

INTELLIGENT TRAFFIC SYSTEM: COOPERATION OF MANET AND IMAGE PROCESSING

This paper presents a model for intelligent traffic system. Proposed methods describe ideas of information computing from images and its distribution to different nodes in traffic system. Two algorithms for foreground extraction forwarded by classification of extracted objects are discussed in the paper. These approaches serve as a basis for information computing. Communication framework over MANET is presented for efficient information delivery. Appropriate protocol is designed at first, and then some algorithms for efficient data delivery are discussed.

1. Introduction

The objective of this paper is to provide a view on an intelligent traffic system. It consists of information computing from images and distribution of this information to other nodes in the traffic network. Such system can be used in areas of safety transportation, reliability, cost reduction of transport etc. In the first part a possibility of obtaining suitable information is analyzed. In the second part a framework for efficient data delivery (how the message looks like, what the context is, etc.) is proposed. A possibility of integrating these different parts into one common system is also analyzed. The type of information, which is computed in the first part, depends mostly on the system deployment, but its general character could be:

- *Surveillance and security* – the protection of objects from accidents, vehicle crashes, pedestrian injuries, or just monitoring a region of interest (ROI) from unauthorized access. A situation which could be solved is, for example, the monitoring of rail crossing [1], entrances into companies etc.
- *Monitoring systems* – providing early information about current traffic situation (traffic jams, weather conditions, etc).

The process of information computing is based on image processing. Inputs are obtained from a camera positioned in a hover view (light pool). The camera is focused on ROI in which an action occurs).

Images or video sequences are processed with our system (software demo) and the result provides valuable information about the state of system in the given ROI. These procedures can be generalized as a task of *foreground extraction*. The information obtained from the image processing part needs to be delivered to the final customer and it should be accomplished in an effective way. For this purpose the communication framework for a mobile ad-hoc network is designed. The protocol satisfying all the necessary require-

ments for efficient data delivery (size, effectiveness, universality, etc.) is introduced.

Generalization can be made as follows: information is computed from a video sequence (type of information depends on the system deployment, for example: “alert → there is an object on railway crossing”). This data is encapsulated into the message and then it is distributed over a mobile ad-hoc network (MANET). Each receiving node initializes a decision process based on the importance of a received information (see Fig. 1).

In the second chapter of this work a system for foreground extraction is described. Two different methods are analyzed. The first method is a background modeling and the second one is a key point detector. The results from classification process are summarized in the third chapter. The 4th chapter proposes a protocol for information exchange between all the objects of C2X¹⁾ system. In the last chapter some important points for communication establishing and maintaining are outlined.

2. Foreground Extraction

Extraction of foreground objects is a fundamental task in the image processing (to distinguish object from a scene). If this is accomplished with a sufficient result, object recognition and classification, event detection could be performed. There are different methods. An overview can be found in [2], some of them work with one video frame – known are point detectors like SIFT [3], SURF (Speed-UP Robust Features) [4], or Moravec and Harris detectors. Other methods utilize a sequence of images to model background – Eigenbackground, Mixture of Gaussians [5], Background comparison etc. Generally all approaches need to cope with:

- Light and weather changes.
- Noise caused by trees, grass.

* Stefan Badura, Anton Lieskovsky

Faculty of Management Science and Informatics, University of Zilina, Slovakia, E-mail: Stefan.Badura@fri.uniza.sk

¹⁾ Car to Car. Car to Infrastructure

- Permanent changes in background (new car in parking), etc.

The SURF which will be discussed first is an invariant key point detector. We will estimate some key points that could represent the foreground. Next, we will discuss some methods for background modeling.

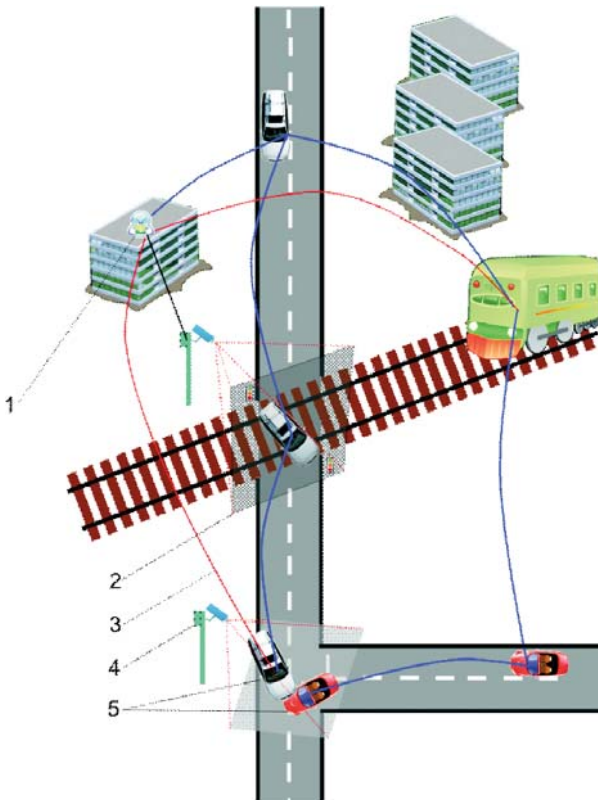


Fig. 1 Design of system deployment.

- 1 - Computational center (image processing).
- 2 - Possible representation of ROI. 3 - Communication between objects.
- 4 - Camera focused on affected ROI. 5 - Objects participated in communication

Invariant Key Point Detector

The SURF is a robust invariant key point detector. It is used for foreground extraction and its description with a set of feature vectors.

The algorithm is a three-step process:

1. Detection and localization of interest points.
2. Construction of feature vector - descriptor.
3. Descriptors matching.

We utilize just first step in our work - *interest point detection* and *localization*. This satisfies the problem of foreground extraction.

²⁾ All the algorithms were tested on a dataset which was captured from the entrance to a company. ROI covered an area where vehicles and persons passed. In the region there was also a barrier which brought noise into the system, therefore, our methods are adapted to this specific problem.

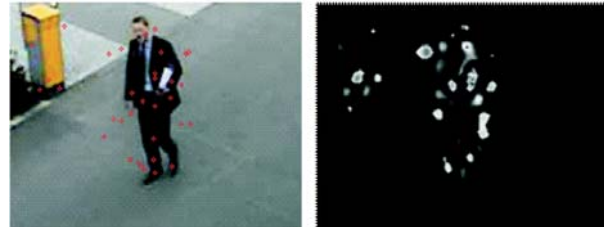


Fig. 2 Left image represents results from invariant key point detector (red circles are detected key points). Right image is an example of detected maxims in determinant matrix.

Detected interest points are considered as candidates for foreground.

If this method is applied on the surveillance system, the following consideration can be made: if the number of detected interest points is increasing, there is probably an object in the scene. This process uses Fast - Hessian features detector and it is based on determinant $D(H)$ of Hessian matrix $H(f)$. The Hessian matrix is the matrix of partial derivatives of two dimensional function $f(i,j)$ (see eq. 1)[4].

$$H(f) = \begin{bmatrix} \frac{d^2f}{di^2} & \frac{d^2f}{didj} \\ \frac{d^2f}{didj} & \frac{d^2f}{dj^2} \end{bmatrix} \quad \text{Hessian matrix} \quad (1)$$

$$D(H) = \frac{d^2f}{di^2} \cdot \frac{d^2f}{dj^2} - \left(\frac{d^2f}{didj} \right)^2 \quad \text{Deter. of } H \quad (2)$$

Equations 1 and 2 are defined for the continuous function f , but we work with an image, which means a discrete space. Each point represents pixel intensity for this image at (i,j) coordinates. Approximated derivatives are then computed by convolution with an appropriate kernel (H is also approximated - H'). For the determinant approximation (D') the following formula is used:

$$D'(H') = D_{ij} \cdot D_{ij} - (0.9D_{ij})^2 \quad (3)$$

D_{ii}, D_{jj}, D_{ij} - are block responses (derivatives) in image for pixels at (i,j) , convolved with suitable mask.

Determinant computation is necessary for interest point detection (which are maxims in determinant matrix - see Fig. 2). The next step of the proposed method is construction of the scale space by producing image pyramids. We'll stop here because we have information about the foreground (or better said - we have candidates). If we wanted to continue with other stages, we can obtain better results, but the computation time increases with any other step.

Background modeling

An example of background modeling is the Mixture of Gaussians. In our experiments this method did not seem very robust for the given problem²⁾.

We used a model based on a simple principle:

$$res = f(t) - f(t - n) \tag{4}$$

res - result, *f(t)* - image intensity in time *t*; *f(t - n)* - image intensity in past.

It is the *background comparison model*, simple and easy to compute. To ensure better resistance more background models are maintained. The *first* example (eqs. 5 and 6) is a modification of the basic model from eq. 4.

$$res = f(t) - ref(t - 1) \tag{5}$$

$$ref(t) = f(1 - \alpha) \cdot ref(t - 1) + \alpha \cdot f(t) \tag{6}$$

Ref(t) - reference image (running average of background pixel intensity); α - remembering constant

Another model: if pixel is executed *n*-times in a sequence as foreground, then it is moved to background. We designed a model where two simple approaches are mixed together for background maintenance. If pixel in a given time *t* is decided to be a foreground in *both* models, then it is executed as foreground for a given time, otherwise it belongs to background. Such an approach should cope with long term and also some temporary changes. This model is not such a strong tool as, for example, Mixture of Gaussian which can adapt to weather or light changes very quickly. But with our approach, we can, firstly, better manage background models and, secondly, our dataset was collected in a specific (real) situation, so it needs a unique model.

Two different methods for the problem of foreground extraction were compared each other. Background modeling methods and invariant key point detector. The primary criterion was the rate of correctly identified objects passing the ROI. Better results were obtained with background modeling - background comparison model. Errors occurred especially in the situation when there were shadows. The Mixture of Gaussians (MoG) is another well known method for background maintaining. Our model is in advance in remembering objects for longer time compared to the MoG. MoG and background comparison models work with video sequences, SURF does not when it works just with one frame (time dimension is not used). The SURF does not need to maintain the background model (there is no problem with forgetting objects or light changes). Weakness of this method is that low contrasted objects are recognized very poorly, for small regions the SURF does not have to recognize any objects (see Fig. 4), the whole process is also time consuming against background comparison model. SURF is more reliable in situation where known objects are searched.

3. Classification Process

The main goal of this part is to provide additional information about the type of objects in given ROI. This information is then added to the information computed in the 2nd part. This part is not so important; it is just an illustration of another way of computing some information from image that can increase an informational character of the system. Generally, in this part we are interested in three types of objects (pedestrians, personal cars, other type of

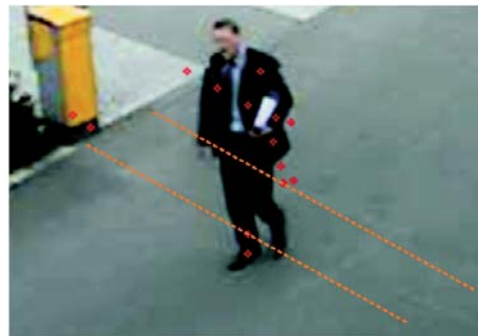


Fig. 3 An example when in ROI was an object, but it wasn't detected (ROI is the area between lines)

vehicles and false positives) - together four classes are distinguished. The highest score of positive object's identification moved up-to 92.5 %. Before classification, a successful foreground recognition has to be accomplished (previous part). If foreground candidates are known, ROI is shrunk just to a small region around object.

A simple classifier is used in our work based on the vector distance (MSE - Mean Square Error or Hamming distance). Each feature vector is computed from appropriate images. Two different inputs were compared. First input is the result from the foreground extraction process (frame subtraction, see Fig. 4) and the second one is obtained from edge images (see Fig. 6). Next, the vectors are computed from these images, their sizes move around 10-15 values. The character of these values is:

1. From *subtracted* images - eccentricity, percentage of pixels, centroids, compactness...
2. From *edge* images - number of single pixels, couples, triples, tetras, number of lines "-", "|", "/", "\", number of crosses "+", "x"...

These two types of vectors are used as input for our classifier³⁾. From the experiments it was clear that better results were achieved using the second type of inputs (edge images).

In previous text methods for information computing from images were introduced. At first foreground objects were extracted, secondly these objects were classified into classes. An example of a final information could be:

³⁾ All values need to be normalized.



Fig. 4 Results from background comparison model are shown, for various objects.

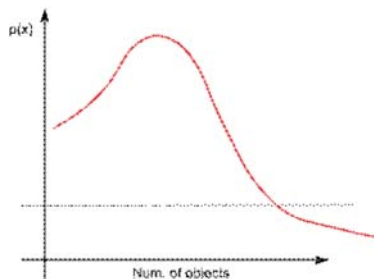


Fig. 5 Probability dependence of optimal message delivery and number of objects in communication is shown.

ALERT: On rail crossing is an object- it could be a car.

Next step is a delivery of this information to other traffic objects (a train approaching the affected rail crossing). The information delivery is discussed in next part.



Fig. 6 The image shows the process of edges extraction. Blur is applied on original image, then Sobel edge detector is used. Detected edges serve as input for classification process.

COMMUNICATION PROTOCOL

In this part we present framework for information representation. Common communication protocol is introduced. MANET (Mobile Ad-Hoc network) is used for information exchange. An important task in this type of network is to solve the problem how to represent information, what the message looks like and how to deliver the information. When the message is created, it is sent to MANET nodes which extract suitable information for themselves. The information from the message serves as basis for decision

process in different kinds of collision situations (as mentioned before – traffic jam, crashes etc.). The designed message should provide:

- Content for Emergency Warning Messages (content for information obtained from an image processing part).
- Generating next hop messages (messages that need to be transmitted to other vehicles).
- Ability to provide a message life cycle (registration, transfer, removal, etc.).
- Simplicity, universality etc.

Designed message is also restricted by physical characteristics of the system. There are limitations for example in a transmission rate, available bandwidth and transmission power of devices. Some implementation problems to be considered are:

- Message size (in kB).
- Insurance against unauthorized access to the message (it should be considered that a secure message can increase the message size).
- A small number of objects incorporated in communication can cause that the message will not be properly distributed.
- A lot of objects involved into communication can lead to congestion (see Fig. 6).

Content of the message

The common structure of the message in C2X –systems must provide:

- Autentification and identification (IP, MAC, secure certificate).
- Messages encoding.
- The size and the structure of the message must allow smart processing (compression).
- The possibility to use certification authorities.
- Information about state of the objects (position, velocity, direction, time stamp...).
- Information about *occurred events* (accident, emergency car passing, mist, etc.).
- Information about multimedia description of the message (image path or path to sound file).
- Covering information description:
- Source of the message (car, operation center, *computation center* hydrometeorological institute, etc.).
- Type of a distributed message (public, private).
- Communication language of the message (English, Slovak, etc.).
- Priority of the message (probability for certainty of given event).

Protocol definition

Possible solution is to define the message according the XML standards. This format ensures simple access regardless of used platform at which it runs. The message of this type (XML) fits the characteristic for a common communication protocol described earlier. The following list contains the definition of designed protocol:

- *identifier* – Unique number, or text string that identifies the message assigned to a sender.
- *sender* – Unique address identifying the creator of the message.
- *password* – Sender verification (secure function implementation).
- *source* – Identification of operator or device (vehicle).

- *sent* - Time and date of the sent message. Format is standardized in ISO 8601 (example: "2002-05-24T16:49:00-07:00").
- *scope* - Code for type of message distribution (Public, Restricted, Private).
- *msgType* - Type of the message (Alert, Update, Cancel, Ack, Error).
- *language* - Code of used language for the informational content. Code values standardized in RFC1766.
- *category* - (Security - military, legal, government, private, security, Geo - geophysics, Met - meteorological, etc.)
- *instruction* - Text of recommended action, for receiver.
- *headline* - Title of the message record.
- *GPS* - GPS data.
- Other data necessary to identify events.

Basic structure of the message is inspired from Common Alert Protocol [6]. Names of elements and their values satisfy communication standards to make the whole system functional.

Protocol implementation

The algorithm for message processing should meet the following requirements:

- Control of the environment.
- Receiving and evaluating messages.
- Process management and displaying messages.



Fig. 7 Example of simulation, for communication protocol

```
<alert xmlns = "http://www.incident.com/cap/1.0">
  <identifier>KSTO1055887203</identifier>
  <sender>KSTO@NWS.NOAA.GOV</sender>
  <sent>2010-01-17T14:57:00-07:00</sent>
  <status>Actual</status>
  <msgType>Alert</msgType>
  <scope>Public</scope>
```

```
...
<info>
  <category>Road</category>
  <event>RAIL CROSSING COLLISION</event>
  <urgency>Immediate</urgency>
  <severity>Severe</severity>
  <certainty>VeryLikely</certainty>
  <expires>2010-06-17T14:59:00-07:00</expires>
  <senderName>NATIONAL TRAFFIC SERVICE</senderName>
  <headline>ALERT : On rail crossing is an object.</headline>
  <description>VEHICLE STOPS ON RAIL
  CROSSING</description>
  <instruction>POSSIBLE DANGEROUS SITUATION. BE
  CAREFULL (SLOW DOWN)</instruction>
  ...
</info>
</alert>
```

The above XML description is an example of the message (a part of the message). This protocol accomplishes all the necessary requirements for C2X systems defined at the beginning of the 4th chapter. For this purpose a demo application was developed to acquire reliable information about its network abilities. The simulation model was created to analyze communication between various objects and for different types of traffic.

4. Distribution of the information

Both static and mobile objects participate in communication. When the result of image processing is transmitted the holder of the original information is a computational center. It is a static object and, therefore, it is important to use the model of "off-line replication". At first it is necessary to grasp properties of on-line and off-line replication.

Off-line replication [7]

- Data are originally stored in one node then they are replicated to another one.
- Suitable for standard access methods.
- Issues:
- Users get only meta-data information concerning created replicas, e.g., by email or through the web-site.

On-line replication

Receivers of messages from the computational center are mobile users. Each of them forwards message to another user. In this case it is called on-line replication.

- Data replicas are created parallel by the node and shared between mobile users.
- Limitations:
- Suitable for 'custom' access methods, incompatible with 'standard' ones.
- Hard to implement, possible performance delays.

Where to send a message?

The question is which node should be addressed with a new message and how to deliver the message. It is already known what the message looks like. Next step is to deliver it. The holder of the original message is the computational center. The message needs to be transmitted over Ad-Hoc network from this center. The receiver of the message is a mobile object (vehicle, train). The algorithm has to be implemented, which can administrate sudden changes in the network topology (connecting disconnecting objects etc.).

Different algorithms for the mentioned purposes are analyzed.

1th Algorithm – simple approach

- The computational center creates a set of available clients (S_i which is regularly updated – in very short intervals).
- The computational center broadcasts short messages (ack.) about available information to an object from S_i . This information is broadcasted repeatedly (during the time it is valid).
- Time period is established for client to respond.
- If during the time period, one or more clients respond, the communication is considered as successful and the message with information is sent to the client.
- If some nodes from S_i do not respond to the computational center during a given period of time, they are disqualified for the next time period (disqualification period). They are moved to DQL – DisQualification List (if nobody answers, time period can be increased).

The algorithm described above ensures availability of original data, but it does not consider requirements of a single client. Clients do not have to demand all the data (just part of them – replica). The proposed Simple algorithm does not provide mechanism for information confirmation. Another algorithm is introduced.

2th Algorithm – algorithm based on confirmation

Let us suppose a set of mobile clients S_i . Each of them has a request for updating the original data. Let us further suppose R_i as a set of such requests (R_i is a subset of original data). Then the algorithm for replication request (R_i) from the client S_i can be:

- The computational center (holder of original data) emits its identifier to environment (id + time stamp of the last update).
- The client S_i sends a request for replication R_i from the computational center.
- The computational center processes the request and sends a response to the client.
- If the client receives a complete message, it sends acknowledgement to the computational center.

- The computational center erases R_i from the list of requirements.
- If the client during the time period t does not answer, the computational center checks availability of this client, t is increased and the message is repeated.
- If the client does not respond for a certain amount of efforts, the communication is stopped and the client's request R_i is erased from the request list.

5. Conclusion

In this paper we presented a model for detection and delivery of information about unwanted events in traffic systems (recognition objects in given ROI). Two demos were created to prove the functionality of the model. The first model is for image processing. It is designed for recognition of foreground objects and for providing information about unwanted events. Another demo is a simulation of data delivery and protocol utilization. For image recognition and classification we worked with about 20 video sequences with different objects approaching ROI. It was about 28 minutes of records. We obtained up to 97% of correct assignments. It means that 3 % of objects passing given ROI were identified negatively. For classification part the highest score of correct assignments objects into classes moved around 92, 5%. From the protocol simulation the following conclusion can be made: the protocol meets many of the requirements and could be effectively used for common tasks in mobile Ad-Hoc networks. XML standard ensures its universality, simplicity etc.

6. Future work

We would like to continue with a design of foreground recognition system (general image processing tasks). There is a need to cope with shadows, sudden light changes and also with high level of details in a specific situation. We would like to test the proposed methods in different real situations (detecting weather conditions to inform drivers, detecting traffic jams etc.). For the part of information delivery, it is necessary to test algorithms for effective data replication. It is also important to have in mind a power range of devices which will be used (e.g. NEC LinkBird-MX).

Acknowledgement:

This contribution is the result of the project implementation: **Centre of excellence for systems and services of intelligent transport**, ITMS 26220120028 supported by the Research & Development Operational Programme funded by the ERDF.



Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

"Podporujeme výskumne aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EU."

References

- [1] BADURA, S., LIESKOVSKY, A.: *Advanced Utilization of Image Processing from Industrial Cameras in Traffic*. In: Proc. EURO-Zel Zilina 18th International Symposium, ISBN: 978-80-554-0198-0, pp. 153–160, 2010.
- [2] YILMAZ, A., JAVED, O., SHAH, M.: Object Tracking: A Survey. In: *ACM Comput. Surv.* 38, 4, Article 13, p. 45, 2006.
- [3] LOWE, D. G.: *Object Recognition from Local Scale-invariant Features*. In: Proc. of the International Conference on Computer Vision, pp. 1150–1157, 1999.
- [4] BAY, H., ESS, A., TUYTELAARS, T., VAN GOOL, L.: SURF: Speeded Up Robust Features. In: *Computer Vision and Image Understanding (CVIU)*, Vol. 110, No. 3, pp. 346–359, 2008.
- [5] STAUFFER, CH., GRIMSON, W.E.L.: *Adaptive background mixture models for real-time tracking* [online].
- [6] OASIS: Common Alerting Protocol, v. 1.0. OASIS Standard 200402, March 2004. http://www.terena.org/activities/tf-storage/storage/meetings/1/NDS_PSNC_collaboration_on_storage_services.ppt [online].