Novel ITS based on Space-Air-Ground Collected Big-data

Gang Xiong, Fenghua Zhu, Haisheng Fan, Xisong Dong, Wenwen Kang, and Teng Teng

Abstract— Based on the big-data collected from Space-Air-Ground, novel ITS (Intelligent Transportation System) is researched, including key technologies such as sensor data acquisition, dynamic data transmission, mass data storage, multi-source data fusion, data mining & analysis etc. On this basis, novel intelligent transportation cloud platform is designed, which includes Space-Air-Ground big-data acquisition & transmission sub-system, cloud computing supporting platform, intelligent transportation application & service sub-system. With the help of the data visualization, data prediction, and decision making, the whole big-data including people (passenger, driver), car, road traffic environment, can create the core values. To provide transportation data services for enterprise users and business users, such as customized mining, and specific industry analysis; To provide accurate transportation information services for the public; To provide business model for all levels of users, such as data visualization and customizing service.

I. INTRODUCTION

Most traditional methods of traffic data collection rely mainly on ground data. In fact, big-data collected from Space-Air-Ground can improve current ITS (Intelligent Transportation System) further. Aerial imagery sensors can provide sufficient resolution to sense vehicle location and movements across broader spatial and temporal scales. Digital video, Global positioning systems (GPS), and automated image processing are used to improve the spatial coverage, accuracy and cost effectiveness of data collection [1]. By remote sensors, high-resolution images are available and distribute their merit for traffic problems including illegal parking [2]. With high-resolution monochrome images collected from helicopter, 98% of the vehicles could be detected and tracked automatically [3]. Traffic Information Management System (TIMS) was designed and developed, which can be applied to traffic control, vehicle guidance and dispatching, route planning and infrastructure construction [4]. An image fusion method is proposed with Bilinear Resampling Wavelet (BRW) transform, it has good performance for preserving the spectral and spatial resolutions for remote sensing images, with lowest loss of spectral information [5]. Zhiqiang Xiao proposed a method to update road network in GIS automatically by use of the remote sensing imagery [6]. Qingquan Li proposed a method to detect and track vehicles based on airship video and calculate traffic parameters in real-time [7]. The fixed point theory is a

favorable theoretical background for the network organization, dynamic characteristic and regularity of Space-Air-Ground integrated network [8]. A multifunctional transceiver for future ITSs is proposed with two operation modes, namely, radar (sensing) mode and radio (communication) mode [9].

Based on big-data collected from Space-Air-Ground, the novel ITS is researched in this paper. In Section II, their current situation and future trend of related R&D areas are analyzed. In section III, its main research contents are summarized. In section IV, its technical solution is designed and main research contents are summarized. In section V, its potential benefits are analyzed, and conclusions are drawn out.

II. RELATED R&D AREAS: THEIR CURRENT SITUATION AND FUTURE TREND

A. Big-data

In recent years, the rapid growth of data has been both a great challenge and chance for many industries. It brings the information society into the big-data era. Big-data are generally referred to those data sets on which the perception, acquisition, management, processing and serving are out of the capabilities of common machines and hardware & software tools in limited time. IDC Company reports that the total amount of global data is 1.8ZB, and it is predicted to be 35ZB in 2020.

Big-data has its special functions in protecting digital sovereignty, maintaining society stability, and promoting sustainable development of the society and economy. In information era, the national competitiveness is partly reflected by the scale and liveness of big-data and the ability to explain and utilize the data. For any one country, falling behind in area of big-data means falling off the strategic commanding heights of the related Hi-tech industries. For this reason, the American government concord six departments and investigated 200 million dollars to launch the "Big-data R&D Plan". In this plan the NSF (National Science Foundation) proposed to form a unique subject covering mathematics, statistics and computation algorithms.

To prepare for the development trends of big-data and provide better data analysis service for enterprises and personal users, it is urgently needed to build different kinds of big-data platforms to meet the various requirements of different users. Unlike traditional data platforms, during the building of a big-data platform, the characteristics of big-data, including immense scale, varied types, rapid fluxion, dynamic hierarchy and huge value, should be taken into special considerations. Besides, problems like sorted storage of data, openness of data platforms, intelligent processing of the data, and interactivity between users and the data platforms make great challenges. Despite its difficulty, a few companies have made their representative progress in this area, e.g. the Freebase of Google, the Probase from the Microsoft, and

Gang Xiong and Fenghua Zhu are with the Qingdao Academy of Intelligent Industries，Qingdao, Shandong, 266109, China Prof. Gang Xiong is the corresponding author (email: gang.xiong@ia.ac.cn).

Xisong Dong is with the Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.

Wenwen Kang is with the Cloud Computing Center, Chinese Academy of Sciences, Songshan Lake, Dongguan, 523808, China

Haisheng Fan and Teng Teng are with the ChinaRS Geoinformatics (Guangdong) Co., Ltd., Dongguan, 523808, China.

CNKI (China national knowledge internet). As to commercial big-data platforms, there are three typical ones: the big-data analysis platform Infosphere from IBM, the uniform data management platform Teradata from the Teradata Corporation, and the Chinese first electric commerce cloud platform "Jushita" which is jointly built by the Tmall, Aliyun cloud engine and HiChina..

B. Remote Sensing Spatial Information

With the development of remote sensing technology from satellites, its application is becoming broader and deeper in many industries. It is developing towards those characteristics such as multi-sensor, multi-platform, multi-angle and high resolution of space, spectrum, time and radiation. Miniaturization of the earth observation systems, networking of the satellites, all-time and all-weather observation of the earth have been the three main development directions. America plans to execute 17 new satellite projects between 2010 and 2020, covering all the fields of earth sciences. China will send about 14 high-resolution satellites as a major special project of building the high-resolution earth observation system, which will be integrated with other observation methods, enabling all-time, all-weather and global covering observation of the earth.

Nowadays remote sensing applications have been switched from qualitative analysis to quantitative analysis. Satellite remote sensing data have been the main data source for updating of topographic map of 1:50,000 or even smaller measuring scale. Products of remote sensing data with different resolving and spectral characteristics coexist, which provides information insurance for resource management and disaster response. As to processing of remote sensing data, there exist a few parallel processing systems, for example the Pixel Factory System from France and the GeoImaging Accelerator GXL System from the PCI Company of America.

III. MAIN RESEARCH CONTENT OF NOVEL ITS

The main aspects of this section are shown in Fig. 1, including ITS big-data collecting, ITS cloud computing supporting platform and ITS big-data application & service platform.
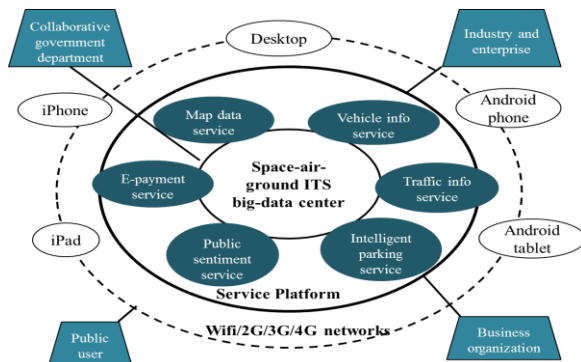

Figure 1. Main contents of novel ITS.

A. ITS Big-data Collecting

ITS big-data are collected from multiple sources, including remote sensing data from satellites, aerial photography, unmanned aerial vehicles, and ground ITS data from videos, IC cards, inductive loops and cellphones. These data are related to the three elements — people, vehicle and road — of transportation systems. People data mainly describe driving and paying behavior of drivers and travel behavior of passengers. Vehicle data mainly describe the basic information, real-time location, operation, and crowdsourcing road conditions. Road data mainly describe the geometry of road networks and characteristics of road infrastructures.

1) Public transit operation data

Real-time passenger data about the time and station of onboard/off-board can be obtained through on-bus QR code payment system based on 3G/WIFI wireless network. Based on these data, real-time intelligent bus schedule and scientific bus line planning can be achieved. And in return passengers can benefit from the more safe, comfortable and convenient public transportation service. Since more passengers are attracted to the public transit, the traffic congestion will be relieved. In summary, the public transit operation data have great potential of social and economic benefits.

2) On-vehicle terminal data

Firstly intelligent wireless on-vehicle terminals are developed. Efforts are focused on the development of functional, logical and physical system architectures by integrating the monitor, navigation, sensor, communication and control units. These terminals are mainly used to the modern management of buses, taxis and other types of vehicles, including management of driving safety, operation, quality of service, centralized intelligent schedule, and electronic station boards etc. The on-vehicle terminal data can also meet different requirements of different users. The government needs the data to help with traffic control, traffic law enforcement, transit organization and emergency responses. While enterprises need the data for logistics information and passengers for trip guidance.

3) Crowdsourcing road condition data

Improving traffic conditions and finally implementing ITS need huge and complex transportation data. Among them the tremendous, accurate, crowdsourcing data compose an important part. With the popularity of social media like online communities, blogs, microblogs and social networks, netizens are joined together to share the real world information. This provides an opportunity of obtaining road condition data from crowdsources. Practices indicate that crowdsourcing road condition data have become an important supplementary part of real-time traffic information.

4) Intelligent parking data

With acceleration of urbanization and rapid increasing of car ownership, many problems arise. Besides bad traffic congestion and environment population, parking and parking management problems becomes more and more outstanding. The problems can be solved from a management prospective. Information of parking space on road and in communities and parking lots is percept and shared through Internet of Things, 3G communication technology and cloud computing. Intelligent parking space perception and guiding platforms are developed on the background of building smart cities.

5) Spatial data

Spatial data are among the richest ones since they can be collected by earth observation satellites at much larger scale. Those satellites are equipped with broad spectral observation technologies and capable of high resolution photography. The highest resolution can reach 0.06 meter, making it available of highly refined spatial data. For areas requiring emergency response or timely protection, unmanned aerial vehicles are used to collect the local spatial data. By this means the emergency response time can be within 3 hours and spatial data collection can be finished within 3 days. Image data are mainly collected by ground photography and aerial photography, which are the original data for three-dimensional (3D) modeling. An example image can be seen in Fig. 2.

Figure 2. An example of aerial photography image.

B. ITS Cloud Computing Supporting Platform

The space-air-ground big-data service and other related services are implemented through cloud computing. The system hierarchy of the ITS cloud computing supporting platform is shown in Fig. 3. Based on cloud computing, the space-air-ground big-data are used to provide on-demand services. The system hierarchy is divided into three parts: the physical resource management, the logic resource management and the data center operation and maintenance.
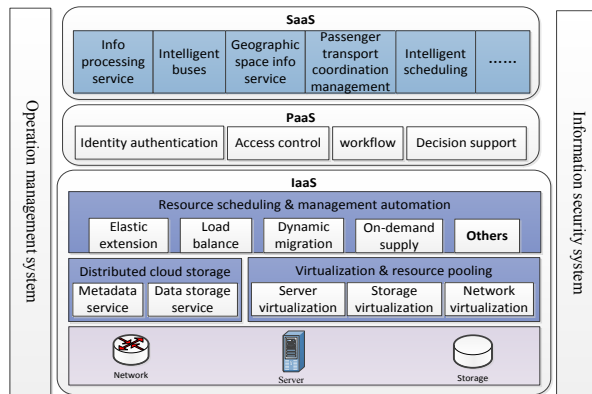

Figure 3. Hierarchy of the ITS cloud computing supporting platform.

The whole platform is composed of several sub-systems, among which each one provides specific services of information, management and surveillance. The cloud computing virtualization platform is introduced to ensure all-time stable and efficient services. By virtualization the application system and physical machine are separated, thus system service interruption resulting from physical breakdown will be reduced. The physical resources can also be deleted, upgraded or changed without negative impact on the users.

C. ITS Big-data Application & Service Platform

The ITS big-data application & service platform is built based on the ITS big-data and ITS cloud computing supporting platform, and adopts the "central data storage and processing" and "local application & service" mode. The real-time transportation information is grabbed from ITS big-data and analyzed together with the historical data. Intelligent prediction will be made on the platform to provide decision supports for users.

The ITS big-data application & service platform is mainly used to provide services for the government, enterprises and public users. The government can use the platform for management of transportation law enforcement. Multiple services are available, including accurate geo-information, traffic management, emergency response, on-road parking space management, and public transit supervision. Enterprises can use the platform for accurate geo-information, assistant decision making and commercial data analysis. Besides, mobile applications are developed for the public users to provide them transportation information services. The public trip behavior data can be gathered through the APPs. And in return the public users get improved services like accurate geo-information, real-time traffic condition, driving and parking guidance etc.

Based on this platform, the industry data, computation resources and characterized intelligence analysis results can be shared by different kinds of users. In this way, much system resource and cost will be saved, and at the same time the efficiency will be highly improved.

IV. TECHNICAL SOLUTION OF NOVEL ITS

A. The Space-air-ground Big-data Collection and Transmission Technology and On-vehicle Terminals

1) The Beidou/GPS dual-mode position & navigation technology

Not like GPS, the Beidou position & navigation system uses a kind of active dual-direction two-dimensional navigation technology. Accurate three dimensional position data are resolved by ground control center and then transmitted to users through the satellites, instead of directly resolved by the user devices. The position data include parameters like three dimensional location, velocity, time and posture and so on. To navigate a car, firstly the car's space location is detected by the dual-mode position system. And then the space location is mapped to the city map location. Furthermore, the control center can also provide user-defined distance reporting service with the differential distance algorithm. On the other hand, the GPS is used to supplement and enhance the Beidou signal resources, thus improving the speed and accuracy of the position system.

2) Integrated intelligent on-vehicle terminal

As shown in Fig. 4, the intelligent on-vehicle terminal is composed of hard disk, sound pick-up, monitor and positioner and so on. The driver can start up the terminal by either starting the engine or setting a timer. The wireless access point function is integrated with the terminal, which makes it possible to organize all the terminals in a wireless local area network (WLAN) based on Wi-Fi.
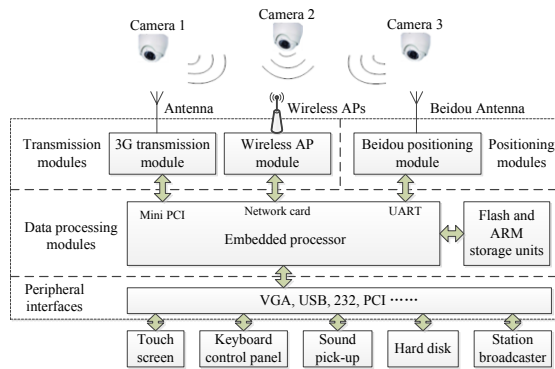
Figure 4. Structure of the intelligent on-vehicle terminal.

Vehicles are connected to the big-data cloud computing center through 3G network. Real-time information about the locations, operations and statuses of multiple vehicles is shared based on the big-data center. On the other hand, real-time video and audio data are compressed and transmitted to the surveillance center for further processing and analysis. A copy of the data is also stored in the on-vehicle hard disk, in case of communication interruption.

B. The Space-air-ground Big-data Fusion and Mining

1) Data fusion

The space-air-ground big-data come from multiple sources. Data from different sources must be fused first. The basic process of data fusion includes multi-source data collection, data preprocessing, data fusion and target parameters estimation. The data fusion is accomplished at three levels. The first is data level fusion, which finishes data preprocessing and association. The second is characteristic level fusion, which is supposed to predict the traffic parameters. The third is state level fusion, which is responsible to determine the transportation state based on the current traffic flow information.

2) Data mining

The knowledge we want is achieved through data mining technologies. Data mining is a multi-step process, including problem definition, data prepare, pattern recognition and model evaluation. The transportation data mining model is shown in Fig. 5.
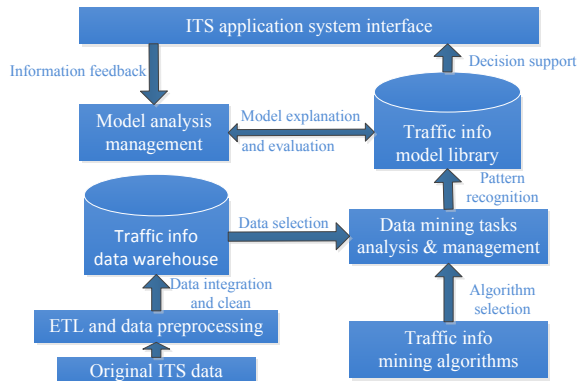

Figure 5. Transportation data mining model.

The ETL and data preprocessing provides a clean, consistent, integrated and reduced data set for pattern recognition of the transportation big-data. The transportation data mining model is hierarchical with four layers: the application layer, the analysis & logic layer, the algorithms & tools layer, and the data layer. The application layer provides the entrances for users to call functions of the lower layers. While the analysis capability of the system is reflected by the analysis & logic layer.

The algorithms & tools layer is a set of different algorithms and tools needed for traffic flow analysis, including data mining techniques, statistics methods and similarity measurement methods and so on. Besides the data mining techniques, some assistant methods are also very important. For example, the similarity measurement method is critical to the performance of a cluster method. Different similarity measurement methods should be implemented to accommodate to different data sets or analysis targets.

Prediction models and classification models are adopted as the data mining models. The prediction models have three layers: the basic data layer, the characteristic attributes layer and the state description layer. Different layers are used for different prediction targets. The basic data layer is used for traffic flow prediction. The characteristic attributes layer can be used for predictions of traffic flow properties, traffic accidents and their types, and traffic congestion. While the state description layer is used for predictions of network-wise level of service, as well as development trends and impact evaluation of traffic accidents.

Traffic flow can be classified by different characteristics. Thus before classification, traffic data should first be reduced to only keep those related characteristics. There exists research addressing the traffic flow classification problem by fuzzy logic method or by artificial neural network (ANN) method. The fuzzy logic classification model is described by a set of "if… then" rules, which are understandable. While by ANN method, a networked is trained to classify the data. But the trained ANN is often unexplainable.

C. The Space-air-ground Big-data Processing

1) Parallel receiving of massive data

The ITS contains massive data collection terminals. The servers are challenged with massive data receiving work. Efficient and stable solution can be achieved by parallel data receiving strategy. The terminals and collectors communicate to the masters regularly to report their states.

[1]. The terminal reports its online state and connection status with the collector to the master via TCP. If a terminal cannot connect to the collector, it will report the problem to the master, and the master will allocate another available collector to it.

[2]. The collector reports its load status and online state to the master. If the workload is too much, it will request the master to share some workload to other collectors.

[3]. If the master doesn't receive the reports of a collector for a specific interval, the collector is thought to be breaking down. Then the related terminals will be notified to send their data to other collectors.

[4]. When the data storage function breaks down, the collectors will cache the data temporarily. The cached data will be sent to the GT-data on the storage function recovers.

2) Segmental storage of massive small files

The GT-Data provides distributed storage for both small and large files. But the bottom implementations are different.

For large files, they are split into segments, which are stored in different nodes of the storage clusters. The file information and segmentation information are stored in the metadata. When reading a file, the indices of the segments of the file are firstly read from the metadata so the system knows where to read the segments. Finally these segments are merged to form the entire file.

While for small files, the same approach is not practical. Since the amount of small files is much larger, the size of the metadata will be too huge to be loaded in the memory. Besides, considerable storage space will be wasted. As is known, the file system has minimum space (e.g. 4KB) occupancy for each file. For files smaller they will still occupy the minimum space. So it's necessary to merge small files to large files before they are stored. In current practice, the GT-Data organizes small files as binary stream of key-value pairs into large Hfiles. The index information is built in the "-ROOT-" and ".META" index tables.

3) Duplication storage of massive data

All the data blocks of the GT-Data files have their copies. Parameters about the size and duplication coefficient of data blocks are user-defined. A GT-Data file is written as a whole at once. It is ensured that there is only one writer at any time.

The NameNode is in full charge of data block duplication. It receives heartbeats and block status reports from each DataNode periodically. A block status report includes a list of all the data blocks on the DataNode.

4) High-performance I/O of massive data

Within the GT-Data, data exchange is finished based on the wide-band network. The process of a specific data exchange is divided into data reading and data writing.

The process of data reading can be described as following:

- Clients or users open the needed file through the open() method of the file system.

- The file system calls the related NameNodes through predefined protocol to find the locations of the first few blocks of the file.

- The related Namenodes return an address including the data information to the client. Then the client starts data reading by creating a FSDataInputStream instance.

- According to the DataNode addresses of the first few blocks, the FSDataInputStream instance connects to the nearest DataNodes and reads data from the beginning of the file. The client calls the read() method repeatedly, reading data from the DataNodes as binary steams.

- When encountering the end of a data block, the FSDataInputStream instance will close the connection to the related DataNode, and find the optimal DataNode of the next data block.

- When data reading is finished, the client calls close() method to close the FSDataInputStream instance.

The process of data writing is similar to that of data reading, thus needs not be repeated here.

D. ITS Application of the Space-air-ground Big-data

1) Traffic infrastructure data extraction and real-time updating

Road elements, like the road surface, the green belt, ghost islands, and blind turning for large vehicles, can be extracted from images of resolution smaller than 1 meter. While from aerial images and obliquely viewed images with resolution smaller than 10 centimeters, information about the street lamps, well lids and guideboards can be extracted. These road elements could have all kinds of appearances in the images. Different roads have different structures and expressions. Roads of different grades have big differences in the scale. What's more, a lot of noises could appear in the landscape.

We decide to extract the road elements form high-resolution remote sensing images. The images are analyzed from the perspective of spatial resolution. The object oriented classification method for remote sensing images is adopted. Firstly the images are segmented to suppress the noises and objects in the images are detected. Then the target scale is carefully selected, to which the images will be transmitted. After the transmission, the object hierarchy in the image is built. After that, the road information is extracted based on the pre-built road knowledge-base, in which characteristics of different roads are described. Finally, the road lines are automatically or semi-automatically recognized based on the shapes of the green belts, plant areas, and mathematical morphology. The areas of the road surfaces will also be calculated through GIS.

The real-time aerial photography will capture any changes to the urban roads. The changes will be notified and reflected immediately by updating the database in the ITS data center automatically. In this way it is ensured that the users can always get the most accurate road information.

2) Live-action 3D navigation and intelligent prewarning

In intelligent transportation navigation systems, the traditional two-dimensional virtual navigation is substituted by live-action 3D navigation, where the virtual scenes are placed with real screen ages. An example of the live-action navigation map is shown in Fig. 6. The 3D navigation map is not simply the three dimension display on the two-dimensional navigation map. It is implemented through the information communication technology based on the 3D spatial data obtained. According to the collected real world image data, the scene details like the color, material, texture and lightness are carefully rendered. The live-action 3D scenes of the roads are displayed on the navigator. The

live-action 3D navigation map can break the constraints of traditional two-dimensional map on the expression forms. The real world now can be expressed in more details and in a more comprehensive way.

Figure 6. The live-action navigation map.

For areas with a higher accident rate, like crossing and curves, an intelligent prewarning system is built. The blind corners for large vehicles at the curves are calculated based on the high definition cameras and geometric calculation models. According to the length of the vehicle and the performance index, the safety evaluation result is presented to the driver before he arrives at the curve. The areas with high accident rate are highlighted in the live-action 3D navigation map to warn the drivers. With the intelligent prewarning system, drivers can go through the dangerous areas safely and many traffic accidents will be avoided.

3) Driver behavior analysis and prewarning based on the driving big-data

The received data packages are processed differently according to their classes such that we can achieve all-time surveillance, alarming, commanding and operation of the on-road vehicles. The collected spatial coordination data and timestamps of the vehicles are mapped on the digital map, so that we can know the real-time locations and statues of the vehicles. The trajectories of the vehicles are shown in the GIS, and vehicle tracking can be achieved.

Based on the massive driving data and driver behavior data, the drivers' behavior can be efficiently modeled with some statistical analysis methods. The modeling results can be applied to a lot of industries, including the driving behavior correction system for new drivers, the automatic diagnostic system of driver behaviors etc.

V. CONCLUSIONS

In the big-data era, massive data of different kinds are available now, including the remote sensing data, traffic video data, all kinds of perception data and public sentiment data and so on. These data are collected from decentralized sources and used separately in traditional applications. Since each dataset can only describe one aspect of the real world, these traditional applications are not able to utilize the value of the data to the most. We propose to build the space-air-ground big-data insurance system in this paper. The data are able to be updated from perspectives of scale, accuracy and time phase. An accurate and timely data service of dynamic surveillance system can be further implemented.

An ITS service of the mobile internet age is supported by the cloud computing platforms and high speed network transmission. The service is implemented through data collection by mobile clients, analysis and knowledge mining of big-data, high speed network transmission, and intelligent push services. Based on the real time data, users can benefit a lot from the more accurate navigation and parking services, as the new interconnected transportation service mode does. The administrative barriers are eliminated thanks to the information technology. The transportation data from different sources can be integrated and shared on a uniform platform to provide transportation administrative service for the government, and transportation information services for the enterprises and public users. With the help of Novel ITS, the level of service of the transportation system will be further improved.

REFERENCES

[1] Alejandm ANGEL, Mark HICKMAN, Pitu MIRCHANDANI, Dmesh CHANDNANI, "Methods of traffic data collection using aerial video," The IEEE 5th International Conference on Intelligent Transportation Systems, 3 - 6 September 2002, Singapore.

[2] S. Pantavungkour, R. Shibasaki, "Three line scanner, modern airborne sensor and algorithm of vehicle detection along mega-city street," 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, pp. 263-267, 2003.

[3] M. Schreuder.S. P. Hoogendoorn, H. J. Van Znylen, B. Gorte, and G. Vosselman, "Traffic data collection from aerial imagery," International Conference on Intelligent Transportation Systems, vol.1, pp. 779 - 784 2004.

[4] X. W. Chen, L. Ma, F. Z. Zhang and D. K. Yang, "Traffic data collection and management based on spatial information technology," International Conference on Intelligent Transportation Systems, vol.1,pp. 789-792. 2003.

[5] Q. X. Zhou, Z. L. Jing, S. Z. Jiang, "Remote sensing image fusion for different spectral and spatial resolutions with bilinear resampling wavelet transforms," International Conference on Intelligent Transportation Systems, pp. 1206 - 1213. vol.2, 2003.

[6] Z. Q. Xiao, X. Q. Jiang, G. S. Bao, J. X. Huang, L. Lu, "Update of road network in GIS by fusing multi-sensor imagery," IEEE International Conference on Geoscience and Remote Sensing Symposium, pp. 4969 - 4971, 2005.

[7] Q. Q. Li, B. Lei,Y. Yu, R. Hou, "Real-time highway traffic information extraction based on airborne video," 12th International IEEE Conference on Intelligent Transportation Systems, 2009.

[8] H. Liu, J. Zhang, L. L. Cheng, "Application examples of the network fixed point theory for space-air-ground integrated communication network," International Congress on Ultra Modern Telecommunications and Control Systems and Workshops. 2010.

[9] L. Han, K. Wu, "Multifunctional transceiver for future intelligent transportation systems," IEEE Transactions on Microwave Theory and Techniques, VOL. 59, NO. 7, JULY 2011 1879.