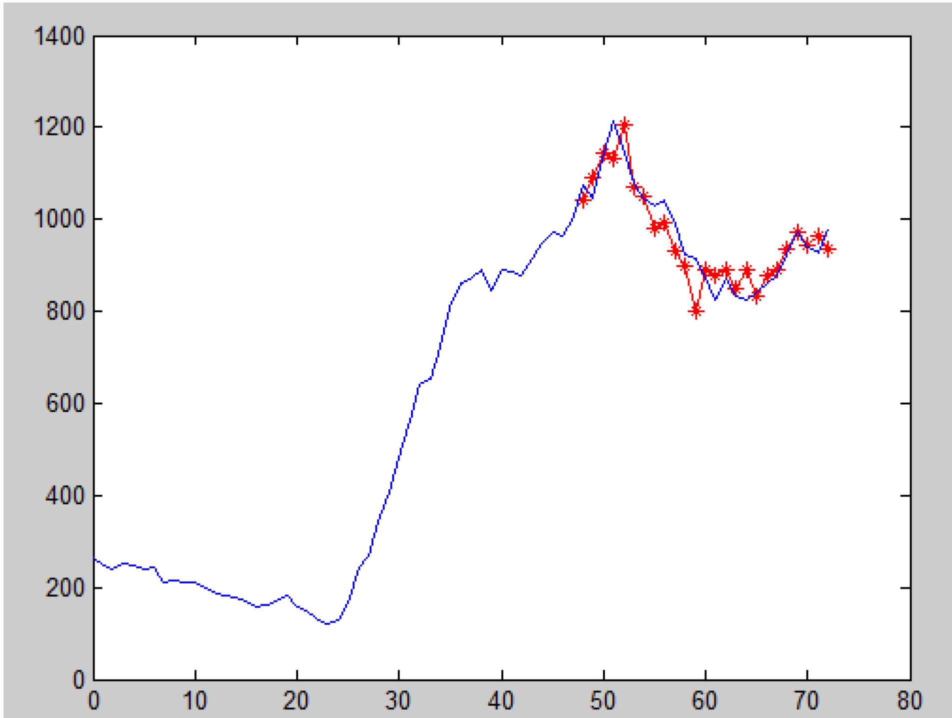

FIGURE 19: The first prediction of Day 10

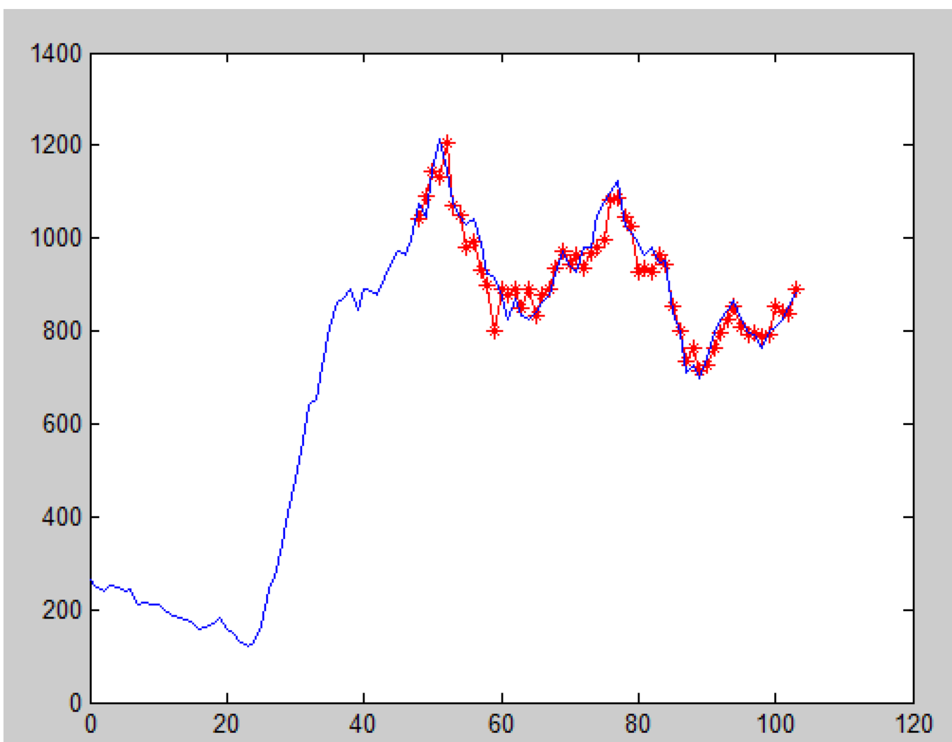

FIGURE 20: The second prediction of Day 10

the minimum value is $\mu_{10_8} = 30.13$. After checking the information from the remaining sharp variation points in Day 8, the only sharp variation point is at $t = 132$ (10:00 p.m.). Using the data from $t = 104$ (5:30 p.m.) to $t = 144$(12:00 a.m.) in Day 8, the final prediction is completed as illustrated in Figure 21.

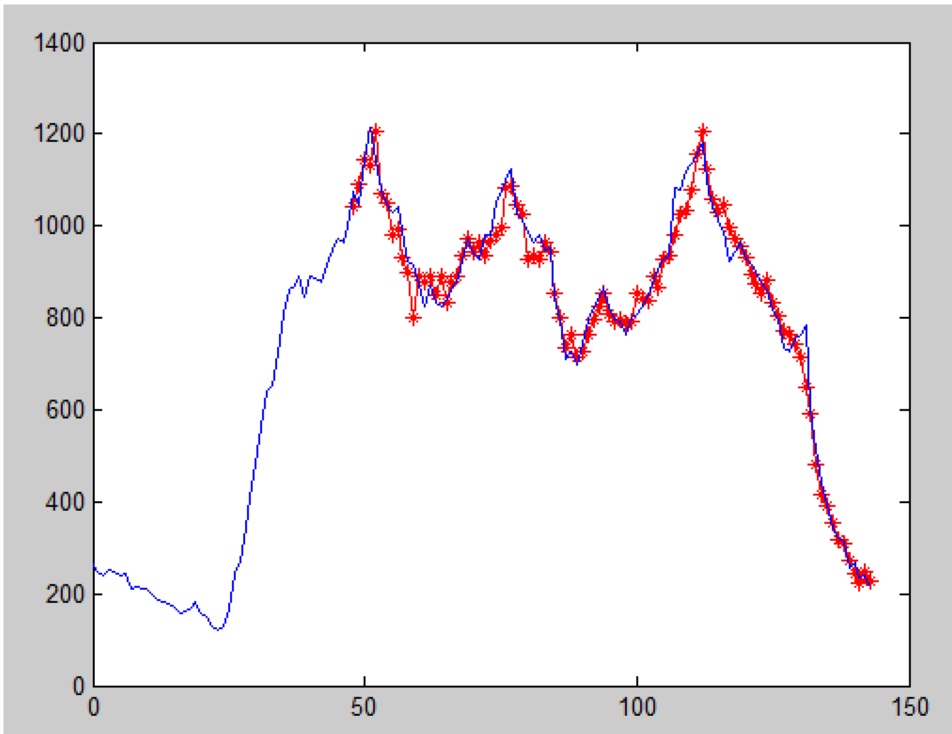

FIGURE 21: Final prediction of Day 10

The last step is to calculate the accuracy of the prediction for Day 10, and the result is 96.58%.

3 Comparison and Analysis

The accuracy for traffic flow prediction has been calculated by using the formula at the bottom of page 21 for both data-driven[10] and model-driven[13] methods. In the experiment presented in this chapter, the same formula is used to calculate the accuracy as well. In the experiment, the predictions is based on sharp variation points, and the results show that it is an effective method for forecasting. The accuracy for test sets is over 96%(Day 9 is 96.34% and Day 10 is 96.58%). The results gained

Model-driven	5-min	10-min	15-min
CATA	95.0%	94.3%	94.0%
CATB	87.3%	87.4%	87.3%
Slip-road	92.2%	91.5%	92.1%

TABLE 4: Accuracy of Model-driven methods[13]

Data-driven	2h	4h
New Volerra	96.02%	92.80%
RBFNN	94.79%	89.96%
Volerra	92.18%	92.07%

TABLE 5: Accuracy of Data-driven methods[10]

from fully calculation from training sets instead of samples, so level of confidence is not needed for the predictions. All of the predictions were made under the same category($\bar{X}$ is around -32, M is about 2795 and σ is arund 65).

The accuracy is high in comparisons with previous methods. While Table 4 and Table 5 show the accuracy of prediction with typical model-driven and data-driven methods, Table 6 shows the accuracy of prediction for Day 9 and Day 10 using the new method. In experiments presented in this thesis, the average accuracy is over 96% and the highest accuracy is over 98%. With the increase of the time interval, the accuracy still shows the characteristic of stability.

For all data, the time interval is ten minutes. It means that the factor of scale a is ten. If the data was collected in five or two minutes intervals, the value of a will

Proposed Method	1st Prediction	2nd Prediction	Final Prediction
Day 9	240 min	560 min	960 min
Accuracy	98.12%	97.74%	96.34%
Day 10	240 min	550 min	960 min
Accuracy	96.60%	96.91%	96.58%

TABLE 6: Accuracy of Proposed Method

be smaller. As a result, more sharp variation points will be obtained. In order to ensure the real-time characteristics of the traffic, the model only makes a prediction in a length of two sharp variation points. If there are more samples of traffic data and a is small enough, the accuracy of prediction with the new method will be even higher.

CHAPTER VII

Augmented Reality System Design

Another objective of this thesis is to make prediction become visible. In order to meet this goal, it is necessary to build a virtual world using augmented reality. When building an augmented reality system, the engine should be considered firstly. Unity 3D is an excellent choice because it is a powerful game engine that can provide all kinds of assets for the virtual environment.

This chapter will discuss the steps for building an augmented system with Unity 3D. The first part is modelling, which covers the basic modules used for the system. The second part discusses how to make system work using scripts and simulation. The last part gives an overview to the system.

1 Modelling

Modelling is a significant part of an augmented reality system. The first step to create a good augmented reality system is the modelling of 3D objects. The main tools for modelling are 3Ds Max and AutoCAD. Additionally, some models or assets can be downloaded from the official asset store. To save time in rendering, this traffic prediction system will not use large-sized models. All 3D models for this research are either created originally or downloaded free from the Unity asset store(www.assetstore.unity3d.com).

Figure 22 shows four different layers[22]:

- Terrain Layer: Using Google Earth is a simple way to make terrain for this layer. It can provide high-precision images. For example, if the system needs

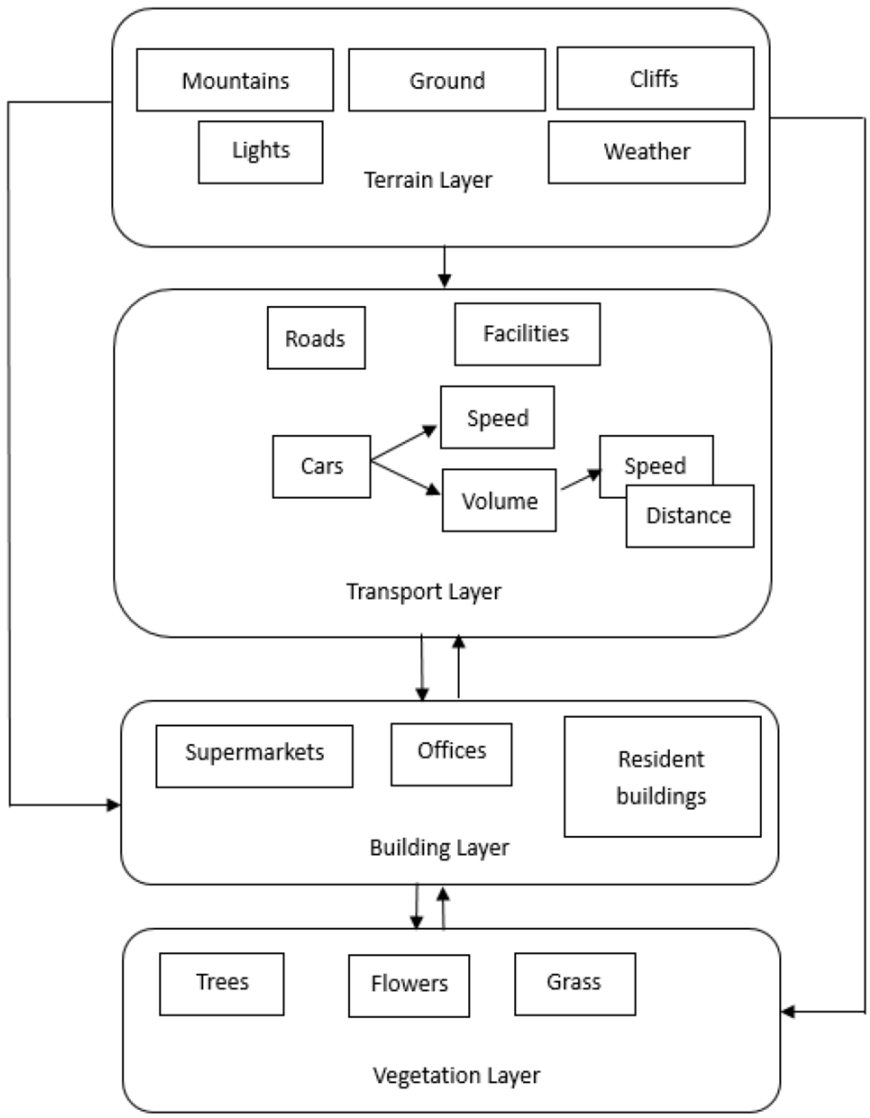

FIGURE 22: System architecture

to build the traffic network of highways from Windsor to Toronto, it can access Google Earth by using a tool named Google SketchUp in order to obtain high-resolution information about terrain. With the help of the software, the 3D terrain can be edited in 3Ds Max and saved in a .3ds format. In addition, GPS can help to get high-resolution information. In this system, terrain layer is created in Unity 3D with default materials.

- Transport Layer: After finishing the terrain layer, the next step is to build the transport layer. In order to describe traffic flow, the main work in this layer is to build roads and other important facilities according to the real location. In addition, some materials in this layer, such as texture and lighting, can be edited in 3DMax. This layer will make the traffic network more realistic. Any type of roads can be done with plugins like EasyRoad or download from the Unity 3D asset store. In this system, cars and roads are the main objects in this layer.

- Building Layer: In this layer, buildings will be added to the scene, because they are important for the final visual effect. All kinds of buildings can be edited in advance using some tools like 3Ds Max and AutoCAD. In order to gain better effects for simulation, the parameters and coordinates of objects in this layer can be obtained from Google Map to help creating models in the same location according to the real world. Several building models are added to this layer in the system.

- Vegetation Layer: The aim of this layer is to build and decorate the scene with some vegetation. This layer is the most difficult one, because the trees and grass have more details to describe. However, the vegetation layer does not have much effect in this system. If some advanced effect is required, materials and textures can be downloaded from the official asset store. Trees and grass are the main objects in the layer of this system.

All the models used in different layers downloaded from the asset shop as shown in the table. All the links are free to download.

Model	Link
Car	https://www.assetstore.unity3d.com/en/#!/content/19264
Car	https://www.assetstore.unity3d.com/en/#!/content/2619
Car	https://www.assetstore.unity3d.com/en/#!/content/2608
Car	https://www.assetstore.unity3d.com/en/#!/content/12133
Building	https://www.assetstore.unity3d.com/en/#!/content/1042
Building	https://www.assetstore.unity3d.com/en/#!/content/13925
Vegetation	https://www.assetstore.unity3d.com/en/#!/content/2936
Vegetation	https://www.assetstore.unity3d.com/en/#!/content/44

TABLE 7: Model links

When all the models were completed, they could be imported to Unity 3D. Animations, textures, scripts, and sounds can be found in an asset file in the project. They are added as new gameobjects in the process of creating a scene. Figure 23 shows the overview of the scene.

FIGURE 23: Overview of the scene

2 Scripts and Simulation

After modelling, the next step is to add scripts for components in the system. In this simulation system, the work should be done for the cars and weather in a way as illustrated in Figure 24. In the figure, Time and Weather are used to control the speed and volume for the cars in four lanes. So there will be two parameters in the system to be controlled by the scripts.

FIGURE 24: Objects need to be scripted in system

2.1 Simulation for Cars

Before adding the scripts to cars in this system, some preparation must be done. The car models should be defined as prefabs. Each time the system generate a car, the model should be obtained from prefabs so that the scripts are the same for the new ones. The change of time in the system is Time.deltatime, which is the smallest unit of time in Unity 3D. This function of time is used to make the system's frame rate independent.

- Speed Control: The speed cannot be controlled directly by scripts. The only way to show the different speeds is defining the changes of velocity along X-axis in the system. So the script will be written as follow:

transform.Translate (a*Time.deltaTime,0,0), where a is one of the parameters.

- Volume Control: Volume is another important part in traffic prediction. However, it cannot be defined by only one parameter in this system. The way to change the volume in this system is to combine the speed of the cars and the distance between cars. The cars will be generated by models in the prefabs in the same time intervals. The time should be controlled by one parameter, and the script should be written as follow:

  newCar = Instantiate (Car, transform.position, transform.rotation) as Rigidbody;

When the simulated time is different, the two parameters should also be different in order to control the speed and volume of the cars.

2.2 Simulation for Weather

In an augmented reality system, weather is a necessary component. This prediction system needs to combine multiple sources of information including weather.

The weather conditions can be generated with the particle system in Unity 3D. Figure 25 shows the interface of the particle system.

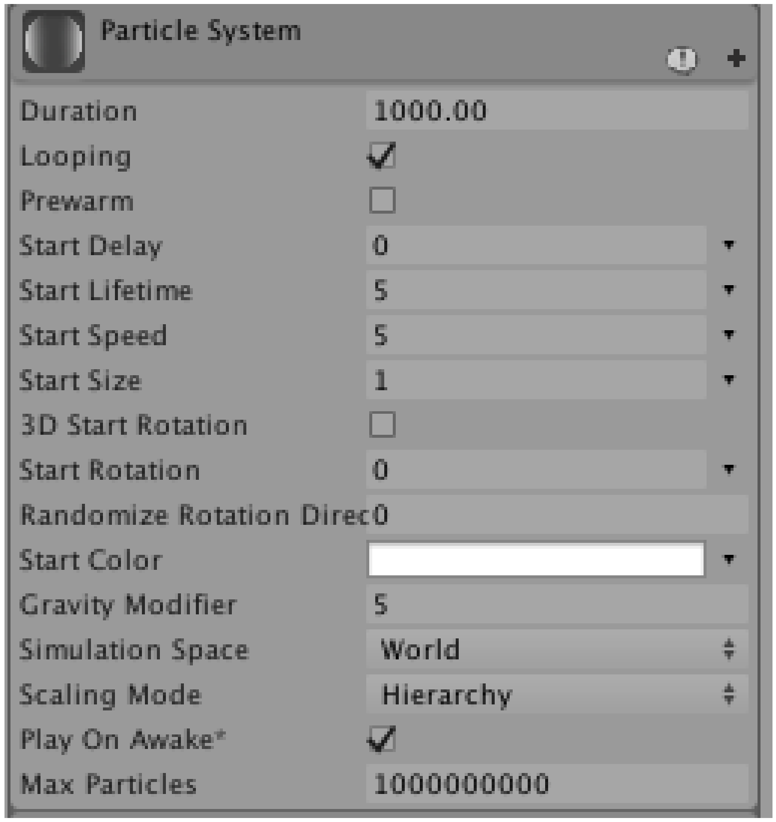

FIGURE 25: Particle system

In order to simulate different weather effects, some parameters should be changed into different values. Start Size means the initial size of particle when it is emitted. Gravity Modifier means the value of gravity when emitting particles.

- Rainy: On rainy days, the Start Size should be set to a relatively lower value, and the Gravity Modifier is relatively higher. The speed of cars should be lower than that in sunny days and the distance should be larger.

- Snowy: In comparison to rainy days, the Gravity Modifier should be set to a lower value while the Start Size should be set to a larger value; the speed of cars should be lower and the distance should be larger than those in rainy days.

- Sunny: This condition is simple because the Start Size and Gravity Modifier are set to zero. The speed of cars is the highest and the distance is lowest.

After initialization, the whole scene of traffic combined with weather conditions can be rendered.

3 System Overview

After all the preparation work, simulation may start. The prediction data was obtained from the experiments presented in Chapter VI, and the values were scaled in the system. Figure 25 shows the system running at 6:00 A.M. on a snowy day.

FIGURE 26: Overview of the simulation system

By clicking the buttons on the screen, parameters can be changed. For example, if the time changes to 6 P.M., the speed of the cars must be slower and the distance between each car must be smaller. If the weather changes to sunny, the snowflakes will disappear, and the cars will move faster than they do on snowy days.

In order to make the prediction more realistic, the models can be created to with more details, and buildings and cars may have more styles and sizes. In addition, the system should allow the specification of a particular time in the near future for simulation to show what will likely to happen in, for example, three hours.

CHAPTER VIII

Conclusion and future work

Traffic jam is the one of the most difficult problems of human's civilization, and all people can do is to reduce its level of harm. Real-time traffic flow prediction is a good way to manage traffic congestion. According to previous research, traffic can be described as a dynamic system and can be predicted by using models. In order to study the fluctuation of traffic flows, wavelet transform is useful and effective.

In this thesis, the method based on sharp varation points has made some contributions to traffic flow prediction.

- Model: A new model using wavelet transform has been developed in this thesis. This model focuses on the variation trends of traffic flows, so it can be used in all kinds of traffic conditions, including urban streets or highways. More parameters can be added in order to improve the model.

- Accuracy: Because the model makes predictions under the same condition of traffic flow, the accuracy is high. The more data that becomes available, the more conditions will be classified and the more accurate the prediction will be. With shorter time intervals in the data, more information will also contribute to higher accuracy, and the accuracy will be more consistent when time periods increase.

- Visualization: After the experiment using the new method, the results are shown in Unity 3D in a virtual world. Combined with some realistic models, the predicted traffic flow becomes visible on a screen. With more data under different conditions, more changes can be introduced to the Augmented Reality system.

In addition, more realistic models can be used to replace the original models.

In the future, improvements should focus on the AR system. Different models should be added to the system with specific details. For example, various kinds of cars and buildings will make the scenes more realistic. In addition, the locations of buildings can be made to be the same as in the real world when building the AR system. If necessary, the system can provide a view for each car, which will allow the traffic to be observed from the perspective of a driver. The control module should also allow the user to control the car at the scene. This will help users to see what will happen in different conditions. Addtionally, more data can support the changes of weather and other factors. Another way to improve model's prediction ability is to collect more data for training sets. In order to gain higher accuracy, the developed algorithms need to be further enhanced by, for example, using fuzzy logic for data classification and matching.

REFERENCES

[1] Baher Abdulhai, Himanshu Porwal, and Will Recker. Short-term traffic flow prediction using neuro-genetic algorithms. *ITS Journal-Intelligent Transportation Systems Journal*, 7(1):3–41, 2002.

[2] Haitham M Al-Deek, Sherif Ishak, and Morgan C Wang. A new short-term traffic prediction and incident detection system on i-4. *The National Academics of Sciences, Engineering, and Medicine*, 1:1–153, 2001.

[3] Richard Arnott, Andre De Palma, and Robin Lindsey. Does providing information to drivers reduce traffic congestion? *Transportation Research Part A: General*, 25(5):309–318, 1991.

[4] Ronald T Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.

[5] György Barna and Ichiro Tsuda. A new method for computing lyapunov exponents. *Physics Letters A*, 175(6):421–427, 1993.

[6] Mark Billinghurst, Adrian Clark, and Gun Lee. A survey of augmented reality. *Foundations and Trends in Human-Computer Interaction*, 8(2-3):73–272, 2015.

[7] Keith Briggs. An improved method for estimating liapunov exponents of chaotic time series. *Physics Letters A*, 151(1):27–32, 1990.

[8] Manoel Castro-Neto, Young-Seon Jeong, Myong-Kee Jeong, and Lee D Han. Online-svr for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert systems with applications*, 36(3):6164–6173, 2009.

[9] Zhang Hong-Bin, Sun Xiao-Duan, and He Yu-Long. Analysis and prediction of complex dynamical characteristics of short-term traffic flow. *Acta Physica Sinica*, 63(4):1–8, 2014.

[10] Meng Hui, Lin Bai, YanBo Li, and QiSheng Wu. Highway traffic flow nonlinear character analysis and prediction. *Mathematical Problems in Engineering*, 2015:1–7, 2014.

[11] H_S Kim, R Eykholt, and JD Salas. Nonlinear dynamics, delay times, and embedding windows. *Physica D: Nonlinear Phenomena*, 127(1):48–60, 1999.

[12] Yisheng Lv, Yanjie Duan, Wenwen Kang, Zhengxi Li, and Fei-Yue Wang. Traffic flow prediction with big data: a deep learning approach. *Intelligent Transportation Systems, IEEE Transactions on*, 16(2):865–873, 2015.

[13] Wanli Min and Laura Wynter. Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, 19(4):606–616, 2011.

[14] Luis Leon Ojeda, Alain Y Kibangou, and Carlos Canudas De Wit. Adaptive kalman filtering for multi-step ahead traffic flow prediction. In *American Control Conference (ACC), 2013*, pages 4724–4729. IEEE, 2013.

[15] Iwao Okutani and Yorgos J Stephanedes. Dynamic prediction of traffic volume through kalman filtering theory. *Transportation Research Part B: Methodological*, 18(1):1–11, 1984.

[16] Soh K Ong and Andrew Yeh Chris Nee. *Virtual and augmented reality applications in manufacturing*. Springer Science & Business Media, 2013.

[17] Brian L Smith and Michael J Demetsky. Short-term traffic flow prediction: neural network approach. *Transportation Research Record*, 1(1453):98–104, 1994.

[18] Dimitrios I Tselentis, Eleni I Vlahogianni, and Matthew G Karlaftis. Improving short-term traffic forecasts: to combine models or not to combine? *IET Intelligent Transport Systems*, 9(2):193–201, 2014.

[19] Mascha Van Der Voort, Mark Dougherty, and Susan Watson. Combining kohonen maps with arima time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies*, 4(5):307–318, 1996.

[20] Eleni I Vlahogianni, Matthew G Karlaftis, and John C Golias. Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach. *Transportation Research Part C: Emerging Technologies*, 13(3):211–234, 2005.

[21] Eleni I Vlahogianni, Matthew G Karlaftis, and John C Golias. Short-term traffic forecasting: Where we are and where were going. *Transportation Research Part C: Emerging Technologies*, 43:3–19, 2014.

[22] Sa Wang, Zhengli Mao, Changhai Zeng, Huili Gong, Shanshan Li, and Beibei Chen. A new method of virtual reality based on unity3d. In *Geoinformatics, 2010 18th International Conference on*, pages 1–5. IEEE, 2010.

[23] Billy M Williams and Lester A Hoel. Modeling and forecasting vehicular traffic flow as a seasonal stochastic time series process. Technical report, 1999.

[24] Zhang Yong and Li Shi-Gao. Analysis of scale-free characteristic on sharp variation point of traffic flow. *Acta Physica Sinica*, 63(24):1–6, 2014.

[25] Zuduo Zheng, Soyoung Ahn, Danjue Chen, and Jorge Laval. Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations. *Transportation Research Part B: Methodological*, 45(2):372–384, 2011.

[26] Zuduo Zheng, Soyoung Ahn, Danjue Chen, and Jorge Laval. Freeway traffic oscillations: microscopic analysis of formations and propagations using wavelet transform. *Transportation Research Part B: Methodological*, 45(9):1378–1388, 2011.

VITA AUCTORIS

NAME:	Minxuan Zhang
PLACE OF BIRTH:	Anqing, Anhui province, China
YEAR OF BIRTH:	1991
EDUCATION:	Fuzhou University, Information Securtiy, Math and Computer Science , Fuzhou, China, 2013
	University of Windsor, M.Sc in Computer Science, Windsor, Ontario, 2016