International Journal of Multimedia and Ubiquitous Engineering
Vol. 2, No. 3, July, 2007

Visual Traffic Noise Monitoring in Urban Areas

Andrzej Czyzewski and Piotr Dalka

Multimedia Systems Department,
Gdansk University of Technology, Gdansk, Poland
ac@pg.gda.pl, dalken@sound.eti.pg.gda.pl

Abstract

The paper presents an advanced system for railway and road traffic noise monitoring in metropolitan areas. This system is a functional part of a more complex solution designed for environmental monitoring in cities utilizing analyses of sound, vision and air pollution, based on a ubiquitous computing approach. The system consists of many autonomous, universal measuring units and a multimedia server, which gathers, processes and presents data obtained from the distributed measuring units. The results are visualized on numerical maps. The paper contains a functional and technical description of the monitoring system. It describes also the algorithm for moving vehicle detection in video sequences based on a pixel-level difference among the image frames and a continually updated background model utilizing mixtures of Gaussians. The experiments carried out involve the implemented algorithm to the detection of vehicles in the recorded video sequences. The results obtained are illustrated with some examples and discussed.

1. Introduction

Nowadays, effective monitoring of environmental hazards in metropolitan areas becomes a very important issue. Such threats involve air pollution caused by combustion gases and particles and excessive noise. It is also necessary to monitor and to control the road and the railway traffic effectively in order to ensure a best possible vehicles flow. Amongst entities that are highly interested in implementing and utilizing the complex system for metropolitan areas environmental monitoring are administrative authorities, which are obligated by a statutory duty to fight out noise and other environmental threats. The outcome of the system would have positive influence on increasing the global safety of citizens, which is the main task for the police and other safety agencies. Furthermore, an access to the results of the system operation would be very useful for emergency and sanitary agencies, railway companies, hydrological and meteorological services and research entities. An open access to the selected data presented in an efficient way (e.g. with a Web browser) would be appreciated and recognized by citizens as an important and practical prove of usefulness of modern IT technologies developed by joint academic and industrial resources together with administrative authorities and agencies.

The paper presents the idea and implementation of an advanced multimedia system for metropolitan areas monitoring. It contains functional and technical description of one component of the system, which is intelligent traffic noise monitoring. The key part of this solution is an algorithm for moving vehicle image detection in video recordings; its effectiveness is studied and sample results of moving vehicle and train detection and segmentation are presented in the paper. The last section concludes the studies carried out and formulates a scope for the future work in this area.

2. Urban area environmental monitoring

There are two main objectives leading to the development of the environmental monitoring system. The first one is to test new technical solutions in model devices in order to use them in future production prototypes. The second, more general objective is related to the research methodology and assumes using the pilot infrastructure to verify usefulness of various environmental pollution indicators for an effective visualization of the environment condition in the city.

Environmental pollution indicators need to be divided into two groups. The first one contains typical indicators such as noise levels [1], concentration and composition of combustion gases and particulates, or electromagnetic radiation. The second group of environmental pollution indicators should include new measures, which need to be defined and examined whether they are useful for generating layers of the geoinformatic system based on numerical models of metropolitan areas. Generating layers of such a system requires an innovative approach based on experimental studies, because definitions of existing indicators and measures are not sufficient for effective visualization of environmental threats on dynamic numerical maps. For example, noise annoyance indicators require long-term averaging of noise levels (separately during day and night), meanwhile the current (calculated in different way with the use of short-term averaging) noise level might inform more effectively about current problems in the monitored metropolitan area.

The basic principle of the environmental monitoring system, i.e. creating dynamic, emissive maps illustrating the occurrence intensity of various harmful environmental factors, instead of static maps, provides an original approach to the methodics of environmental condition visualization. Furthermore, the important value added is the possibility to aggregate and to create synchronously the data gathered in the database of the teleinformatics system. A cumulative analysis and a synchronous representation of data have many innovative applications, such as:

– combined imaging of hazard factors accumulation (i.e. synchronous monitoring of the carbon monoxide/dioxide and temperature could be used to detect fire threats),
– detecting interrelationships between some harmful factors based on large data representation, which makes it possible to study the influence of environmental condition on the health of large populations,
– contributing to the evolution of general methodics for generating geoinformatic system layers, especially in the context of methods for producing clear visualizations of multidimensional, dynamically evolving data on numerical maps.

The crucial part of the environmental monitoring system consists of a collection of measuring units located in various places inside a city. The measuring units are characterized by the following features:

– autonomy (the measuring unit may operate at any place and it can use the main, sun or battery cell power source),
– versatility (the unit can measure many environmental factors, depending on the selection of sensors connected to its inputs; measuring data can be transmitted using any kind of available telecommunication medium).
– maintenance-free (configuring and monitoring of the measuring unit and new firmware uploading are performed remotely),

- water and environmental factor resilience and protection against theft,
- small dimensions facilitating installation of the unit and limiting its visibility after installation,
- lower costs comparing to the prices of hitherto existing devices used for environmental monitoring.

The second part of the environmental monitoring system is formed by its network layer. The server equipped with the specialized software is adjusted to perform the following tasks:

- automatic data acquisition (transmission is controlled by the server),
- data filtration and storing in relative databases,
- giving an access to the metropolitan area numerical model as the basis for creating specialized layers of the geoinformatic system,
- determining correlations and discovering connections in data sets related to all measured factors,
- hierarchical data access control,
- advanced data visualization in the form of geoinformatic system layers based on the numerical metropolitan area model,
- statistical and sectional data presentation on demand,
- automatic functionality and safety monitoring of measuring units,
- emergency reporting (discovered automatically through data analysis),
- on-demand data transmission to other devices, e.g. other servers, PC computers, PDA devices,
- storing historical data, including operations performed by users on the server, system and data utilization etc.

The system defined above forms scalable and innovative solution which is universal enough to supply practically unlimited amount of modern services from the domain described as information society technology. For example, placing monitoring units in production plants would make possible to study the risk connected with the emission of toxic factors and installing measuring units in schools would allow monitoring noise levels during breaks, which is one of the main factors responsible for children hearing deterioration. Monitoring vicinity of music clubs and discos would improve silence disturbance prevention and decrease the number of complaints coming from nearby residents. Equipping the server with the software able to detect the noise of breaking glass or a scream, provided that there are many measuring units placed around the city, would significantly improve the safety of citizens and their possessions.

3. Road and railway traffic monitoring

Traffic monitoring, its analysis and efficient visualization on numerical maps are ones of the main functionalities of the metropolitan area environmental monitoring system. Estimating traffic noise based on quantity and type of road and railway vehicles in various point of the city provides an important contribution to noise prevention. Furthermore, the possibility to watch the traffic in the real time and to plan the route taking into account current traffic jams can make an invaluable assistance for emergency and safety agencies, public transport companies and drivers.

The traffic monitoring system is build of measuring units located in many places around the city and a server equipped with a database and specialized software. The measuring unit consists of a camera capturing video images and the measuring unit itself, which analyses video input from the camera, detects moving vehicle images in the real time and sends results to the server using any available telecommunication link.

The necessity of video image processing locally (by the measuring unit and not by the server) results from two reasons. First of all, a video stream of the quality sufficient for effective vehicle detection (VGA resolution, 15 fps, the minimal compression ratio) requires a broadband communication channel, whose accessibility in every required place in the city can be problematic and its usage – expensive. Furthermore, the server able to process raw video data from several dozen cameras has to be very powerful, which would drastically increase its cost. It would also limit the scalability of the entire system, because adding another measuring station to the system could be impossible because of the limited server performance.

The camera used for traffic monitoring must meet a few special requirements. It must be placed in a protective housing which is equipped with a cooling blower for hot, summer days and a heater for cold, winter nights. The lens of the camera needs to be very bright and equipped with an automatically controlled aperture. This would enable the camera to operate effectively during a whole year, both during day and night and to adjust itself to the changing lighting conditions.

It is also necessary to choose the type of the camera depending on its output interface. It is possible to utilize typical CCTV (Closed Circuit Television) cameras, which are equipped with an analogue output, usually using a BNC connector and coaxial cable. The second possibility is to use an IP camera with various network servers embedded, which is connected with other devices with a wired (e.g. Ethernet interface) or wireless (e.g. Wi-Fi) link. The latter solution seems to be more adequate because of two reasons. First, direct connection between the camera and the measuring unit with a twisted pair cable significantly simplifies video data acquisition by the measuring unit – various network protocols, such as HTTP or RTSP may be used for this purpose. Therefore, there is no need to build any additional module for A/D conversion of the signal. The video compression used by IP cameras does not pose any problems because it may be set to a minimum ratio. It is only required that the camera is able to compress video data with the MJPEG codec, which provides a much better quality of individual video frames than the MPEG4 standardized protocol. The utilization of network cameras has also an additional advantage. It is possible to watch video in the real time from any place in the world (with the quality depending on parameters of transmission channels), without engaging measuring units to this kind of activity. This makes possible to use the measuring unit equipped with an IP camera as a part of a typical surveillance system.

The video image processing in the real time requires a significant amount of computational power. Therefore the measuring unit should be build out of some modern and effective components. At the same time, it should be characterized by the minimal power requirements because of the battery power supply usage. These two conditions may be met using a modern, single-board computer with components (especially the processor) optimized for low power consumption as the base of the measuring unit.

The main task of the measuring unit used for traffic noise monitoring is to process and to analyze video data from the camera. The unit detects locations and determines parameters of all vehicles in every video frame. Parameters may be static (e.g. shape, texture) or dynamic (e.g. velocity vector). Collective results of detection are sent to the server at given time intervals. Additionally (if there is enough computational power) the measuring unit may identify the kind of detected vehicles (e.g. a truck, bus or car). Optionally, it can also detect and report unusual and dangerous events, i.e. a vehicle moving on the pavement, car accident or speed limit violation.

A computational centre of the traffic (and environmental) monitoring is formed by a server with a relative database. All results sent by measuring units and data related to the units themselves (i.e. their locations, firmware versions, diagnosis data) are stored in the database. The main task of the server is to process the data gathered in the database and to present them to users in a clear and efficient way.

For a proper visualization and analysis of results, it is essential to utilize a numerical model of the city, which is a part of the Geographic Information System (GIS). The measurement results (e.g. the traffic) may be shown as layers placed directly on the city map. It is possible to present the traffic density in the real time or to demonstrate cumulative results regarding any given time period. The server functionality is supplemented with an option to generate statistical, sectional and summary reports. An access to the server and measurement results is possible from any computer in the world (provided that a user has a sufficient access rights) using a Web browser.

Utilizing the numerical model of the metropolitan area enables users to perform complex analyses which would not be possible in any other way. Especially, implementing noise propagation model and taking into account local land relief and heights of buildings it is possible to assess the noise level caused by the traffic in any place of the city.

An open architecture allows the server to integrate and visualize on the maps any data coming both from autonomous measuring units and directly from safety or administration agencies (i.e. information regarding road closures and detours caused by emergencies or reconstructions). This creates the opportunity to manage public transportation systems in the metropolitan area effectively.

4. Moving vehicle image detection

Moving object detection and segmentation is an important part of video based applications, including content based video summarization and coding, computer vision, videoconferencing, digital entertainment and video monitoring. In the last application, results of vehicle detection and segmentation in video streams can be used for inexpensive and accurate calculating traffic statistics or real-time visualizing the traffic on numerical maps.

Most video segmentation algorithms usually employ spatial and/or temporal information to generate binary masks of objects [3]. Spatial segmentation is basically image segmentation, which partitions the frame into homogenous regions with respect to their colours or intensities. This method can be typically divided into three approaches. Region-based methods rely on spatial similarity in colour, texture or other pixel statistics to identify separate objects while boundary-based approaches use primarily a differentiation filter to detect image gradient information and extract edges [4]. In the third, classification-based approach, a classifier trained with a feature vector extracted from the feature space is employed to combine different cues such as colour, texture and depth [5].

Temporal segmentation is based on change detection followed by motion analysis. It utilizes intensity changes produced by moving objects to detect their boundaries and locations. Although temporal segmentation methods are usually more computationally effective than spatial approaches, they are sensitive to noise or lighting variations [3].

There are also methods combining both spatial and temporal video characteristics, thus leading to spatio-temporal video segmentation. In this case, visual properties (e.g. variation of luminance and variation of motion) are considered both spatially and temporally. The final outcome is a 3-D surface encompassing an object position through time called object tunnel [6].

The solution presented in the paper utilizes spatial segmentation to detect moving objects in video sequences. The most popular region-based approach is background subtraction [7][8], which generally consists of three steps. First, a reference (background) image is calculated. Then, the reference image is subtracted from every new image frame. And finally, the resulting difference is thresholded. As a result, binary images denoting foreground objects in each frame are obtained. The simplest method to acquire the background image is to calculate a time averaged image. However this method suffers many drawbacks (i.e. limited adapting capabilities) and cannot be effectively used for traffic monitoring.

A popular and promising technique of adaptive background subtraction is modelling pixels as mixtures of Gaussians and using an on-line approximation to update the model [9][10]. This method proved to be useful in many applications as it is able to cope with illumination changes and adapt the background model accordingly to the changes in the scene, e.g. motionless foreground objects eventually become a part of the background. Furthermore the background model may be multi-modal, allowing regular changes in the pixel colour. This makes it possible to model such events as trees swinging in the wind or traffic light sequences. Thus this method was used for moving object detection in the visual traffic noise monitoring system.

In this method, each image pixel is described separately by a mixture of K Gaussian distributions [12]. The probability that a pixel has value x_t at the time t is given as:

$$p(x_t) = \sum_{i=1}^{K} w_t^{\,i} \eta\left(x_t, \mu_t^{\,i}, \Sigma_t^{\,i}\right) \tag{1}$$

where w_t^i denotes the weight and μ_t^i and Σ_t^i are the mean vector and the covariance matrix of ith distribution at the time t, and η is the normal probability density function. The number of distributions K is usually a small number (3 or 5) and is limited by the available computational power. For simplicity and to reduce memory consumption it is assumed that RGB colour components are independent.

Each Gaussian distribution represents a different background colour of a pixel. The longer a particular colour is present in the video stream, the higher value of the weight and lower values in the covariance matrix of the corresponding distribution are.

With every new video frame, the parameters w, μ, Σ of distributions for each pixel are updated according to the previous values of the parameters, current pixel value x_t and the model learning rate α [12]. The weights of distributions are modified according to the equation:

$$w_t = (1-\alpha)w_{t-1} + \alpha M_t \tag{2}$$

where M_t is equal 1 for the first matching distribution and 0 for other distributions. After adjusting, the weights of distributions are normalized.

If there is a matching distribution its mean μ_t and variance σ_t^2 values for every RGB component are adjusted as follows:

$$\mu_t = (1-\alpha)\mu_{t-1} + \alpha \cdot x_t \tag{3}$$

$$\sigma_t^{\,2} = (1-\alpha)\sigma_{t-1}^{\,2} + \alpha(x_t - \mu_t)^2 \tag{4}$$

The higher α the faster model adjusts to changes in the scene background (e.g. caused by gradual illumination changes), although moving objects remaining still for a longer time (e.g. vehicles waiting at traffic lights) would become a part of the background quicker.

During determining whether the current pixel is a part of a moving vehicle, only distributions characterized by high weights and low values in covariance matrices are used as the background model. If the current pixel matches one of the distributions forming the background model, it is classified as the background of the scene; otherwise it is considered as a part of a foreground object.

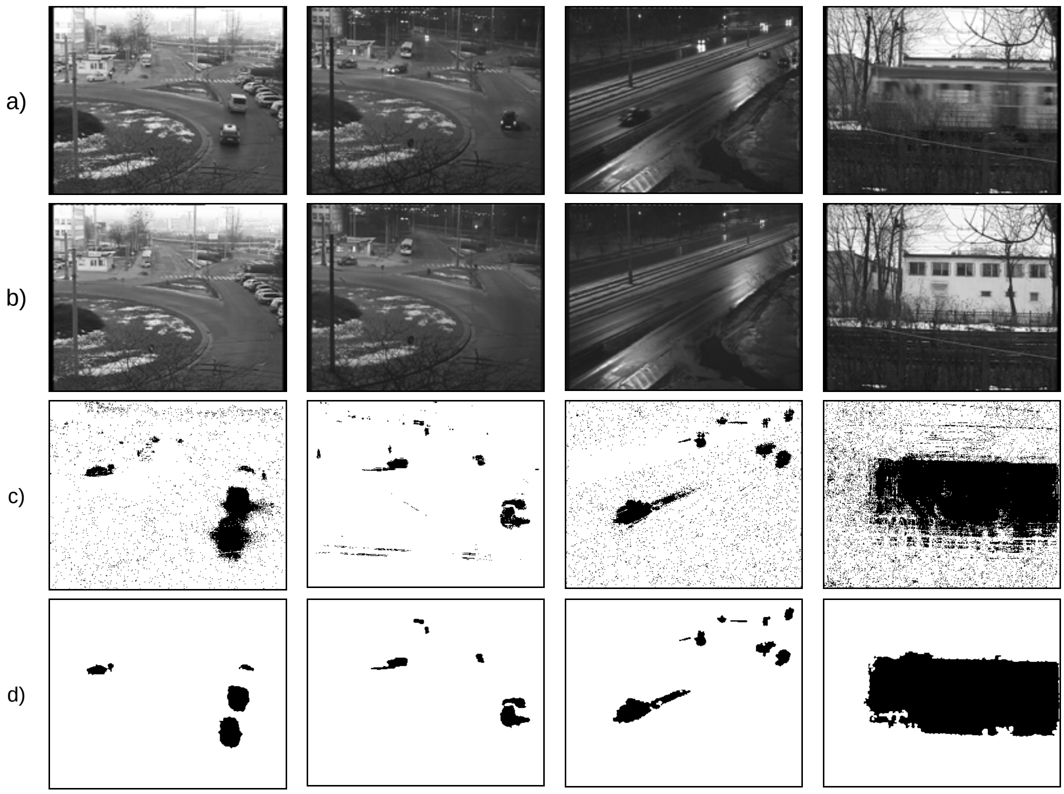

Figure 1. Sample results of vehicle and train detection; a) original frames from recorded video sequences; b) scene background for the current frame; c) raw results of background removal without any further processing; d) final results of vehicle segmentation

Moving vehicle image detection is supplemented with the shadow detection module which is required for every outdoor video processing application, especially in the field of traffic monitoring. The shadow of a moving vehicle is always present, moves together with a vehicle and as such is detected as a foreground object by a background removal application.

The shadow detection method is based on the idea that while the chromatic component of a shadowed background object remains generally unchanged, its brightness is significantly lower. The algorithm used to estimate chromatic and brightness differences between the current pixel colour and the background model is based on the colour model described in [11]. It makes possible to separate the RGB colour space used in the model into chromatic and brightness components. Only pixels recognized as a part of a foreground object during the background subtraction process are checked whether they are part of a moving shadow.

A binary mask denoting pixels recognized as belonging to foreground objects in the current frame is the result of the background subtraction. The mask is morphologically refined in order to allow vehicle segmentation [12]. Morphological processing of a binary mask consists of removing regions (connected components) having too few pixels, morphological closing and filling holes in regions.

5. Experiments

Some sample results of vehicle and train detection and segmentation are presented in Fig. 1. The implemented algorithm is able to determine the scene background and mark locations of all vehicles correctly, both during a day and a night. The first column in Fig. 1 demonstrates usefulness of the moving shadow detection and morphological processing of binary masks which made possible to separate two vehicles originally labelled as one region. The algorithm is also able to detect vehicles in night sequences (the second and the third column in Fig. 1). There is one major drawback of the night environment. Car headlights illumine the road ahead of a car and nearby buildings which causes that illuminated areas are classified as foreground objects. Supplementary decision layer needs to be added to the algorithm to prevent such false detections and obtain exact vehicle shapes. An intelligent decision algorithm implementation (based on rough sets theory) is planned to solve this problem.

Some sample results of detection of trains passing in front of the camera are shown in the fourth column in Fig. 1. The segmentation algorithm successfully removed swinging traction wires and traction posts in the foreground from the resulting images. Unfortunately, a very small colour difference between tracks and a train chassis makes it impossible to separate these two regions perfectly.

a)

b)

Figure 2. Determining usual movement paths; a) sample backgrounds from video recordings; b) images illustrating movement density in analyzed scenes (the darker pixel the higher density)

The algorithm presented can be additionally used to determine typical movement paths in analyzed scenes automatically. This task is performed by summarizing binary masks in the given period of time; areas of the scene with the highest movement density are characterized by the highest values in the summary image. Sample results are presented in Fig. 2. Movement paths obtained with this method may be directly used in a surveillance system to detect unusual events, such as a vehicle moving outside the road surface.

6. Conclusions

This paper presented the idea of the advanced system for the metropolitan area environmental monitoring, including traffic noise monitoring. There are numerous advantages resulting from the system implementation which will facilitate everyday life of various emergency, safety and administration agencies, as well as citizens. The possibility of the road and railway traffic monitoring in the real time and dynamic management of the traffic inevitably leads to significant economical profits.

The key part of the system is the algorithm for the moving vehicle image detection, which utilizes mixtures of Gaussians. The effectiveness of the algorithm is tested and the results hitherto achieved prove that the model describing the scene background can adapt itself to changes in the scene and the algorithm is able to detect moving vehicles with a good accuracy. The outcome may be directly used by the server to visualize the traffic on numerical maps and to generate statistical rapports.

Future work will be divided in two tasks. The first one will be focused on building a prototype metropolitan area monitoring system consisting of a set of measuring units and a server. The system will be used for implementing and testing hardware and software solutions for all component devices. The aim of the second task is to improve the algorithm for vehicle image detection. Planned modifications include dynamic adjusting the learning rate, depending on the current scene change rate, and implementing a supplementary module which would facilitate detection in night sequences employing intelligent reasoning. Such improvements should improve accuracy of moving vehicle image detection further, especially in difficult conditions and would increase the effectiveness of traffic monitoring.

Acknowledgements

Research subsidized by the Polish Ministry of Science and Higher Education within the Grant No. R02 010 01.

References

[1] Czyzewski A., and Kotus J., “Universal System for diagnosing environmental noise”, *Management of Environmental Quality: An International Journal*, vol. 15, no. 3, 294-305, 2004.

[2] Barron J., Fleet D., Beauchemin S., “Performance of optical flow techniques”, *Int. J. of Comp. Vision*, vol. 12, no. 1, 42-77, 1994.

[3] Li H., Ngan K., “Automatic Video Segmentation and Tracking for Content-Based Applications”, *IEEE Communication Magazine*, vol. 45, no. 1, 27-33, 2007.

[4] Salembier P., Marques F., “Region-Based Representations of Image and Video: Segmentation Tools for Multimedia Services”, *IEEE Trans. Circuits and Syst. For Video Tech*., vol. 9, no. 8, 1147-1169, 1999.

[5] Liu Y., Zheng Y., “Video Object Segmentation and Tracking Using y-Learning Classification”, *IEEE Trans. Circuits and Syst. For Video Tech.*, vol. 15, no. 7, 885-899, 2005.

[6] Konrad J., “Videopsy: Dissecting Visual Data in Space-Time”, *IEEE Communication Magazine*, vol. 45, no. 1, 34-42, 2007.

[7] KaewTrakulPong P., Bowden R.: “A real time adaptive visual surveillance system for tracking low resolution colour targets in dynamically changing scenes”, *Image and Vision Computing* 21, 913-929, 2003.

[8] Yang T., Li S., Pan Q., and Li, J., “Real-Time and Accurate Segmentation of Moving Objects in Dynamic Scene”, *ACM Multimedia-2nd International Workshop on Video Surveillance and Sensor Networks*, New York, October 10-16, 2004.

[9] Stauffer C., and Grimson W., “Learning patterns of activity using real-time tracking”, *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 22, No. 8, 747–757, 2000.

[10] Elgammal A., Harwood D., and Davis L., “Non-Parametric Model for Background Subtraction”, *ICCV Frame-rate Workshop*, September 1999.

[11] Horprasert T., Harwood D., Davis L., “A statistical approach for real-time robust background subtraction and shadow detection”, *Proc. of IEEE Frame Rate Workshop*, 1-19, Kerkyra, Greece, 1999.

[12] Dalka P., “Detection and Segmentation of Moving Vehicles and Trains Using Gaussian Mixtures, Shadow Detection and Morphological Processing”, Int. Conf. on Computer Vision and Graphics, September 25-27, Warsaw, Poland, 2006.

Authors

Andrzej Czyzewski

Received his M.Sc. degree in Sound Engineering from the Gdansk University of Technology in 1982, his Ph.D. degree in 1987 and his D.Sc. degree in 1992 from the Cracov Academy of Mining and Metallurgy. Prof. Czyzewski joined the staff of the Sound Engineering Department of the Gdansk University of Technology in 1984. In December 1999 Mr. President of Poland granted him the title of Professor.

Piotr Dalka

Received his M.Sc. degree in Sound Engineering from the Gdansk University of Technology in 2003. His thesis was related to the problem of automatic recognition of musical instrument sounds. In 2006 he graduated from the Faculty of Management and Economics of Gdansk University of Technology. He is interested in audio and video processing and artificial intelligence algorithms. Now he is a candidate to Ph.D. His thesis is devoted to novel methods of vehicle image detection and classification.