USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No 002OY01

# Estimating AADT from combined air photos and ground based data: System design, prototyping, and testing

By

Mark R. McCord, Principal Investigator
Professor of Civil and Environmental Engineering and Geodetic Sciences
Ohio State University
mccord.2@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
Ohio State University
goel.1@osu.edu

Report Submission Date: October 15, 2009

# DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

USDOT Region V Regional University Transportation Center Final Report

# TECHNICAL SUMMARY

NEXTRANS Project No 002OY01 Final Report, October 2009

## Estimating AADT from combined air photos and ground based data: System design, prototyping, and testing

### Introduction

Average annual daily traffic (AADT) is perhaps the most fundamental measure of traffic flow. The data used to produce AADT estimates are largely collected by in-highway traffic counters operated by traffic monitoring crews who must cover thousands of segments in their statewide systems on a continual basis. In addition to being costly, dangerous, and disruptive, the combination of limited resources and the large number of highway segments spread across the expansive geographic regions of the state requires that the state DOTs collect short-term sample volumes on a multi-year cycle. We have developed a method that combines the older, ground-based traffic data with traffic information contained in recent air photos in a statistically justified manner to produce more accurate estimates of AADT. To take advantage of this promising method in practice, it is necessary to develop an efficient way to use it on a widespread, repeated basis in an operational setting.

The proposed work builds on previous efforts that led to conception, development, and preliminary testing of the estimation method. We designed the components of a software system that can be used to efficiently produce the improve AADT estimate, conducted empirical tests of the performance of the estimate, and worked toward gaining insitutional acceptance for this novel estimation approach.

### Findings

We were successful in developing the components of a software system to digitize the appropriate information from an aerial image and combine the information with information contained in traffic monitoring data bases to produce our proposed, improved AADT estimate. We digitized information on roadway segments in twelve air photos and combined the information with traditionally collected ground-based data to produce the proposed AADT estimates in approximately one minute per image. That is, if imagery collected for other purposes is available, an updated, improved AADT estimate could be obtained in minimal time. It is important to note that, for this study, we stored important input parameters in such a way that they were exogenously linked to the imaged segment and that the empirical work was conducted in a “laboratory” setting where the information was readily available. In an operational system, the ground-based information would need to be linked to the imaged segment by automatically registering the imaged segment to a geospatial database that contains the appropriate representation of the segment in the traffic monitoring databases. This is a subject for our proposed following year project.

We also used the developed software components to produce output estimates that were part of a carefully designed and novel empirical study in which our proposed AADT estimate outperformed traditional AADT estimates on three important measures – the mean absolute relative error in the estimate, the proportion of times that the estimate produced less than 10% absolute relative errors (a criterion often considered a target in practice), and the proportion of times that our proposed estimate outperformed the traditional estimate.

Finally, we discussed our concepts and presented our results on several occasions to the heads of traffic monitoring, aerial engineering, and GIS at the Ohio Department of Transportation (ODOT). These interactions were considered successful in that they led to a decision to take the next step toward implementation by developing and integrating our software components in a prototype system at the Ohio Department of Transportation in a follow-on project.

## Recommendations

Based on our findings, we recommend that the components of the software system we developed be integrated in a prototype system at a state department of transportation to examine and overcome issues involved with developing an operational system. We also recommend that additional empirical tests be conducted to further demonstrate the improved accuracy of the proposed AADT estimate in an effort to gain acceptance for operational use.

## Contacts

### *For more information:*

**Mark R. McCord**
Principal Investigator
Civil and Environmental Engineering and Geodetic Sciences
Ohio State University
mccord.2@osu.edu

**Prem Goel**
Co-Principal Investigator
Statistics
Ohio State University
goel.1@osu.edu

**NEXTRANS Center**
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

NEXTRANS Project No 002OY01

# Estimating AADT from combined air photos and ground based data: System design, prototyping, and testing

By

Mark R. McCord, Principal Investigator
Professor of Civil and Environmental Engineering and Geodetic Sciences
The Ohio State University
mccord.2@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
The Ohio State University
goel.1@osu.edu

Report Submission Date: October 15, 2009

# DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

# ACKNOWLEDGMENTS

The authors acknowledge the assistance and feedback from Dave Gardner, John Ray, Dave Blackstone, and Chuck Bernthold**,** of the Ohio Department of Transportation, and Zhoujun Jiang, of the Mid-Ohio Regional Planning Commission. Mr. Gardner and Dr. Jiang are members of the project advisory committee. We thank Yufang Zhang and Cheng Chen for their assistance in preparing this report and acknowledge the financial support offered by the sponsor and by The Ohio State University. The views and opinions contained in this report are those of the authors and do not represent the views, opinions, or policies of any other individuals, agency or group.

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# CHAPTER 1. INTRODUCTION, PROBLEM, AND APPROACH

## *1.1. Introduction*

Average annual daily traffic (AADT) is perhaps the most fundamental measure of traffic flow. The data used to produce AADT estimates are largely collected by in-highway traffic counters operated by traffic monitoring crews who must cover thousands of segments in their statewide systems on a continual basis. In addition to being costly, dangerous, and disruptive, the combination of limited resources and the large number of highway segments spread across the expansive geographic regions of the state requires that the state DOTs collect short-term, sample volumes on a multi-year cycle.

We have developed a method that combines the older, ground-based traffic data with traffic information contained in recent air photos in a statistically justified manner to produce more accurate estimates of AADT. If implemented, data collection procedures could conceptually be adjusted so that fewer costly and dangerous ground counts would be required. As such, the method could produce more accurate AADT estimates at lower cost and a higher level of safety. To take advantage of this promising method in practice, it is necessary to develop an efficient way to use the method on a widespread, repeated basis in an operational setting.

This work sets out to move this promising method toward implementation.

## *1.2. Problem*

The overall problem to be addressed is that of working toward the development, implementation, and use of a system and a process in which aerial imagery, primarily collected for non-traffic monitoring purposes by state DOTs, is used to improve AADT estimates

## *1.3. Approach*

To work toward an operational system used by state DOTs, our approach consisted of three major thrusts.

*Thrust 1*: Develop the components of a prototype software system that can efficiently extract and combine the relevant information from air photos and traffic monitoring data bases at state DOTs to produce the proposed AADT estimates

*Thrust 2*: Build on previous analytical and empirical work to further demonstrate the improved accuracy of the proposed AADT estimates

*Thrust 3*: Work toward gaining the institutional acceptance of the approach for operational use by state DOTs

# CHAPTER 2. METHODOLOGY

## *2.1. Background Methodology*

To facilitate the understanding of the work documented in this report, we first describe the traditional approach to AADT estimation and the proposed method for combining information in air photos with traditionally collected ground counts to provide an improved AADT estimate.

### *2.1.1. Traditional method of estimating AADT*

AADT is the average of all daily volumes on a particular highway segment, where the average is taken across an entire year. Conceptually, AADT is calculated by obtaining the daily traffic volume (the number of vehicles that pass a point on the segment over the entire day) for every day in the year, and averaging these volumes. These daily volumes could conceivably be collected by Automatic Traffic Recorders (ATRs), which are recorders such as magnetic loop detectors that are permanently installed in the roadway pavement and collect volume data continually throughout the year. However, cost considerations limit the number of roadway segments that can be installed with ATRs. For example, one estimate (Jiang et al., 2006) places the number of ATR-equipped segments at only about 3% of the segments for which a state DOT must estimate AADT.

AADT on a roadway segment not equipped with an ATR is typically estimated from sampled coverage counts. Coverage counts are daily volumes that are usually collected by temporary “in-the-road” counters on two consecutive days. To produce a “coverage count-based” estimate of AADT on segment *s*, the daily coverage counts are first adjusted (“deseasonalized”) by seasonal factors that account for systematic temporal patterns and then averaged. The seasonal factors are usually associated with the day-of-the-week and the month-of-the year and are produced from ATR data on segments expected to follow the same temporal pattern in traffic as the coverage count segment. Letting $V^{24}_s(\delta, \gamma)$ and $V^{24}_s(\delta+1, \gamma)$, respectively, denote the daily (24-hour) volumes obtained on segment *s* on the $\delta^{th}$ and $(\delta+1)^{st}$ days of year $\gamma$, and $F^{md}_s[m, d; \gamma]$ denote the seasonal factor in year $\gamma$ associated with month-of-year *m* and day-of-week *d,* the coverage-count based AADT estimate $AADT^C_s(\gamma)$ on segment *s* in year $\gamma$ derived from the daily volumes is:

$$\begin{aligned} AADT^C_s(\gamma; \delta, \delta+1) = (1/2) \times \{ & V^{24}_s(\delta, \gamma) \times F^{md}_s[M(\delta, \gamma), D(\delta,\gamma)] \\ & + V^{24}_s(\delta+1, \gamma) \times F^{md}_s[M(\delta+1, \gamma), D(\delta+1, \gamma)] \end{aligned} \quad (1)$$

where $M(\delta, \gamma)$ and $D(\delta, \gamma)$represent, respectively, the month-of-year and day-of-week corresponding to the $\delta^{th}$ day of year $\gamma$.

Because of the large number of segments that are not equipped with ATRs, it would be time-consuming and expensive to obtain coverage counts every year on all segments not equipped with ATRs. Therefore, in practice, coverage counts are only obtained on a subset of segments in one year. In the following years, the coverage counts are obtained on other segments. As such, the traffic monitoring crew returns to a specific segment after *n* years, leading to what is called an *n*-year collection cycle. For example, a state may conduct its traffic count program so that traffic monitoring crews sample a specified set of high priority segments on a three-year cycle (*i.e.,* once every three years) and the remainder of the segments on a six-year cycle (once every six years).

Growth factors (FHWA, 2001) can be used to adjust the AADT estimate obtained from coverage counts taken in a year $\gamma$ to an estimate for a later year $\gamma'$ in the cycle. The growth factor for a pair of years $\gamma$ and $\gamma'$ on segment *s* is defined as:

$$\mathrm{GF}_s(\gamma, \gamma') = \mathrm{AADT}^{\mathrm{true}}{}_s(\gamma')/\ \mathrm{AADT}^{\mathrm{true}}{}_s(\gamma) \qquad (2)$$

where $AADT^{true}{}_s(\ .)$ is the true (not estimated) AADT on segment *s* in the specified year. Since the true AADT is not known on the segment for which a coverage count is taken, the $GF_s$ is estimated using ATR data from ATR-equipped segments whose growth in traffic is expected to be similar to that of segment *s* (see*, e.g.,* Jiang et al., 2006; FHWA, 2001). We use $GF^{est}{}_{=}(.,.)$ to indicate the estimated growth factor to be used on segment *s*.

Using the coverage count-based estimate $AADT^C(\gamma)$ of AADT in year $\gamma$ obtained from equation (1) and the estimated growth factor $GF^{est}(\ \gamma, \gamma')$between years $\gamma$ and $\gamma'$, an estimate of AADT in year $\gamma'$, which is denoted by $AADT^{CG}(\gamma', \gamma)$, is calculated as:

$$\mathrm{AADT}^{\mathrm{CG}}{}_s(\gamma', \gamma) = \mathrm{AADT}^{\mathrm{C}}{}_s\,(\gamma) \times\ \mathrm{GF}^{\mathrm{est}}{}_s(\gamma, \gamma') \qquad (3)$$

where the explicit indication of the days on which the coverage count were taken is omitted for notational convenience.

### *2.1.2. Proposed method of estimating AADT using image- and ground-based data*

In addition to equipment (measurement) errors, the traditional $AADT^{CG}{}_s(\gamma', \gamma)$ estimate has two major sources of noise. The first source of noise is that caused by estimating the year $\gamma$ AADT from a limited sample (the two daily volumes) to determine the yearly average. The other source of noise is that of using the estimated growth factor to represent AADT growth on segment *s.*

McCord et al. (2003) proposed a means to estimate AADT from traffic information in contemporary images. The image-based AADT estimate in year $\gamma$, denoted by $AADT^{img}(\gamma)$, is determined by estimating an hourly volume based on the traffic information in the image obtained in year $\gamma$ , converting the hourly volume to a daily volume using hourly factors, and then adjusting the daily volume to an AADT estimate using seasonal factors.

As presented in Jiang, et al. (2006), the image-based AADT estimate obtained for a highway segment $s$ of imaged length $L$, where the image obtained during hour $h$ ($h$=1, …, 24) on day δ of year $\gamma$ contains $N^{veh}$ vehicles, would be:

$$AADT^{I}(\gamma)= (N^{veh}/L) \text{ x } U_s[N^{veh}/L] \text{x } 24 \text{ x} F^{h}_{s}[h,D(\delta,\gamma)] \text{ x } F^{md}_{s}[M(\delta,\gamma),D(\delta, \gamma)] \qquad (4)$$

where $U_s[N^{veh}/L]$ is the average speed on segment $s$ when the traffic density is $N^{veh}/L$, $F^{h}{}_{s}[h, d]$ is the hourly factor used to convert an hourly traffic volume occurring during hour $h$ and the day-of-week $d$ on segment $s$ to an average hourly volume for the day, and the rest of the notation is as presented above. (Like the seasonal factors $F^{md}$, the hourly factors $F^{h}$ can be obtained from ATR data on segments assumed to have similar temporal traffic patterns as segment $s$ (*e.g.,* Jiang, et al, 2006).)

Compared to the traditional, ground-based AADT estimate $AADT^{iCG}$ of equation (3), the image-based estimate $AADT^{I}$ of equation (4) has the advantage of providing more contemporary information on the segment. However, the estimate can be shown to correspond to a ground-based traffic count of very short duration (McCord et al., 2003). Jiang et al. (2006) suggested combining the recent, shorter duration image information with older, longer duration coverage count information through a weighted combination of $AADT^{I}$ and $AADT^{CG}$, where the weights are determined from the sample variances of the ratios of the estimated-to-true AADTs. Specifically, the proposed estimate of AADT in year γ' based on coverage counts obtained in year $\gamma$ and an image in year $\gamma'$, denoted $AADT^{CGI}{}_{s}(\gamma', \gamma)$, is:

$$AADT^{CGI}{}_{s}(\gamma', \gamma) = w \text{ x } AADT^{I}{}_{x}(\gamma') + (1-w) \text{ x } AADT^{CG}{}_{s} (\gamma', \gamma) \qquad (5a)$$

where $AADT^{I}{}_{s}(\gamma')$ is determined from equation (4), $AADT^{CG}{}_{s}(\gamma', \gamma)$ is determined from equation (3), and the weight $w$ is calculated as:

$$w = [(\sigma^{C})^2 + (\sigma^{G})^2] / [(\sigma^{C})^2 + (\sigma^{G})^2 + (\sigma^{I})^2] \qquad (5b)$$

where $(\sigma^{C})^2$ is the variance of the ratio the coverage count-based AADT estimate in year $\gamma$ (the year in which the coverage counts were obtained) to the true AADT in year $\gamma$, $(\sigma^{G})^2$ is the variance of the ratio of the estimated growth factor for segment $s$ between years $\gamma$ and $\gamma'$ to the true growth factor for the segment, and $(\sigma^{I})^2$ is the variance of the ratio of the image-based AADT estimate in year $\gamma'$ (the year in which the image was obtained) to the true AADT in year $\gamma'$. As explained in Jiang *et al.* (2006), the variances can be estimated from available data or empirical studies.

In Jiang *et al*., (2006), the $AADT^{CGI}{}_{s}(\gamma', \gamma)$ estimate is argued to be theoretically more accurate than the traditional estimate $AADT^{CG}{}_{s}(\gamma', \gamma)$ and shown to perform better in a simulation study. As mentioned above, the overriding objective of the work reported in this document is to move toward implementing the use of this estimate in practice.

## 2.2. Methodology for Present Effort

### Thrust 1: Develop a prototype software system

We developed the overall design of a software system to produce the proposed estimate $AADT^{CGI}$ from an air photo and ground-based data stored in formats similar to those expected to be used at a state DOT. The software implementation consists of three components: (i) a component that produces the image-based estimate $AADT^{I}$ in a manner compatible with equation (4); (ii) a component that produces the ground-based estimate $AADT^{CG}$ in a manner compatible with equation (3); and (iii) a component to determine the weight *w* and combine the image- and ground-based estimates to produce $AADT^{CGI}$ in a manner compatible with equation (5). We developed MATLAB modules to implement each of these components. Detailed information on this software and a user's manual are presented in Zhou (2008).

The first module *digitizing.m* allows the user to interactively digitize a .jpeg image to obtain the imaged segment length *L* and the number of vehicles $N^{veh}$ of equation (4) in a semi-automatic manner. This program can be easily operated by someone with minimal technical background. The three output files from this program include locations of digitized points along the segment, the locations of cars, and the locations of trucks. The program linearly connects the digitized points to determine *L* and uses the digized car and truck locations to determine $N^{veh}$. (Recording car and truck locations separately allows the possibility of determining separate estimates for car and truck AADT in future work.) The coordinates of these points are automatically stored in the output files in MATLAB data format.

The flowchart of the digitization program *digitizing.m* is shown in Figure 1. As indicated in the flowchart, the image and an overview window with instructions appear when beginning. There may be times when the user must stop before completing the digitization process. The program allows digitized points to be saved and the user to continue digitizing later on the saved file, which then shows the previously digitized points on the loaded image. The program also allows deletion of a "misclicked" (erroneously digitized) point. Users can zoom in or zoom out by changing the percentage at the left bottom of the image window. Pressing "Enter" in the overview window starts or restarts the digitizing process. An example of such a window is shown in Figure 2.

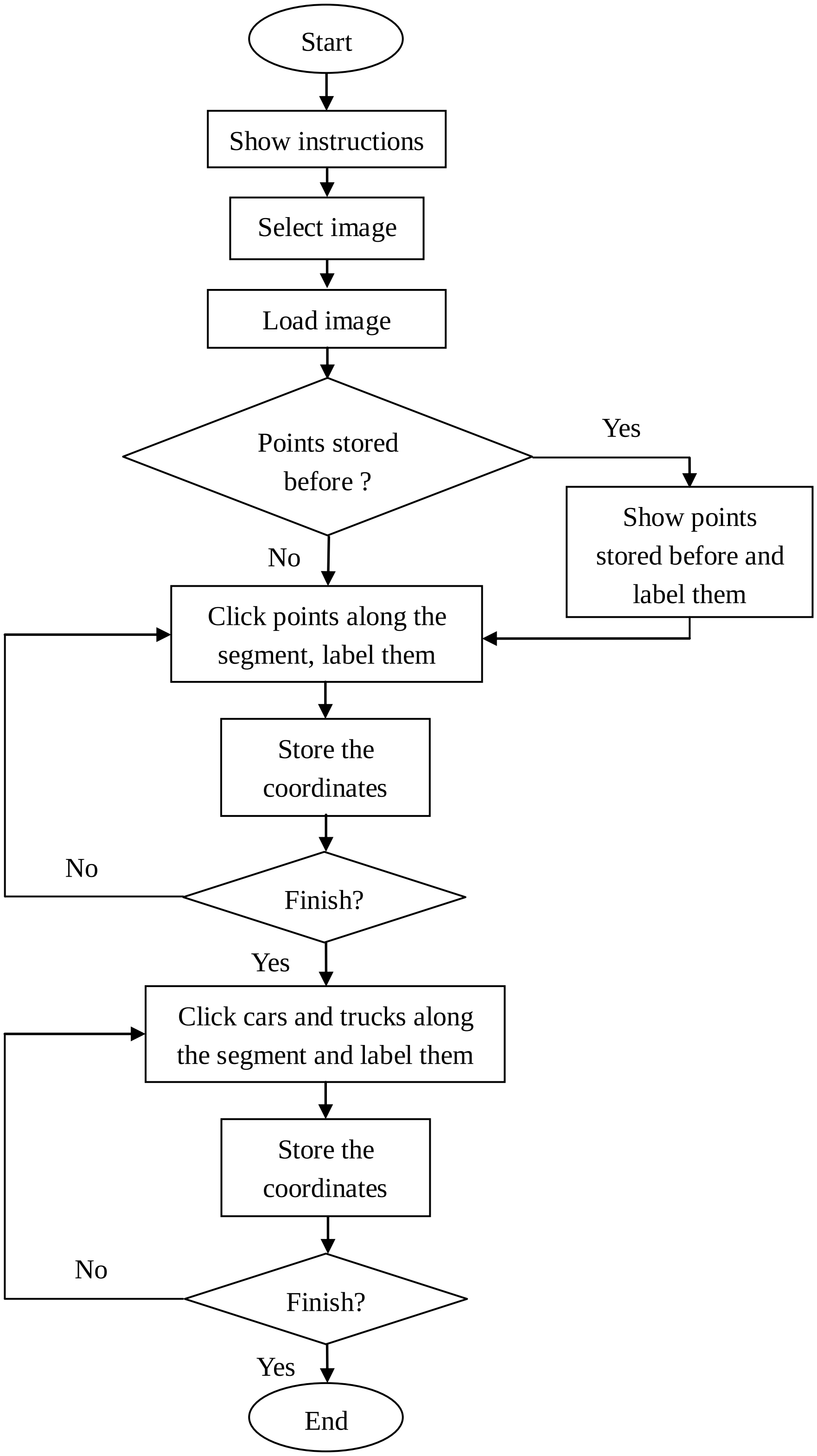

Figure 1. Flowchart of digitization program, *digitizing.m*

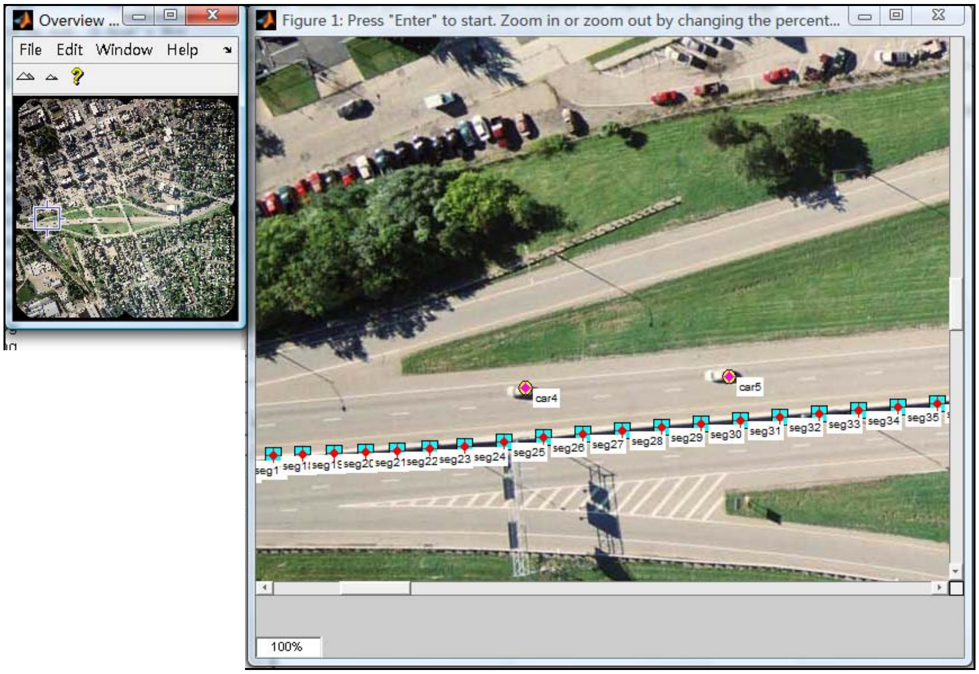

Figure 2. Example of an intermediate step in *digitizing.m*

Once the digitizing step is completed, the image-based estimate *AADT*$^I$ is produced by a MATLAB program *aadtimg.m.* The inputs to this program are the output data files of *digitizing.m* (the digitized points delineating the roadway segment, the locations of cars, and the locations of trucks) and a file named *imageinformation.mat*. The *imageinformation.mat* file includes the functional classification of the segment in the image, the average speed at which vehicles would travel on the segment in uncongested conditions, and image meta data, specifically, the scale of the image and the year, month, day of week, and time when the image was taken. Additional inputs to this program are the hourly and seasonal factors of equation (4) associated with the segment, hour of image, and date of image. In an operational setting the segment-referenced data will need to be called automatically after the specific segment in the image is identified, likely through geo-spatial registration. This will be a major focus of our year 2 project.

The MATLAB data file *AADTImgResults.mat* contains an array *AADTImgResults* of the *AADT*$^I$ values that have been calculated from digitized images. This array is used as an input to the MATLAB program *AADTcomb.m*, where the *AADT*$^I$ and *AADT*$^{CG}$ values are combined to produce the *AADT*$^{CGI}$ estimate. The flowchart of this program and the user guide can be found in Zhou (2008).

In an operational implementation, the most recent coverage-count based AADT estimate for an imaged segment will need to be pulled from the traffic monitoring database. However, to permit our Thrust 2 study (see below), we obtained images on segments that contained ATRs. Therefore, rather than seeking coverage count-based estimates in traffic monitoring data bases, we simulated coverage counts from ODOT ATR data. The MATLAB program *threecardload.m* converts the ATR data file to a MATLAB data file used in the estimation of simulated coverage count-based estimates $AADT^C$. The calculations are carried out by a MATLAB program *AADTc_func.m*. The output is then used in the MATLAB program *AADTcomb.m*, which adjusts the coverage count-based estimate $AADT^C$ for traffic growth according to equation (3). In an operational application, the growth factor is expected to be available from tables already produced by traffic monitoring sections. However, to determine the weighting factor $w$ of equation (5), the variance of the growth factor ratios – $(\sigma^G)^2$ of equation (5b) – must also be available. Therefore, we have developed a MATLAB program *growthf_func.m* to calculate growth factors and variance of growth factor ratios by functional class from ATR data. The outputs of the program *growthf_func.m* include the estimated growth factor *growth_factor* and sample variance of growth factor ratios *growth_factor_var*, which are used in the MATLAB program *AADTcomb.m*.

The $AADT^{CG}$ value of equation (3) is produced in the *AADTcomb.m* program by multiplying the $AADT^C$ value, output from the *AADTc_func.m* program, by the appropriate growth factor. To obtain the weight $w$ of equation (5b), we need to estimate values of $(\sigma^I)^2$, $(\sigma^G)^2$, and $(\sigma^C)^2$. The estimate of $(\sigma^I)^2$ is provided exogeneously as an input value, as proposed in Jiang, et al. (2006). The estimate of $(\sigma^G)^2$ is produced in the *growthf_func.m* program, as described above. The estimate of $(\sigma^I)^2$ is produced in a MATLAB program *coverage_variance_func.m* using the process described in Jiang et al. (2006).

The simple calculations of $w$ and of $AADT^{CGI}$ occur in *AADTcomb.m.* To prepare for further operational implementation, the inputs to this program include the year $\gamma$ (four digits) in which the coverage counts were obtained, the year $\gamma'$ (four digits) in which the segment was imaged, ATR data in MATLAB data format, the station ID of the segment for which the AADT is to be estimated, the $AADT^I$ and $AADT^C$ results produced from the programs described above, and a reference between the image identifier and the segment ID. (Providing this reference automatically will be the subject of the year 2 project.) Stepwise instructions for the *AADTcomb.m* program are provided in Zhou (2008).

### *Thrust 2: Further demonstrate improved accuracy of the proposed estimate $AADT^{CGI}$*

Prior to this project, the Ohio DOT supplied us with twelve digital air photos of six different roadway segments that contained ATRs. The twelve photos, which appear in Appendix A, were taken in 2005. We also received ATR data for the segments from 2003, 2004, and 2005. We used these items to illustrate the accuracy of the proposed approach. The details of the methodology are as follows.

From the ATR data on segment $s$, $s = 1,2, \ldots, 6,$ in the appropriate year, we obtained estimates $AADT^{true}_s(2005)$, $AADT^{true}_s(2004)$, and $AADT^{true}_s(2003)$ of the true AADT on the segment in the indicated year. We call these *estimates* of the true AADT, since there could be error in the ATR counts, and there were a few times when the ATRs malfunctioned to the extent that no data were provided for certain time intervals. (We used the AASHTO method (FHWA, 2001) to account for missing data.).

The ATR data allowed us to determine daily volumes $V^{24}_s(\delta)$ on segment $s$ for every day $\delta$ on which the ATR functioned for the entire day. From these daily volumes we could determine the coverage counts that would have been obtained if the segment had not contained an ATR but, rather, was sampled with a portable counter on the corresponding day. (We note that we are assuming that the counter used to obtain the coverage count would have produced the same count as the ATR on the day.) From two consecutive daily volumes $V^{24}_s(\delta,\gamma)$ and $V^{24}_s(\delta+1,\gamma)$ in year $\gamma$, we could obtain what would have been the corresponding coverage count estimate $AADT^C_s(\gamma\,;\delta,\,\delta+1)$ of the AADT in year $\gamma$ using equation (1). (We obtained the seasonal factors $F^{md}$ from the ATR data using the procedure presented in Jiang, et al., (2006).) Using this approach we produced $AADT^C_s(\gamma\,;\delta,\,\delta+1)$ values for years $\gamma$ = 2003 and 2004 for every pair of consecutive days $\delta$ and $\delta+1$ in the year for which we could obtain a 24-hour volume from the ATR data. The numbers of coverage counts obtained in this way for the six segments in 2003 and 2004 are provided in Table 1. We also indicate the specific segment by the ATR number and remark that data were not available in 2003 for the segment containing ATR 752. In the table, we also indicate the functional class of the segments considered.

| Image # | ATR # | Functional Class | # of Generated Coverage Count Pairs | |
|---|---|---|---|---|
| | | | 2003 Data | 2004 Data |
| 11273-1-1 | 767 | 12 | 296 | 335 |
| 11273-1-2 | 767 | 12 | 296 | 335 |
| 11273-2-3 | 707 | 1 | 277 | 318 |
| 11273-2-4 | 707 | 1 | 277 | 318 |
| 11273-4-5 | 601 | 11 | 136 | 295 |
| 11273-4-6 | 601 | 11 | 136 | 295 |
| 11273-5-7 | 752 | 11 | n/a | 339 |
| 11273-5-8 | 752 | 11 | n/a | 339 |
| 11273-3-9 | 121 | 11 | 158 | 316 |
| 11273-3-10 | 121 | 11 | 158 | 316 |
| 11273-6-11 | 140 | 11 | 340 | 339 |
| 11273-6-12 | 140 | 11 | 340 | 339 |

Table 1. Number of coverage count AADT estimates used in empirical study, by year and segment

We also used the ATR-derived true AADT estimates to determine growth factors $GF^{est}(\gamma,2005)$ between year $\gamma$ and 2005, for $\gamma=2003,\ 2004$. Specifically, to determine $GF^{est}_s(\gamma,2005)$ that would be used when estimating the 2005 AADT from a coverage count estimate in $\gamma$ on segment *s*, we used the arithmetic average of the $GF_{s'}(\gamma,\ 2005) = AADT^{true}_{s'}(2005)/\ AADT^{true}_{s'}(\gamma\ )$ for the five segments *s'* different from *s*. (We note that in practice there would only be one growth factor used for all segments in a roadway category, *e.g.,* functional class. However, there would not be ATR data on the coverage count segment, and we "held out" ATR data from segment *s* to more accurately reflect this concept.) For each coverage count AADT estimate $AADT^C_s(\gamma\,;\delta,\,\delta+1)$ in year γ on segment *s*, which we determined as described above, we used equation (3) with the corresponding growth factor to determine the corresponding "growth factor-based" estimate $AADT^{CG}_{s}(2005; \gamma, \delta, \delta+1)$ of the year 2005 AADT on segment *s.* As proposed in Jiang et al. (2006), we estimated the variance $(\sigma^{G}_{s})^{2}$ of the growth factors to be used with estimates for segment *s* by considering the set of all segments *s′* different from *s* (*i.e.*, "holding out" data from segment *s*), forming the ratios of the estimated growth factor $GF^{est}_{s}$ for *s* to the true growth factor $GF_{s'}$, and calculating the sample variance $(S^{G}_{s})^{2}$. The estimated growth factors $GF^{est}$ and sample standard deviations $S^{G}$ of the growth factors calculated in this way appear in Table 2.

| ATR # | $GF^{est}_{s}(\gamma, 2005)$ | | $S^{G}_{s}(\gamma, 2005)$ | |
|---|---|---|---|---|
| | $\gamma = 2003$ | $\gamma = 2004$ | $\gamma = 2003$ | $\gamma = 2004$ |
| 767 | 0.979 | 0.969 | 0.0252 | 0.0225 |
| 707 | 0.970 | 0.962 | 0.0285 | 0.0246 |
| 601 | 0.980 | 0.959 | 0.0237 | 0.0206 |
| 752 | n/a | 0.971 | n/a | 0.0185 |
| 121 | 0.965 | 0.962 | 0.0200 | 0.0248 |
| 140 | 0.972 | 0.964 | 0.0293 | 0.0256 |

Table 2. Growth factor estimates $GF^{est}_{s}(\gamma, 2005)$ and sample standard deviations $S^{G}_{s}(\gamma, 2005)$ used in empirical study

For each of the twelve images, $i = 1,2,\ldots, 12,$ we used the software described in Thrust 1 to determine the image-based AADT, $AADT^{I}_{s}(2005)$, $i = 1,2,\ldots, 12,$ for 2005, the year in which the image was taken. As mentioned above, the twelve images consisted of two separate images on each of the six segments *s.* (The two images of the same segment would not be considered independent images, since they were taken only a few seconds apart and overlapped by approximately 60%. Therefore, the two images would contain several of the same vehicles.)

For each coverage count-based AADT estimate $AADT^{C}_{s}(\gamma ; \delta, \delta+1)$ in year γ on segment *s* and for each of the two images taken on segment *s*, we produced the combined (image- and ground-based data) estimate $AADT^{CGI}_{s}(2005; \gamma, \delta, \delta+1)$ of the AADT in 2005 on segment *s.* To determine the weight *w* used in combining the image- and ground-based AADT estimates in equation (5a), we used equation (5b), estimating $\sigma^{C}$ as suggested in Jiang et al. (2006), $\sigma^{G}$ by the estimated values in Table 2, and $\sigma^{I}$ with the default estimates of 0.2 suggested in Jiang et al. (2006). For each of these coverage count estimates, we also considered two other estimates of the 2005 AADT on segment *s,* namely $AADT^{C}_{s}(\gamma; \delta, \delta+1)$ and $AADT^{CG}_{s}(2005; \gamma, \delta, \delta+1)$. The $AADT^{C}_{s}(\gamma ; \delta, \delta+1,)$ estimate represents the 2005 AADT estimate for segment *s* from ground-based coverage count data if no growth factor is used. The $AADT^{CG}_{s}(2005; \gamma, \delta, \delta+1)$ represents the 2005 AADT estimate for the segment when using the growth factor to adjust the coverage count-based estimate. (Ignoring potential traffic growth by using $AADT^{C}$ may be expected to outperform $AADT^{CG}$ if, for example, the growth on the segment *s* is relatively small and the error in estimating growth factors for segment *s* from the ATR data on other segments *s′* assumed to have similar growth as that on *s* is relatively large.)

In this way, we had three estimates of the 2005 AADT – $AADT^{CGI}_s(2005; \gamma , \delta, \delta+1)$, $AADT^{CG}_s(2005; \gamma , \delta, \delta+1)$, and $AADT^{C}_s(\gamma ; \delta, \delta+1)$ – for each of the hundreds of coverage counts generated from the 2003 and 2004 AADT. We compared each estimate to $AADT^{true}_s(2005)$, the true 2005 AADT on the segment. We summarized the comparisons, across the set of coverage counts, in three ways, each of which can be thought of as being based on the absolute relative error (ARE). The ARE of the estimate is the absolute value of the difference between the estimated and true AADT, divided by the true AADT. The three summary measures we used were:

- the mean average relative error (MARE), that is the average of the AREs
- the proportion of times that the ARE was less than 10%, a target that has been suggested for AADT estimation (AASHTO, 1992)
- the proportion of times that the ARE of one of the estimates ($AADT^{CGI}$, $AADT^{CG}$, or $AADT^{C}$) was lower than the ARE of another specified estimate (which is equivalent to the proportion of times that one estimated AADT was closer to the true AADT than the other estimated AADT); this proportion of times was conducted for each pair of estimates -- $AADT^{CGI}$ versus $AADT^{CG}$; $AADT^{CGI}$ versus $AADT^{C}$; and $AADT^{CG}$ versus $AADT^{C}$.

The results of the summary measures are presented in the Findings section.

### *Thrust 3: Gain the institutional acceptance of the approach*

This project was motivated by previous work that provided strong evidence that our proposed $AADT^{CGI}$ estimate would produce better accuracy in AADT estimation. In addition, by using $AADT^{CGI}$, the present level of accuracy obtained in AADT estimation could conceivably be achieved with fewer costly and dangerous ground counts. Since the concept is one of taking advantage of already existing imagery, the proposed approach could potentially both improve accuracy and reduce cost. However, the approach would have to be used on a regular basis to achieve these benefits. Regular use would require acceptance by both state DOTs, who would use the approach, and FHWA, who would be willing to accept the results.

In an attempt to improve the likelihood of use, we attempted to interest the Ohio DOT in our approach to the extent that we would be able to develop our prototype system at the agency for testing and demonstration in future years. We also wished to present our numerical results and vision for the operational system at technical conferences. In addition, the Mid-Ohio Planning Commission, the local Metropolitan Planning Agency in the Central Ohio region, is developing a system that accepts traffic counts from multiple sources and allows designated users access to these counts. We wished to investigate the potential of this system for accepting our AADT estimates.

# CHAPTER 3. FINDINGS AND CONCLUSIONS

## *3.1. Findings*

We present the findings by thrust in this section.

### *Thrust 1: Develop a prototype software system*

We were successful in developing the components of the software system to digitize the appropriate information from a digital aerial image and combine the information with information obtained from traffic monitoring data bases to estimate the $AADT^{CGI}$ measure. We digitized the information in an image, determined the $AADT^{I}$ estimate, and integrated the estimate with the ground-based information to produce the $AADT^{CGI}$ estimate in approximately one minute per image. That is, the proposed, improved AADT estimate was obtained very efficiently. However, we stored important input parameters – such as segment speeds, hourly, seasonal, and growth factors, and estimated variances – in such a way that they were exogenously linked to the imaged segment. In an operational system, the information would be linked to the imaged segment by automatically registering the imaged segment to a geospatial database that contains the appropriate representation of the segment in the traffic monitoring databases. This is a subject for our proposed following year project.

### *Thrust 2: Further demonstrate improved accuracy of the proposed $AADT^{CGI}$ estimate*

As discussed in the methodology section, we quantified the performance of the estimate $AADT^{CGI}$ that combines image-based information with traditional ground-based data by comparing its absolute relative error (ARE) to the AREs of two traditional estimates produced using ground-based data only: $AADT^{CG}$, which accounts for growth between the time when the coverage counts were obtained and the year in which the AADT is to be estimated, and $AADT^{C}$, which ignores the potential growth. We summarized the comparisons by the mean absolute relative error (MARE), the proportion of times the AREs were less than 10%, and the proportion of times $AADT^{CGI}$ was better than the other estimate considered.

The results are presented in Table 3a and 3b. The tables indicate the year in which the coverage counts were obtained and present the results by image. To illustrate the table, the first two rows of data correspond, respectively, to images 11273-1-1 and 11273-1-2. These were two images, taken a few seconds apart, on the roadway segment containing ATR 767. Similarly, consecutive pairs of rows continue to correspond to pairs of images on the same roadway segment, with every new consecutive pair corresponding to a new roadway segment. The "# CC Pairs" column indicates the number of coverage counts pairs obtained from the ATR data, as explained above. The “$\text{AADT}^{I}(2005)$” column provides the AADT estimate obtained from the image using equation (4)). The “$\text{AADT}^{true}(2005)$” column provides the estimated true AADT, obtained from the ATR data, for 2005, the target year of estimation in this study.

| Image # | ATR # | # CC Pairs | $AADT^{I}(2005)$ | $AADT^{true}(2005)$ | Mean Average Relative Error, MARE | | |
|---|---|---|---|---|---|---|---|
| | | | | | $AADT^{C}$ | $AADT^{CG}$ | $AADT^{CGI}$ |
| 11273-1-1 | 767 | 296 | 24012 | 31517 | 0.072 | 0.067 | 0.054 |
| 11273-1-2 | 767 | 296 | 33214 | 31517 | 0.072 | 0.067 | 0.063 |
| 11273-2-3 | 707 | 277 | 27324 | 30695 | 0.066 | 0.063 | 0.058 |
| 11273-2-4 | 707 | 277 | 25057 | 30695 | 0.066 | 0.063 | 0.059 |
| 11273-4-5 | 601 | 136 | 43141 | 48902 | 0.060 | 0.055 | 0.043 |
| 11273-4-6 | 601 | 136 | 45948 | 48902 | 0.060 | 0.055 | 0.046 |
| 11273-5-7 | 752 | n/a | n/a | n/a | n/a | n/a | n/a |
| 11273-5-8 | 752 | n/a | n/a | n/a | n/a | n/a | n/a |
| 11273-3-9 | 121 | 158 | 137436 | 122854 | 0.049 | 0.053 | 0.042 |
| 11273-3-10 | 121 | 158 | 132055 | 122854 | 0.049 | 0.053 | 0.043 |
| 11273-6-11 | 140 | 340 | 72721 | 59118 | 0.067 | 0.066 | 0.062 |
| 11273-6-12 | 140 | 340 | 53252 | 59118 | 0.067 | 0.066 | 0.060 |

Table 3a. Empirical results of empirical study for 2003 ATR Data

| Image # | Proportion ARE Less than 10% | | | Pairwise Comparisons of ARE | | |
|---|---|---|---|---|---|---|
| | $AADT^{C}$ | $AADT^{CG}$ | $AADT^{CGI}$ | CG vs C | CGI vs C | CGI vs CG |
| 11273-1-1 | 0.720 | 0.767 | 0.841 | 0.787 | 0.723 | 0.699 |
| 11273-1-2 | 0.720 | 0.767 | 0.780 | 0.787 | 0.936 | 0.649 |
| 11273-2-3 | 0.801 | 0.816 | 0.852 | 0.588 | 0.578 | 0.617 |
| 11273-2-4 | 0.801 | 0.816 | 0.856 | 0.588 | 0.567 | 0.534 |
| 11273-4-5 | 0.853 | 0.853 | 0.897 | 0.897 | 0.890 | 0.882 |
| 11273-4-6 | 0.853 | 0.853 | 0.882 | 0.897 | 0.897 | 0.890 |
| 11273-5-7 | n/a | n/a | n/a | n/a | n/a | n/a |
| 11273-5-8 | n/a | n/a | n/a | n/a | n/a | n/a |
| 11273-3-9 | 0.905 | 0.873 | 0.968 | 0.386 | 0.968 | 0.620 |
| 11273-3-10 | 0.905 | 0.873 | 0.956 | 0.386 | 0.772 | 0.703 |
| 11273-6-11 | 0.794 | 0.815 | 0.812 | 0.524 | 0.571 | 0.447 |
| 11273-6-12 | 0.794 | 0.815 | 0.862 | 0.524 | 0.506 | 0.524 |

Table 3a (cont.)

| Image # | ATR # | # CC Pairs | $AADT^{I}(2005)$ | $AADT^{true}(2005)$ | Mean Average Relative Error, MARE | | |
|---|---|---|---|---|---|---|---|
| | | | | | $AADT^{C}$ | $AADT^{CG}$ | $AADT^{CGI}$ |
| 11273-1-1 | 767 | 335 | 24012 | 31517 | 0.070 | 0.058 | 0.043 |
| 11273-1-2 | 767 | 335 | 33214 | 31517 | 0.070 | 0.058 | 0.054 |
| 11273-2-3 | 707 | 318 | 27324 | 30695 | 0.073 | 0.066 | 0.054 |
| 11273-2-4 | 707 | 318 | 25057 | 30695 | 0.073 | 0.066 | 0.056 |
| 11273-4-5 | 601 | 295 | 43141 | 48902 | 0.041 | 0.043 | 0.044 |
| 11273-4-6 | 601 | 295 | 45948 | 48902 | 0.041 | 0.043 | 0.041 |
| 11273-5-7 | 752 | 339 | 89270 | 92313 | 0.082 | 0.068 | 0.056 |
| 11273-5-8 | 752 | 339 | 83386 | 92313 | 0.082 | 0.068 | 0.049 |
| 11273-3-9 | 121 | 316 | 137436 | 122854 | 0.046 | 0.038 | 0.037 |
| 11273-3-10 | 121 | 316 | 132055 | 122854 | 0.046 | 0.038 | 0.035 |
| 11273-6-11 | 140 | 339 | 72721 | 59118 | 0.067 | 0.061 | 0.061 |
| 11273-6-12 | 140 | 339 | 53252 | 59118 | 0.067 | 0.061 | 0.054 |

Table 3b. Empirical results of empirical study for 2004 ATR Data

| Image # | Proportion ARE Less than 10% | | | Pairwise Comparisons of ARE | | |
|---|---|---|---|---|---|---|
| | $AADT^{C}$ | $AADT^{CG}$ | $AADT^{CGI}$ | CG vs C | CGI vs C | CGI vs CG |
| 11273-1-1 | 0.743 | 0.812 | 0.925 | 0.800 | 0.701 | 0.645 |
| 11273-1-2 | 0.743 | 0.812 | 0.851 | 0.800 | 0.842 | 0.591 |
| 11273-2-3 | 0.733 | 0.796 | 0.865 | 0.619 | 0.604 | 0.642 |
| 11273-2-4 | 0.733 | 0.796 | 0.852 | 0.619 | 0.566 | 0.509 |
| 11273-4-5 | 0.939 | 0.942 | 0.939 | 0.441 | 0.393 | 0.315 |
| 11273-4-6 | 0.939 | 0.942 | 0.946 | 0.441 | 0.437 | 0.431 |
| 11273-5-7 | 0.732 | 0.835 | 0.894 | 0.929 | 0.944 | 0.941 |
| 11273-5-8 | 0.732 | 0.835 | 0.914 | 0.929 | 0.929 | 0.912 |
| 11273-3-9 | 0.905 | 0.962 | 0.965 | 0.649 | 0.870 | 0.386 |
| 11273-3-10 | 0.905 | 0.962 | 0.968 | 0.649 | 0.741 | 0.487 |
| 11273-6-11 | 0.785 | 0.829 | 0.805 | 0.596 | 0.614 | 0.398 |
| 11273-6-12 | 0.785 | 0.829 | 0.891 | 0.596 | 0.558 | 0.543 |

Table 3b (cont.)

The next three columns provide the MARE values when using the $AADT^{C}$, $AADT^{CG}$, and $AADT^{CGI}$ estimates. We see that the proposed $AADT^{CGI}$ estimate produced the lowest MARE for all cases considered except when using the first image on ATR segment 601 (*i.e.,* image 11273-4-5) and considering coverage counts obtained in 2004. (It is interesting to note that $AADT^{C}$ produced a lower MARE than $AADT^{CG}$ on this segment, indicating that on this segment it would have been better not to have used a growth factor. Of course, one would not be able to know this when dealing with real coverage counts in practice, and this observation highlights the variable nature of the estimation.) We note that the $AADT^{CGI}$ estimate produced the lowest MARE when using the second image of this same segment (*i.e.*, image 11273-4-6) taken only a few seconds later. Moreover, when using the 2003 ATR data to generate the coverage counts, the $AADT^{CGI}$ estimate produced the lowest MARE for all cases, including when considering image 11273-4-5.

As mentioned above, the images of the same segment taken a few seconds apart cannot be considered to produce independent results, since some of the same vehicles would appear in both images. However, the MARE results obtained with image 11273-4-5 and the 2004 generated coverage count data illustrate that the effect of the generated coverage counts is important and that the results obtained from the pair of images taken a few seconds apart, although not considered independent, can be very different. As such, it is noteworthy that, based on MARE, $AADT^{CGI\,I}$ improved estimation in 21 of the 22 “image-coverage count year” cases.

The following three columns provide the proportion of AREs that were less than 10% when using the $AADT^{C}$, $AADT^{CG}$, and $AADT^{CGI}$ estimates. The $AADT^{CGI}$ estimate produced the best results (highest proportion) on this measure in 19 of the 22 “image-coverage count year” cases. The $AADT^{CG}$ estimate produced slightly higher proportions when using image 11273-4-5 in the 2004 coverage count year and when using image 11273-6-11 in both the 2004 and 2003 coverage count years. It is interesting to note that the $AADT^{CGI}$ image produced a higher proportion than the $AADT^{CG}$ estimate for the “companion” image, taken a few seconds later on these segments, indicating that for any one image, the proposed estimate *may* degrade performance, but that quality should improve when considered over the entire system. (One can also speculate that, since images are often taken in a sequence, one could combine the multiple images to improve performance even further.) The $AADT^{C}$ estimate never produced a proportion of AREs less than 10% that was higher than that produced by the $AADT^{CGI}$ estimate, although it produced an equal proportion when using image 11273-4-5 in the 2004 coverage count year.

The final three columns indicate, respectively, the proportion of generated coverage counts for which $AADT^{CG}$ outperformed $AADT^{C}$, $AADT^{CGI}$ outperformed $AADT^{C}$, and $AADT^{CGI}$ outperformed $AADT^{CG}$, where better performance is determined by a lower ARE. (A lower ARE is equivalent to producing an estimate closer to the true AADT, in terms of absolute value of difference from the true AADT.) A value greater than 0.5 represents that the estimate considered outperformed the estimate to which it was being compared more often than it was outperformed by the comparison estimate. Using this measure, the $AADT^{CGI}$ estimate did better than the $AADT^{CG}$ estimate in 15 of the 22 “image-coverage count year” cases and better than the $AADT^{CG}$ estimate in 18 of the 22 “image-coverage count year” cases. Based on this measure, the $AADT^{CGI}$ estimate was again the best estimate, but the number of “image-coverage count year” cases (15 out of 22 cases) in which it performed better than the $AADT^{CG}$ estimate was lower than for the MARE measure (21 out of 22 cases) or for the “ARE less than 10%” measure (19 out of 22 cases). This seems to imply that the magnitude of improved accuracy offered by $AADT^{CGI}$ when an $AADT^{CGI}$ estimate performed better than an $AADT^{CG}$ estimate was greater than the magnitude of the degradation in accuracy accompanying $AADT^{CGI}$ when $AADT^{CGI}$ did worse than $AADT^{CG}$.

*Thrust 3: Gain the institutional acceptance of the approach*

We had several meetings individually and jointly with ODOT heads of traffic monitoring, aerial engineering, and GIS. We presented our concepts and results at these meetings and received their agreement to work with us in implementing our approach for testing on the ODOT system in a year two project.

Through discussions with MORPC personnel, it appears that the database they are developing to store traffic counts from multiple sources will not be appropriate at this time for receiving image based AADT estimates. Our approach is based on *combining* the image-based estimate with a ground-based estimate to produce a better overall estimate, whereas the MORPC system is more of a clearinghouse of count data. The image-based estimate $AADT^I$ is not a count. Although it produces an estimated volume that could be considered equivalent to a traffic count (McCord, *et al.*, 2003), the duration of the equivalent traffic count would be of such short duration to make it incommensurate with other counts in the database.

In the spirit of increasing the potential of institutional acceptance, we presented, "Exploiting Traffic Information in Airborne Imagery to Improve AADT Estimates – Empirical Results and Prototype Software System," at the North American Traffic Monitoring Exhibition and Conference (NATMEC) in Washington, DC, in August 2008.

### *3.2. Conclusions*

The work reported here supports the potential of using air photos that exist in state DOT databases to improve AADT estimates. The empirical studies we conducted with ODOT imagery and ground-based data provides more direct support of the improved accuracy of the proposed $AADT^{CGI}$ estimate, compared to existing estimates based on ground-based data only. The ability to produce the $AADT^{CGI}$ estimates with the software components we developed in this project supports the potential of this approach to be an operational one. To move toward implementation, it would be important to implement the components at a state DOT. We intend to do this in the upcoming year at the Ohio Department of Transportation.

# REFERENCES

AASHTO (1992), *AASHTO Guidelines for Traffic Data Programs,* AASHTO, Washington, D.C.

FHWA (2001). *Traffic Monitoring Guide*. U.S. DOT, Publication No. FHWA-PL-01-021.

Jiang, Z. McCord, M., and Goel, P. (2006), "Improved AADT Estimation by Combining Information in Image- and Ground-Based Traffic Data," *Journal of Transportation Engineering*, Vol. 132, Issue 7, pp. 523-530.

McCord, M., Yang, Y., Jiang, Z., Coifman, B., and Goel, P. (2003a). "Estimating AADT from Satellite Imagery and Air Photos: Empirical Results." *Transportation Research Record*, *1855*, Transportation Research Board, Washington, D.C., pp. 136-142.

Zhou, F., "Software for incorporating image information in AADT estimation," Report submitted in partial fulfillment of M.S. degree, Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, September 2008.

# APPENDIX A. IMAGES USED IN EMPIRICAL STUDY

Image # 11273-1-1; ATR #767

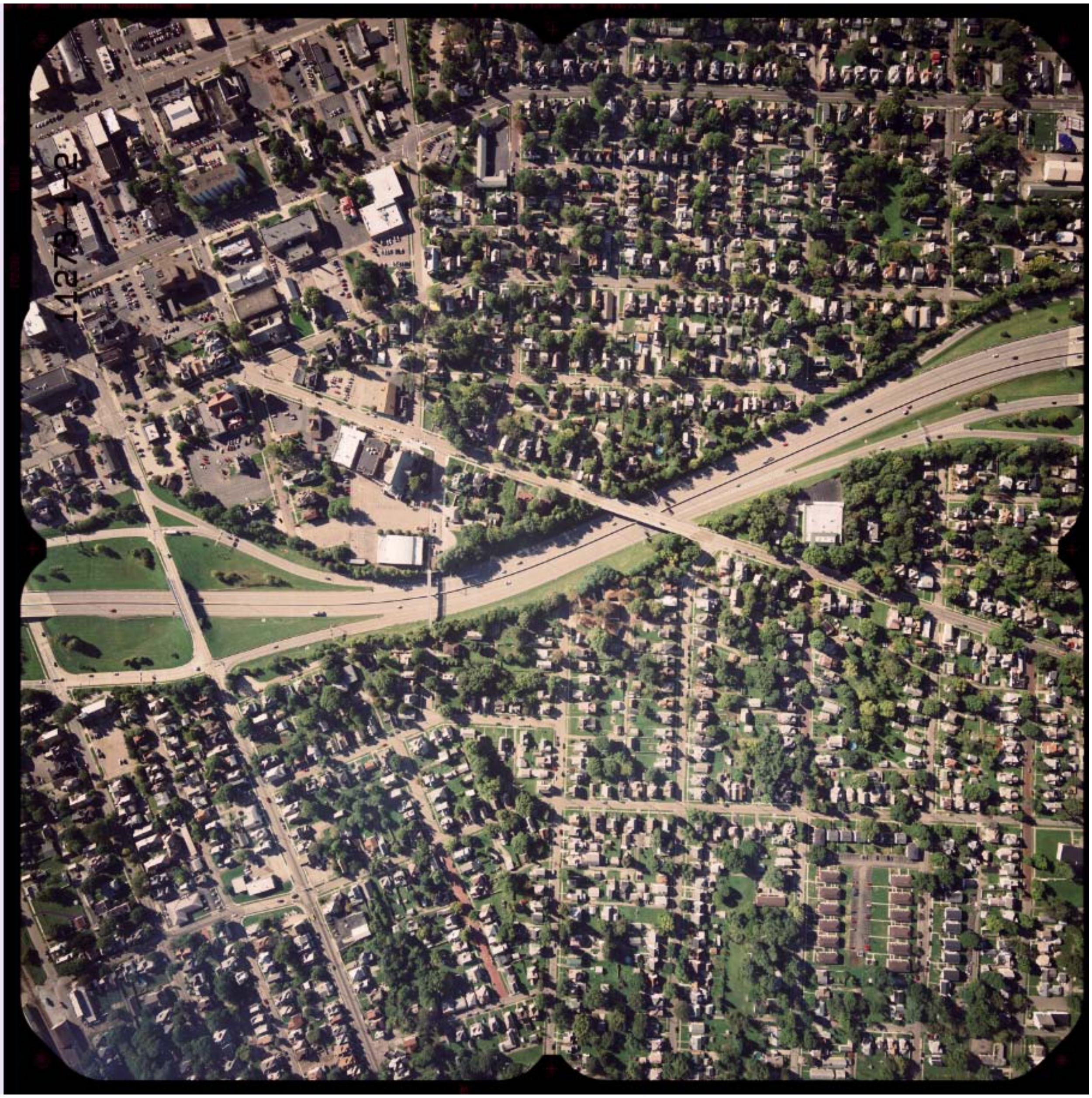

Image #11273-1-2, ATR#767

Image #11273-2-3, ATR #707

Image #11273-2-4, ATR #707

Image #11273-4-5, ATR #601

Image #11273-4-6, ATR #601

Image #11273-5-7, ATR #752

Image #11273-5-8, ATR #752

Image #11273-3-9, ATR #121

Image #11273-3-10, ATR #121

Image #11273-6-11, ATR #140

Image #11273-6-12, ATR #140