Determining the Location of Congestion Hot Spots in Urban Areas Using Image Analysis

Aditya Mokkapati
Louisiana State University and Agricultural and Mechanical College, amokka2@tigers.lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses
Part of the Civil and Environmental Engineering Commons

Recommended Citation
Mokkapati, Aditya, "Determining the Location of Congestion Hot Spots in Urban Areas Using Image Analysis" (2014). LSU Master's Theses. 890.
https://digitalcommons.lsu.edu/gradschool_theses/890

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master’s Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
DETERMINING THE LOCATION OF CONGESTION HOT SPOTS IN URBAN AREAS USING IMAGE ANALYSIS

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering

in

The Department of Civil and Environmental Engineering

by

Aditya Mokkapati
B.E., Osmania University, 2010
May 2014
© [2014/copyright]

[Aditya Mokkapati]

All rights reserved
ACKNOWLEDGEMENTS

I would like to give my deepest and sincerest thanks to my advisor, Dr. Chester Wilmot, for his excellent guidance, encouragement and giving me flexibility to try new ideas. I am very grateful to Dr. Ravindra Gudishala who had given me immense support and the right guidance at every moment during my Master’s degree. The three of us were a great team and it was a pleasure working together. I would like to thank to Dr. Sherief Ishak and Ms. Guo Beibei for serving on my committee.

I thank all my room mates and friends at LSU for taking good care of me, helping me and for all the good times we had together. I also want to thank all my colleagues and friends, Jose Rodriguez, Osama Thabet, Raju Thapa, Eva Rodriguez, Julius Codjoe, and Wan Li for their help in my thesis and support and encouragement during my study at LSU.

Finally, I would like to thank my parents, brother, grandmother and grandfather for all their love and support.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. iii

LIST OF TABLES .......................................................................................................................... v

LIST OF FIGURES ....................................................................................................................... vi

ABSTRACT ..................................................................................................................................... viii

CHAPTER 1. INTRODUCTION ........................................................................................................ 1
  1.1. Problem Statement ............................................................................................................. 2
  1.2. Research objectives .......................................................................................................... 3

CHAPTER 2. LITERATURE REVIEW ............................................................................................ 4

CHAPTER 3. STUDY AREA AND DEVELOPMENT OF METHODOLOGY ......................... 9
  3.1. Study Area ....................................................................................................................... 9
  3.2. Data Collection ................................................................................................................ 10
    3.2.1. Source of data .......................................................................................................... 10
    3.2.2. Amount of data ........................................................................................................ 12
    3.2.3. Capturing and downloading data ............................................................................. 13
  3.3. Performing Image Analysis to Identify Congested Hotspots ......................................... 16
    3.3.1. Image analysis to convert color coded data to number format ............................. 16
    3.3.2. Framework of analysis for 15 minute interval ....................................................... 19
    3.3.3. Framework of analysis beyond 15 minute interval .............................................. 20
    3.3.4. Analysis to produce a congestion index ................................................................. 21
    3.3.5. Methods to represent congestion on a map ......................................................... 24

CHAPTER 4. APPLICATION ......................................................................................................... 29
  4.1. Analysis on one week data .............................................................................................. 29
    4.1.1. Analysis for entire time period between 7 am and 10 pm. .................................. 29
    4.1.2. Analysis for morning peak period and evening peak period. ............................ 32
  4.2. Analysis on one month data ........................................................................................... 35

CHAPTER 5. VALIDATION OF THE METHODOLOGY ......................................................... 37

CHAPTER 6. GRAPHICAL USER INTERFACE ....................................................................... 41

CHAPTER 7. CONCLUSIONS ..................................................................................................... 43

REFERENCES ............................................................................................................................ 45

APPENDIX : MATLAB CODE .................................................................................................. 47

VITA ............................................................................................................................................ 51
LIST OF TABLES

TABLE 1 Costs and Concerns Associated with Different Traffic Measuring Technologies (14, 15) .................................................................................................................. 6

TABLE 2 Analysis of 15 min. interval matrices ......................................................................................... 20

TABLE 3 Analysis of 30 min. interval matrices .......................................................................................... 21

TABLE 4 Congestion index ranges for hotspots as per period of analysis .................................................. 23

TABLE 5 Recommended resolution and scaling factors................................................................................. 26
LIST OF FIGURES

FIGURE 1 Geographical boundary of the study area (3) ................................................................. 9

FIGURE 2 Image of area of interest in Baton Rouge, LA on Thursday at 5:00 PM
depicting three colors on the roads. Snippets show pixels (3) ..................................................... 11

FIGURE 3 Layout of programs and icons before GhostMouse software is run ......................... 15

FIGURE 4 Image toolbox window (left). Inspect pixel value window (right) showing
index values along with RGB values of pixels ................................................................. 18

FIGURE 5 Example of Google image with historic traffic information (left) and matrix
populated with 0 and 1 based on the color of interest .......................................................... 19

FIGURE 6 Coloring surrounding pixels in proportion to the congestion index value of
pixels on the road .............................................................................................................. 24

FIGURE 7 Congestion on Road A east bound is represented using solid red dots .. 25

FIGURE 8 Colormap showing different temperatures in the infrared image (24) .............. 26

FIGURE 9 Colormap depicting different altitudes (25) ............................................................ 27

FIGURE 10 Built-in colormaps in MATLAB (26) ................................................................................ 28

FIGURE 11 Custom colormap .................................................................................................. 28

FIGURE 12 Frequency analysis for 15 min congestion for entire day (7 am to 10 pm)
on 09/03/13 ................................................................................................................... 30

FIGURE 13 Congestion Index and buffer analyses of Congestion Index for entire day
(7 am to 10 pm) for data collected on 09/03/13 ................................................................. 31

FIGURE 14 Congestion Index and buffer analyses for entire day (7 am to 10 pm) for
data collected on 09/03/13 showing roads around downtown and Louisiana State
University in Baton Rouge ................................................................................................. 32

FIGURE 15 Congestion Index and buffer analysis for morning peak period (7 am to 9 am)
for data collected on 09/03/13 ......................................................................................... 33

FIGURE 16 Congestion Index and buffer analysis for evening peak period
(4 pm to 6 pm) for data collected on 09/03/13 ................................................................. 34

FIGURE 17 Congestion Index and buffer analysis for entire month with 3 samples of
data (7 am to 10 pm) ......................................................................................................... 36
FIGURE 18 (a) I-12 Corridor in Baton Rouge that was analyzed along with the Bluetooth detectors that were used, (b) shows part of the analyzed portion of the highlighted section, at 5:00 pm, in figure (a)................................................................................................................................................... 38

FIGURE 19 Bluetooth speeds over I-12 between Essen Ln. and Airline Hwy. between the days 7/22/13 -7/28/13 ................................................................................................................................................... 39

FIGURE 20 Graphical user interface of image analysis........................................................................................................ 42
ABSTRACT

The identification of congestion hot spots plays an important role in congestion monitoring of highway and arterial systems. Traditionally, vehicle detection systems or probe vehicles are used to measure travel time and identify congestion hot spots. This approach is time-consuming and resource intensive. With the advent of new technologies like crowdsourcing, travel time can be collected more cost-effectively using fewer resources. The research presented in this paper develops a new methodology to identify congestion hot spots in an urban area using freely available historical speed data gathered by Google and published as a traffic layer in Google maps. The methodology uses image analysis techniques like image segmentation and pixel analysis to formulate a sequence of steps that identify areas consistently experiencing congestion. As a case study, the developed methodology was applied in the Baton Rouge metropolitan area for the months of August and September 2013 and then compared with speeds and travel time obtained from Bluetooth signal detection devices deployed in the same region. The results showed that the methodology developed in this thesis could identify congestion hotspots that are consistent with data acquired through Bluetooth-driven measurement of travel time. The methodology can be used for areas where Google traffic maps (or other maps with color coded traffic data) are freely available as a quick and inexpensive method of identifying congestion hotspots. Since the method also quantifies the degree of congestion, the method can be used to prioritize corridors for applying congestion mitigation measures. The research presented in this paper is of potential benefit to traffic engineers, State DOT’s and other policy makers who are interested in identifying congestion hotspots in urban areas.
CHAPTER 1. INTRODUCTION

The knowledge about location of congested hotspots in an urban area serves as important input in prioritization of corridors for application of congestion mitigation strategies and road expansion plans. Traditionally, traffic engineers and State Departments of Transportation (DOTs) have used a range of travel time measuring devices like inductive loops, probe vehicles, automatic license plate recognition (ALAR) system, Radio-frequency identification (RFID) tag reader system, geo-located cell phones, etc. to identify congested hotspots in urban areas. With the advent of new travel time measurement technologies like Bluetooth and crowdsourcing the effort and resources required to measure travel time and archive historical travel time data are becoming cheaper and easily accessible.

With the entry of private enterprises like Google, INRIX and NAVTEQ into the travel-time-acquisition business, it is now possible to purchase travel time data integrated with geographical information systems (GIS) data in real time. In addition, it is also possible to buy archived historical travel time data at competitive prices. This new development opens new avenues for traffic engineers, State DOT’s, policy makers and other stakeholders to acquire and run analysis on traffic data by simply initiating a contract with any of the commercial vendors in the market.

Vendors such as INRIX and NAVTEQ gather data from various sources which range from crowdsourcing technology from millions of users who use their mobile apps for navigation, delivery vans, long-haul trucks, navigation systems in consumer vehicles, to traffic and public safety agencies etc. (1, 2). This data can be bought from them. Google also has been collecting data in a similar fashion for the past few years, and they display it on Google Maps in terms of live and average conditions. However, the data in this case is represented only in a visual manner.
with colors unlike the other data sources which provide information in numeric values (3). On the positive side, this data is free for public use.

1.1. Problem Statement

As mentioned in the previous section, transportation professionals currently use a variety of traffic monitoring devices to collect traffic data to determine congestion characteristics on roads. However, these methods are resource intensive and/or time consuming to identify congestion areas using these technologies for the following reasons:

1. In order to identify congested roads in an urban area data has to be collected for the entire urban area. Therefore, a network of these traffic monitoring devices should be deployed over the entire area. However, capital investment for such large scale projects would run into several hundreds of thousands of dollars considering the cost and number of devices that will be required.

2. City wide studies could be done with smaller investments as well. A small set of devices could be bought and sampling studies could be performed using this small set, on each of corridors of the city. But the problem here is that this strategy is very time consuming since the same set of devices needs to be uninstalled and redeployed over each road for a considerable time, to get a confident sample.

In addition to cost and time, many of the traditional traffic monitoring devices have limited capabilities.

This research satisfies the need to identify congested corridors in urban areas in a very inexpensive, quick and reliable way. This is performed by harvesting traffic data from Google Maps which Google has been gathering for the past few years in a manner similar to INRIX and NAVTEQ. Google represents this data only in a visual manner with colors. This research
provides a method to capture this color coded data and identify congested corridors that are frequently and severely congested. This research is performed not only to identify corridors but also to develop tools that allow one to tap into freely available resources that help in identifying traffic parameters of interest.

1.2. Research objectives

The objectives of this research are as follows:

- Develop a mechanism to extract color coded traffic data from Google Maps.
- Develop a new congestion index for Google Maps, to identify corridors by the degree of congestion experienced.
- Represent the degree of congestion, on a map using an indicator. The size of the indicator must be proportional to the degree of congestion as in desire lines for O-D data.
- Identify hot spots in Baton Rouge, LA where further detailed analysis could be performed using conventional traffic monitoring devices.
- Develop a graphical user interface to process the data and run the analysis easily.
CHAPTER 2. LITERATURE REVIEW

There have been several studies that investigated the possibility of applying image analysis techniques to extract pertinent information from GIS maps, satellite images, raster maps, and videos acquired through recording devices like Autoscope. Zhang, Sheng and Li (4) performed adaptive image color segmentation analysis to identify traffic signs from images. For this two techniques were used, 1) CPT (Central Projection Transformation) and 2) PNN (Probabilistic neural networks). CPT technique converts 2-D images to 1-D vectors. This vector is used in pattern recognition using PNN to identify the presence and type of traffic signs from images of road taken vehicle borne imaging systems and video cameras. Likewise, researchers Song and Civco (5) used support vector machine and Image segmentation analysis techniques to derive accurate road geometry information from images obtained from remote sensing.

In the past, several studies have investigated bottlenecks and their congestion characteristics. Chen, et al. (6) developed an algorithm that identified the location of bottlenecks along with their activation and deactivation time on the basis of speed differentials between upstream and downstream detector locations. A similar system was developed by Zhang and Levinson which is based on traffic occupancy and counts (7). Bertini et al. (8) developed an approach that detects the location and duration of active bottlenecks, as well as bottleneck related congestion upstream using volume and speed data. Researchers in Portland developed a tool to identify recurrent freeway bottlenecks, using historical data such as flow rate of traffic, speed of vehicles and proportion of trucks and passenger cars (9, 10). In all of the above studies, traffic data was acquired from loop detectors. Sisiopiku (11) performed a congestion analysis in Michigan for which she observed traffic condition through a series of field trips, surveyed commuters and analyzed historical traffic and incident records. These approaches aided in
identifying bottlenecks and its related congestion. However, their data sources were primarily traffic sensors on roads or surveys.

In addition to loop detectors, transportation professionals have used several techniques in the past to measure travel time on arterials and freeways to identify congested roads, such as: probe vehicles, automatic license plate recognition (ALAR) system, Radio-frequency identification (RFID) tag reader system, geo-located cell phones, and traffic information by INRIX and Bluetooth monitoring devices.

The above mentioned technologies have limited usability. Conventional loop detectors capture vehicle counts, occupancy and spot speed but cannot calculate travel times. Data from GPS units on board fleet vehicles are only available on limited routes that they travel frequently. Moreover, these devices are on a limited number of truck fleets. In addition, speeds are limited on some trucks since governors are installed in them (12). License plate readers have privacy issues, since they capture images of individual license plates. In addition, they are expensive (13). Toll tag readers are a good source, but they are only available on toll roads. Geo-located cell phones is a technology where approximate location of cell-phone users is determined by triangulating cell-phone signals between three cell phone towers. While the trip ends are known for this data (Ex: Home and Walmart on College Dr.); it is not possible to assign trips to a particular route. Mobile network operators initially provided data but have since refused to provide data for analysis in most cases. They are afraid they would lose customers if customers thought their privacy was being compromised. A comparison between these technologies is present in Table -1.
### TABLE 1 Costs and Concerns Associated with Different Traffic Measuring Technologies (14, 15)

<table>
<thead>
<tr>
<th>Technology</th>
<th>Cost</th>
<th>Privacy</th>
<th>Travel Time /Speed Accuracy</th>
<th>Freeways</th>
<th>Arterials</th>
<th>Non-location specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluetooth</td>
<td>$2,000-$4,000 per mile</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Traffic information by INRIX</td>
<td>Arterials - $200/mile/year Freeways - $800/mile/year</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>GPS Fleets</td>
<td>$500-$1,000/mile/year</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Cell Phone Location</td>
<td>$500-$1,000/mile/year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loop detectors</td>
<td>$7,500-$20,000 per mile</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Toll Tags</td>
<td>$20,000 per mile</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
</tbody>
</table>

Bluetooth and INRIX provide reliable information, however they are not viable options since large scale deployment over an entire city for prolonged periods of time render costs to be several hundreds of thousands of dollars.

Google maps provide a wide variety of information. They provide the ability to find addresses, driving directions and duration, traffic information, phone numbers of businesses,
reviews of restaurants, etc. Their services are used by several government agencies and private enterprises for various purposes such as, situation awareness for disaster response, data visualization, in a taxi enterprise to connect drivers to passengers, decision making for economic development, etc. (16).

In the field of transportation, Google Maps are often used as a base map by various agencies to display information such as, travel advisories, traffic information, transit route and schedule information, incident reports, variable message signs, etc. Iowa’s Dot provides detailed statewide real-time traffic information maps laid over a base map from Google Maps. The data includes: current traffic speeds, traffic camera footage, electronic roadway sign messages, lane closure, snow conditions, road maintenance activities, etc. (17). Similarly, Delaware DOT provides interactive traffic maps for their state which lets users view average speeds, volumes, occupancy measurements along with alerts. They use Google Maps to help them display this information (18). New Jersey displays incident data, roadwork, traffic cameras positioned along the highway, etc. on Google Maps (19). Caltrans provides Google traffic data along with lane closures, message signs, chain controls, etc. on their website using Google Maps (20). Researchers in Washington State developed a traffic information system called GATI (Google maps based arterial traffic information) which allows them to display locally collected traffic information and query data of interest on Google Maps (21). Similarly, in Jakarta, Indonesia a researcher extracted traffic updates posted on Twitter by the Traffic Management Center and converted it into visual data and presented it on Google Maps. This data was presented in three different colors to depict 3 different levels of traffic conditions (22). DOTs, other agencies and researchers have been using Google Maps, but only to display the information they have collected using their own data sources.
Based on the review, no research was identified that developed a methodology to identify congestion hot spots using Google traffic layer images. Thus, to the author’s knowledge it was a worthwhile endeavor to develop a methodology using image analysis techniques to identify congestion hotspots using freely available historical traffic data on Google Maps.
CHAPTER 3. STUDY AREA AND DEVELOPMENT OF METHODOLOGY

3.1. Study Area

The study area for this thesis is limited to Baton Rouge metropolitan area, Louisiana. However, this technique can be employed in any area. Google provides traffic data for highways and arterials. Hence the data collected was only for these types of roads and no local roads were included. FIGURE 1 shows the geographical area included in the study.

FIGURE 1 Geographical boundary of the study area (3)
In the following sections it is described how traffic speed information is stored in Google Map’s traffic layer, the method used to access and download historical traffic data from Google Maps in the form of images, the steps involved in analyzing these images to extract information required to identify congestion hotspots, description of a new congestion index and, finally, the technique used in identifying and displaying congestion hotspots.

3.2. Data Collection

3.2.1. Source of data

Google Maps has a popular layer called the traffic layer and this layer is capable of displaying both current traffic conditions on roads and historical traffic data. The historical traffic data can be accessed by opening Google Maps in a browser like Google Chrome, with the traffic layer switched on and ‘Traffic by day and time’ option chosen in the box on the bottom left corner of the browser window. The historical data represents traffic conditions on different days of the week and the data is available at 15 minute time intervals. It is color coded using five different colors to represent varying traffic conditions. For freeways the colors and their respective speeds are as shown below.

- Green: more than 50 mph (80 kmph).
- Yellow: 25 - 50 mph (40 - 80 kmph).
- Red: less than 25 mph (40 kmph).
- Red/Black: very slow, stop-and-go traffic.
- Gray: no data currently available.

For arterials roads, these speeds do not apply. The colors only give an indication of the severity of the traffic. Green means that traffic conditions are good, yellow means fair, and red or red/black means poor traffic conditions (23).
The Red/Black color is often seen in live traffic data. However, the chances for it to appear in 15-minute interval historical data, which is an average of several days’ data, are very slim on either road types. Therefore, effectively, there are 4 colors that appear in the historical traffic data maps on Google, i.e. green, yellow, red and grey. Among these red is of importance, since it signifies poor traffic conditions on arterials and slow speeds on highways. FIGURE 2 displays all the 4 colors in a map with historical traffic data at 4:45 p.m. on a Thursday, in Baton Rouge. It is important to note that traffic conditions of each time period was assumed to represent the average traffic conditions of the succeeding 15 min. Accordingly, in this case the 4:45 p.m. image represents the average traffic conditions between 4:45 p.m. and 5:00 p.m.

FIGURE 2 Image of area of interest in Baton Rouge, LA on Thursday at 5:00 PM depicting three colors on the roads. Snippets show pixels (3).
All of the maps displaying traffic conditions are freely available, but there is no direct way to access the traffic data that is used in building the maps. Since Google Maps displays this historical data on its website, images of past traffic conditions are available. These images are made up of thousands of pixels as seen in FIGURE 2. The pixels which are red act as data points that represent congestion on the roads. Therefore, this pixel data can be used to identify congested corridors. In order to use this pixel data from images, image analysis techniques must be employed.

3.2.2. Amount of data

To identify corridors on which congestion occurs frequently, historical data from all days in the week at all time intervals are needed. This is because congestion can occur on corridors at different times on the same day and/or over different days at the same time. However, it is common knowledge and also was confirmed on Google Maps that in Baton Rouge, congestion generally does not occur before 7 am or after 10 pm on any day. So capturing images can be limited to this time frame. As mentioned earlier, this data is available at 15 min interval for every day of the week. Hence for one day, 60 images (15 hours * 4 images/hour) will be acquired for the city. Similarly, 60 images are acquired for each day of the week, making a total of 420 images for the week. A two month observation showed that Google’s historical data maps are updated in a cycle of approximately between a week and 10 days. To capture historical data over a longer period, the author downloaded new historical data which was captured every 7-10 days over a period of a month (i.e. 3-4 sets of historical data were downloaded to represent conditions over a month).

Live traffic data was observed to be consistent at different zoom levels. However, the zoom level impacted the way in which historical traffic data was represented. At any particular
15-minute period in a day, the freeways on Google Maps more or less showed the same color on road segments at different zoom levels. But in the case of arterials, colors were observed to change by varying zoom levels. However, this phenomenon was observed only at isolated locations. Effort was made to understand this phenomenon by contacting Google Map’s technical team but no reply was received from them. Consequently, it is proposed to collect images at a scale of 1: 178,816. This scale is chosen since it covers the entire Baton Rouge Metropolitan area in one image with traffic color data on all arterials and freeways visible.

3.2.3. Capturing and downloading data

Before image analyses were performed, all the images required to perform an analysis are to be downloaded from the Internet by taking screenshots of them from an Internet browser like Google Chrome. To collect this data on a long term basis, i.e. to take screenshots and store the 420 image manually in a systematic order every week or 10 days, would be a very tedious and error-prone task. Therefore, an automatic procedure needs to be developed where mouse movements and key strokes were to be recorded using software called Ghost Mouse which is linked to a standalone screenshot capturing software called FastStone Capture.

The GhostMouse software in which mouse movements and key stokes are to be recorded and coded to capture Google Maps images should have the computer environment set up in a specific way. This means the operating system of the computer, location of icons, windows on the screen, size and relative position of additional monitor should be set up before the software is run, to automatically capture images. The entire setup is explained in detail below and with the layout shown in FIGURE 3:

1. There should be two monitors

   a. Monitor-1 present left side (resolution 1600 x 900)
b. Monitor-2 present right side (resolution 1280 x 1024)

2. In monitor-1 the following setup should be done.
   a. The FastStone Capture software’s window should be present in the right top corner.
   b. MS Excel sheet with path names to store the images in chronological order should be maximized.
   c. The Ghost Mouse window should be present in the yellow box over the excel sheet.
   d. The program icons of FastStone Capture and MS Excel should be pinned to the taskbar right next to the start menu.

3. In monitor-2 the following setup should be done.
   a. Google Maps should be opened by going to https://maps.google.com in Google Chrome browser. The browser window should be maximized.
   b. In the search bar ‘Baton Rouge, LA’ is typed and the left information panel is hid by clicking on the hide panel button.
   c. Google Maps is then set to ‘Map’ style, if it is not already set.
   d. The traffic layer is switched on which pops up a small traffic options box in the left bottom corner.
   e. The ‘Traffic at day and time’ option is selected in from the option. The time and day below it is set to 7 am, Monday.

Note: Approximately half of the above-mentioned arrangements need to be made only the first time the program is set up on a computer; the other half need to be set each time a new sample is collected.
FIGURE 3 Layout of programs and icons before GhostMouse software is run.
Once the above mentioned setup is performed, repetitive actions of the mouse and keyboard strokes performed by the user are captured onto a file by the GhostMouse software. This task needs to be performed only once. When the user requires a new set of data to be downloaded this coded GhostMouse file can be run with after setting up the computer as shown in FIGURE 3. Mouse movements in combination with the keyboard strokes interact with the windows, icons, programs as recorded in the initial setup and data is saved onto the hard drive. Stable internet connection typically loads traffic data on maps within 1 second. However, the pre-coded GhostMouse file was set to accommodate any internet lag (up to 3 sec) that could occur due to increased internet latency or decreased bandwidth.

All captured images will be stored in 7 different folders, one for each day. They will be in GIF (Graphics Interchange Format) format with 256 colors with a size of 1280x1024 pixels. Automating this process will make data collection reliable and error free.

3.3. Performing Image Analysis to Identify Congested Hotspots

3.3.1. Image analysis to convert color coded data to number format

The rationale behind the image analysis technique lies in exploiting the pattern in which an image’s digital representation is stored in a computer memory. There are different formats that can be used to store an image in computer memory. The GIF format was chosen in this case because image data is ‘indexed’ rather than presented in ‘true color’ as in the very widely used JPEG (Joint Photographic Experts Group) format. An image in GIF format has an array that is the size of the image (1024x1280, in this case), where each cell of the array has an index value that is between 0 and 255. Each of the indices refers to colors in a predefined palette of 256 colors to display the color in that cell. This makes GIF format more flexible to use compared to a true color image has color values for each pixel stored as RGB (Red, Green, Blue) triplets.
There are several software packages that are commercially available and can be used to perform image analysis. However, MATLAB was chosen to perform image analysis because the basic data structure in MATLAB is the array, an ordered set of real or complex elements. Moreover, MATLAB stores most images (such as GIF) as two-dimensional arrays or matrices in which each element of the matrix corresponds to a single pixel in the displayed image. Furthermore, MATLAB provides several built-in functions that can be used to write custom scripts to perform image analysis.

In the first step images are imported into MATLAB. MATLAB stores the image as a 1280 by 1024 matrix with each cell of the matrix storing information related to a single corresponding pixel on the image. This convention makes working with images in MATLAB similar to working with any other type of matrix data, and makes the full power of MATLAB available for image processing applications.

In the second step, after importing each image into MATLAB one needs to use MATLAB’s matrix processing capabilities to identify cells corresponding to pixels of interest in identifying congestion hotspots. It was known a priori that Google Maps color code congestion on highways and arterials with a red color. However, in the bottom snippet in FIGURE 2, it can be seen that congested corridors are not represented by one red color but a range of red colors. Thus, it is required to identify this range of pixels that represent congestion.

In order to identify the entire range of congestion representing red colors, index values and RGB (Red, Green, Blue) values should be used together. In addition to each color being represented by an index value as described before, they are also described in terms of the three base colors RGB each of whose values ranges from 0 to 1 (e.g. black has an RGB of [0,0,0], white has an RGB of [1,1,1] and parrot green has an RGB around [0.16,0.68,0.01]). However,
this information is not directly available. MATLAB’s image processing toolbox has a tool called the “inspect pixel values” tool which can be used to obtain this background information. This tool provides a means to manually inspect each pixel and provides 4 types of information, 1) lets the user know if a particular red pixel is present on road or not 2) the color of the pixel in a visual form 3) its associated index value and 4) its associated RGB values as shown in the pair of windows in FIGURE 4.

FIGURE 4 Image toolbox window (left). Inspect pixel value window (right) showing index values along with RGB values of pixels.

Using the “inspect pixel values” tool the author inspected several maps and came up with the best suitable range for RGB values for red pixels that represent congestion on roads. They happen to be R>0.45, G<0.38 and B<0.32. Using MATLAB’s programming interface index values of colors associated with the above RGB value ranges can be isolated to form a list. This list will contain congestion-representing index values. Using a conditional function, the matrix
cells that contain these congestion-representing index values can then be populated with a value of one (1) and the rest of the cells with zero (0). So if congestion is present on a road, a value 1 is inserted into all the cells in the matrix that represent congested portions of the road. FIGURE 5 shows a pictorial representation of this concept with an initial image along with its analytical matrix. The MATLAB code for this presented in APPENDIX A.

![Matrix representation of congestion](image)

FIGURE 5 Example of Google image with historic traffic information (left) and matrix populated with 0 and 1 based on the color of interest.

It is important to note that the congestion-representing index values change from image to image. Therefore, the analysis to convert an image to numbers should be done after acquiring these index values in each case using the fixed RGB range mentioned above.

3.3.2. Framework of analysis for 15 minute interval

The process of populating matrix cells based on congestion needs to be repeated for each 15-min. time period, every day for a single week. If one then sums up these matrices over a seven-day period a new set of 60 15-min. matrices are produced. These represent aggregate traffic condition over a week (between 7 am to 10 pm). A road experiencing congestion throughout the week in a particular time interval would get a value of 7 in the matrix for that period.
time interval. The process of summing up the matrices for each time interval is shown schematically in TABLE 2.

TABLE 2 Analysis of 15 min. interval matrices

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
<th>Final matrix for 15 min interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>= Matrix for 15 min interval</td>
</tr>
<tr>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>= Matrix for 7:15 am</td>
</tr>
<tr>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>= Matrix for 7:30 am</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

The process of analyzing images in order to identify congestion becomes tedious if one were to analyze one month’s worth of images manually. Recognizing this aspect, a script was written in MATLAB that automatically performs steps required to analyze images. The MATLAB code for this and the following sections of the analysis is presented in APPENDIX A 3.3.3. Framework of analysis beyond 15 minute interval

The results obtained from 15 min interval analysis might not be enough to identify severely congested corridors. This is because, it examines if corridors were congested for only 15 min. It would be useful to analyze if corridors are congested for 30 min., 45 min. and 1 hr. intervals, since duration of congestion is a factor that characterizes congestion problem on a corridor. The results from these might give a totally different perspective of congestion trends in the city, from that of 15 min interval analysis.

For the above purpose, the presence of congestion in consecutive 15 min. intervals is to be analyzed. For instance, in case of a 30 min. interval analysis, 7:00 am and 7:15 am data on Monday, analysis is to be performed for each cell of the matrix to check if they have congestion in both intervals. The cells which have congestion in both the images will be presented, in a new matrix named ‘7:00 to 7:30 am matrix’, with the value 1. This procedure will be performed for all contiguous pairs of 30 min interval matrices for all 7 days of the week. These matrices should
then be summed up over the entire week to produce a new set of 59 30-min. matrices. The resultant set of 59 matrices would represent aggregate 30 min. congestion in the entire week. There would be 59 matrices since there were 59 consecutive 30 min. intervals between 7 am and 10 pm. A schematic representation of 30 min. interval analysis is shown in TABLE 3 for a sample pixel where a red highlighted cell indicates congestion in the matrix cell represented by the pixel for the 15-min time period shown.

**TABLE 3 Analysis of 30 min. interval matrices**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Final matrix for 30 min interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>Tuesday</td>
<td>Wednesday</td>
<td>Thursday</td>
<td>Friday</td>
<td>Saturday</td>
<td>Sunday</td>
</tr>
<tr>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>7:00 AM</td>
<td>= Matrix for 7:00 AM to 7:15 AM</td>
</tr>
<tr>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>7:15 AM</td>
<td>= Freq. of congestion=2 Matrix for 7:15 AM to 7:30 AM</td>
</tr>
<tr>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>7:30 AM</td>
<td>= Freq. of congestion=3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similarly, this procedure is performed for 45 min and 1 hr. intervals to obtain their respective final matrices (58 and 57 in number respectively). The range of values in the final matrix in all four cases would range from a minimum of 0 to a maximum of 7.

3.3.4. Analysis to produce a congestion index

Duration of congestion is one of the factors that characterize the severity of a congestion problem. For instance, if a road is congested for 15 min., that road is not as affected by congestion as the one that is continuously congested for 1 hr. Thus, the notion is proposed that smaller intervals have a smaller effect on congestion than longer intervals of congestion. Accordingly, weights of 1/10, 2/10, 3/10 and 4/10 will be assigned to the 15 min., 30min., 45 min., and 1 hr. time interval values, respectively. These weights were chosen such that they add up to 1 and can be allocated in a linear fashion to each of the 4 analysis time period according to the number of 15min each has.
Secondly, the research suggests that frequency of congestion also characterizes the severity of the congestion problem. To quantify this, two parameters are required in each of the 4 time intervals. The first is the total number of final matrices in each time interval (denominator) and the second is the weekly average daily congestion frequency in each time interval of each pixel (numerator).

Using the two congestion characteristics mentioned above, a congestion index can be calculated for each pixel on roads using the following equation.

Congestion Index for pixel \([X, Y]\) = \(\frac{1}{10} \left( \frac{A_1}{A_2} \right) + \frac{2}{10} \left( \frac{B_1}{B_2} \right) + \frac{3}{10} \left( \frac{C_1}{C_2} \right) + \frac{4}{10} \left( \frac{D_1}{D_2} \right)\)

where,

\(A_1 = \) Weekly average daily 15 min congestion frequency = \(\frac{\sum_{T=7\text{ am}}^{T=10\text{ pm}} \sum_{D=\text{Monday}}^{D=\text{Sunday}} T_{15D}}{7}\)

\(A_2 = \) Total no. of final 15 min. interval matrices in a week.

\(B_1 = \) Weekly average daily 30 min congestion frequency = \(\frac{\sum_{T=7\text{ am}}^{T=10\text{ pm}} \sum_{D=\text{Monday}}^{D=\text{Sunday}} T_{30D}}{7}\)

\(B_2 = \) Total no. of final 30 min. interval matrices in a week.

\(C_1 = \) Weekly average daily 45 min congestion frequency = \(\frac{\sum_{T=7\text{ am}}^{T=10\text{ pm}} \sum_{D=\text{Monday}}^{D=\text{Sunday}} T_{45D}}{7}\)

\(C_2 = \) Total no. of final 45 min. interval matrices in a week.

\(D_1 = \) Weekly average daily 60 min congestion frequency = \(\frac{\sum_{T=7\text{ am}}^{T=10\text{ pm}} \sum_{D=\text{Monday}}^{D=\text{Sunday}} T_{60D}}{7}\)

\(D_2 = \) Total no. of final 1hr. interval matrices in a week.

\([X, Y] = \) The coordinates of the pixel.

\(T_{15D} = 15\text{ min matrix}\)

\(T_{30D} = 30\text{ min matrix}\)

… and so on.
Since the time between 7 am to 10 pm is considered for analysis, the denominators will be fixed. Their values will be A2=60, B2=59, C2=58 and D2=57. Each of four pairs ((A1, A2), (B1, B2) (C1, C2) and (D1, D2)) represent frequency characteristic and the 
\[
\frac{\text{Number of 15 min. intervals}}{10}
\]
represents the duration characteristic of congestion.

The computation of congestion index CI should be performed for every pixel, which will result in each one of them having a value between 0 and 1. Pixels with congestion index (CI) of zero have no congestion while positive values of CI indicate the presence of congestion. For the congestion index of a cell to attain the value of 1, the road must be congested during all time periods between 7 am and 10 pm for a week. Obviously, no road would experience that level of congestion, so CI values computed were always less than 1.

The CI formula does not represent congestion problems on a road section at a specific point in time of a day or a week, but represents a general level of congestion experienced according to the time period analyzed; the minimum is at least a week. However, with regard to a specific time period in a day analysis can be performed for morning peak, evening peak or special event time etc.

Based on CI, the CI ranges shown in the TABLE 4 are recommended for categorization as congestion hotspots. The categorization was developed based on observed relationship between CI and recurrent congestion over one month’s time.

**TABLE 4** Congestion index ranges for hotspots as per period of analysis

<table>
<thead>
<tr>
<th>Time period of analysis</th>
<th>Congestion Index of Hotspots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning peak period/Evening peak period – 2hr.</td>
<td>0.26 to 1</td>
</tr>
<tr>
<td>Entire day (7am to 10 pm) -15 hr.</td>
<td>0.039 to 1</td>
</tr>
</tbody>
</table>
3.3.5. Methods to represent congestion on a map

3.3.5.1. Representation of congestion index using buffer analysis

In the previous section, a pixel-level congestion index was proposed. The congestion index can be used to represent congestion in a visual manner. The idea is to have solid red circles on the map proportional in size to the congestion index values obtained for each pixel in the previous step.

This representation is performed by populating the cells around a pixel with the same congestion index value to a radius proportional to the congestion index of that pixel. Then, all these pixels with congestion index are colored red in MATLAB, thus forming a red circle proportional to the level of congestion experienced. This is shown schematically in FIGURE 6 and FIGURE 7 where eastbound traffic on road A experiences congestion over its entire length but with the most prevalent congestion occurring in advance of the intersection with road B.
FIGURE 7 Congestion on Road A east bound is represented using solid red dots. In this case two aspects can be varied manually when representing data, namely:

1. Resolution factor: CI values for pixels are usually very small numbers and typically occupy a large number of decimal places. If the resolution factor is too coarse, only large values of CI will be recognized and it will be difficult to distinguish between CI values which are very similar, being only different in the second or third decimal place (e.g. 0.239 and 0.234). This will render the buffer analysis to provide imprecise and very aggregate results. If a fine resolution value (e.g. by considering six decimal places of CI values) is used, the differences between congested areas are more clearly distinguished. However, trials showed the computing time to exponentially increase with an increase in each decimal place and the computer could run out of memory as well. Therefore, the person who runs the program should be prudent when selecting the resolution value to represent data. A resolution factor of $4 \times 10^n$ is to be used, where $n$ is the number of decimal places that one prefers to use.
2. Scaling factor: The scale factor is used to increase/decrease the size of the solid red circles that represent the CI values of the pixels. Larger scaling factor will produce larger red circles and smaller scaling factors will produce smaller red circles.

As per observations the recommended values for resolution and scaling are present in TABLE 5

<table>
<thead>
<tr>
<th>Time period of analysis</th>
<th>Resolution</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire day (7am to 10 pm) -15 hr.</td>
<td>4000</td>
<td>40</td>
</tr>
<tr>
<td>Morning peak period/Evening peak period – 2hr.</td>
<td>400</td>
<td>30</td>
</tr>
</tbody>
</table>

3.3.5.2. Representation of Frequency of 15 min. congestion on a Map

The 15 min. frequency of congestion matrix is to be calculated for the required time period of the day. Each cell will contains a number which represents the number of times it appeared red in each of the 15 min. interval images. MATLAB’s built-in colormaps can be used to color these cells. For example, FIGURE 8 shows a thermal map of a group of people captured

FIGURE 8 Colormap showing different temperatures in the infrared image (24)
using an infrared camera. The colormap of the image provides a simple yet effective representation to distinguish between warm and cold regions in the image. Similarly, FIGURE 9 shows an elevation map of North America and South America using a colormap. The colormap helps in easily distinguishing between higher and lower elevated areas.

**FIGURE 9** Colormap depicting different altitudes (25)

FIGURE 10 shows a list of MATLAB’s built-in colormaps that can be used. In addition, a custom colormap may be developed as seen in FIGURE 11 to depict intensities in a limited number of ranges.
FIGURE 10 Built-in colormaps in MATLAB (26)

FIGURE 11 Custom colormap
CHAPTER 4. APPLICATION

4.1. Analysis on one week data

4.1.1. Analysis for entire time period between 7 am and 10 pm.

The data for this analysis was collected on 09/03/13. Invalid red pixel data in the image that did not represent congestion was removed by applying screening procedures and visual confirmation. These red pixels were from features in the image such as: the legend, browser extension icons, Google’s logo, the top part of the interstate sign present on interstates, the red balloon in the center of the image and other isolated pixels from letters of labels on roads and places. Then a 15-min frequency analysis and Congestion Index analyses were performed (Note: The background in all the maps was only used as a base map; the colors in the maps except red have no congestion/traffic significance).

By applying the second method to represent congestion in the downtown area of Baton Rouge the congestion map present in FIGURE 12 was obtained. The four ranges of frequencies of 15 min. congestion are shown using a custom colormap of 4 colors. This map gives a perspective of congestion in the city as a function of frequency of congestion. From the map it is clear that I-10 east bound direction is relatively more congested than other roads, especially at its intersection with I-110.

This map gives a perspective of 15 min. congestion in the city. However, it does not provide any information on congestion duration and moreover it is difficult to detect and decipher congestion frequency of isolated pixels at intersections.

The first method using Congestion index with the buffering technique provides more information about congestion since it takes into account both frequency and duration of congestion into account. Here the diameter of the solid red circle about the pixels is in proportion
to its congestion index as seen in FIGURE 13. FIGURE 14 is a result of the same analysis. It shows congestion over roads near downtown and Louisiana State University in greater detail.

From the congestion index analysis it can be observed that the I-10 over the Mississippi River, arterials next to the I-10 before the I-10/I-12 split (Perkins Road, College Drive), two intersections (Old Hammond Highway - Jefferson and Brightside dr. – Nicholson dr.) and a major arterial (Airline Highway) crossing the I-12 on the right-hand side of the diagram, experienced heavy congestion with CI values in the range of 0.0393 to 0.2331. Hence they are considered hotspots. These observations are consistent with common opinion among motorists of sites in Baton Rouge with high levels of congestion.

FIGURE 12 Frequency analysis for 15 min congestion for entire day (7 am to 10 pm) on 09/03/13
FIGURE 13 Congestion Index and buffer analyses of Congestion Index for entire day (7 am to 10 pm) for data collected on 09/03/13
4.1.2. Analysis for morning peak period and evening peak period.

In Baton Rouge the morning peak occurs between 7 am and 9 am and evening peak between 4 pm and 6 pm. In both cases the CI was calculated and buffer analysis was performed. The resultant maps are shown in FIGURE 15 and FIGURE 16.

From the images it can be concluded that for the 1 week data collected on 9/03/13, morning peak period is significantly less congested than evening peak period. During the evening peak period, Airline Hwy., I-10, I-110, Perkins Rd., and College Dr., Siegen Lane and two intersections (Old Hammond Highway - Jefferson and Brightside dr. – Nicholson dr.) are severely congested. These were observed to have CI values in the range of 0.261 to 0.665. Hence, they are considered hotspots.
FIGURE 15 Congestion Index and buffer analysis for morning peak period (7am to 9 am) for data collected on 09/03/13
FIGURE 16 Congestion Index and buffer analysis for evening peak period (4 pm to 6 pm) for data collected on 09/03/13.
4.2. Analysis on one month data

The same analysis was performed for an entire month. Google updated traffic data only 3 times for the month in which analysis was done. These three unique data sets were collected on 8/12/13, 8/19/13, and 9/03/13 (8/26/13 data was not included since it was the same as 8/19/13) and arithmetically averaged using the following formula

\[
\text{Average matrix}_{(i,j)} = \frac{\text{Sample one}_{(i,j)} + \text{Sample two}_{(i,j)} + \text{Sample three}_{(i,j)} + \cdots + \text{Sample }k_{(i,j,k)}}{k}
\]

Where:

i = time of the day

j = day in the week

k = samples that represent a month’s traffic

This equation produced a set of 420 average matrices, representative of the entire month, which was used in the rest of congestion analysis for the month. CIs were calculated for all pixels and buffer analysis was performed which produced the results shown in FIGURE 17.

From the congestion index analysis it can be observed that I-10 to the east of the Mississippi River, arterials next to the I-10 before the I-10/I-12 split (Perkins Road, College Drive), Siegen Ln., two intersections (Old Hammond Highway - Jefferson and Brightside Dr. – Nicholson Dr.), experienced heavy congestion with CI values in the range of 0.0393 to 0.1207. Hence, they are considered to be congestion hotspots based on the definition established in TABLE 4.
FIGURE 17 Congestion Index and buffer analysis for entire month with 3 samples of data (7 am to 10 pm)
CHAPTER 5. VALIDATION OF THE METHODOLOGY

To demonstrate the validity of the proposed methodology an image analysis was first conducted using images downloaded from Google Maps of the I-12 corridor between Essen Ln. and Airline Hwy in Baton Rouge. After performing the image analysis, travel time and average travel speed data were gathered from a network of Bluetooth devices installed along the I-12 corridor. The frequency of congestion computed through image analysis for specific time period was then compared with average travel speed for the same time period from observations from the Bluetooth devices. The procedure used to perform the analysis is described in the following paragraphs.

For the I-12 corridor images were captured from Google Maps on 7/31/13 in GIF format with 256 colors at a resolution of 1280x1024 pixels, just as had been done with the earlier images. FIGURE 18 (a) shows the part of corridor that was the subject of this analysis. These images were captured for all seven days of the week between 3 p.m. and 7 p.m. The time period between 3 p.m. and 7 p.m. was used because in the preliminary analysis it was observed that the I-12 experiences congestion only between the above noted timeframe. Sixteen images (4 images/hour x 4 hours) were captured per day which resulted in a total of 16 x 7 = 112 images for the entire week.

The MATLAB script was then used to analyze the captured images. However, in this case both red and yellow colors were used to compare with the Bluetooth data, because both together represent fall in speeds below 50 mph. After analyzing, the 15 min. interval matrices which contain 0s and 1s, they were added across all days of the week. This yielded one matrix for each of the 15 min intervals. Therefore, in total, a total of 16 matrices were produced which represented the weekly traffic conditions existing between 3 p.m. and 7 p.m. The number in each
cell represents how often congestion conditions (speeds <50 mph) occurred at the locations represented by the pixel, in a 7 day period, at a specific 15 minute time period. FIGURE 18 (b)

FIGURE 18 (a) I-12 Corridor in Baton Rouge that was analyzed along with the Bluetooth detectors that were used, (b) shows part of the analyzed portion of the highlighted section, at 5:00 pm, in figure (a)

shows the resulting matrix after adding matrices representing 5:00 pm for seven days. The cells that contain value 4 indicate the presence of congestion on 4 days and the cells coincided with
the pixels on the I-12 corridor shown in FIGURE 18 (a). Similarly, the same congestion frequency of 4 was observed for the remaining 15 min. time intervals until 6 pm.

In Baton Rouge 11 BlueToad Bluetooth monitoring devices were deployed over I-12 between Essen Ln. and Walker. These devices capture and display, speeds and travel time in real time and all captured data is archived as well. Among the eleven devices, two devices which were installed at Essen Ln. interchange and Airline Hwy. interchange were used to gather required data for validation. The device locations are shown in FIGURE 18 (a) with a blue balloon.

Speed data collected between 7/23/13 and 7/28/13 aggregated at 15 min. was then downloaded from the devices. The following graph displays, the speeds over the week.

![Speed Graph](image)

**FIGURE 19 Bluetooth speeds over I-12 between Essen Ln. and Airline Hwy. between the days 7/22/13 -7/28/13**

From the graph in FIGURE 19 it is clear that between 5 and 6 p.m. there was a drastic fall in speeds to below 50 mph, which implies congestion. Free flow speeds on the corridor resumed soon after that. This happened for 4 consecutive days of the week i.e. from Monday through Thursday while Friday through Sunday did not experience any congestion. The image
analysis on Google Maps traffic data precisely demonstrates the same trend. On this section of
the corridor for the same time period the congestion frequency on the corridor was 4, as
represented by cell values in the matrix shown in FIGURE 18 (b). This implies congestion
occurred 4 out of 7 possible times in a week on this corridor between 5 p.m. and 6 p.m. Thus two
things have been validated from this analysis a) Google reports speeds that are consistent with
real time speed reports and b) the robustness of the proposed methodology.
CHAPTER 6. GRAPHICAL USER INTERFACE

A graphical user interface (GUI) was developed in MATLAB to run the entire analysis with minimum effort. FIGURE 20 shows a screenshot of the GUI. The way this program works is as follows:

1. The directory in which the images are stored day-wise using GhostMouse and FastStone Capture is loaded using the “Load Directory” push button.
2. Next, the “Process Images” push button is clicked to process the images.
3. The following drop down menu is used to select the time period of analysis i.e. morning peak, evening peak or full day (7 am to 10 pm). Since this project was conducted with Baton Rouge as the application site, the morning and evening peak time periods have been set for it.
4. The scale and resolution can be set as preferred by the user or recommended values can be used.
5. Finally, the “Generate Buffer Analysis Image” button should be clicked to generate the congestion map.
FIGURE 20 Graphical user interface of image analysis.
CHAPTER 7. CONCLUSIONS

1. The unique contribution of this thesis is a new methodology that is capable of analyzing color coded GIS maps representing traffic speed information to determine recurrent congestion locations.

2. In addition, a new congestion index was defined that can be used to detect different intensities of recurrent congestion. The application of the developed methodology in the Baton Rouge metropolitan area resulted in a congestion index (CI) that ranged from 0.26 to 1 for morning/evening peak period and CI that ranged from 0.039 to 1 for 15 hour time period.

3. The methodology developed in this research is automated to a large extent using various software capable of capturing data and performing image analysis, and by programming to automate the sequential execution of the individual analytical procedures. This lends itself to application over large geographical areas and extended time periods.

4. The methodology was validated by comparing congestion identified through image analysis with one identified from travel time measured using Bluetooth detection devices.

5. This research fills an important gap that exists in image analysis that is capable of tapping into freely available resources like Google traffic data.

6. Based on the 1-month analysis I-10 to the east of the Mississippi River, arterials next to the I-10 before the I-10/I-12 split (Perkins Road, College Drive), Siegen Ln., two intersections (Old Hammond Highway - Jefferson and Brightside dr. – Nicholson dr.) are found to have severe congestion. These roads and intersections are potential candidates for further analysis using traffic monitoring devices like Bluetooth traffic monitoring devices, Autoscope (video detection), license plate readers, loop detectors etc.
7. This research is of potential benefit to state DOTs, practicing traffic engineers, policy makers and researchers working in traffic engineering field.
REFERENCES


15. INRIX Sales Dept., telephone communication with Pete Costello, Director of Business Development, Feb 2013.


APPENDIX : MATLAB CODE

Part-1 Image analysis to convert pixel data to number format and compute 15, 30, 45 and 60 min. interval congestion matrices

Values 1,2,3 and 4 respectively are plugged in the place of ‘sumR’ to produce the 4 congestion index matrices.

```matlab
function finalMeas = gif_redT_new_with_intervals_fix_allred (sumR)

%% Read the image
days = {
'Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday'};

for daysItr = 1:size(days,2),
    cd (days(daysItr));

    files = ls;
    files(1:2,:) = [];
    %files = regexp(fileList, '\s', 'split');
    %files(end) = [];
    numOfIm = size(files,1);

    % Create memory for the final measure
    finalMeasSz = numOfIm-sumR+1;
    if (daysItr==1)
        finalMeas = cell(finalMeasSz,1);
        imSz = size(imread(files(1,:)));
        for i=1:finalMeasSz
            finalMeas{i} = zeros(imSz);
        end
    end

    for imItr=1:finalMeasSz
        finalT = zeros(imSz);
        for i=imItr:imItr+sumR-1
            Fname = files(i,:);
            [imDat imMap] = imread(Fname);
            tempMeas = zeros(imSz);
            redInd = imDat(redYpos,redXpos);

            % assign all the pixels with red indices to '1'
            a=[];
            p=1;
            for z=1:256
                if imMap(z,1)>0.45
                    if imMap(z,2)<0.38
                        if imMap(z,3)<0.32
                            a(p,1)=z-1;
                            p=p+1;
                        end
                    end
                end
            end
```
Part 2: Congestion index computation for 15 hours of the day

%Adding matrices for congestion index
function CI_mat = Adding_for_congestion_index (Min15, Min30, Min45, Min60)

Final15_sum=zeros(1024,1280);
for i=1:60
    Final15_sum=Final15_sum+Min15{i,1};
end
Final15=Final15_sum./7;

Final30_sum=zeros(1024,1280);
for i=1:59
    Final30_sum=Final30_sum+Min30{i,1};
end
Final30=Final30_sum./7;

Final45_sum=zeros(1024,1280);
for i=1:58
    Final45_sum=Final45_sum+Min45{i,1};
end
Final45=Final45_sum./7;

Final60_sum=zeros(1024,1280);
for i=1:57
    Final60_sum=Final60_sum+Min60{i,1};
end
Final60=Final60_sum./7;

CI_mat=
1/10*(Final15/60)+2/10/2*(Final30/59)+3/10/3*(Final45/58)+4/10/4*(Final60/57) ;

Part 3: Removal of invalid data

for i=1:1024;
    for j=1:1280;
        if index_mat(i,j)>0.24;
            index_mat(i,j)=0;
        end;
    end;
end;

%clean data on top, top right, left and right bottom
index_mat(1:153,:)=0;
index_mat(1:265,1203:1280)=0;
index_mat(:,1:282)=0;
index_mat(1009:1024,963:1280)=0;
%clean data for highway signs, 12 of them
for i=1:2:24
    index_mat(hsm(i,2):hsm(i+1,2),hsm(i,1):hsm(i+1,1))=0;
end

% clean data for Letter A and 'Baton rouge'
for i=1:2:12
    index_mat(red_peg_area(i,2):red_peg_area(i+1,2),red_peg_area(i,1):red_peg_area(i+1,1))=0;
end

Part-4 Analysis to produce buffer analysis using CI

function output = buffering_of_cong(index_mat)

resul_val=4000; % you may change this to 4000 and the next one to 20
scaling_val=40;
index_mat_round = round(index_mat*resul_val); % since outliers > 0.1 have been removed. Need a fixed highest index value present in the index_mat

set_of_values = unique(index_mat_round);

temp_mat=zeros(size(index_mat));
final_mat=zeros(size(index_mat));
output_temp=zeros(size(index_mat));

for i=2:size(set_of_values,1)
    temp_mat(index_mat_round==set_of_values(i,1))= set_of_values(i,1);
    scaling_factor = round(set_of_values(i,1)*scaling_val*(1/resul_val));
    str_obj = strel('disk', scaling_factor, 0);
    finalM = imdilate(temp_mat,str_obj);
    final_mat = final_mat + finalM;
    temp_mat=zeros(size(index_mat));
end

output_temp(final_mat>0)=2;
output=output_temp;
cmap = [1,1,1 ; 1,0.0,0.0];
clims = [0 1];
y=imagesc(output,clims);
colormap(cmap);
imwrite(output,cmap,'redmap_Full_new_x10.gif'); % output is fixed

Part-5 Procedure to overlay buffer analysis image on a base map

% image overlay for Full

imAlphaData=zeros(1024,1280); % change as per image size
% creating Alphadata matrix, with a condition that all red pixels in the top
% layer have opacity and all white pixels have an opacity - zero
imAlphaData(output>0)=0.9;  % changed
figure()
bg=imread('Back_ground.gif');
[cData colmap]=imread('Back_ground.gif');

imshow(bg,colmap);
axis  off
hold on

% top=imread('redmap_Full_new_x10.gif');
% top=imread('redmap_lines_Full.gif');
top=imread('full.png');
iim2 = image(top);
set(iim2,'AlphaData',imAlphaData);
% imwrite(bg,colormap,'final_map.gif')
hold off
VITA

Adiyta Mokkapati was born in Hyderabad, India. After graduating high school from Vignan Vidyalayam, Hyderabad, he pursued his intermediate studies in Sri Chaitanya Junior College. He then attended Vasavi College of Engineering in Hyderabad. Here he completed a Bachelor in Engineering degree in civil engineering. After working as a transportation engineer in India for a year, he entered graduate school under the direction of Dr. Chester G. Wilmot. Aditya is now a candidate for the degree of Master of Science in The Department of Civil and Environmental Engineering at Louisiana State University.