



**FACULTY
OF ELECTRICAL
ENGINEERING
CTU IN PRAGUE**

Faculty of Electrical Engineering
Department of Radioelectronics

Habilitation thesis

**Applications of high-speed
camera systems**

Stanislav Vítek

Prague, November 2019

Copyright

This work is a compilation of papers published throughout the course of my research. Some of the articles included in this habilitation thesis are protected by the copyright of SPIE, and IET. They are presented and reprinted in accordance with the copyright agreements with the respective publishers. Further copying or reprinting can be done exclusively with the permission of the respective publishers.

Acknowledgement

I would like to thank my family for their continuous and unparalleled love, help and support. I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am.

This thesis would not have been possible without the inspiration and support of a number of wonderful individuals, my friends Petr Páta, Martin Jelínek, David Bursík, and Stanislav Zvánovec.

Contents

1	Introduction	1
2	Camera systems in astronomy	3
2.1	Optical Transients	3
2.1.1	Electron-multiplying CCD	6
2.1.2	Image stacking	8
2.2	Meteors	9
2.2.1	Image intensifier	10
2.2.2	Image processing in the intensified camera system	11
2.3	Author's contribution	14
3	Camera systems for visible light communication	15
3.1	VLC for Vehicle-to-Vehicle Communication	15
3.2	VLC for Indoor Navigation	16
3.3	Author's contribution	17
4	Wireless smart cameras	19
4.1	Cameras for assistive technologies	19
4.1.1	Video analysis	19
4.1.2	Person identification in low bitrate video-sequences	21
4.2	Management of parking lots	29
4.3	Author's contribution	30
5	Conclusions and further research	31
A	Appendix A	40
B	Appendix B	51
C	Appendix C	67
D	Appendix D	76
E	Appendix E	84
F	Appendix F	87

1 Introduction

Throughout its existence, people have admired the rapidly changing natural phenomena of different scales, which are usually accidental discoveries and then disappears. Prehistoric man was fascinated with fire, from the perspective of today's schoolchildren simple chemical reaction – exothermic oxidation of combustible gasses. Falling stars, meteors leaving bright strikes in the atmosphere, were symbols of coming disasters. Without a recording, it was, however, difficult to further investigate these phenomena. When the Danish astronomer Tycho Brahe discovered the "new star"¹ on November 11, 1572, it was possible only due to that the light remained visible to the unaided eye until March 1574.

Naked-eye observation brings the observer a fantastic immediate experience. However, the observer cannot explore the natural phenomenon even after the end; he cannot look for new connections and meanings. Development of photographic techniques in the 19th century naturally supported the general desire for further exploration and a deeper understanding of the optical effects. Film cameras, and later scanning using the strobe, enabled capturing of very fast phenomena. Astrophotography was made on photographic plates, i.e., glass planes covered with photosensitive emulsions. The advantage of photographic material is a high dynamic range of the recordings; the problems are namely difficult analysis, limited usability, material consumption and the gradual degradation of light-sensitive layers while aging. Despite all these negative properties, records on the photographic plates were often the basis of the great discoveries. We can name the meteorite fell on April 7, 1959, in Czechoslovakia near to Příbram. It was the first meteorite whose trajectory was tracked by multiple cameras recording the associated fireball. Ceplecha then calculated the trajectory and four pieces were found, the largest having a mass of 4.425 kilograms.

Although the photographic plates are still used to record the trajectories of the meteorites, almost all modern astronomical research is carried out with photo-electronic equipment, by which we mean instrumentation that converts radiant energy (such as light) into electrical signals which can be digitized; first it was vacuum tubes, then Charged-couple devices (CCD) invented in the 1970s. CCD based cameras provide high sensitivity including photon counting modes and high frame-rates. Nowadays, the miniaturization of the electronic component has made it possible to design and fabricate portable devices that contain low-noise Complementary Metal–Oxide–Semiconductor (CMOS) cameras and powerful processors, and thanks to penetration of smartphones can now literally every human being have supercomputer in his pocket and use it daily.

The habilitation thesis presents the advances in the application of real-time camera systems. It focuses on selected areas in this field. Different circumstances of the image data acquisition are taken into account, and the various methods of the subsequent analysis are studied. Specifically, the habilitation thesis addresses the following issues:

High-speed observing of optical transients and meteors This part is dedicated to my activities in the field of astronomical observation. At first, it was observing

¹Nowadays we know, that "new star" was one of the few recorded supernovas in the Milky Way Galaxy, type Ia supernova, which occurs when a white dwarf star accretes material from a companion star, and that material explodes in a thermonuclear reaction that destroys the white dwarf.

of Gamma-Ray Bursts (GRBs) with the network of robotic telescopes (project BOOTES, Burst Optical Observer and Transient Exploring System), at second it is an observation of faint meteor showers by the use of wide-field unique scientific instrument employing 60fps CMOS camera and an image intensifier (project MAIA, Meteor Automatic Imager and Analyser).

Camera based VLC communication This part is dedicated to experimental work in the field of low-bitrate car-to-car communication and also using light emitting diodes for VLC. Experiments were carried with consumer-level low-cost cameras with framerates between 30 and 50fps.

Low-power camera systems This parts of the habilitation thesis deals low-power wireless camera systems, designed to classify perimeter, process visual information locally and communicate with the central server through the Internet of Things (IoT) network.

2 Camera systems in astronomy

Astronomical observation is typical example of the projection of an object in infinity through an optical system having image sensor in its focus.

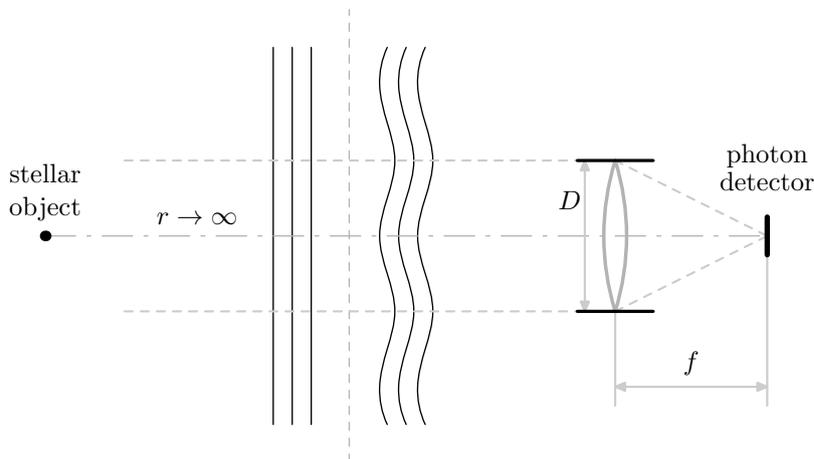


Figure 2.1: Schematic arrangement of considered imaging system.

Figure 2.1 shows an arrangement of such an imaging system: the stellar object is observed at a large distance, i.e. it is essentially a point source of light. In the interstellar space, the light is distributed evenly; the waveform is distorted when the light passes the atmosphere and the distorted light is then transmitted through the optical system to the image sensor.

In the case of ideal imaging system coupling aberrations free lens (i.e. without coma, astigmatism or similar visual defects) and ideal sampling device, the image of a distant point source of light is a point source. In practice, the image is described by a point spread function (PSF) that shows how the light from a point source (like a star) is spread over multiple pixels in the image; the PSF can vary across the field of view and its shape depends on which factors are limiting image quality. This phenomenon, known as spatially variant PSF, occurs mainly in wide-field imaging systems.

Astronomical observations exhibits some features that make it impossible to use standard camera systems and postprocessing methods. Spatial distribution of pixel intensity is nonhomogeneous, and pixels are typically grouped in isolated spots – stellar objects. When a camera system for astronomical purposes is designed, it is necessary to carefully consider suitable exposure time, acceptable signal-to-noise ratio and detector gain.

2.1 Optical Transients

A transient astronomical event is an astronomical object or phenomenon whose duration may be from seconds to days, weeks, or even several years. The term is used for violent deep-sky events, such as supernovae, novae, dwarf nova outbursts, gamma-ray

bursts, and tidal disruption events, as well as gravitational microlensing, transits and eclipses. These events are part of the broader topic of time domain astronomy.

Author of this thesis focused in his work to prompt automated observation of gamma-ray bursts (GRBs), mostly done with robotic telescopes involved the BOOTES (The Burst Observer and Optical Transient Exploring System) project [1]. GRBs were discovered by the Vela satellites in 1967, as one of the most energetic phenomena in the universe [2]. GRBs show up as bright brief pulses (from a few milliseconds to several hundred seconds in duration) of high-energy gamma ray photons (in MeV range). The impossibility of focusing gamma-rays made their precise localization challenging, and the deficiency in observational data created space for an abundance of theoretical models.



Figure 2.2: BOOTES-2 60cm telescope, equipped with EMCCD camera. La Mayora (CSIC), Málaga, Spain.

The examination of GRB durations as detected by BATSE revealed a bimodal distribution with short GRBs of mean duration less than 2s, and long, lasting more than 2s [3]. Nowadays it is generally accepted that the long GRBs (LGRB), are associated with the collapse of massive stars and the birth of a black hole [4]. This theory is supported by observations of supernovae superposed to the afterglow emission, and is in agreement with their redshift distribution [5]. Studies of host galaxies revealed a strong affinity to star-forming regions, additionally supporting this theory [6, 7].

While the high-energy component of a GRB emission can only be observed from space, the optical part of their spectra is commonly followed-up by ground-based telescopes. In this concept, a satellite instrument detects and localizes the event and relays

the coordinates to ground-based instruments. Such optical counterparts are typically seen as rapidly fading point sources superimposed on a very faint distant galaxy (the latter visible only by large telescopes).

The majority of later observations of these optical afterglows can be explained by what we call a forward shock – synchrotron radiation originating from the front surface of the shock between some ejecta produced by the GRBs internal engine and the surrounding circumstellar material. Parameters such as electron distribution factor p or circumstellar medium density profile can be derived from such observations. Also, it is possible to obtain a "radiogram" of the universe along the line of sight, as the synchrotron spectrum is simple and without intrinsic line features - any such in the transient spectrum is due to intervening matter. Obtaining such spectra (and determining the redshift) has recently become a relatively common practice, given a bright enough transient and a localization certain enough to spend large telescope time on it.

However, if the optical afterglow is observed within a few minutes of the event, a distinctly different situation appears. The familiar forward shock radiation is not yet dominant (it might be still rising to its peak), and signatures of the inner engine or more exotic radiation sources such as reverse shock (i.e. the rear side of the ejecta-circumburst material surface) may be detected. Such observations exist, but are scarce, and often it is impossible to discern what kind of radiation source might have been responsible for which component due to lack of color or spectral information.

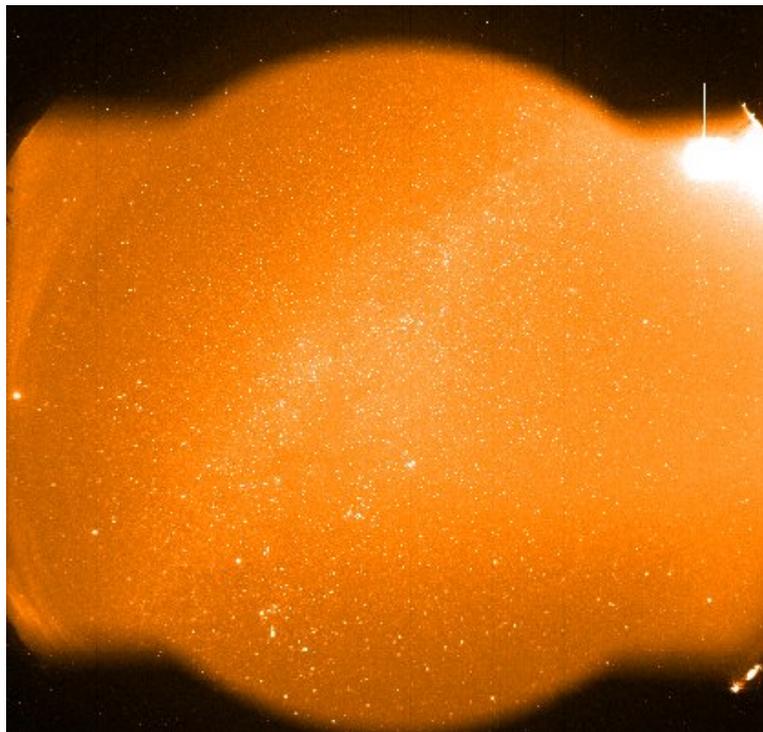


Figure 2.3: A typical allsky image taken on a dark, cloudless night. Limiting magnitude is 10 towards the zenith and 8 close to the horizon.

Automated methods of optical transient's follow-up can be typically thought of as two different levels:

Wide-field cameras watch large fractions of the sky, with the goal of obtaining optical data simultaneous to the high-energy event. Examples include the robotic system Pi of the Sky and the Palomar Transient Factory. These systems, in general, are constructed so that they could detect transient objects independently and provide alerts to other systems. Castro-Tirado et. al reported [8] device consisting of with full-frame CCD ($4,096 \times 4,096$ pixels of size 9×9 microns) combined with 16 mm f/2.8 fisheye lens. An integration time of 30s still allows to record point-like images from stars. With a readout time in the range 15-20s (depending on the hardware used to readout the CCD), the dead-time is not higher than 40%. The performance of this camera allows to record transient events on the sky like optical transients related to GRB, flaring stars, meteors, night clouds, aircraft and artificial satellites. The limiting magnitude under good observing conditions (no moon, absence of clouds, etc.) ranges from R-10 (towards the zenith) to R-8 (close to the horizon). The average pixel scale is 1.5/pix. Fig. 2.3 shows a typical 45s image obtained by the BOOTES-1 all sky camera. No filter is used. A prototype camera was running at the BOOTES-1 astronomical station in Huelva (Spain) since December 2002 and a second one is working at the BOOTES-2 station in Malaga (Spain) since July 2004. Later were all stations of the BOOTES robotic telescope network equipped with similar devices.

Small to medium scale robotic telescopes react automatically to GCN or other alerts and perform a pre-programmed script of exposures. Some of these have image processing to automate transient detection, typically a form of pre-processing to present a human observer with material for a decision. All these systems aim to react as quickly as possible, with some (including the BOOTES telescopes) able to begin observing within a few seconds. While their detection limits are much shallower than large telescopes, they can provide very good results at the early stage of GRB follow-up. The early results of these telescopes often serve as a decision material for whether to trigger a large telescope. 0.6m BOOTES or 0.5m ROTSE belong to this group.

Besides, the GRB phenomenon is very random, and as it is difficult to determine its location, it is also a very distant phenomenon, so the light is weak, and observers need to use a suitable image amplifier to get an acceptable signal-to-noise ratio. The sensitivity of a conventional CCD is limited by the noise introduced by the charge to voltage conversion process [9]. Furthermore, the readout noise increases with pixel rate. So the best read-noise performance can be achieved when readout speed is reduced considerably. Currently, as the most suitable detector to use with small to medium scale robotic telescopes is considered electron-multiplying CCD (EMCCD) cameras, able to overcome these limits to deliver high sensitivity with high speed.

2.1.1 Electron-multiplying CCD

EMCCD were developed to allow detection of very faint transient sources, down to a single photon, which would normally be lost in the CCD amplifier noise. They operate as ordinary CCDs, except for an electronic multiplier through which pixel charges

can be sent prior to the output amplifier [10]. The electron multiplier is made up of many low-gain stages, producing a large (but statistically uncertain) and tunable gain, with a corresponding loss of dynamic range [11]. Noise introduced by the gain process was extensively studied by Basden [12]. Recognising the potential advantages of EMCCDs for astronomical spectroscopy lead to construction of QUCAM2 on the ISIS spectrograph of the 4.2-m William Herschel Telescope [13] and ULTRASPEC on the EFOSC2 spectrograph of the 3.5-m New Technology Telescope [14]. Deeper details on using EMCCD for astronomical spectroscopy may be also found in [15]. Harpsøe et al. [16] introduced novel methods of photometric reduction for EMCCDs.

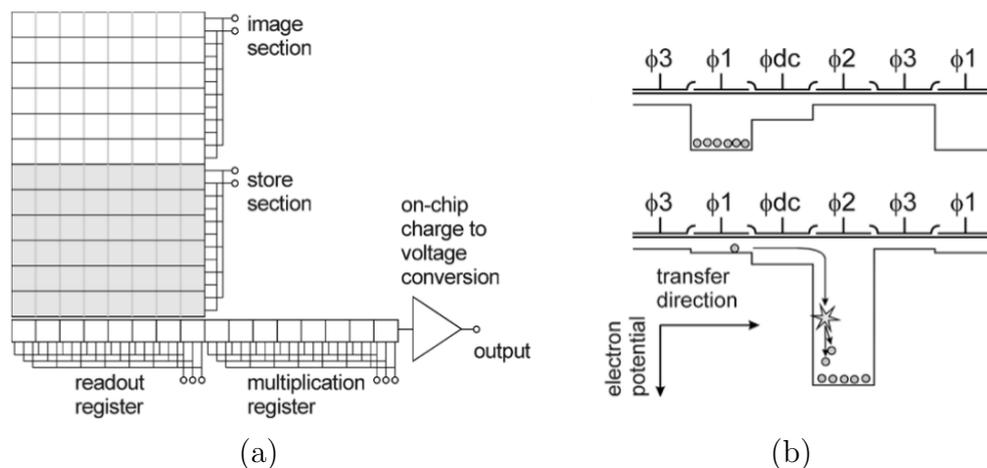


Figure 2.4: EMCCD principles. (a) Schematic of the EMCCD arrays. (b) Transfer of charge through a multiplication element.

EMCCDs have the same basic structure as CCDs. The principal difference between an EMCCD and a traditional CCD is the presence of a special extended serial register, known as a multiplication register (see Figure 2.4 [9]). Here two of the gates (ϕ_1 and ϕ_3) are clocked with normal amplitude drive pulses (~ 10 V) and can use the same pulses as those applied to two phases of the readout register. The pulses applied to ϕ_2 of the multiplication element have higher amplitude, typically 40–45 V [9]. A gate is placed prior to ϕ_2 , which is held at a low *dc* level. The potential difference between ϕ_{dc} and the high level of ϕ_2 can be set sufficiently high so that signal electrons can undergo impact ionization processes as they are transferred from ϕ_1 to ϕ_2 during the normal clocking sequence. Thus, the number of electrons in the charge packet increases as it passes through a multiplication element. Although the probability of impact ionization and thus the mean gain per stage is very low, the number of stages can be high.

Noise components of electron multiplying CCD (EMCCD) include photon shot noise, dark current noise, clock induced charge, multiplication noise and readout noise.

Photon shot noise is the fundamental quantum nature of light and constitutes the theoretical noise limitation of any low light level imaging application. So this type of noise is the same as in microchannel plate (MCP).

Dark current noise is noise results from the dark signal due to the electrons being

generated randomly by the photosensor. This process is described by Poisson distribution as the random arrival of electrons [17].

Clock induced charge occurs as a result of impact ionization during charge transfer. As dark current noise is neglected when the EMCCD is cooled enough, clock induced charge becomes the main noise source. It is integration time independent and can be minimized by carefully controlling of clock amplitude, rising edge speeds and parallel transfer rates [18].

Multiplication noise is same as in MCP appears due to loss mechanisms and electron multiplication statistics.

Readout noise is generated by the output amplifier and subsequent electronic circuitry, including kTC noise, $1/f$ noise, and quantization noise. It increases with readout frequency and becomes dominant noise at high frame rates for conventional CCDs. In an EMCCD, readout noise is reduced to sub-electron level when multiplication gain is high enough [18].

At high framerates, there may be little or no information in a single image, requiring smart image coaddition in sets and source extraction [19, 20] and astrometric registration of both single images and products of image coaddition, and possibly advanced techniques like noise reduction (EMCCD, non-Poissonian noise distribution), lucky imaging [21] and resolution improvement [22].

2.1.2 Image stacking

Image stacking has been widely adopted in astronomy as method to combine multiple noisy datasets (both imaging and non-imaging observation) in order to achieve higher signal-to-noise ratio and allow average signals to be pulled out of the data; for example Zibetti et al. detected intracluster light by stacking 683 Sloan Digital Sky Survey (SDSS) clusters [23], Hogg et al. stacked Keck IR data to get faint galaxy colors [24], or White et al. stacked images of quasars [25]. Deeper explanation of the technique can be found in [26].

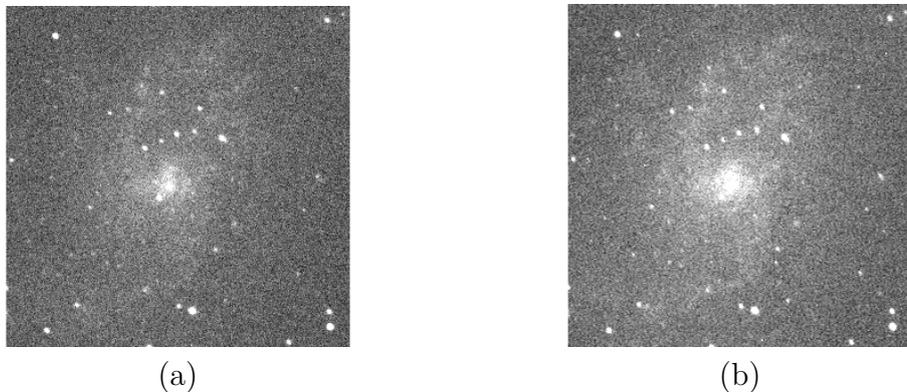


Figure 2.5: Examples of astronomical images. (a) exposure time of 32 s. (b) exposure time of 64 s.

Basic image stacking can be simple done with averaging of the pixels over set of images, or better by using median. More advanced method is $\kappa - \sigma$ clipping [27] iteratively rejecting deviant pixels. For each iteration the mean and standard deviation of pixels in the stack are computed and pixels which value is farthest from the mean than more than $\kappa \cdot \sigma$ is rejected. The mean of the remaining pixels in the stack is computed for each pixel distribution of pixel values. There is also more robust variant of this method replacing rejected pixels by the median value. More advanced methods of stacking are based on the entropy [28] or are using image weighting [29].

Besides of the methods co-adding frames with the same exposure time, it is also possible to combine frames with different exposure times by the use of high dynamic range imaging methods (HDRi) ² Especially objects like nebulae or galaxies often hide other stars that due to the low dynamic range of the image sensor disappear in many times brighter object. However HDRi methods are often non-linear, Vitek and Pata showed that HDRi can add scientific value and could help observers to discover unexpected attractions in their image data ³

The evaluation methods used in [Vitek] were based on the standard astronomical procedures like aperture and PSF photometry. Nasyrova and Vitek later introduced another approach of co-added images evaluation based on analysis of noise models in source and co-added images [30].

2.2 Meteors

Meteors are streaks of light that appear in the sky when an interplanetary dust particle ablates in the Earth's atmosphere. The word meteor is also sometimes used for the particle, which should be called meteoroid, according to proper astronomical terminology. Meteor observations are a relatively inexpensive way to better understand the distribution of material in our solar system. The study of meteors and meteoroids provides clues about their parent objects: comets [31] and asteroids [32]. Meteor observations are typically performed using radar [33], all-sky photographic [34] and CCD (charge coupled device) cameras [8], or television (TV) cameras optionally equipped with image intensifier. Since the light curve of the meteor contains information about the mass of the original particle, camera-based systems are more common. Moreover, both the shape of this curve as well as the height interval where the meteor radiates correspond to the structure of the parent meteoroid. Therefore, combinations of multiple ways of observations are also commonly used [35].

All-sky cameras with huge spatial resolution and long exposure times are great to detect intense light phenomena, like bolides or fireballs. Video data has the advantage that if the meteor is recorded with high time resolution from at least two stations simultaneously, its atmospheric trajectory can be calculated. Moreover, the heliocentric orbit can be determined if we know the exact time of the event, which is common for video observation. It was shown that the properties of systems with image intensifiers

²MCCANN, John J.; RIZZI, Alessandro. The art and science of HDR imaging. John Wiley & Sons, 2011.

³VÍTEK, Stanislav; PÁTA, Petr. Realization of High Dynamic Range Imaging in the GLORIA Network and Its Effect on Astronomical Measurement. *Advances in Astronomy*, 2016, 2016.

enable the detection of meteors with masses down to fractions of a gram.

Meteors are considerably much more luminous than GRBs, so it is not necessary to have such an amplification of the image; observation focused on the highest possible frame rate and the highest possible field of view with reasonable angular resolution. For that reason, for example project MAIA employs in its devices two main components: a second-generation image intensifier XX1332 and a GigE camera JAI CM-040GE. The image intensifier has a large diameter input (50 mm) and output (40 mm) apertures, high gain (typically 30,000 to 60,000 lm/lm), and a spatial resolution of 30 lp/mm [36]. Since the diameter of the photocathode in image intensifier is 50 mm, and the angle of view for meteor observation should be about 50° , then the most suitable focal length of the input lens comes at about 50 mm. The MAIA system uses a Pentax SMC FA 1.4/50 mm—this lens offers the angle of view of 47° . Due to the large aperture, a high input signal-to-noise ratio is achieved at the intensifier.

2.2.1 Image intensifier

An image intensifier is based on MCP. It includes three main components (see Figure 2.6 [18]): a photocathode, a MCP and a phosphor screen. The photocathode converts incident photons into electrons. The electrons are accelerated toward the MCP. Under the high voltage applied to the MCP, the incident electrons gain sufficient energy to knock off additional electrons and hence amplify the original signal [18].

MCP is an array of miniature electron multipliers oriented parallel to one another; typical channel diameters are in the range $10 - 100 \mu\text{m}$ and have length to diameter ratios between 40 and 100 [37].

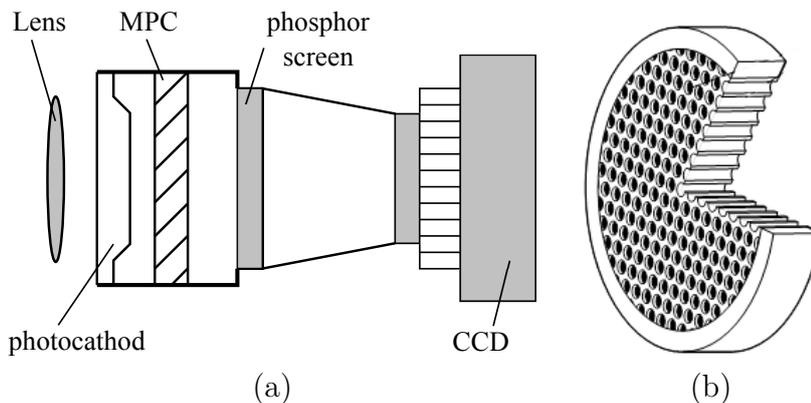


Figure 2.6: MCP principles. (a) Components of an image intensifier tube. (b) A microchannel plate.

MCP is usually fabricated from lead glass. Two sides of the plate are covered by a metal layer that has the function of electrodes. Electrons are emitted on the surface of the photo-emission material – cathode and move to the anode. Walls of channels are covered by the material with a high secondary emission coefficient [38]. Each channel can be considered to be a continuous dynode structure. Here the secondary emission allows to electron getting the multiplication factors of $10^4 - 10^7$ [37].

In the intensified image acquisition system, the image intensifier is a main source of noise:

Photocathode noise includes photon shot noise and Equivalent Background Illumination (EBI). It is two dominant noise contributions to the system [18]. The process of photon capturing has an uncertainty that arises from random fluctuations when photons are collected by the photodiode. Such uncertainty leads to photon shot noise and is described by the Poisson process [17]. EBI is a measure of photocathode dark current and is specified as illumination on the photocathode in microlux of illumination of 2856 K light that is necessary to generate image intensifier tube dark current [18].

MCP noise appears due to loss mechanisms and electron multiplication statistics, an additional noise component arises in the MCP. For second-generation image intensifiers, the noise factors tend to lie between 1.6 and 2.2. For intensifiers of the third generation, the noise factors range between 2 and 3.5 [18].

Phosphor screen noise – phosphor screen usually emits green light and are made of rare earth oxides or halides, with decay times of a few hundred nanoseconds to a few milliseconds. Both the decay times and uncertainty in phosphor screen quantum efficiency cause phosphor screen noise [18].

2.2.2 Image processing in the intensified camera system

It is almost impossible to find a general and reliable noise model of image intensifier; since it contains automatic gain control (AGC), the input-output conversion function is highly nonlinear. However, AGC helps to accommodate high dynamic range; on the other hand, it also causes prominent speckle noise and strong fluctuations in bright pixels in the acquired images. These fluctuations can be easily confused with the variable object and could be a source of false detection.

A typical meteor track is comprised of a streak lasting up to several video frames propagating linearly across space and time. For longer exposure times, typically used in all-sky systems, those streaks can be relatively long, so relatively simple methods like Hough transform can be employed. In the high frame-rate camera systems it is necessary to choose more advanced method. One of the most common methods is to calculate the difference of two consecutive frames to remove static stellar objects and determine the sum of the pixels of potential neighboring objects in different directions [39]. The object is considered to be a meteor if one or more of those sums exceeds a certain threshold. Popular are also methods based on matching filters and methods employing neural networks.

Based on the previous experiences with algorithms mentioned above, Vitek proposed an algorithm of frame classification based on the comparison between temporal statistical characteristics of a pixel and the model built on the relation between the mean and the standard deviation of the pixel [40]. To reduce the impact of high luminance objects that may appear in the field of view and change gain of the intensifier, the model is updated in a relatively short window.

In [41] Vitek and Nasyrova introduced novel method to fast tracking of meteors in noisy video-sequences. Described approach builds on properties of Discrete Pulse Transform [42]. It represents any discrete signal as a sum of pulses, where a pulse is a signal which is zero everywhere except for a certain number of consecutive elements which have a constant nonzero value. Unlike the discrete Fourier and wavelet transforms, the DPT is not a discretization of an underlying continuous model but is inherently discrete. The DPT is composed of non-linear morphological filters based only on the order relations between elements of the discrete signal [43]. Assuming that a meteor leaves a larger track in the single frame than a star or a fluctuating pixel, we can experimentally determine how many pulses represent meteor streaks in the decomposed image and use simple thresholding to obtain binary mask identifying meteor candidate. To remove a stellar object from the list of meteor candidates, we can compare the position of binary mask in two or three consecutive frames.

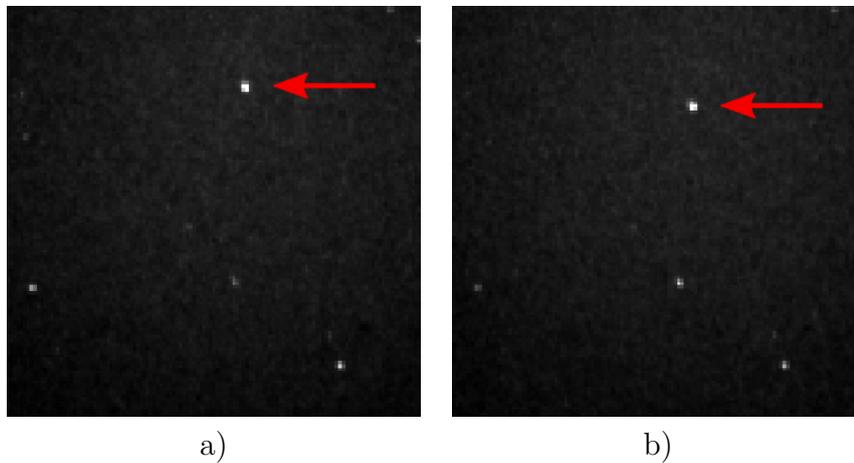


Figure 2.7: Cutouts of two frames from the sequence with meteor.

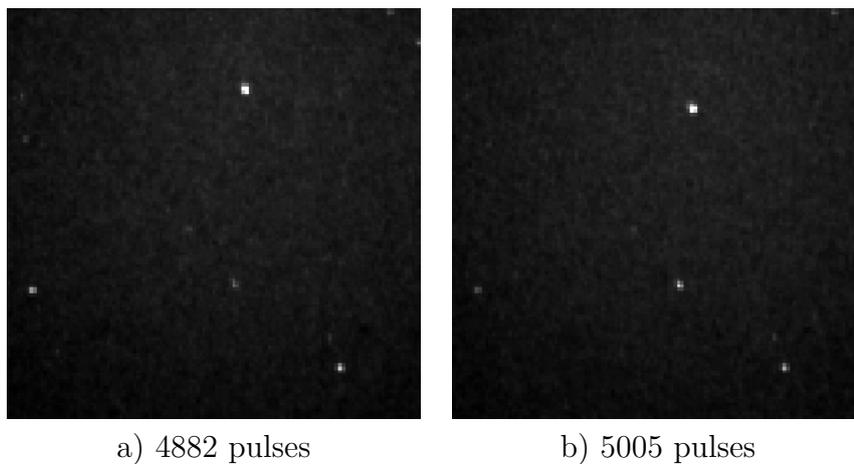


Figure 2.8: Inverse DPT using all pulses.

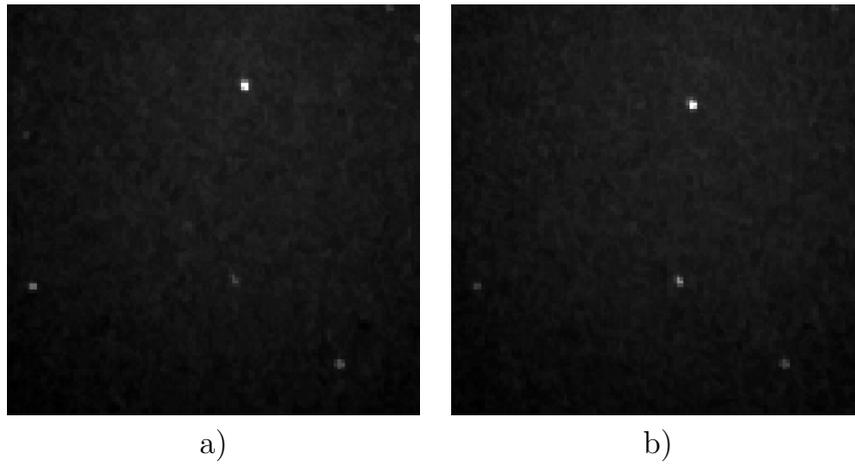


Figure 2.9: Inverse DPT using pulses with area ≥ 5 .

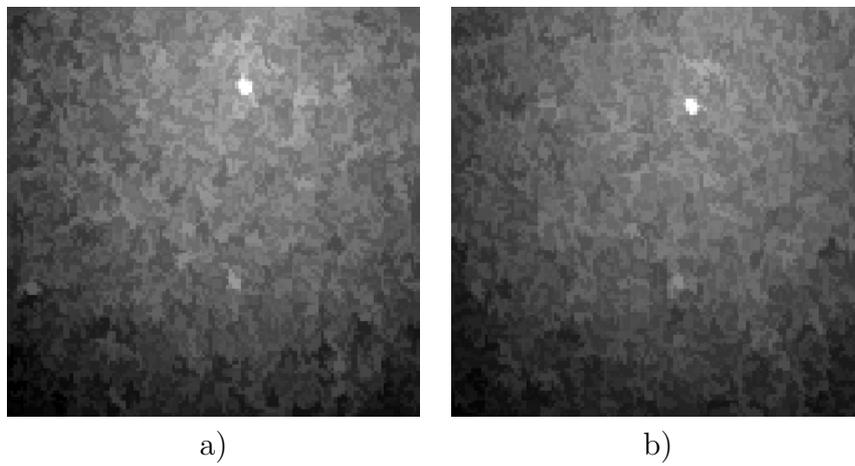


Figure 2.10: Inverse DPT using pulses with area ≥ 10 .

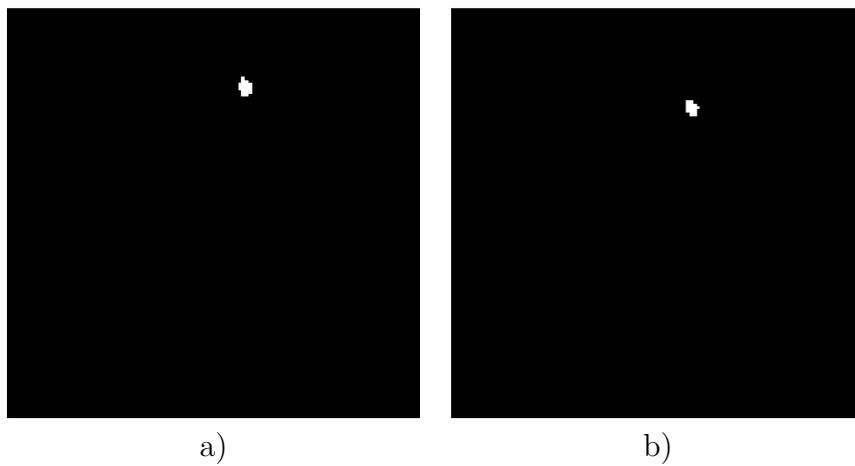


Figure 2.11: Binary mask after thresholding ($T=220$).

2.3 Author's contribution

Dr. Stanislav Vítek is a member of BOOTES collaboration since 2002. He is one of the developers of an open-source software package RTS2, dedicated to control robotic observatories. Within this project, he focuses on image processing routines (focusing, searching for new objects), and device control. Apart from BOOTES, he collaborated on numerous telescopes, like BART, D50 (Astronomical Institute in Ondřejov), or TBT (ESA).

Since 2010 he is the leading software developer of the MAIA project, responsible for the control system and image processing. Dr. Vítek proposed and implemented novel algorithms of fast meteor tracking in noisy video sequences.

Publications related to this section

1. VÍTEK, Stanislav, et al. Long-term continuous double station observation of faint meteor showers. *Sensors*, 2016, 16.9: 1493. (Appendix A, p. 40)
2. VÍTEK, Stanislav; NASYROVA, Maria. Real-time detection of sporadic meteors in the intensified tv imaging systems. *Sensors*, 2018, 18.1: 77. (Appendix B, p. 51)

3 Camera systems for visible light communication

Visible light communication (VLC) is a wireless data transmission technology that builds on the idea of using light source for both illumination and data communications. It uses light emitting diodes (LEDs) or liquid crystal display as a light source, which gives rise to some inherent advantages: low power consumption, a long lifetime, and rapid blinking speed [44]. The dual functionality provided by VLC (i.e., lighting and data communication from the same high-brightness LEDs) has created a whole range of interesting applications, including but not limited to home networking, high-speed data communication via lighting infrastructures in offices, vehicle-to-vehicle communication, vehicle-to-everything communication, mobile attocells, high-speed communication in aeroplane cabins, in-trains data communication, traffic lights management and communications [45]. Recent research in VLC has successfully demonstrated data transmission at over 500 Mbps over short links in office and home environments [46].

VLC systems employ a photodiode (PD) or a camera sensor (CS) as the receiving module. Since as a receiver in the latter case can be used daily use devices like a smartphones, camera-based optical communication systems (Optical camera communication, OCC) can be considered a convenient and versatile short-range communication technology within the framework of optical wireless communications. In OCC, the camera captures two-dimensional data in the form of image sequences, thus enabling multidimensional data transmission over the free space channel. However OCC systems offer relatively low data transmission rate, it can be enough for many of different applications like car-to-car communication, indoor navigation, or all-optical Internet of things (OIoT).

In an LED-based illumination and data communication environment, the data modulation at the transmitter must provide a wide range of dimming level and exhibit no flickering. A common human cannot perceive light source flickering when frequency is higher than approximately 75Hz - 100Hz; this so called Critical Flicker Frequency (CFF) depends on light source luminance and viewing angle. Dependency between visual response of human eye and luminance logarithm is described by Ferry-Porter law [47]. Consumer level and low-cost cameras are often limited to maximum frame-rate between 30 fps and 60 fps, so OCC only offers a low data rate, typically tens of bits per second. In order to prevent loss of unsampled data and poor signal detection, OCC often employs special techniques like undersampled on-off keying.

3.1 VLC for Vehicle-to-Vehicle Communication

Almost all modern vehicles have already LED-based head and brake lights, and indicator lights, thus being possible for the concept of vehicular VLC (VVLC) as a new cost-effective way to implement vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications [48, 49, 50]. Furthermore, LED-based roadside units can also be used for both signaling and broadcasting safety related information to vehicles on the road, thus the the V2V communications may play an important role in enhancing vehicle safety and should be reliable and efficient in transmitting traffic related information under various weather conditions on the road.

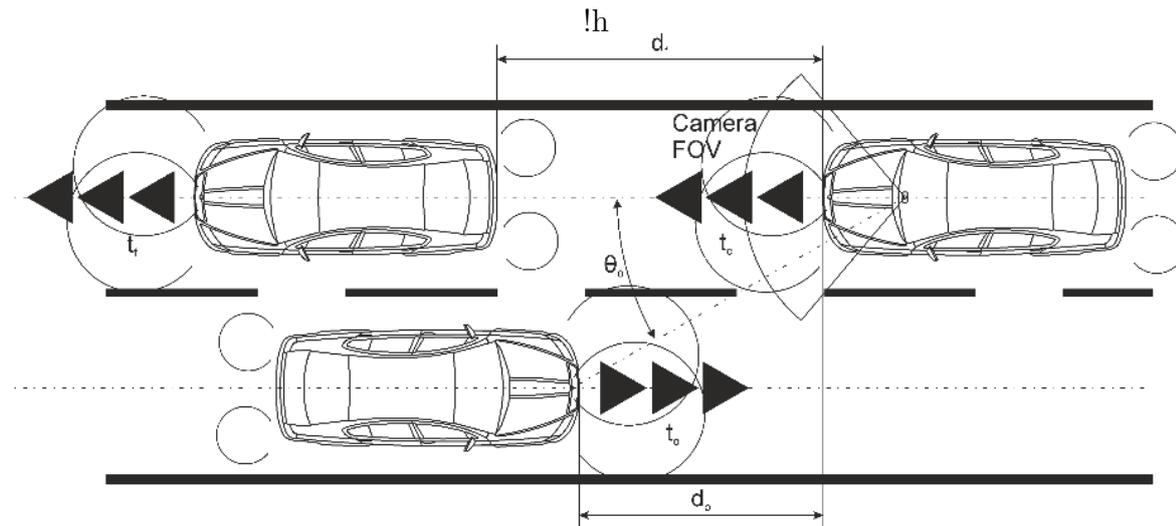


Figure 3.1: Typical road traffic situation for a two-lane road.

Several test use-cases and experimental results have been published for VVLC networks consisting of onboard units, vehicles, and road side units (i.e., traffic lights, street lamps, digital signage, etc.) [51, 52]. Recent studies reported on V2V communications use either PDs or a CS to detect oncoming vehicles and subsequently control the illumination pattern of the head lights to avoid the glaring. VLC systems based on the LED transmitter (Tx) and a camera based Rx were proposed for automotive applications in [53], where a signal reception experiment was performed for static and moving camera Rxs with up to a 15 Mb/pixel/s error-free throughput under fixed condition. In [54], it was shown that under the driving condition the Rx could detect and accurately track an LED Tx array with an error-free communication over a transmission range of 25 – 80 m. Contrary to a typical VLC communications scenario (with a data source, driver, Tx, PD-based Rx, and processing units), the camera-based VLC system can receive and separate multiple light sources within its field of view (FOV) and extract the information using image sequence processing. This detection technique also offers a unique feature, for example, utilization of the multiple-input and multiple-output (MIMO) capabilities supporting parallel data transmission [55].

In V2V and other outdoor VLC systems, the transmission is strongly subjected to weather conditions over the optical channel, such as rain, snow and fog, paths including as well interference due to lights from the Sun, other vehicles, street light, etc. which can reach the camera based receiver positioned within the cars.

3.2 VLC for Indoor Navigation

Another important application of VLC technique is indoor navigation. As the present mainstream in positioning, a satellite-based radionavigation systems like GPS, Galileo or GLONASS are widely used in order to provide real-time information about position. However, in challenging environments, such as urban canyons and indoors like large shopping malls and complex venues, satellite-based positioning and navigation is inaccurate and discontinuous since the signals transmitted by satellites are usually degraded

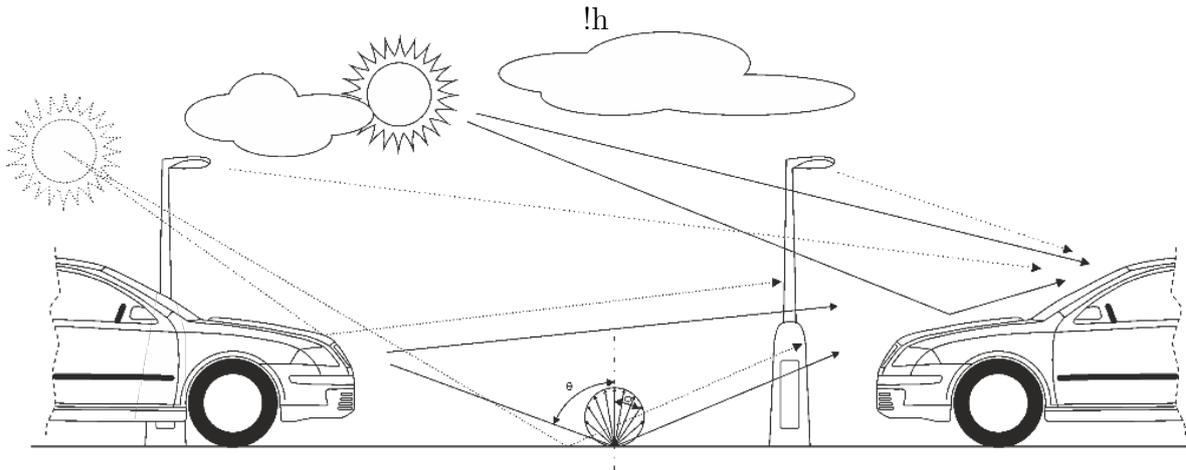


Figure 3.2: A C2C VLC link with possible noise sources.

and interrupted by clouds, ceilings, walls, and other obstructions. Consequently, indoor positioning systems (IPS) using indoor wireless signals (e.g., WiFi [56], Bluetooth [57], radio frequency identification (RFID) [58], and ZigBee [59]) have been proposed to fill the gap of satellite signals to improve the performance of indoor positioning.

Compared with the traditional indoor positional methods, visible light positioning (VLP) is advantageous not only with its stability, convenience and immunity from electromagnetic interference, but also with its positioning accuracy. Several methods exist to collect information from environment and analyze them to localize target [60]. Received signal strength (RSS), time of arrival (TOA), time difference of arrival (TDOA) and angle of arrival (AOA) are some methods which are used to localize target.

The VLP proposed in [61] used wearables with embedded cameras (e.g., smart glasses, watch) and polarized the beacons. Then the orientation and location of the lights were extracted from the camera's video and the device's position was calculated by implementing an AOA algorithm. The average location's error was 30 cm with a 1.8 s delay working at 300 MHz. An experimental demonstration of indoor VLP using image sensors was presented in [62], the position of the camera was determined from the geometrical relations of the LEDs in the images. These LEDs constantly transmitted their coordinates using Phase Shift Keying (UPSK) modulation. The mean positioning error of this VLP reached 5 cm for a distance of 1.20 m, and it increased to 6.6 cm for a height of 1.80 m.

3.3 Author's contribution

Dr. Stanislav Vitek is involved in the European project H2020 Marie Skłodowska-Curie Innovative Training Network (ITN) project No. 764461 "Visible Light Based Interoperability and Networking". He participated in the application for this project and is named in it as one of the key members of the CTU research team I lead. As part of the project implementation, he significantly participates in the management of 2 CTU doctoral students (Shivani Rajendra Teli, Zahra Nazari Chaleshtori) and research activities of another 2 doctoral students from Northumbria University, Newcastle upon

Tyne (Elizabeth Eso) and the University of Las Palmas Grand Canaria, Spain. (Vicente Matus).

Publications related to this section

1. VITEK, Stanislav, et al. Influence of Camera Setting on Vehicle-to-Vehicle VLC Employing Undersampled Phase Shift On-Off Keying. *Radioengineering*, 2017, 26.4: 947. (Appendix C, p. 67)
2. CHVOJKA, Petr, et al. Analysis of nonline-of-sight visible light communications. *Optical Engineering*, 2017, 56.11: 116116. (Appendix D, p. 76)
3. CHAVEZ-BURBANO, P., et al. Optical camera communication system for Internet of Things based on organic light emitting diodes. *Electronics Letters*, 2019. (Appendix E, p. 84)

4 Wireless smart cameras

Recent technological progress enabled closed circuit (CCTV) to become the regular part of wireless sensor networks (WSN). In some applications, such as environmental monitoring, automatic license plate recognition or management of parking spaces, it is possible to process visual information locally and transmit only reduced text information. On the other hand, when a camera system is used, for example, to prevent crime or to identify offender, highly compressed video stream is transmitted to the server where further post-processing is applied.

This chapter deals with low-power smart camera systems which are using limited computing resources. It focuses on camera systems mentioned above which employ image and video processing algorithms optimized to speed and have the ability to communicate with the central server through both short and long-range wireless network like 4G or now popular IoT (Internet of Things) networks SigFox and LoRa. First section is dedicated to selected aspect of camera systems used for assistive purposes, second part focuses to design of low-power camera system to manage parking slots.

4.1 Cameras for assistive technologies

In 2060, there would be more than twice as many elderly than children. In 2008, there were about three and a half times as many children as very old people (above 80). In 2060, children would still outnumber very old persons, but by a small margin: the number of very old people would amount to 80% of the number of children. Those numbers means that more money will be spent on care for the sick and elderly people. In this situation, remote monitoring can reduce the amount of recurring admissions to hospital, facilitate more efficient clinical visits with objective results, and may reduce the length of a hospital stay for individuals who are living at home. Telemonitoring can also be applied on long-term basis to elderly persons to detect gradual deterioration in their health status, which may imply a reduction in their ability to live independently.

4.1.1 Video analysis

In remote video monitoring for assistive purposes, video analysis generally means recognition of basic human activities and triggering actions in case of unexpected behavior of the monitored person. The problem of human action recognition is quite complicated but with adequate choice of image processing methods is possible to find model of articulated non-rigid body. We aim to recognize five types of human daily activities: lying, sitting, standing, walking and other movements including transitions between sitting and standing or lying, and some leg movements when the human subject is sitting or lying – these movements are not assumed to be comparable to walking.

Generally it's possible to split problem of video analysis into following five points:

Feature extraction – the goal of feature extraction is to reduce a variable sized image to a fixed set of visual features. Such features cover wide range of indicators from relatively simple (colour patterns, edges, corners, histogram) to more complex

(blobs, foreground/background estimation, segmentation, optical flow). Among the representatives of the low-level feature extraction algorithms, we consider methods such as Canny edge detector, Harris or Hessian corner detectors, or some shape based methods like template matching and Hough transform. At the present time, also deep convolutional networks has become one most advanced and effective method of feature extraction, especially if the available computing power is sufficient.

Feature tracking – tracking visual features in video allows for an estimate of pixel-level correspondences and pixel-level changes among adjacent video frames. It is a key to providing critical temporal and geometric information for object motion/velocity estimation, camera self-calibration and visual odometry. Typical methods to employ are Kalman filters or particle filters.

Objects localization and classification – predicting the type or class of an object in an image including variations within one class of objects (for example different human poses). For the object recognition can be used for example pre-trained Haar classifiers, linear SVM (Support-Vector Machine) classifier model or CNNs (Convolutional Neural Nwtworks).

Recognition of spatio-temporal patterns – like above mentioned lying, sitting, etc. For the tracking of human motion are used points or blobs based models – see Fig. 4.1.

Storing of obtained information as metadata linked to the original image or video data.

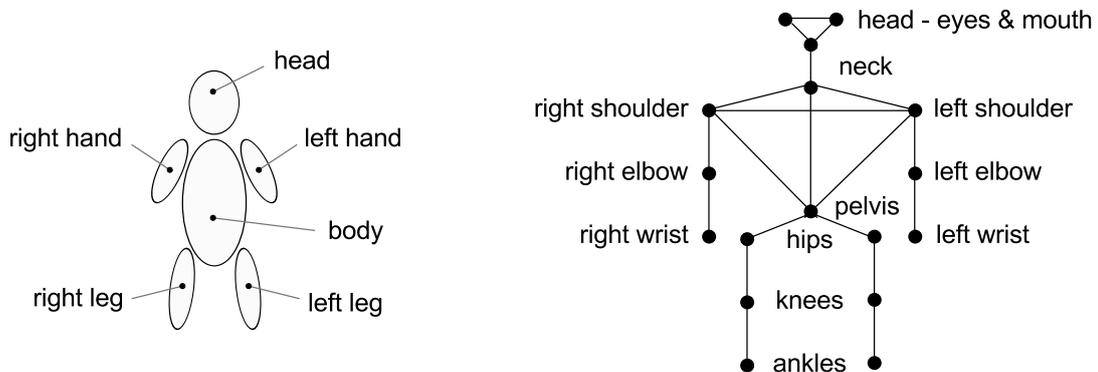


Figure 4.1: Models of human body. (a) Blob-based human model. (b) Stick-figure human model.

For the detection of various kind of objects could be successfully used Viola-Jones object detection framework Viola & Jones (2001), which is able to provide competitive object detection rates in real-time. It can be trained to detect a variety of object classes: human face, hand, upper body etc. During detection phase of the method a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. (Simple Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle, which can be at

any position and scale within the original image. Each feature type can indicate the existence or absence of certain characteristics in the image.

4.1.2 Person identification in low bitrate video-sequences

Typical wireless cameras can encode captured frames into low bitrate video sequences. However, facial recognition methods are usually tasks demanding high computational power. Since it is also necessary to continually update the face image database for comparison purposes, it preferable to perform face tasks on more powerful machines. However, highly effective compression algorithms will, of course, affect the efficiency of face recognition methods.

Some of the compression techniques, used in the field of security, were evaluated by Klima et al. [63]. The impact of compression of CCTV videos on the ability to identify the person's face was investigated by Kovesi [64] or Keval and Sasse [65], who proposes the use of DCT-based compression over wavelet-based for these purposes. Kovesi also points out that the color information is distorted due to quantization and therefore the importance of pigmentation is lost. Study of the impact of the degree of compression by HEVC on observers' ability to detect certain events in videos obtained by outdoor CCTV cameras was also performed [66]. Apart from standard DCT-based compression algorithms, Elad et al. proposed special purpose low bit-rate compression of face images [67].

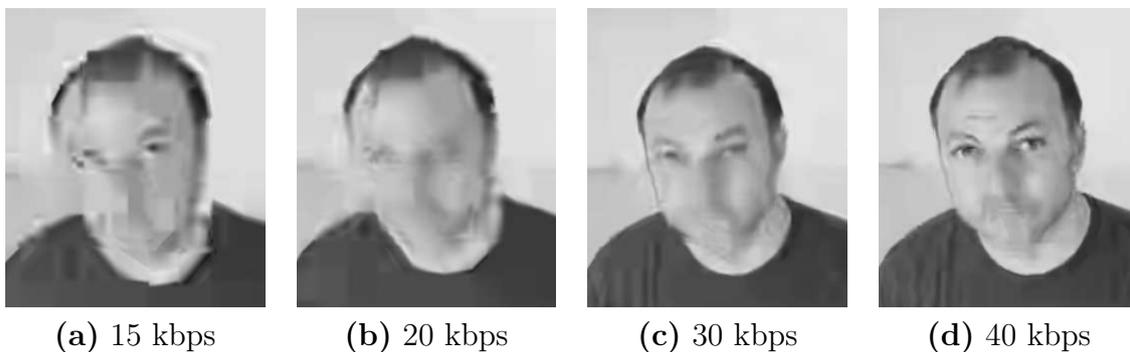


Figure 4.2: Example of HEVC extremely low bitrate compression.

Since H.264 has been the most widely accepted and adopted format in past years, its successor H.265/HEVC became de facto standard in an online and broadcast domain for content compression and distribution, mostly due to Ultra High Definition Television (UHDTV) penetration in the market. HEVC standard brings the promise of huge bandwidth savings of approx. 40-45% over H.264 [68] encoded content with similar subjective quality [69]. HEVC replaces macroblocks, used in previous standards, with Coding Tree Units (CTUs), able to use a larger block structure of up to 64x64 pixels and to better sub-partition the picture into variable sized structures [70]. HEVC initially divides the picture into CTUs which can be 64x64, 32x32, or 16x16 with a larger pixel block size usually increasing the coding efficiency. However some aspects of the H.265 design require more processing than in previous H.264/AVC (Advanced Video Coder) standards, some other aspects have been simplified, and software encoding and decoding is very feasible on current devices [71].

In this chapter are presented results of the study of the reliability of the reliability of human observers and an automatic facial recognition algorithm when identifying an unknown person in the CCTV footage under different levels of HEVC compression. For the study were prepared seven video-sequences simulating two scenarios:

1. person to be identified passes through the corridor, in about half-way, the person looks directly into the camera. This scenario can represent cases when CCTV camera is placed in the shop window.
2. person to be identified enters the room and stays in the reserved area, looking into the camera for a short moment. This scenario represents the CCTV footage from the bank, or at the post-office.

The videos in the dataset are of resolution 768×576 pixels, 25 frames per second and with YUV 4:2:0 color sampling. They were compressed by x265 encoder⁴, an open source free software, and library for encoding video using HEVC, using Main profile with default settings (hierarchical encoding, without deblocking and adaptive loop filter).

The most important aspects influencing the quality of surveillance video systems are the illumination conditions of the site, the camera, video compression, viewing angle, and angular resolution [72]. In the scope of this study, only the compression is taken into account, all other parameters affecting identification of the person are kept constant in order to eliminate their influence. Following results of previous work [66], test videos have bitrates of 70 kilobits per second (kbps) and lower, taking into account that average human observer begins to perceive the influence of compression on bitrates below 40 kbps. Five different degrees of HEVC compression were applied to the videos, resulting in average bitrates 70, 60, 50, 40, 30, 20, and 15 kbps.

To compare human performance with automated facial recognition methods, following experiments were designed: at first, subjective tests with a group of human observers were performed, at second OpenFace framework [73] is used to identify previously detected faces in the videos. The experimental design has been inspired by the real-world application, where security footage is shown to a subject, and he/she is supposed to identify the unknown person in the video based on the standard ID picture. The objective of the experiment is also to estimate how the video compression level affects the observer's decision if he has to choose from more than one possibility. The results of the experiment can be used to determine the smallest possible bitrate that can be used to identify a person in the video sequence. Another application can be optimization of the bitstream of security cameras; in the majority of the time there is no activity in the field of view, so it is possible to save bandwidth, but if a human face is detected, bitrate of the stream can be adaptively increased.

Subjective tests

The subjective tests were performed in subjects' home environment. The sequences included seven videos from CCTV cameras with a different person in each of them.

⁴<http://x265.org>



Figure 4.3: One frame from the testing video, second scenario. Reference faces are distorted for the purpose of publication.

The first video in the sequence was considered as training for the observers and the votes were not reflected in the results. Videos with duration of five seconds were displayed twice in a row with the one-second interval of the mid-gray background between them. Simultaneously, photographs of the actor and three other persons selected to look similar to the actor (e.g. similar haircut, a shape of the face, etc.) were displayed on the left side of the screen (see Figure 4.3).

Participants were supposed to write the number of the position of the person in the video, and also were asked to state the level of their certainty with the decision (1 – not sure at all, 2 – not really sure, 3 – almost sure, 4 – entirely sure). Lastly, they wrote the main reasons that had driven them to the particular decision. This information provided us with more insight into the observers behaviour and the importance of particular features for the identification.

One hundred and thirty-eight observers participated in the tests. Participants were mostly students of bachelor study programme at Czech Technical University in Prague, i.e., in the age between 19 and 21. Students did not receive any training before the experiment. They were all provided with the instruction sheet to follow the above-described procedure. Analyzing the reasons behind the decisions showed that the subjects are highly influenced by the person's hair and shape of the face, if visible enough. They were also likely to take advantage of the differences among the particular possibilities, such as different age or some distinctive features like eyebrows, ears, etc. Generally speaking, the people with dark hair were easier to distinguish because the background was white and the compression artifacts are not that strong in the regions with higher contrast. The hardest to recognize were, therefore, the bald people. The success of the recognition is, of course, also dependent on the set of possibilities given

to the observer.

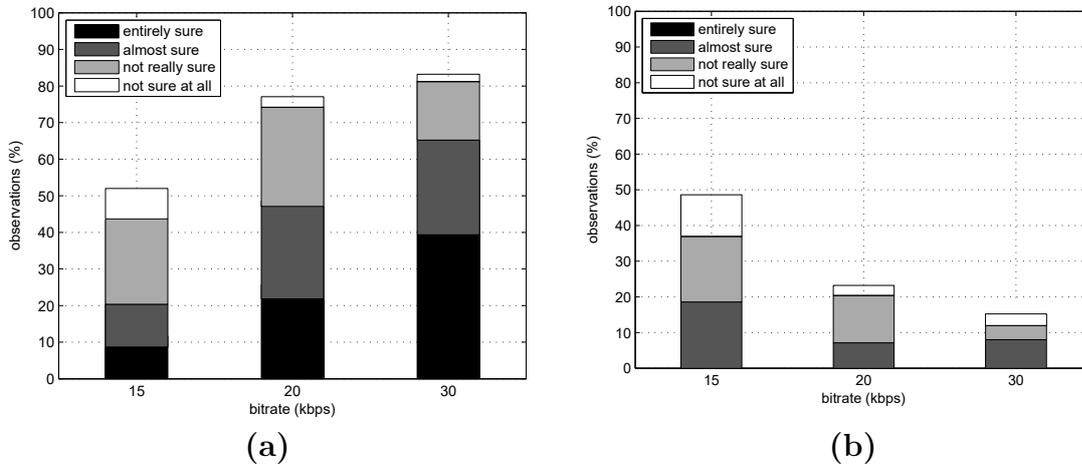


Figure 4.4: Certainty histogram. (a) hits. (b) misses.

The participants were able to correctly recognize 82% of the people in videos with 30 kbps, 79% in videos with 20 kbps, and 55% in videos with 15 kbps. Surely, these values can be affected by the limited number of videos/observers/tasks but some conclusions can be obtained even from this limited set. The histograms of subjects' certainty levels for correct – hits (see Figure 4.4a) and false recognitions – misses (see Figure 4.4b) are depicted. This allows us to take a closer look at the observers' behaviour since the hit with lowest certainty level represents a "lucky" guess. On the other hand, the miss with high confidence suggests the overly self-confident observer and could be the base for a screening of subjects. However, this would have to be dealt with high caution because the error could be also an honest mistake (e.g. typo, etc.).

The histograms for hits (Figure 4.4a) show a very low number of lucky guesses but the considerable amount of recognitions with certainty level 2, especially for low bitrates. The observers were therefore not very confident with their correct recognitions in these cases. The number of confident misses for videos with 15 kbps is alarming and suggests that this level of compression is not suitable for the security applications.

video	framerate [kbps]		
	30 kbps	20 kbps	15 kbps
1	1	1	0.9
2	0.72	0.72	0.27
3	0.63	0.36	0.09
4	0.72	0.54	0.27
5	1	0.90	0.72
6	0.63	0.54	0.18
7	1	1	1
average	0.81	0.72	0.49

Table 1: The success rate of identifying the correct person by the human observer.

video	1	2	3	4	5	6	7
$L2_{min}$	1.13	0.67	1.14	0.73	1.18	1.07	1.13
$L2_{mean}$	1.33	1.12	1.42	1.43	1.25	1.25	1.37

Table 2: $L2$ distances between reference faces.

Table 1 summarizes success rate of face detection achieved by human observers. Table also includes training video-sequence (first row). Values in Table 1 do not take any account of an observer’s certainty.

We also evaluated the similarity of faces in the set of possible actors. Table 2 displays $L2$ distances of the sets, where $L2_{min}$ is minimal $L2$ distance between the proper face and any other, and $L2_{mean}$ represent average $L2$ distance in the set. One can see that poor efficiency of a human observer may be caused by the existence of very similar faces in the set, especially in the cases of videos 2 and 4. In the cases of videos 3 and 6, the efficiency of the human observer is affected mostly by the behavior of coder; both above-mentioned videos shows actors with very short or even no hair.

Automatic face recognition

As a representative of the facial recognition methods based on deep learning networks was selected OpenFace [73]. This pipeline, built on the foundation of the scientific computing framework Torch,⁵ can detect and track facial landmark, estimate head poses, estimate eye gaze and recognize facial action. For our purpose, we took the opportunity to compare two photographs, i.e. compressed face found in the single frame and picture of the possible actor. For facial landmark detection, OpenFace uses Conditional Local Neural Fields (CLNF [74]) which learns the nonlinearities and spatial relationships between pixel values and the probability of landmark alignment [74] and detects 68 landmarks including eyes, lips, and eyebrows. The output is the predicted similarity score of two faces computed as the squared $L2$ distance between their representations.⁶ For more details about end to end learning for the task of the face recognition using Convolutional Neural Networks (CNN), a reader is encouraged to refer to [75, 76, 77].



Figure 4.5: Compared faces. (a,b) Person A. (c,d) Person B.

⁵<http://torch.ch/>⁶<https://cmusatyalab.github.io/openface/demo-2-comparison>

Following supportive experiment explores OpenFace abilities – the mentioned framework is used to compare two faces depicted in Figure 4.5. Table 3 shows $L2$ distances between two faces of two people; a lower score indicates two faces are more likely of the same person. According to [73], a $L2$ distance threshold of 0.99 would distinguish two faces, which is also demonstrated in the Table 3.

Image 1	Image 2	$L2$ distance
Person A1 (Figure 4.5a)	Person A2 (Figure 4.5b)	0.533
Person A1 (Figure 4.5a)	Person B1 (Figure 4.5c)	2.266
Person A1 (Figure 4.5a)	Person B2 (Figure 4.5d)	1.957
Person A2 (Figure 4.5b)	Person B1 (Figure 4.5c)	2.677
Person A2 (Figure 4.5b)	Person B2 (Figure 4.5d)	2.245
Person B1 (Figure 4.5c)	Person B2 (Figure 4.5d)	0.505

Table 3: Similarity scores between two faces.

In the second part of the experiment, only those frames of videos under test, which contain at least approximately 60% of the face in the uncompressed version of the video (i.e., faces with both two eyes visible) were selected. The number of selected frames is then represented by number 100 on the vertical axes in the Figure 4.6a and Figure 4.6b, respectively.

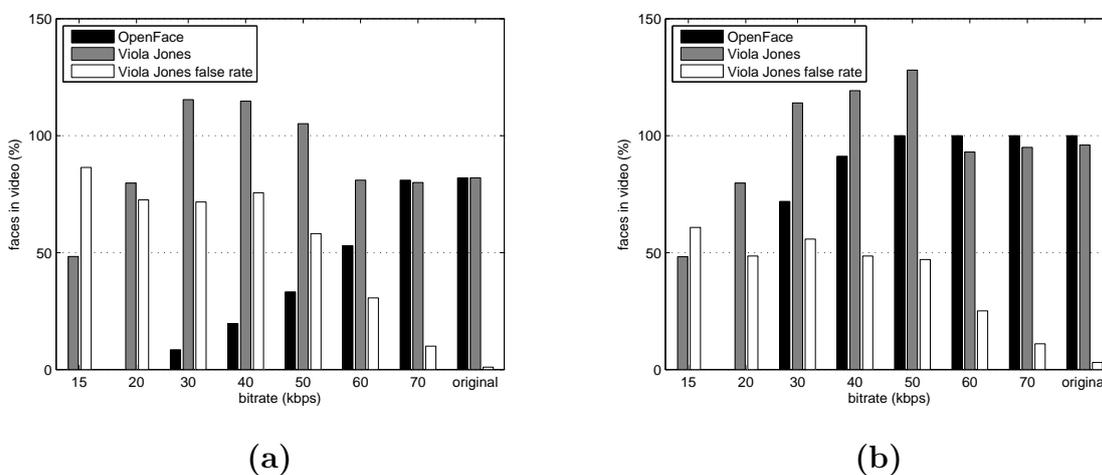


Figure 4.6: Detections of faces. (a) The first scenario. (b) The second scenario.

To evaluate the performance of the automatic face recognition, we employed a methodology based on Receiver Operating Characteristics (ROC) analysis [78]. ROC analysis is a popular way to determine abilities of a classification system. It quantifies the separability of probability distributions under two hypotheses. In our case, one distribution is created by the distances for the detected face and the portrait belonging to the correct person in each frame, while the second distribution is formed by the distances for the other three portraits.

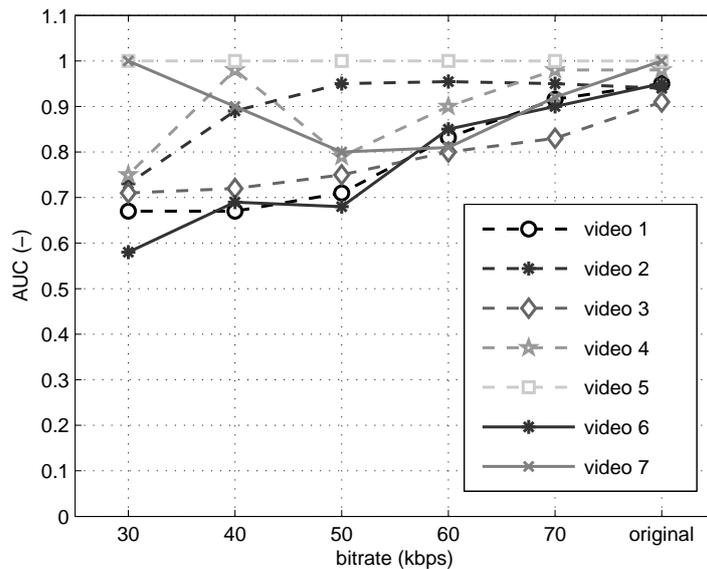


Figure 4.7: AUC values from the ROC analyses for each video.

An outcome of each ROC analysis and a good quantifier of the performance is a Area Under ROC Curve (AUC) [79], calculated as

$$AUC = \frac{1}{N_C \times N_W} \sum_{i=1}^{N_C} \sum_{j=1}^{N_W} \mathcal{H}(W(j) - C(i)), \quad (1)$$

where C and W are the vectors of distances for the two cases (correct and wrong), N_C and N_W are the numbers of samples in each of the vectors, and $\mathcal{H}(\cdot)$ is a Heaviside function defined as

$$\mathcal{H}(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0. \end{cases} \quad (2)$$

AUC value of 1 is reached when the two distributions are completely separated, while the value of 0.5 means that the classification is equivalent to random guessing.

We expect that, if the classifier works correctly, the distances for the incorrect portraits should be larger and the distributions should therefore be well separated. The AUC values for all of the videos are depicted in Figure 4.7.

It can be noticed that some AUC values are lower than 0.5. In these cases, the distances for the correct person are actually *larger* than for the wrong portraits. This represents a very dangerous error.

Overall, the performance drops significantly for the bitrates below 70 kbps and is very poor for 40 and 30 kbps where the human observers only start to have problems with recognition. This outcome proves that the automatic systems are definitely not

video	original	70kbps	60kbps	50kbps	40kbps	30kbps
1	0.90	0.90	0.82	0.66	0.66	0.62
2	0.53	0.52	0.52	0.50	0.50	0.30
3	0.75	0.75	0.76	0.75	0.72	0.69
4	0.97	0.96	0.77	0.76	0.79	0.75
5	1.00	1.00	1.00	1.00	1.00	0.98
6	0.78	0.78	0.70	0.65	0.65	0.40
7	0.77	0.77	0.76	0.76	0.88	1.00
average	0.81	0.81	0.76	0.73	0.74	0.67

Table 4: $CC_{0.99}$ values for each video.

yet ready to substitute human observers for such recognition tasks and require much higher quality input to be reliable.

As demonstrated in the supportive experiment, the faces are considered to be of the same person if the distance is equal to or lower than 0.99. We therefore decided to also calculate the percentage of correctly classified points for the threshold $T = 0.99$ (see Figure ??). Formally, the value $CC_{0.99}$ is defined as

$$CC_{0.99} = \frac{1}{N_C + N_W} \left[\sum_{i=1}^{N_C} \mathcal{H}(0.99 - C(i)) + \sum_{j=1}^{N_W} \mathcal{H}(W(j) - 0.99) \right]. \quad (3)$$

The value indicates to what extent the assumption of the distinguishing ability by this threshold is fulfilled for each video.

Table 4 summarizes values of $CC_{0.99}$ calculated for each of eight video-sequences. Note that videos 1, 2, 3, and 4 belong to the second scenario and videos 5, 6, and 7 belong to the first scenario. Clearly, selection of the scenario does not affect the results of this statistic. Naturally, the best results are obtained for uncompressed video-sequences, which are equivalent to the bitrate of about 4000 kbps.

It should be noted that the bitrate of 70 kbps provides almost the same performance as the original, meaning that such compression is virtually transparent with respect to the classification abilities of the algorithm.

An interesting behavior can be seen in case of video no. 2. The Table 4 shows that only 53% of values are correctly classified, however, the AUC values are high. This means that the algorithm is mostly able to correctly provide larger distances for the incorrect faces but the distances for the correct portrait are often higher than the threshold 0.99.

Unexpectedly high values of $CC_{0.99}$ for low bitrates (30 kbps of videos 5 and 7) are caused by very small numbers of frames, where faces are detected – for details about

efficiency of the automatic face detector see Figure 4.6.

Apparently, there is a link between values in the Table 1 and Table 4. Both human observers and automatic face detector exhibit low success rate for video 2 and video 3. Although it can be seen that the performance of the human observers for 30 kbps is comparable to the performance of the automatic algorithm for 60 kbps.

4.2 Management of parking lots

Another problem, which can be covered with wireless cameras, is the management of parking spaces, i.e., system which can determine occupancy of the parking space based on the information from multiple cameras. The importance of detection of parking space availability is still growing, particularly in major cities. As recent population growth in urban areas, finding a vacant space in parking lots during peak hours may be almost impossible. Numerous studies shown that drivers spend on average eight minutes finding a vacant space [80].

This situation is, of course, widely open to application of some system aims to help driver to found vacant parking space. Systems able to manage this problem can be categorized into counter-based, sensor-based, and image or video based. The first two categories have a couple of drawbacks: counter-based systems could help only with information about a total number of vacant spaces, a sensor-based system costs a lot because of the number of sensors required to cover the entire parking lot. However, the third category is usually considered as quite expensive and producing a significant amount of data, which are unable to transmit over the wireless network, the growth in low-cost, low-power sensing and communication technologies enables a wide range of physical objects and environments to be monitored in fine spatial and temporal detail. A network of dedicated low-power devices connected to the cloud then could be part of the Internet of Things (IoT) platform for smart cities [81].

The vision-based method commonly employs two steps

1. hypothesis generation – it detects objects, and outputs either an image of a vehicle or non-vehicle, can be classified into three basic categories; (1) knowledge-based, (2) stereo-vision-based, and (3) motion-based methods
2. the hypothesis verification – the objects within the image are classified into either vehicles or non-vehicles.

The author of this thesis deals mostly with knowledge-based vision systems [82] using the histogram of oriented gradients (HOG) as a feature extractor and support vector machine (SVM) as a classifier. Feature extraction and namely length of the feature vector is an essential factor affecting both processing time and accuracy. A long feature vector can consume more time and energy in the classification stage. In-vehicle detection, there are several feature extraction methods, including Scale Invariant Feature Transform (SIFT), Haar-like, Gabor filter, log-Gabor. However, HOG remains a popular feature extraction method since it is robust in various conditions like low-light, low-quality images, blurred images, color variation, multi-scales of an image, etc. All the conditions mentioned above are typical for low-cost wireless outdoor cameras.

4.3 Author's contribution

Dr. Stanislav Vitek is a leading developer and manager of the project aiming to prepare a commercially available version of the system described in Appendix F. He proposed a fast and reliable algorithm to classify images of parking spaces, able to run on low-power embedded systems (Raspberry Pi Zero). This project is currently under the support of the Operational Programme Prague – Growth Pole of the Czech Republic.

He is also developing wireless cameras as a regular part of assistive systems in smart homes. He focuses on the reconstruction of video sequences transmitted through low bitrate communication channels.

Publications related to this section

1. VÍTEK, Stanislav; MELNIČUK, Petr. A Distributed Wireless Camera System for the Management of Parking Spaces. *Sensors*, 2018, 18.1: 69. (Appendix F, p. 87)

5 Conclusions and further research

Following our research and findings so far, the future path for my research can be divided into three main categories:

- **Image processing in astronomy** – real-time machine learning-based classification of detected objects, both optical transients and meteors; the information available in flares may be of various origin and nature, so short term predictions of the evolution of the color and light curve may be essential to choose the optimal observational strategy.
- **Optical camera communication** – exploring new potential of optical wireless communication systems based on CMOS cameras. I see a great opportunity in the evolution of complex and personalized VLP indoor navigation systems, which could also help to people with disabilities.
- **Wireless cameras** – evolution of 5G networks will trigger the avalanche. Many problems, such as on-site repairs, monitoring of Industry 4.0 assembly lines, etc. will be solved by the use of AI-powered wireless cameras, enabling untrained workers to do almost any job.

The future is here. And I want to contribute.

References

- [1] Martin Jelínek, Alberto J Castro-Tirado, Antonio de Ugarte Postigo, Petr Kubánek, Sergey Guziy, Javier Gorosabel, Ronan Cunniffe, Stanislav Vítek, René Hudec, Victor Reglero, et al. Four years of real-time grb followup by bootes-1b (2005–2008). *Advances in Astronomy*, 2010, 2010.
- [2] Ray W Klebesadel, Ian B Strong, and Roy A Olson. Observations of gamma-ray bursts of cosmic origin. *The Astrophysical Journal*, 182:L85, 1973.
- [3] Chryssa Kouveliotou, Charles A Meegan, Gerald J Fishman, Narayana P Bhat, Michael S Briggs, Thomas M Koshut, William S Paciesas, and Geoffrey N Pendleton. Identification of two classes of gamma-ray bursts. *The Astrophysical Journal*, 413:L101–L104, 1993.
- [4] SE Woosley. Gamma-ray bursts from stellar mass accretion disks around black holes. *The Astrophysical Journal*, 405:273–277, 1993.
- [5] Chris L Fryer, SE Woosley, and Dieter H Hartmann. Formation rates of black hole accretion disk gamma-ray bursts. *The Astrophysical Journal*, 526(1):152, 1999.
- [6] Javier Gorosabel, AJ Castro-Tirado, S Guziy, A de Ugarte Postigo, D Reverte, A Antonelli, Stefano Covino, Daniele Malesani, D Martín-Gordón, A Melandri, et al. The short-duration grb 050724 host galaxy in the context of the long-duration grb hosts. *Astronomy & Astrophysics*, 450(1):87–92, 2006.
- [7] AS Fruchter, AJ Levan, L Strolger, PM Vreeswijk, SE Thorsett, D Bersier, I Burud, JM Castro Cerón, AJ Castro-Tirado, C Conselice, et al. Long γ -ray bursts and core-collapse supernovae have different environments. *Nature*, 441(7092):463, 2006.
- [8] Alberto J Castro-Tirado, Martin Jelínek, Stanislav Vítek, Petr Kubánek, Josep M Trigo-Rodríguez, Antonio de Ugarte Postigo, Tomas J Mateo Sanguino, and Igor Gomboš. A very sensitive all-sky ccd camera for continuous recording of the night sky. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 70191V–70191V. International Society for Optics and Photonics, 2008.
- [9] Mark Stanford Robbins and Benjamin James Hadwen. The noise performance of electron multiplying charge-coupled devices. *IEEE Transactions on Electron Devices*, 50(5):1227–1232, 2003.
- [10] Jaroslav Hynccek. Impactron-a new solid state image intensifier. *IEEE Transactions on Electron Devices*, 48(10):2238–2241, 2001.
- [11] Mark Stanford Robbins and Benjamin James Hadwen. The noise performance of electron multiplying charge-coupled devices. *IEEE transactions on Electron Devices*, 50(5):1227–1232, 2003.
- [12] Alastair Basden. Analysis of emccd and scmos readout noise models for shack-hartmann wavefront sensor accuracy. *arXiv preprint arXiv:1506.07929*, 2015.

-
- [13] Simon Tulloch. Ing technical note 132. 2010.
- [14] SM Tulloch, P Rodríguez-Gil, and VS Dhillon. Radial velocity study of the post-period minimum cataclysmic variable sdss j143317. 78+ 101123.3 with an electron-multiplying ccd. *Monthly Notices of the Royal Astronomical Society: Letters*, 397(1):L82–L86, 2009.
- [15] SM Tulloch and VS Dhillon. On the use of electron-multiplying ccds for astronomical spectroscopy. *Monthly Notices of the Royal Astronomical Society*, 411(1):211–225, 2011.
- [16] Kennet BW Harpsøe, Michael I Andersen, and Per Kjægaard. Bayesian photon counting with electron-multiplying charge coupled devices (emccds). *Astronomy & Astrophysics*, 537:A50, 2012.
- [17] Mikhail Konnik and James Welsh. High-level numerical simulations of noise in ccd and cmos photosensors: review and tutorial. *arXiv preprint arXiv:1412.4031*, 2014.
- [18] Wenwen Zhang and Qian Chen. Signal-to-noise ratio performance comparison of electron multiplying ccd and intensified ccd detectors. In *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*, pages 337–341. IEEE, 2009.
- [19] Barak Zackay and Eran O Ofek. How to coaad images. i. optimal source detection and photometry of point sources using ensembles of images. *The Astrophysical Journal*, 836(2):187, 2017.
- [20] Barak Zackay and Eran O Ofek. How to coaad images. ii. a coaddition image that is optimal for any purpose in the background-dominated noise limit. *The Astrophysical Journal*, 836(2):188, 2017.
- [21] Nicholas M Law, Craig D Mackay, and John E Baldwin. Lucky imaging: high angular resolution imaging in the visible from the ground. *Astronomy & Astrophysics*, 446(2):739–745, 2006.
- [22] Emmanuel J Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 2013.
- [23] Stefano Zibetti, Simon DM White, Donald P Schneider, and Jon Brinkmann. Intergalactic stars in $z \sim 0.25$ galaxy clusters: systematic properties from stacking of sloan digital sky survey imaging data. *Monthly Notices of the Royal Astronomical Society*, 358(3):949–967, 2005.
- [24] David W Hogg, G Neugebauer, Lee Armus, K Matthews, Michael A Pahre, BT Soifer, and AJ Weinberger. Near infrared imaging of the hubble deep field with the keck telescope. *arXiv preprint astro-ph/9611070*, 1996.
- [25] Richard L. White, David J. Helfand, Robert H. Becker, Eilat Glikman, and Wim de Vries. Signals from the noise: image stacking for quasars in the FIRST survey. *The Astrophysical Journal*, 654(1):99, 2007.

- [26] Peter Kurczynski and Eric Gawiser. A Simultaneous Stacking and Deblending Algorithm for Astronomical Images. *The Astronomical Journal*, 139(4):1592, 2010.
- [27] Günter Dietmar Roth and Günter D Roth. *Handbook of practical astronomy*. Springer, 2009.
- [28] Andrew German, MR Jenkin, and Yves Lespérance. Entropy-based image merging. In *The 2nd Canadian Conference on Computer and Robot Vision (CRV'05)*, pages 81–86. IEEE, 2005.
- [29] James Annis, Marcelle Soares-Santos, Michael A Strauss, Andrew C Becker, Scott Dodelson, Xiaohui Fan, James E Gunn, Jiangang Hao, Željko Ivezić, Sebastian Jester, et al. The sloan digital sky survey coadd: 275 deg² of deep sloan digital sky survey imaging on stripe 82. *The Astrophysical Journal*, 794(2):120, 2014.
- [30] M Nasyrova and S Vitek. Evaluation of co-added astronomical images. *Contrib. Astron. Obs. Skalnaté Pleso*, 47:208–214, 2017.
- [31] Josep M Trigo-Rodriguez, José M Madiedo, Iwan P Williams, and Alberto J Castro-Tirado. The outburst of the κ cygnids in 2007: clues about the catastrophic break up of a comet to produce an earth-crossing meteoroid stream. *Monthly Notices of the Royal Astronomical Society*, 392(1):367–375, 2009.
- [32] Vladimír Porubčan, IP Williams, and L Kornoš. Associations between asteroids and meteoroid streams. *Earth, Moon, and Planets*, 95(1-4):697–712, 2004.
- [33] James S Hey and GS Stewart. Radar observations of meteors. *Proceedings of the Physical Society*, 59(5):858, 1947.
- [34] Pavel Spurný and Jirí Borovicka. The autonomous all-sky photographic camera for meteor observation. In *Asteroids, Comets, and Meteors: ACM 2002*, volume 500, pages 257–259, 2002.
- [35] Robert J Weryk and Peter G Brown. Simultaneous radar and video meteors—i: metric comparisons. *Planetary and Space Science*, 62(1):132–152, 2012.
- [36] Karel Fliegel, Jan Švihlík, Petr Páta, Stanislav Vitek, and Pavel Koten. Meteor automatic imager and analyzer: current status and preprocessing of image data. In *SPIE Optical Engineering+ Applications*, pages 81351S–81351S. International Society for Optics and Photonics, 2011.
- [37] Joseph Ladislav Wiza et al. Microchannel plate detectors. *Nucl. Instrum. Methods*, 162(1-3):587–601, 1979.
- [38] Bahaa EA Saleh and Malvin Carl Teich. *Základy fotoniky 3*. Matfyzpress, 1994.
- [39] Sirko Molau. The meteor detection software metrec. In *Proceedings of the International Meteor Conference, 17th IMC, Stara Lesna, Slovakia, 1998*, pages 9–16, 1999.

- [40] Stanislav Vitek and Maria Nasyrova. Real-time detection of sporadic meteors in the intensified tv imaging systems. *Sensors*, 18(1):77, 2018.
- [41] Stanislav Vitek and Maria Nasyrova. Fast meteor tracking in noisy video sequences. *Astronomische Nachrichten*, 2019.
- [42] Roumen Anguelov and Inger Fabris-Rotelli. Lulu operators and discrete pulse transform for multidimensional arrays. *IEEE Transactions on Image Processing*, 19(11):3012–3023, 2010.
- [43] CH Rohwer and LM Toerien. Locally monotone robust approximation of sequences. *Journal of computational and applied mathematics*, 36(3):399–408, 1991.
- [44] Toshihiko Komine and Masao Nakagawa. Fundamental analysis for visible-light communication system using led lights. *IEEE transactions on Consumer Electronics*, 50(1):100–107, 2004.
- [45] Parth H Pathak, Xiaotao Feng, Pengfei Hu, and Prasant Mohapatra. Visible light communication, networking, and sensing: A survey, potential and challenges. *IEEE communications surveys & tutorials*, 17(4):2047–2077, 2015.
- [46] Lubin Zeng, Dominic C O’Brien, Hoa Le Minh, Grahame E Faulkner, Kyungwoo Lee, Daekwang Jung, YunJe Oh, and Eun Tae Won. High data rate multiple input multiple output (mimo) optical wireless communications using white led lighting. *IEEE Journal on Selected Areas in Communications*, 27(9):1654–1662, 2009.
- [47] Christopher W Tyler and Russell D Hamer. Analysis of visual modulation sensitivity. iv. validity of the ferry–porter law. *JOSA A*, 7(4):743–758, 1990.
- [48] SeokJu Lee, Jae Kyun Kwon, Sung-Yoon Jung, and Young-Hoon Kwon. Evaluation of visible light communication channel delay profiles for automotive applications. *EURASIP journal on Wireless Communications and Networking*, 2012(1):370, 2012.
- [49] Alin Cailean, Barthélemy Cagneau, Luc Chassagne, Suat Topsu, Yasser Alayli, and Jean-Marc Blosseville. Visible light communications: Application to cooperation between vehicles and road infrastructures. In *2012 IEEE Intelligent Vehicles Symposium*, pages 1055–1059. IEEE, 2012.
- [50] Cen B Liu, Bahareh Sadeghi, and Edward W Knightly. Enabling vehicular visible light communication (v2lc) networks. In *Proceedings of the Eighth ACM international workshop on Vehicular inter-networking*, pages 41–50. ACM, 2011.
- [51] Shun-Hsiang Yu, Oliver Shih, Hsin-Mu Tsai, Nawaporn Wisitpongphan, and Richard D Roberts. Smart automotive lighting for vehicle safety. *IEEE Communications Magazine*, 51(12):50–59, 2013.
- [52] Shun-Hsiang You, Shih-Hao Chang, Hao-Min Lin, and Hsin-Mu Tsai. Visible light communications for scooter safety. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 509–510. ACM, 2013.

- [53] Isamu Takai, Shinya Ito, Keita Yasutomi, Keiichiro Kagawa, Michinori Andoh, and Shoji Kawahito. Led and cmos image sensor based optical wireless communication system for automotive applications. *IEEE Photonics Journal*, 5(5):6801418–6801418, 2013.
- [54] Tom Nagura, Takaya Yamazato, Masaaki Katayama, Tomohiro Yendo, Toshiaki Fujii, and Hiraku Okada. Tracking an led array transmitter for visible light communications in the driving situation. In *2010 7th International Symposium on Wireless Communication Systems*, pages 765–769. IEEE, 2010.
- [55] Toru Nagura, Takaya Yamazato, Masaaki Katayama, Tomohiro Yendo, Toshiaki Fujii, and Hiraku Okada. Improved decoding methods of visible light communication system for its using led array and high-speed camera. In *2010 IEEE 71st Vehicular Technology Conference*, pages 1–5. IEEE, 2010.
- [56] Yuan Zhuang, Zainab Syed, You Li, and Naser El-Sheimy. Evaluation of two wifi positioning systems based on autonomous crowdsourcing of handheld devices for indoor navigation. *IEEE Transactions on Mobile Computing*, 15(8):1982–1995, 2016.
- [57] Yuan Zhuang, Jun Yang, You Li, Longning Qi, and Naser El-Sheimy. Smartphone-based indoor localization with bluetooth low energy beacons. *Sensors*, 16(5):596, 2016.
- [58] Antonio Ramón Jiménez Ruiz, Fernando Seco Granja, José Carlos Prieto Honorato, and Jorge I Guevara Rosas. Accurate pedestrian indoor navigation by tightly coupling foot-mounted imu and rfid measurements. *IEEE Transactions on Instrumentation and measurement*, 61(1):178–189, 2012.
- [59] Shih-Hau Fang, Chu-Hsuan Wang, Ting-Yu Huang, Chin-Huang Yang, and Yung-Sheng Chen. An enhanced zigbee indoor positioning system with an ensemble approach. *IEEE Communications Letters*, 16(4):564–567, 2012.
- [60] Ali H Sayed, Alireza Tarighat, and Nima Khajehnouri. Network-based wireless location: challenges faced in developing techniques for accurate wireless location information. *IEEE signal processing magazine*, 22(4):24–40, 2005.
- [61] Zhice Yang, Zeyu Wang, Jiansong Zhang, Chenyu Huang, and Qian Zhang. Wearables can afford: Light-weight indoor positioning with visible light. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, pages 317–330. ACM, 2015.
- [62] Bangjiang Lin, Zabih Ghassemlooy, Chun Lin, Xuan Tang, Yiwei Li, and Shihao Zhang. An indoor visible light positioning system based on optical camera communications. *IEEE Photonics Technology Letters*, 29(7):579–582, 2017.
- [63] Milos Klima and Karel Fliegel. Image compression techniques in the field of security technology: examples and discussion. In *38th Annual 2004 International Carnahan Conference on Security Technology, 2004.*, pages 278–284. IEEE, 2004.

- [64] Peter Kovesi. Video surveillance: Legally blind? In *2009 Digital Image Computing: Techniques and Applications*, pages 204–211. IEEE, 2009.
- [65] Hina U Keval and M Angela Sasse. Can we id from cctv? image quality in digital cctv and face identification performance. In *Mobile Multimedia/Image Processing, Security, and Applications 2008*, volume 6982, page 69820K. International Society for Optics and Photonics, 2008.
- [66] Stanislav Vítek, Lukáš Krasula, Miloš Klíma, Vojtěch Hvězda, and Marcelo Herrera Martínez. Influence of hevc compression on event detection in security video sequences. In *2013 47th International Carnahan Conference on Security Technology (ICCST)*, pages 1–6. IEEE, 2013.
- [67] Michael Elad, Roman Goldenberg, and Ron Kimmel. Low bit-rate compression of facial images. *IEEE Transactions on Image Processing*, 16(9):2379–2383, 2007.
- [68] Ray Garcia and Hari Kalva. Subjective evaluation of hevc and avc/h. 264 in mobile environments. *IEEE Transactions on Consumer Electronics*, 60(1):116–123, 2014.
- [69] Philippe Hanhart, Martin Rerabek, Francesca De Simone, and Touradj Ebrahimi. Subjective quality evaluation of the upcoming hevc video compression standard. In *Applications of digital image processing XXXV*, volume 8499, page 84990V. International Society for Optics and Photonics, 2012.
- [70] Gary J Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand. Overview of the high efficiency video coding (hevc) standard. *IEEE Transactions on circuits and systems for video technology*, 22(12):1649–1668, 2012.
- [71] Frank Bossen, Benjamin Bross, Karsten Suhring, and David Flynn. Hevc complexity and implementation analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12):1685–1696, 2012.
- [72] RA Smith, Ken MacLennan-Brown, John F Tighe, Neil Cohen, Sophie Triantaphillidou, and Lindsay W MacDonald. Colour analysis and verification of cctv images under different lighting conditions. In *Image Quality and System Performance V*, volume 6808, page 68080Y. International Society for Optics and Photonics, 2008.
- [73] Brandon Amos, Bartosz Ludwiczuk, Mahadev Satyanarayanan, et al. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, 6, 2016.
- [74] Tadas Baltrusaitis, Peter Robinson, and Louis-Philippe Morency. Constrained local neural fields for robust facial landmark detection in the wild. In *Proceedings of the IEEE international conference on computer vision workshops*, pages 354–361, 2013.
- [75] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. A discriminative feature learning approach for deep face recognition. In *European conference on computer vision*, pages 499–515. Springer, 2016.

-
- [76] Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, et al. Deep face recognition. In *bmvc*, volume 1, page 6, 2015.
- [77] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.
- [78] John A Swets. *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Psychology Press, 2014.
- [79] Donald Bamber. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of mathematical psychology*, 12(4):387–415, 1975.
- [80] Donald C Shoup. Cruising for parking. *Transport Policy*, 13(6):479–486, 2006.
- [81] Zhanlin Ji, Ivan Ganchev, Máirtín O’Droma, Li Zhao, and Xueji Zhang. A cloud-based car parking middleware for iot-based smart cities: Design and implementation. *Sensors*, 14(12):22372–22393, 2014.
- [82] Stanislav Vitek and Petr Melničuk. A distributed wireless camera system for the management of parking spaces. *Sensors*, 18(1):69, 2018.

A Appendix A

VÍTEK, Stanislav, et al. Long-term continuous double station observation of faint meteor showers. *Sensors*, 2016, 16.9: 1493.

Author's contribution: image processing, data analysis, programming.

Article

Long-Term Continuous Double Station Observation of Faint Meteor Showers

Stanislav Vítek ^{1,*}, Petr Páta ¹, Pavel Koten ² and Karel Fliegel ¹

¹ Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Prague, Czech Republic; pata@fel.cvut.cz (P.P.); fliegek@fel.cvut.cz (K.F.)

² Astronomical Institute of the Academy of Sciences of the Czech Republic, Fričova 298, 251 65 Ondřejov, Czech Republic; koten@asu.cas.cz

* Correspondence: viteks@fel.cvut.cz; Tel.: +420-224-352-232

Academic Editor: Gonzalo Pajares Martinsanz

Received: 13 July 2016; Accepted: 7 September 2016; Published: 14 September 2016

Abstract: Meteor detection and analysis is an essential topic in the field of astronomy. In this paper, a high-sensitivity and high-time-resolution imaging device for the detection of faint meteoric events is presented. The instrument is based on a fast CCD camera and an image intensifier. Two such instruments form a double-station observation network. The MAIA (Meteor Automatic Imager and Analyzer) system has been in continuous operation since 2013 and has successfully captured hundreds of meteors belonging to different meteor showers, as well as sporadic meteors. A data processing pipeline for the efficient processing and evaluation of the massive amount of video sequences is also introduced in this paper.

Keywords: faint meteor shower; meteoroid; CCD camera; image intensifier; image processing

1. Introduction

Modern image sensors used in astronomy provide high sensitivity and high frame-rates, allowing for the detection of weak and rapidly-changing events in the atmosphere. Among these phenomena are meteors, streaks of light that appear in the sky when an interplanetary dust particle ablates in the Earth's atmosphere. The study of meteors and meteoroids provides clues about their parent objects: comets [1] and asteroids [2]. The light curve of the meteor contains information about the mass of the original particle. Both the shape of this curve as well as the height interval where the meteor radiates correspond to the structure of the parent meteoroid. It is possible to investigate the above-mentioned and many other properties of meteors using video records. Video data has the advantage that if the meteor is recorded with high time resolution from at least two stations simultaneously, its atmospheric trajectory can be calculated. Moreover, the heliocentric orbit can be determined if we know the exact time of the event, which is common for video observation. It was shown that the properties of systems with image intensifiers enable the detection of meteors with masses down to fractions of a gram.

The multi-station observation of meteors using two or more video systems first appeared in the 1970s [3,4], and became a standard technique for the measurement of meteoroid trajectories. Current networks vary in the number of cameras and observation locations. Video systems within the framework of the Spanish Meteor Network (SPMN) use three cameras at three different locations [5]. The Cameras for Allsky Meteor Surveillance (CAMS) system operates 60 identical narrow-angle field of view (FOV, 30°) cameras at three locations [6]. The above-mentioned networks employ low-cost 1/3" or 1/2" security cameras [7] with a typical spatial resolution of 720 × 564 pixels and a frame rate between 20 and 25 frames per second (fps). Highly-sensitive E2V CCD still cameras [8] or cameras with custom-made electronic shutter systems [9] are also utilized. GigE (Gigabit Ethernet) cameras

with 30 fps are used in the framework of the French initiative Fireball Recovery and InterPlanetary Observation Network (FRIPON) [10].

Double-station observation using S-VHS (Super Video Home System) camcorders coupled with image intensifiers started at the Ondřejov observatory about two decades ago [11]. The system had a horizontal resolution of 420 lines per picture, and video data were stored on S-VHS tapes, generally unsuitable for scientific purposes due to the mechanical movement of the tape causing video jitter. The MAIA (Meteor Automatic Imager and Analyzer) system introduced in this paper is a technological successor of the original analog one [12]. It consists of two identical stations placed in Ondřejov and Kunžak, Czech Republic (see Figure 1). The distance between stations is 92.5 km. The Ondřejov camera is aiming at the azimuth of 40° , elevation 45° , the Kunžak at azimuth of 120° , elevation 45° .

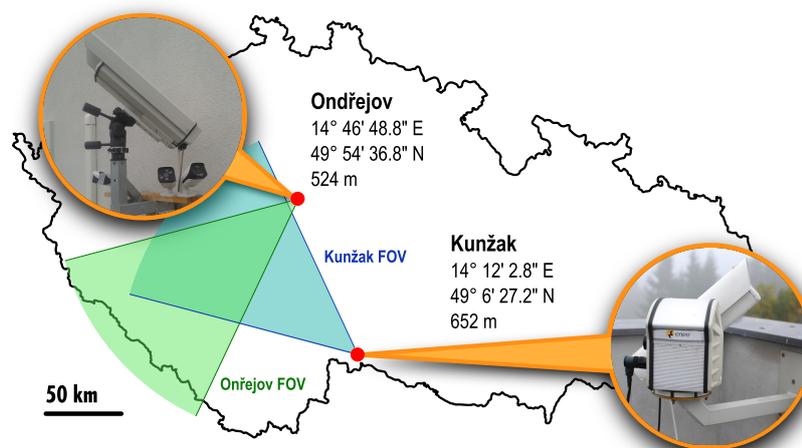


Figure 1. Location of the Meteor Automatic Imager and Analyzer (MAIA) stations. FOV: field of view.

With spatial resolution close to VGA and temporal resolution up to 61 fps, every single MAIA station produces almost 2 TB of raw video data per night. Since there is limited data storage at each MAIA station, image processing algorithms have to make a decision about the importance of the recorded data during daytime and free up the disk space for the following night. Due to restricted Internet access, the problem of the automatic processing of a massive amount of data has to be solved employing local computing power (e.g., Field-Programmable Gate Array, FPGA, or Graphical Processing Unit, GPU). These techniques make high-performance parallel processing tasks feasible. The computing power of a GPU can reach a thousandfold performance of a standard central processing unit (CPU) with an affordable price. The utilization of a GPU enables a substantial improvement in the performance of astronomical data processing algorithms. At this point, it is worth mentioning related solutions of N-body [13], radio-telescope signal correlation [14], adaptive mesh refinement [15], and gravitational microlensing [16]. These examples mostly focus on complicated numerical and cosmological problems, data mining [17], and the visualization of tera-scale astronomical datasets [18]. Classical image processing problems are addressed to a lesser extent.

The paper is organized as follows. Section 2 gives an overview of the technical properties of the MAIA system. Quality and formal aspects of acquired video data (significantly affecting the design of video processing pipeline) are addressed in Section 3. Section 4 concludes the paper.

2. Meteor Automatic Imager and Analyzer

The design of the MAIA system (see Figure 2) is based on our expertise gained with its previous analog version used in Ondřejov for many years. The electro-optical subsystem of MAIA consists of two main components: a second-generation image intensifier XX1332 and a GigE camera JAI CM-040GE. The image intensifier has a large diameter input (50 mm) and output (40 mm) apertures, high gain

(typically 30,000 to 60,000 lm/lm), and a spatial resolution of 30 lp/mm [19]. Since the diameter of the photocathode in image intensifier is 50 mm, and the angle of view for meteor observation should be about 50°, then the most suitable focal length of the input lens comes at about 50 mm. The MAIA system uses a Pentax SMC FA 1.4/50 mm—this lens offers the angle of view of 47°. Due to the large aperture, a high input signal-to-noise ratio is achieved at the intensifier.

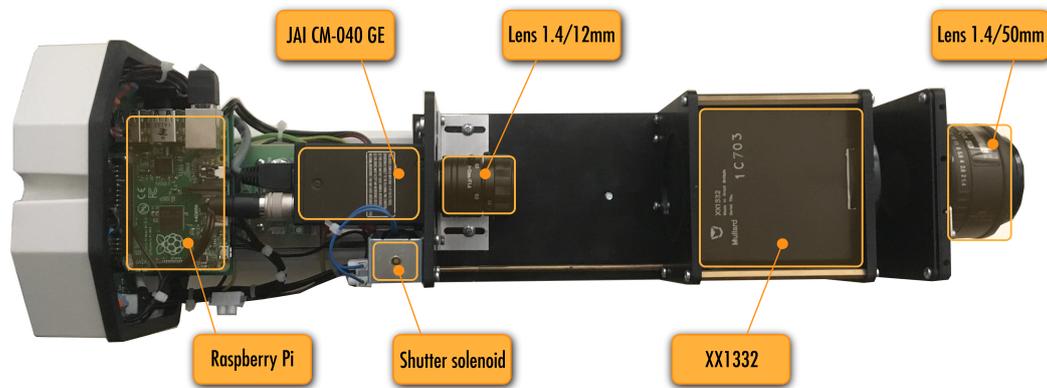


Figure 2. MAIA uncovered.

Camera JAI CM-040GE is equipped with a 1/2" progressive scan CCD sensor. This sensor provides a resolution of 776×582 pixels (i.e., 0.45 megapixels), $8.3 \mu\text{m}$ square pixels, and 10- or 8-bit monochrome output. Maximum framerate of 61.5 fps can be increased if needed with vertical binning and partial scan. The exposure time can vary between 54.874 s to 16.353 ms (or preset electronic shutter 1/60 to 1/10,000 in 10 steps) in full frame scan. The camera has high sensitivity of 1.3 lux (on the sensor, maximum gain, shutter off, 50% of peak video level) and a signal-to-noise ratio greater than 50 dB at 0 dB gain setting. The focal length of the camera lens (Pentax H1214-M 1.4/12 mm) was selected to get a perfect match between the output screen of the image intensifier (diameter of 40 mm), the height of the CCD (4.83 mm), and a suitable distance between the camera and the image intensifier (about 10 cm).

The outer housing of the device was selected taking into account weatherproof requirements. The characteristics of the housing are very similar to those required for video surveillance. The body of the housing is made of extruded aluminum, and the end-cover plates of die-cast aluminum. The weatherproof feature is maintained by the rubber gaskets between the cover plates, and three cable glands. The housing is equipped with a heater kit and a sun shield. One-day exposure of sunlight through the fast lens reliably damages the sensitive layer of the image intensifier. Thus, the crucial part of the device is a mechanical shutter fulfilling the function of protection from the Sun. The solenoid opens the shutter for acquisition while a spring ensures that the shutter is closed in the case of power failure. The electro-mechanical design of the solenoid-operated mechanical shutter was a delicate matter. High reliability was required, but electromagnetic compatibility issues also had to be carefully treated, since the image intensifier is prone to electromagnetic interference.

MAIA has a local computer (Raspberry Pi) to handle data transfers between the main computer and the instrument. The data communication between the main computer and the instrument include video stream from the CCD camera, control signals (shutter, heating, local power supply), and environmental data (temperature, humidity). The distance between the device and the main computer is about 10–15 m. A fiber optic cable is used to ensure uncompromising protection of the instrument during thunderstorms.

2.1. Electro-Optical Characteristics

The XX1332 image intensifier has a highly nonlinear input–output conversion function as a result of the automatic gain control (AGC). This nonlinearity can be characterized by the dependence of the normalized gain (normalized ratio of the output pixel level in the captured image and the input power of the light, measured at a wavelength of 650 nm) on the normalized input power. The curve describing this dependency is depicted in Figure 3a. The image intensifier’s AGC feature helps to accommodate extremely high dynamic range, and also brings high nonlinearity, which is especially critical for photometric measurements.

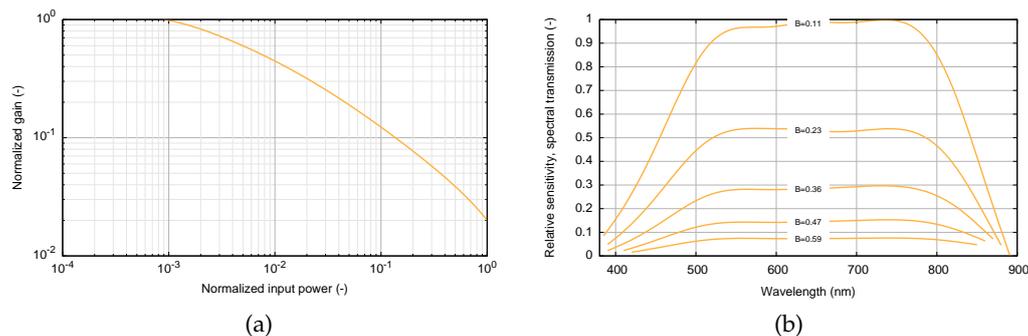


Figure 3. Properties of the MAIA system (a) Gain of the intensifier; (b) Spectral sensitivity.

The overall relative spectral sensitivity can be seen in Figure 3b. This characteristic takes into account the properties of the input lens, the image intensifier, the camera lens, and the camera itself. There are five curves of the relative spectral sensitivity for several digital values in the output image ($B = 1$ means white level; i.e., 255 in 8 bpp representation). This result applies to the particular setting of the camera; i.e., electronic shutter set to 1/100 s exposure time, gain of 0 dB, and zero black level. It is evident that the sensitivity is not constant for a chosen wavelength. This is the impact of AGC, as discussed above. The sensitivity is much higher for low-level light conditions in order to achieve sufficiently bright images on the image intensifier’s output screen. However, the spectral dependence of the sensitivity does not change significantly with the variable gain set by the device’s AGC. The FWHM (Full Width Half Maximum) spectral range of the system is approximately 455–845 nm; i.e., slightly shifted to the near-infrared (NIR) domain. This property is crucial, since meteors radiate significantly in this spectral region [20].

2.2. Spatiotemporal Characteristics

Intensified TV systems exhibit a prominent speckle noise component, caused by the intensifier’s AGC. The level of individual bright spots in the video frame fluctuates significantly, while the overall signal level remains roughly constant (i.e., a couple of bright spots increase their level, while the level is decreased for other bright spots). This phenomenon affects conventional image processing algorithms concerning their scalability and performance. Figure 4a shows the dependence of the pixel intensity standard deviation on the pixel intensity value for a video sequence of 100 frames.

Aside from the high non-linearity discussed in the previous section, MAIA also exhibits features of a shift-variant imaging system. Figure 5 shows the shape of stellar objects in the FOV. It is clearly visible that the object in the middle of the image has a circular shape, whereas objects occurring on the border are heavily distorted. The shape of a stellar object is usually modeled by an asymmetric Gaussian function [21]. Then, the object’s spatial distortion (i.e., its ellipticity) can be described as a ratio of sigma parameters in vertical and horizontal directions. Figure 4b shows the dependence of an object’s ellipticity on the angular distance from the center of the image. The object’s ellipticity plays a significant role in the efficiency of the object detection—for angular distances higher than 40° (i.e., close to the border of the FOV), efficiency decreases significantly.

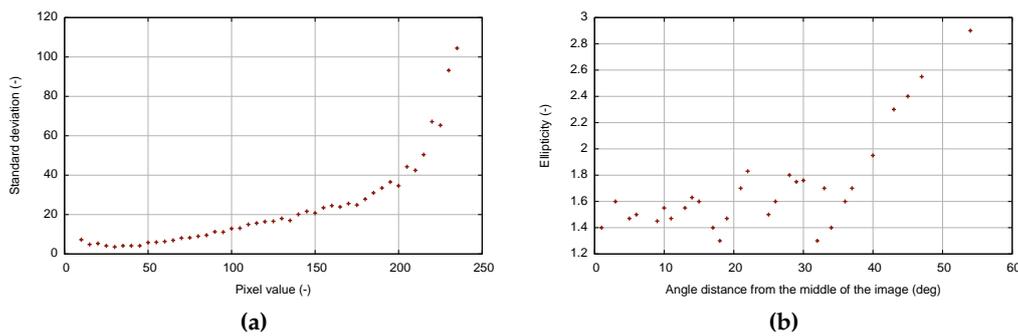


Figure 4. (a) Pixel value standard deviations; (b) Ellipticity of stellar objects.

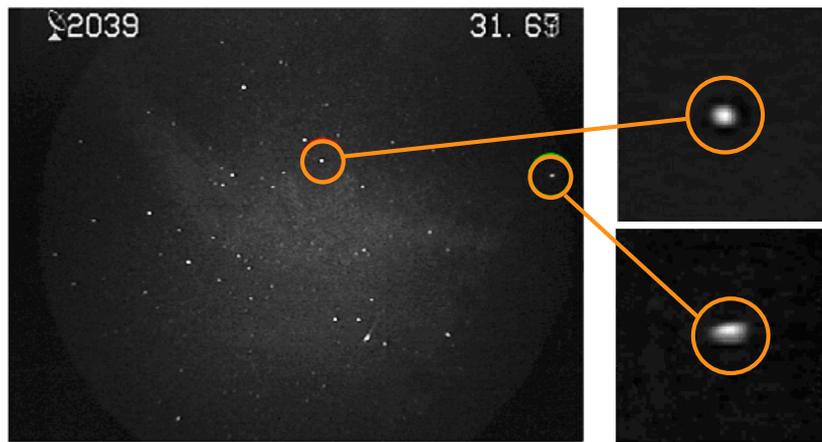


Figure 5. Shift-variant point spread function of the MAIA electro-optical system.

3. Video Data Processing Pipeline

The primary goal of the MAIA video processing pipeline is to find faint meteor showers in the recorded video sequence. A typical video sequence consists mostly of static stellar objects and noise. There is little or no change between the consecutive frames, due to the high frame rate of the camera. Thus, variations in the image data (e.g., meteors or optical transients) are detectable while using relatively simple algorithms based on comparison via image subtraction. Figure 6 summarizes the essential elements of the proposed video processing pipeline. The double-station system deals with simultaneous observation of the same astronomical events observed from two different locations. Therefore, it is important to ensure synchronous execution of the processes running on both stations. The internal clock of the computer is used as the time authority. An NTP (Network Time Protocol) service is employed to synchronize each computer. Achieved precision of approximately one second is sufficient for the successful identification of meteors recorded in double-station configuration. Precise alignment of the appropriate meteor frames is performed through calculation of the meteor's atmospheric trajectory.

As discussed in the previous sections, MAIA's electro-optical characteristics are far from those of an ideal imaging system. Currently, the conventional approach for obtaining the Point Spread Function (PSF) of the space-shift-variant system is based on modeling the wavefront aberrations using Zernike polynomials [22]. An efficient way of bypassing the impact of spatiotemporal fluctuations is described in [23]. The authors employed statistical analysis with nonlinear preprocessing of image intensity using Box-Cox and logarithmic transform. Our pipeline uses a more traditional method based on the detection and classification of the object while taking into account spatial relations between consecutive frames.

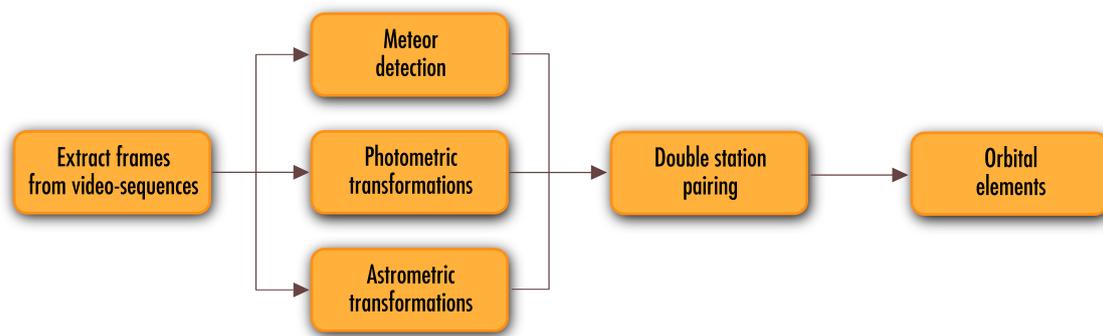


Figure 6. Double-station video processing pipeline.

3.1. Object Detection and Classification

A Canny algorithm [24] is used for object detection in the MAIA system. The algorithm detects edges, while static objects are identified as stars (see Figure 7a). The remaining moving objects can be classified as meteors. Linear motion between consecutive frames is then detected for such moving objects. If an object with linear motion is identified within a certain number of consecutive frames (usually at least five frames), then it is classified as a meteor. Finally, the positions of all detected objects are exported into an external text file (MAIA Object File, MOF). Moreover, the video sequence (approx. 80 frames, i.e. 35 frames before and 35 after the event) containing the meteor candidate is saved into a video file and further processed.

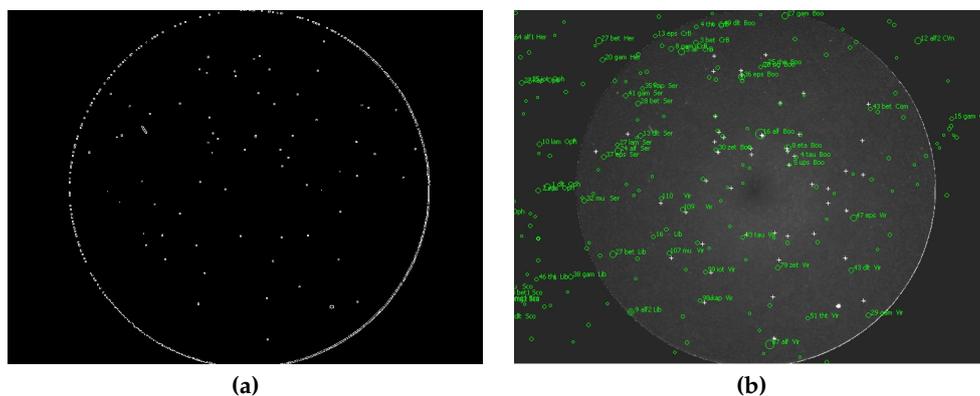


Figure 7. Calibration of the frame. (a) Edges detected by Canny filter; (b) Identified stars.

The first step in the video processing pipeline is the calculation of the calibration image. The calibration image is obtained as a time average of an odd number of video frames (five frames is usually sufficient). Then, the dark frame is subtracted, and the image is flat-fielded. Stars can be detected using values available in the MOF file. Another approach is based on repeated star detection using the Canny method. The star catalog information is loaded using the previously recorded meteor data (i.e., date, time, aiming point of the camera). Then, the catalog stars can be plotted over the acquired image (Figure 7b). Samples of the detected stars (x, y) and catalog stars (α, δ) are aligned. Then, parameters of the transformation between both coordinate systems are determined, and the corresponding pairs are identified. The signal of the stellar object is measured as a sum of the pixels in the box surrounding the star. Finally, the calibration curve is constructed (Figure 8a).

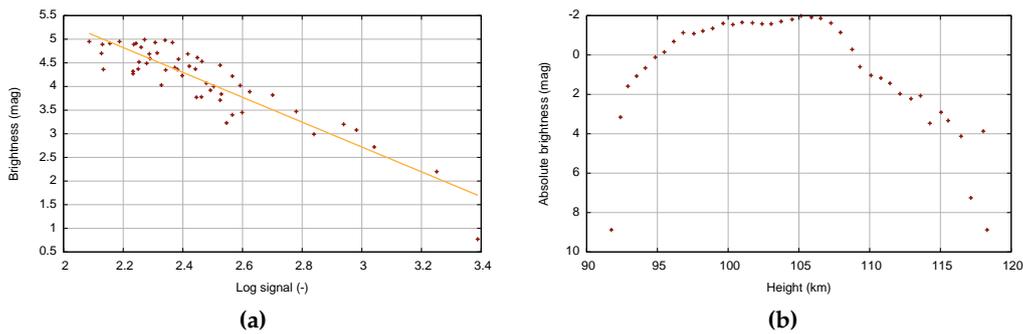


Figure 8. (a) Calibration curve; (b) Light curve of detected meteor.

The second step consists of the meteor's position and brightness measurements. The data stored in the MOF file are used. The particular point can be defined manually, or the MOF positions can be adjusted (Figure 9a). When finished, the boxes around the meteor are set for each frame. The meteor signal is measured as a sum of the pixels within the box (Figure 9b). The brightness of the meteor is determined using the calibration curve. Finally, the light curve of the meteor is calculated (Figure 8b).

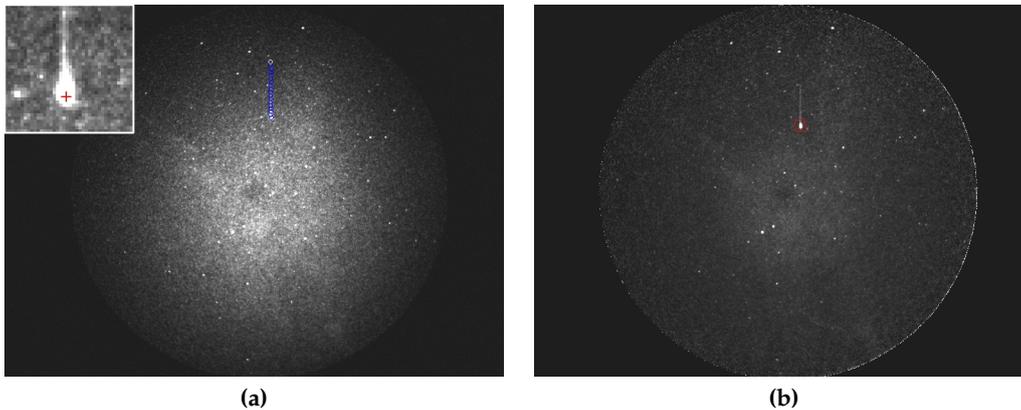


Figure 9. Measurement of meteor parameters. (a) Detected Meteor's path and its magnified detail; (b) Meteor signal measured within the box for each frame.

3.2. GPU Acceleration

The general-purpose CUDA (Compute Unified Device Architecture) GPU is a highly parallel multi-threaded architecture [25]. CUDA became the de facto standard software development kit (SDK) for astronomy computation [26]. One can find numerous studies on the acceleration of image processing for real-time applications, including techniques for real-time moving object detection, a topic related to the subject of this paper. The main bottleneck of GPU acceleration is inefficient data transfer between the host and the device—meaning that implementation of data transfers between the host and the GPU device can negatively affect the overall application performance [27]. The MAIA pipeline solves this issue by batching many small transfers into one larger transfer; i.e., the GPU simultaneously processes more frames.

The most time-consuming operation of the MAIA image processing pipeline is Canny detection. It is used for the detection of objects within particular video frames. This critical part of the pipeline is therefore implemented on the GPU. For the testing, we used video sequences acquired by the MAIA system with a duration of 10 min each (i.e., approximately 36,000 frames at the frame rate of 61 fps). It is worth noting that frame rate is not fully constant and depends highly on the bus workload. The resolution of a single frame is 776×582 pixels. Since a pixel takes 2 B of memory (10b

depth, monochrome), one frame requires approximately 1 MB of memory. Consequently, simultaneous processing of more frames brings substantial performance gain. Figure 10 shows the execution times required for processing, depending on the frame size. An image size of 3104×2328 pixels means that eight regular input frames are processed simultaneously. Assuming the speed of transfer of about 3000 MB/s between the host and the GPU, this transfer takes around 10 ms, including all necessary data (i.e., convolution kernels, both input and output). We compared execution times of the sequential code written in OpenCV, executed using an Intel Core i7-4790 machine with 16 GB of RAM to parallelized code executed using NVIDIA K4200 and NVIDIA Tesla K40 graphic cards.

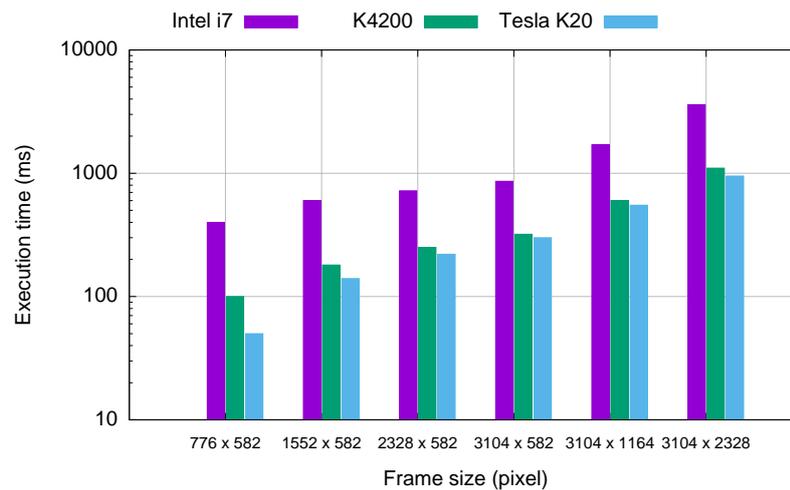


Figure 10. Comparison of the execution time of sequential and parallelized implementation of the Canny detection algorithm.

The time needed for the processing of a 60 s video sequence on the quad-core CPU is about 100 s. The current version of the GPU-optimized code can process the same sequence in 70 s. Our goal is to process a 60 s video sequence in 60 s or faster (i.e., in real time or faster).

4. Conclusions

Automatic double station observations using the MAIA system were carried out continuously during last three years, on each night with good weather conditions. For example, in the year 2015, the camera at Kunžak station was in operation for 202 nights and recorded more than 6500 meteors. The data for the Ondřejov station are slightly lower, since the observational conditions in Ondřejov are worse, mostly due to light pollution. Currently, the observations and meteor detection are fully automatic, while the video data processing is semiautomatic, since the operator's input is still needed.

The comparison of the detection efficiency (which was done for the Perseid campaign) shows that the current MAIA pipeline can detect about 72% of meteors detected by the old system with a S-VHS camera. In the case of the old system, the detection software was usually run several times with parameters tuned manually by the operator. Such an approach is not applicable to the automatic MAIA system. On the other hand, the MAIA system successfully detected some meteors which were not detected by the old system.

The current implementation offers close to real-time video processing; i.e., the processing time is close to the duration of the captured video sequence. Once the development of the MAIA pipeline is finished, it will be released under GPL (General Public License) or a similar license (For information about the current status of the project, please visit <http://maia-project.net>).

Acknowledgments: This work was supported by the grant No. 14-25251S Nonlinear imaging systems with spatially variant point spread function of the Czech Science Foundation. We also would like to thank namely Lukáš Fritsch for his help with the electronics, Petr Janout for his help with the optics, and Filip Kozbrzek for his help with the mechanical part of the system.

Author Contributions: S. Vítek wrote the paper with contribution of all co-authors. All co-authors discussed results. S. Vítek wrote control software and GPU processing. P. Koteň and P. Páta wrote image processing pipeline. K. Fliegel designed electro-optical part of MAIA system.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Trigo-Rodríguez, J.M.; Madiedo, J.M.; Williams, I.P.; Castro-Tirado, A.J. The outburst of the κ Cygnids in 2007: Clues about the catastrophic break up of a comet to produce an Earth-crossing meteoroid stream. *Mon. Not. R. Astron. Soc.* **2009**, *392*, 367–375.
2. Porubčan, V.; Williams, I.; Kornoš, L. Associations between asteroids and meteoroid streams. *Earth Moon Planets* **2004**, *95*, 697–712.
3. Clifton, K.S. Television studies of faint meteors. *J. Geophys. Res.* **1973**, *78*, 6511–6521.
4. Hawkes, R.L.; Jones, J. Two station television meteor studies. In *Symposium-International Astronomical Union*; Cambridge University Press: New York, NY, USA, 1980; Volume 90, pp. 117–120.
5. Madiedo, J.M.; Trigo-Rodríguez, J.M. Multi-station video orbits of minor meteor showers. *Earth Moon Planets* **2008**, *102*, 133–139.
6. Jenniskens, P.; Gural, P.; Dynneson, L.; Grigsby, B.; Newman, K.; Borden, M.; Koop, M.; Holman, D. CAMS: Cameras for Allsky Meteor Surveillance to establish minor meteor showers. *Icarus* **2011**, *216*, 40–61.
7. Samuels, D.; Wray, J.; Gural, P.S.; Jenniskens, P. Performance of new low-cost 1/3" security cameras for meteor surveillance. In Proceedings of the 2014 International Meteor Conference, Giron, France, 18–21 September 2014; pp. 18–21.
8. Oberst, J.; Flohrer, J.; Elgner, S.; Maue, T.; Margonis, A.; Schrödter, R.; Tost, W.; Buhl, M.; Ehrich, J.; Christou, A.; et al. The smart panoramic optical sensor head (SPOSH)—A camera for observations of transient luminous events on planetary night sides. *Planet. Space Sci.* **2011**, *59*, 1–9.
9. Atreya, P.; Vaubaillon, J.; Colas, F.; Bouley, S.; Gaillard, B. CCD modification to obtain high-precision orbits of meteoroids. *Mon. Not. R. Astron. Soc.* **2012**, *423*, 2840–2844.
10. Colas, F.; Zanda, B.; Bouley, S.; Vaubaillon, J.; Vernazza, P.; Gattacceca, J.; Marmo, C.; Audureau, Y.; Kwon, M.K.; Maquet, L.; et al. The FRIPON and Vigie-Ciel networks. In Proceedings of the 2014 International Meteor Conference, Giron, France, 18–21 September 2014; pp. 34–38.
11. Koteň, P. Software for processing of meteor video records. In Proceedings of the Asteroids, Comets, Meteors—ACM 2002, Berlin, Germany, 29 July–2 August 2002; pp. 197–200.
12. Koteň, P.; Stork, R.; Páta, P.; Fliegel, K.; Vítek, S. Simultaneous analogue and digital observations and comparison of results. In Proceedings of the 2016 International Meteor Conference, Egmond, The Netherlands, 2–5 June 2016; pp. 133–136.
13. Hamada, T.; Nitadori, K.; Benkrid, K.; Ohno, Y.; Morimoto, G.; Masada, T.; Shibata, Y.; Oguri, K.; Taiji, M. A novel multiple-walk parallel algorithm for the Barnes–Hut treecode on GPUs—towards cost effective, high performance N-body simulation. *Comput. Sci.-Res. Dev.* **2009**, *24*, 21–31.
14. Wayth, R.B.; Greenhill, L.J.; Briggs, F.H. A GPU-based real-time software correlation system for the murchison widefield array prototype. *Publ. Astron. Soc. Pac.* **2009**, *121*, 857–865.
15. Schive, H.Y.; Tsai, Y.C.; Chiueh, T. Gamer: A GPU-accelerated adaptive mesh refinement code for astrophysics. *Astrophys. J. Suppl. Ser.* **2010**, *186*, 457–484.
16. Thompson, A.C.; Fluke, C.J.; Barnes, D.G.; Barsdell, B.R. Teraflop per second gravitational lensing ray-shooting using graphics processing units. *New Astron.* **2010**, *15*, 16–23.
17. Brescia, M.; Cavuoti, S.; Djorgovski, G.S.; Donalek, C.; Longo, G.; Paolillo, M. Extracting knowledge from massive astronomical data sets. In *Astrostatistics and Data Mining*; Springer: New York, NY, USA, 2012.
18. Hassan, A.; Fluke, C.J.; Barnes, D.G.; Kilborn, V. Tera-scale astronomical data analysis and visualization. *Mon. Not. R. Astron. Soc.* **2013**, doi:10.1093/mnras/sts513.
19. Fliegel, K.; Švihlík, J.; Páta, P.; Vítek, S.; Koteň, P. Meteor automatic imager and analyzer: Current status and preprocessing of image data. *SPIE Opt. Eng. Appl.—Int. Soc. Opt. Photon.* **2011**, doi:10.1117/12.893700.

20. Borovička, J.; Koteň, P.; Spurný, P.; Boček, J.; Štork, R. A survey of meteor spectra and orbits: Evidence for three populations of Na-free meteoroids. *Icarus* **2005**, *174*, 15–30.
21. Stetson, P.B. DAOPHOT: A computer program for crowded-field stellar photometry. *Publ. Astron. Soc. Pac.* **1987**, *99*, 191–222.
22. Anisimova, E.; Bednář, J.; Blažek, M.; Janout, P.; Fliegel, K.; Páta, P.; Vítek, S.; Švihlík, J. Estimation and measurement of space-variant features of imaging systems and influence of this knowledge on accuracy of astronomical measurement. *SPIE Opt. Eng. Appl.—Int. Soc. Opt. Photon.* **2014**, doi:10.1117/12.2061736.
23. Kukul, J.; Klimt, M.; Švihlík, J.; Fliegel, K. Meteor localization via statistical analysis of spatially temporal fluctuations in image sequences. *SPIE Opt. Eng. Appl.—Int. Soc. Opt. Photon.* **2015**, doi:10.1117/12.2188186.
24. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *6*, 679–698.
25. Cook, S. *CUDA Programming: A Developer's Guide to Parallel Computing with GPUs*; Newnes: Sebastopol, CA, USA, 2012.
26. Fluke, C.J.; Barnes, D.G.; Barsdell, B.R.; Hassan, A.H. Astrophysical supercomputing with GPUs: Critical decisions for early adopters. *Publ. Astron. Soc. Aust.* **2011**, *28*, 15–27.
27. Harris, M. How to Optimize Data Transfers in CUDA C/C++. Available online: <https://devblogs.nvidia.com/parallelforall/how-optimize-data-transfers-cuda-cc/> (accessed on 13 July 2016).



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

B Appendix B

VÍTEK, Stanislav; NASYROVA, Maria. Real-time detection of sporadic meteors in the intensified tv imaging systems. *Sensors*, 2018, 18.1: 77.

Author's contribution: main idea, design of the system, data analysis.

Article

Real-Time Detection of Sporadic Meteors in the Intensified TV Imaging Systems

Stanislav Vítek *  and Maria Nasyrova 

Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2,
166 27 Prague, Czech Republic; nasyrmar@fel.cvut.cz

* Correspondence: viteks@fel.cvut.cz; Tel.: +420-224-35-2232

Received: 9 November 2017; Accepted: 27 December 2017; Published: 29 December 2017

Abstract: The automatic observation of the night sky through wide-angle video systems with the aim of detecting meteor and fireballs is currently among routine astronomical observations. The observation is usually done in multi-station or network mode, so it is possible to estimate the direction and the speed of the body flight. The high velocity of the meteorite flying through the atmosphere determines the important features of the camera systems, namely the high frame rate. Thanks to high frame rates, such imaging systems produce a large amount of data, of which only a small fragment has scientific potential. This paper focuses on methods for the real-time detection of fast moving objects in the video sequences recorded by intensified TV systems with frame rates of about 60 frames per second. The goal of our effort is to remove all unnecessary data during the daytime and make free hard-drive capacity for the next observation. The processing of data from the MAIA (Meteor Automatic Imager and Analyzer) system is demonstrated in the paper.

Keywords: sporadic meteor; real-time detection; image intensifier; meteor automatic imager and analyzer; graphical processing unit

1. Introduction

Meteors are streaks of light appearing in the sky when meteoroids ablate in the Earth's atmosphere. Observation of meteors is a cost-effective way to understand the distribution of material in our solar system. Meteor observations are typically performed using radar [1], passive radio detectors [2], all-sky photographic [3] and CCD (charge coupled device) cameras [4], digital video cameras [5] or television (TV) cameras optionally equipped with an image intensifier. While the radio-based detection methods can be performed during the daytime, thus being suitable for estimation of total meteor activity, camera-based methods are limited to night time. Regardless of this limitation, camera-based observations allow building the light curve (i.e., the time-dependent fluctuations of light emitted by a meteor), which may contain information about the mass and structure of the original particle or parent object: comets [6] and asteroids [7]. Wide-band observation with a suitably-designed bank of photometric filters additionally allows obtaining information about the chemical composition of the meteoroid [8,9]. Although camera-based systems are more common, combinations of multiple ways of observations are also used [10]. All-sky cameras with a huge spatial resolution and long exposure times are great for detecting intense light phenomena, like bolides or fireballs. However, for the calculation of atmospheric trajectory, it is necessary to observe meteors simultaneously from at least two different places, optionally with high temporal resolution. Moreover, a higher frame rate brings more data for the modeling of the meteoroid structure [11].

Meteor observation with two or more camera systems is currently a standard technique for the measurement of meteoroid trajectories. There are networks of different scales and technology: the Spanish Meteor Network (SPMN) [12] has about 25 video and CCD stations; Cameras for Allsky

Meteor Surveillance (CAMS) [13] operates more than 60 narrow-field cameras at five locations in the United States (three in California, one in Florida and also on the Mid-Atlantic coast). The concept was later applied by amateur astronomers in the Netherlands and Belgium [14]. The Croatian Meteor Network (CMN) [15] employs more than 30 cameras. The Polish Fireball Network (PFN) consists of 36 continuously-active stations with 57 sensitive analogue video cameras and 7 high-resolution digital cameras [16]. The Canadian Automated Meteor Observatory (CAMO) uses four CCD cameras running at 80 frames per second (fps) and coupled to 18-mm image intensifiers [17]. The Desert Fireball Network (DFA) currently covers one third of Australia (approximately 2.5 million km²) with the use of 49 digital single reflex camera (DSLR)-based stations with nominal spacing of 130 km [18]. The Fireball Recovery and Interplanetary Observation Network (FRIPON) covers all of the French territory and employs 100 all-sky cameras with an average distance of 100 km between the stations [19]. In cooperation with the FRIPON, an initiative is also being developed, the Italian network PRISMA (Prima Rete per la Sorveglianza sistematica di Meteore e Atmosfera), intended to use gamma-ray spectrometers allowing one to reveal the presence of short-lived cosmogenic radioisotopes [20].

Utilization of a high frame rate camera results in a shorter exposure time, and it will naturally reduce the overall sensitivity of the imaging system. The earliest low-level television (LLTV) meteor observations were made in the 1960s with unintensified image orthicon cameras [21]. While this was a significant step forward in terms of sensitivity when compared with photographic techniques, the sensitivity was later improved by coupling with an image intensifier [22]. The image intensifiers are usually one of two basic types: (a) the first generation consisted of a photocathode followed by an accelerating electron lens, which focused the electron image onto the phosphor of the output window; (b) the second and third generation image intensifiers exploited the phenomenon of electron multiplication in a micro-channel plate (MCP). Several stages of first generation image intensifiers may be cascaded with a combined gain of the order of 100,000. Second generation intensifiers have gains approaching those of three-stage first generation device. In combination with fast and low noise CCD cameras, such a high-speed can detect stars down to about +8 to +9 apparent magnitude [23]. Recent development in CMOS technology suggests that MCP will be replaced by CMOS sensors in low-light-level (LLL) applications. Current CMOS sensors are reaching very low electron noise levels. Moreover, the quantum efficiency of CMOS-based devices can be up to 90% [24]. The state-of-the-art devices dedicated to low light level fast imaging are electron multiplying CCDs (EMCCDs), which are, however, still much more expensive in comparison to MCPs.

There are two main tasks for the software for meteor analysis: meteor detection (optionally in real-time) and determination of meteor parameters from double-station or network-based observations. This paper focuses on the first task, fast and reliable detection of meteors. The choice of a suitable method of detection strongly depends on the method of image acquisition, particularly on the frame rate of the image sensor. A typical meteor track is comprised of a streak lasting up to several video frames propagating linearly across space and time. For longer exposure times, typically used in all-sky systems, those streaks can be relatively long. Thus, a couple of works in this field employ the Hough transform [25,26].

Numerous papers deal with the problem of meteor streak detection. One of the most widely-used software applications performing the task of meteor detection TV systems, MetRec [27], works with the difference image to remove static stellar objects and calculates the sum of the pixels of potential neighboring objects in different directions. The object is considered to be a meteor if one or more of those sums exceeds a certain threshold. MeteorScan, a popular software package developed by Pete Gural [28], uses a technique based on frame differencing, variance-dependent thresholding and multi-frame Hough integration. Gural later proposed a method using matching filter [29], where an object's motion is hypothesized for a particular starting point, speed and direction. Another popular package, UFOCapture (<http://sonotaco.com/>) applies a 5 × 5 spatial filter with frame differencing, which is then masked and thresholded. Kozak describes a typical meteor detection pipeline subtracting the average of 40–50 frames from the currently processed one [30]. CAMO employs All Sky and Guided

Automatic Real-time Detection (ASGARD) [31]. CAMS and CMN use the method of detecting meteors in frames of digital video that had been compacted into a single bit-mapped color file [32]. Since this method produces large volumes of false detections (up to 50%), Siladi et al. [33] proposed a method using neural networks and support vector machine (SVM) to reach a 10% false rate. Molau and Gural [34] reported a real-time success-rate of meteor detection better than 80% for both MetRec and MeteorScan, with a false alarm rate less than 20%.

All methods as mentioned above perform well while processing video sequences with less than VGA spatial resolution and a temporal resolution of no more than 25 fps. However, in the TV systems using frame rates of typically 25 fps or faster, meteor paths can be only a couple of pixels; see Figure 1 for example. As is shown in Figure 2, a meteor streak sampled at a high frame rate (here 61 fps) has a similar shape to stars. Figure 3 displays radial profiles of sampled meteor streaks in comparison with the sampled stellar object at the same times. Moreover, when an observing system employs non-linear devices like MPC, the algorithms have to deal with strong noise with a generally unknown distribution. Since we are targeting highly automated camera systems with minimal human interaction, our goal is also to minimize false alerts.

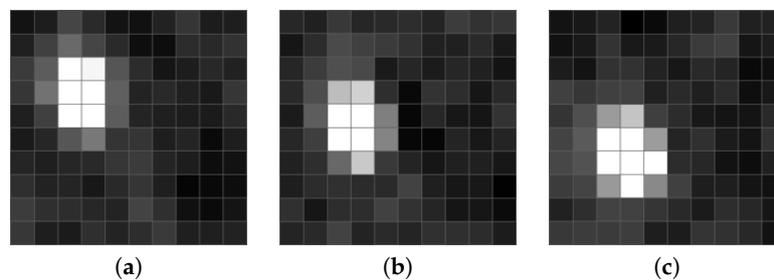


Figure 1. Detail of the sampled meteor streak. (a) Time T. (b) Time T + 30 ms. (c) Time T + 60 ms.

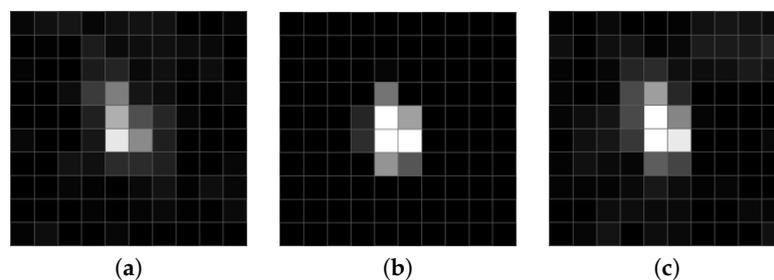


Figure 2. Detail of the static stellar object. (a) Time T. (b) Time T + 30 ms. (c) Time T + 60 ms.

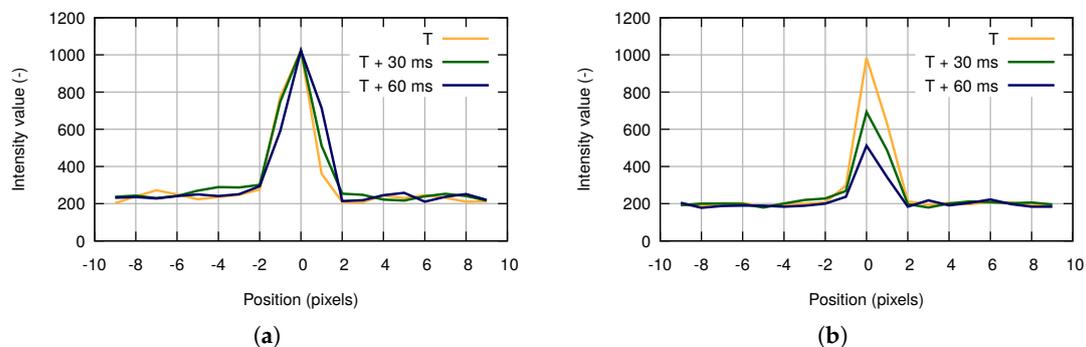


Figure 3. Profiles of the objects present in the frame. (a) Sampled meteor streak. (b) Stellar object.

The paper is organized as follows. Section 2 introduces the characteristics of the second generation image intensifier. Section 3 describes a proposed algorithm for meteor detection in the video sequences

acquired with the TV intensified system. Section 4 gives an overview of graphics processing unit (GPU)-based acceleration of our algorithm. Section 5 presents the results of real data processing, and, finally Section 6 concludes the paper.

2. Characterization of the MCP Image Intensifier

In this section, we will summarize the characteristics of the second generation MCP image intensifier. One of the representatives of this branch of imaging devices is Philips (Photonis) XX1332. The XX1332 image intensifier has a highly nonlinear input-output conversion function as a result of the automatic gain control (AGC). The AGC feature helps to accommodate extremely high dynamic range and also brings high nonlinearity, which is especially critical for photometric measurements. It calculates the mean intensity in the image and adjusts the multiplication gain (which results in higher excess noise) accordingly (increases if less photons are present and decreases for higher overall fluxes).

The AGC feature naturally affects the opto-electronic conversion function (OECF) of the instrument. To cover this characteristic, we used the ISO 14524 chart [35] (see Figure 4) illuminated under laboratory conditions. We used 17 various illuminance levels ranging between 1.6 mlx (mililuxes) and 2.4 lx, which leads to background luminance levels between 125 $\mu\text{cd}/\text{m}^2$ (microcandela per square meter) and 187 mcd/m^2 (milicandela per square meter). From the known illuminance and optical density of the particular patch, it is possible to calculate the patch luminance. Figure 5a displays OECF measured for six of 17 various background levels (gain levels in the image intensifier). It can be seen, however, that the OECF for the fixed gain is not perfectly linear; rather, high linearity is achieved.

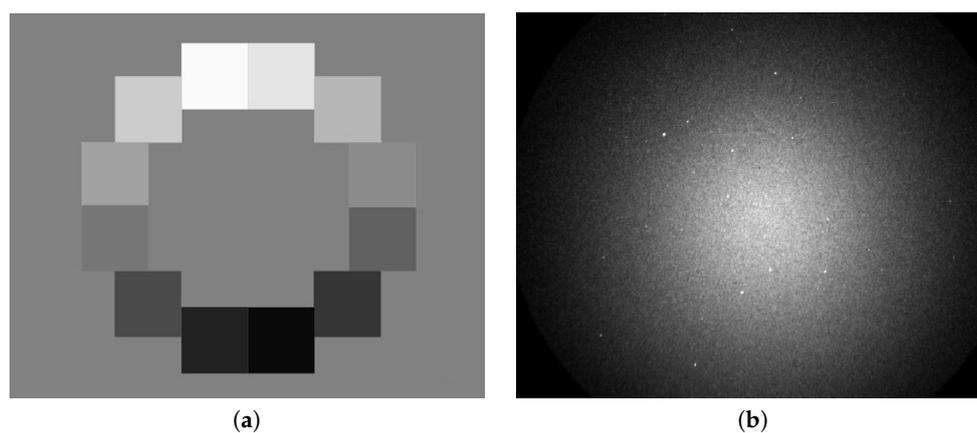


Figure 4. Examples of input data. (a) ISO 14524 test chart. (b) Real image data.

The same ISO 15524 chart allows covering the dependency of noise characteristics on the spatially-variable signal level (represented by patches of the chart) and automatic gain control in the image intensifier. We also investigated the behavior of the image intensifier at different working conditions by the change of the chart illumination [36]. Figure 5b shows the curves for the six chosen illumination levels. Every curve represents the dependency of the standard deviation on the patch luminance. The several curves show that the system is highly signal dependent even in the case of constant illumination level (against the assumption, the standard deviation is growing with growing patch luminance). Furthermore, the standard deviation decreases with the growing background luminance.

The above-mentioned features, typical for the intensifier TV systems, cause the presence of speckle noise components in the acquired video sequences. The level of individual bright spots in the video frame fluctuates significantly, while the overall signal level remains roughly constant (i.e., a couple of bright spots increase their level, while the level is decreased for other bright spots). This phenomenon affects conventional image processing algorithms based on the subtraction concerning their scalability and performance. Together with findings from the measurement of the

noise standard deviation, it naturally leads to the assumption that brighter parts (pixels) of the video sequence have higher variance.

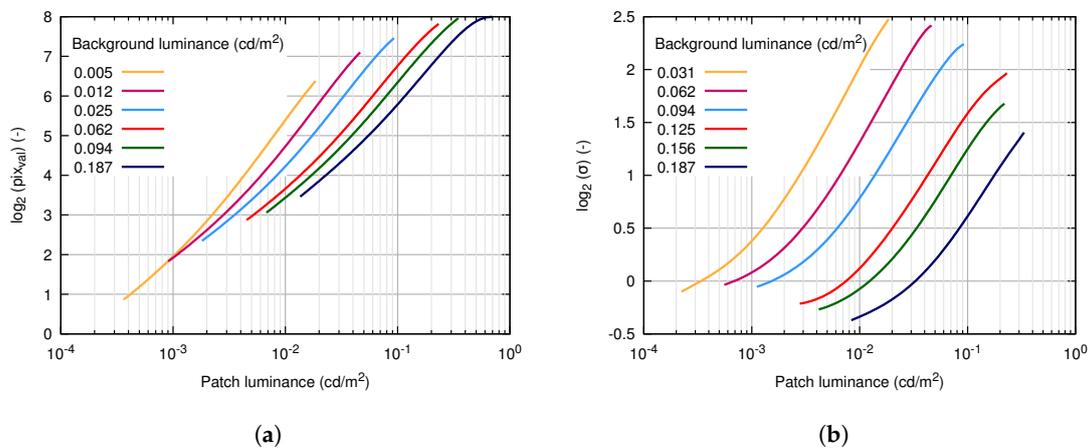


Figure 5. Image intensifier characteristics. (a) Opto-electronic conversion function dependence of the digital level in the output image on patch brightness for various background luminances. (b) Noise standard deviation dependence on patch brightness for various background luminances (gain levels of image intensifier). cd/m²: candela per square meter.

This type of image acquisition system requires new methods of meteor detection. The idea arises from the previous analysis that it is difficult to examine the relationships between the pixels within one frame. We propose an algorithm that takes into account single pixel probability characteristics calculated across a certain number of frames. Figure 6 shows consecutive frames and the sliding window w_i of size N . The value of the mean $\mu_i(x, y)$ and the standard deviation $\sigma_i(x, y)$ of the pixel at spatial position (x, y) in the i -th frame is calculated from the values of the pixel in the window. To detect a meteor, the algorithm searches for the relationship between pixel characteristics valid for the i -th frame and the model calculated from the first M frames of a video sequence.

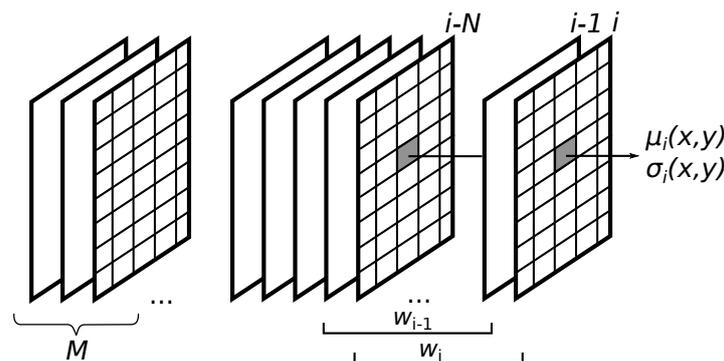


Figure 6. Idea of time-domain single pixel analysis.

The model builds on the relation between the mean value and standard deviation of the pixels in the frame. In Figure 7, the circle marks present this relation in frames without a meteor, and it demonstrates an example of how this relation changes when a sliding window includes frames with a meteor. One can see a deviation in a certain interval of pixel intensity values caused by the temporal appearance of the meteor on the dark background, which increases the standard deviation of pixels with low intensity.

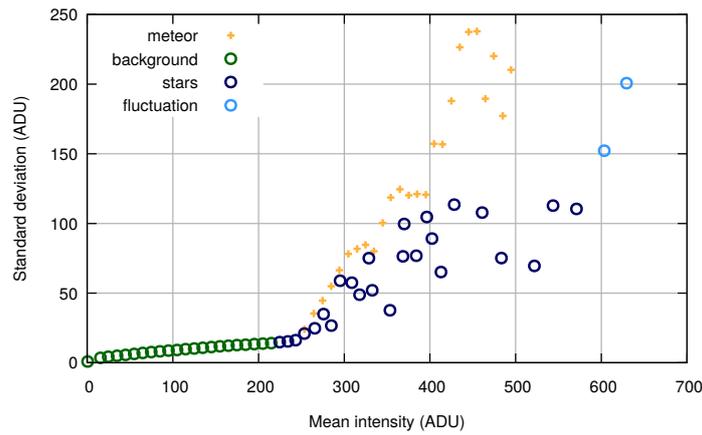


Figure 7. Comparison of statistical characteristics. ADU: analog-to-digital unit.

Both, meteor appearance and the noise can increase the mean and standard deviation of a pixel across a window. Thus, it is not enough to keep these parameters for each pixel. The model has to describe the estimation of the permitted standard deviation depending on the mean intensity. In this case, we propose to construct the model by approximation of the relation between the mean value and the standard deviation in frames without a meteor. We consider video sequences of a duration of 10 min and propose the renewal of the model from the last M frames labeled as frames without a meteor. It compensates variations during the night.

3. Description of the Algorithm

Based on the above-described idea, we propose the frame classification method shown in Figure 8. The statistical analysis block provides the calculation of the mean value $\mu_i(x, y)$ and standard deviation $\sigma_i(x, y)$ of each pixel through N frames. We use the recursive calculation of this characteristic based on known $\mu_{i-1}(x, y)$ and $\sigma_{i-1}(x, y)$. With the model built from the frames with only static objects present, an algorithm can detect the transient (i.e., moving) object. To reduce false detection, we also introduce the post-analysis block exploring how many times the algorithm marked the single pixel and its neighbors as candidate objects.

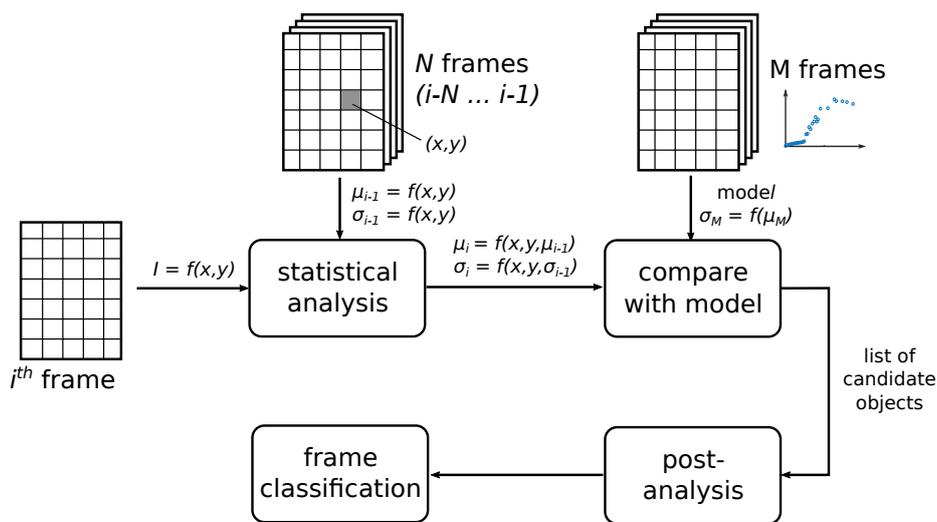


Figure 8. Block diagram of the proposed classifier.

3.1. Statistical Analysis

Widely-used methods of computing the standard deviation require two passes through the data. Since our effort focuses on real-time data processing, more suitable for implementation are single-pass algorithms. Our pipeline uses the robust iterative formula by Welford [37]. Since we have determined the mean μ_0 and variance σ_0^2 of a single pixel for the window included in $(i, i + 1, \dots, i + N - 1)$ frames, we can estimate how parameters μ_1 and σ_1^2 change when we slide the window by one position:

$$\mu_1 - \mu_0 = \frac{\sum_{i=1}^N I_i - \sum_{i=0}^{N-1} I_i}{N} = \frac{I_N - I_0}{N}, \quad (1)$$

where I_i is the pixel intensity in the i -th frame and N is the window size. To evaluate the difference of variances, the unbiased sample variance is used:

$$\sigma^2 = \frac{\sum_{i=1}^N I_i^2}{N-1} - \frac{(\sum_{i=1}^N I_i)^2}{(N-1)N}. \quad (2)$$

Hence, we obtain:

$$\begin{aligned} (N-1)\sigma_1^2 - (N-1)\sigma_0^2 &= \left(\sum_{i=1}^N I_i^2 - N\mu_1^2 \right) - \left(\sum_{i=0}^{N-1} I_i^2 - N\mu_0^2 \right) \\ &= I_N^2 - I_0^2 - N(\mu_1^2 - \mu_0^2) = (I_N - I_0)(I_N + I_0 + \mu_1 + \mu_0). \end{aligned} \quad (3)$$

This means that we can use the iteration formula to calculate the mean value and the variation of a pixel across the window of size N frames. In our algorithm, we use a window size equal to 15 frames, which is enough to follow the changes of the standard deviation and to detect a meteor in a frame.

3.2. Comparison with the Model

To get the list of candidate objects in the i -th frame, we perform the statistical analysis across the moving window of size N . The calculated standard deviation of a single pixel at spatial position (x, y) with a certain mean value is compared with a corresponding value of the a priori model $\sigma_M = f(\mu_M)$. If the standard deviation of the pixel is significantly higher than the model standard deviation, we label this pixel as a candidate object.

The model represents the relationship between the mean of pixel values μ_M and the standard deviation σ_M , and it is constructed from data samples of M frames including static objects only (typically the first 15 frames in a video sequence). To get these data samples, we calculate the mean values and standard deviation of each pixel across M frames. In this case, the number of samples associated with the background is significantly bigger. To get an equal number of samples in different intervals of dynamic range, we average these parameters in single intervals.

The precision of the model is a crucial factor affecting algorithm performance. We found that the model is well described by the formula:

$$\sigma_M = \begin{cases} a_1 \cdot e^{b_1 \mu_M}, & \mu_M \leq \hat{B}; \\ a_2 \cdot e^{b_2 \mu_M}, & \mu_M > \hat{B}, \end{cases} \quad (4)$$

where \hat{B} is an estimation of the background and $a_{1,2}$ and $b_{1,2}$ are parameters approximating data samples from M frames. Accurate background estimation significantly reduces the number of detection errors. There are different methods, for example sigma-clipping [38], multiresolution support [39], modified distance transform [40], etc. The trade-off between efficiency, simplicity and speed leads to the use of the convolution with the averaging filter [41] of size 11×11 pixels for this particular task.

An example of how the model fits the data samples can be seen in Figure 9a. Figure 9b shows the dependency of the estimated background value on the size of the filter.

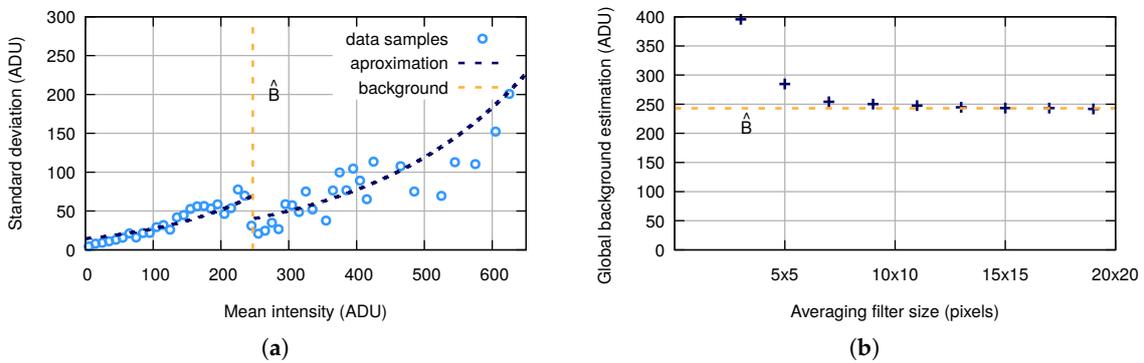


Figure 9. (a) The model of the statistical characteristic dependency. (b) Size of the averaging filter for the global background estimation.

3.3. Post-Analysis

The list of candidate objects includes both true and false positive detections. False candidate objects are typically one-pixel or a small connected area; most of them can be removed efficiently by the use of the morphological transformation [42]. In the proposed algorithm, we apply dilation followed by erosion (Figures 10 and 11). Dilation with 2×2 structuring element allows connecting candidates that are close to each other and ensuring saving a meteorite trajectory following erosion (Figure 10). Erosion with 3×3 structuring element removes all candidates that have no eight-connectivity, which is an effective way to get out of negative candidates (Figure 11). Using bigger structuring elements for morphological transformation can cause the removal of the meteor trajectory.

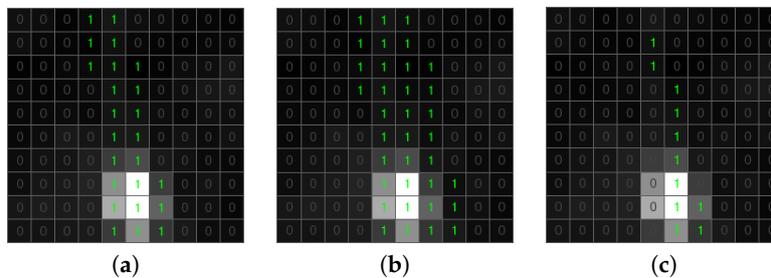


Figure 10. Morphological transformation of a positive candidate object. (a) Candidate object. (b) Dilation. (c) Erosion.

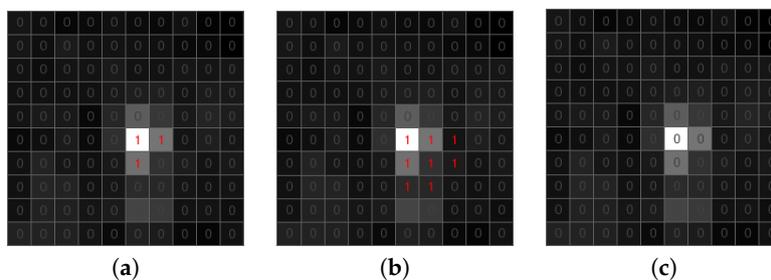


Figure 11. Morphological transformation of a negative candidate object. (a) Candidate object. (b) Dilation. (c) Erosion.

Classification of residual candidate objects requires further analysis. Our algorithm uses a counter calculating how many times the single pixel was marked as a candidate object in previous frames. We analyzed the results of this calculation for a meteor and a static object. In Figure 12, the green path is a meteor trajectory. The counter associated with the positive candidate objects tends to decrease its value in the direction of an object moving smoothly. As we can see, the counters of single pixels have no big difference from their non-zero neighbors. This allows excluding candidates having significant differences in counters associated with the negative candidate (Figure 13). Based on this assumption, we define the difference between a pixel's counter and its non-zero neighbors. If the biggest difference is lower than four, we mark a pixel as an object.

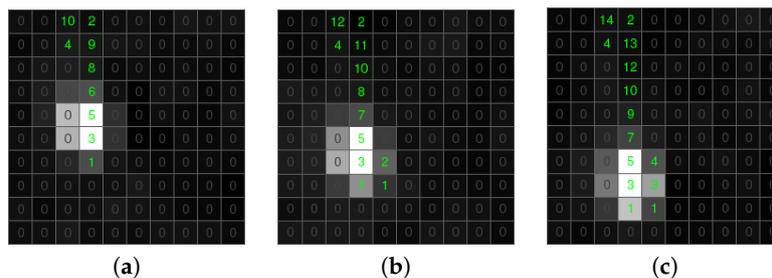


Figure 12. The counters of a positive candidate associated with pixels labeled as a candidate object. (a) Time T. (b) Time T + 30 ms. (c) Time T + 60 ms.

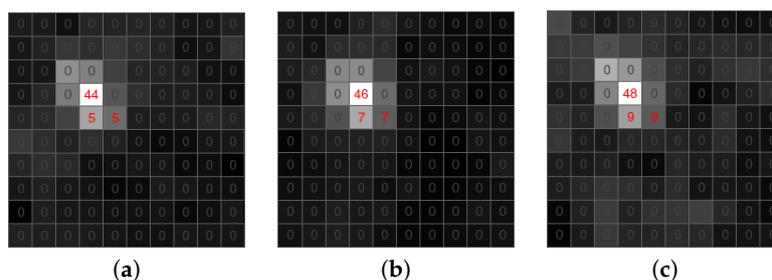


Figure 13. The counters of a negative candidate associated with pixels labeled as a candidate object. (a) Time T. (b) Time T + 30 ms. (c) Time T + 60 ms.

The result of post-analysis is a list of detected meteors in the frame, which is the basis for frame classification. If the list is not empty, we mark the frame as including a meteor.

4. GPU Acceleration

Besides true positive detections of meteors, the second most important parameter of the algorithm is an execution time. The algorithm was designed to be implemented on a GPU using CUDA (Compute Unified Device Architecture), a highly parallel multi-threaded architecture [43]. A block diagram of this implementation is shown in Figure 14. One of the main bottlenecks of GPU acceleration is inefficient data transfer between the host and the device, negatively affecting the overall application performance. Thus, our GPU implementation simultaneously processes several frames, as proposed by Vitek in [44]. We transfer frames $(i, i + 1, \dots, i + N_{gpu} - 1)$ to the GPU global memory where $(i = 1, N_{gpu} + 1, 2 \cdot N_{gpu} + 1, \dots)$, and N_{gpu} is the window size chosen based on the parameters of the GPU. After some experiment, we found that six-frame processes simultaneously represent a good-enough trade-off between accuracy and execution time.

The recursive calculation of the statistic characteristics described in Section 3.1 is used for all transferred frames except the last one. To get $(\mu_i, \mu_{i+1}, \dots, \mu_{i+N_{gpu}-2})$ and $(\sigma_i, \sigma_{i+1}, \dots, \sigma_{i+N_{gpu}-2})$, only one set of referent parameters μ_{i-1} and σ_{i-1} is used. This is the main difference from the CPU implementation, which uses referent parameters for each frame. The statistic characteristics of the last

frame in the window have to be calculated based on all frames in the window without using recursion because it defines the accuracy of detection in the next frames.

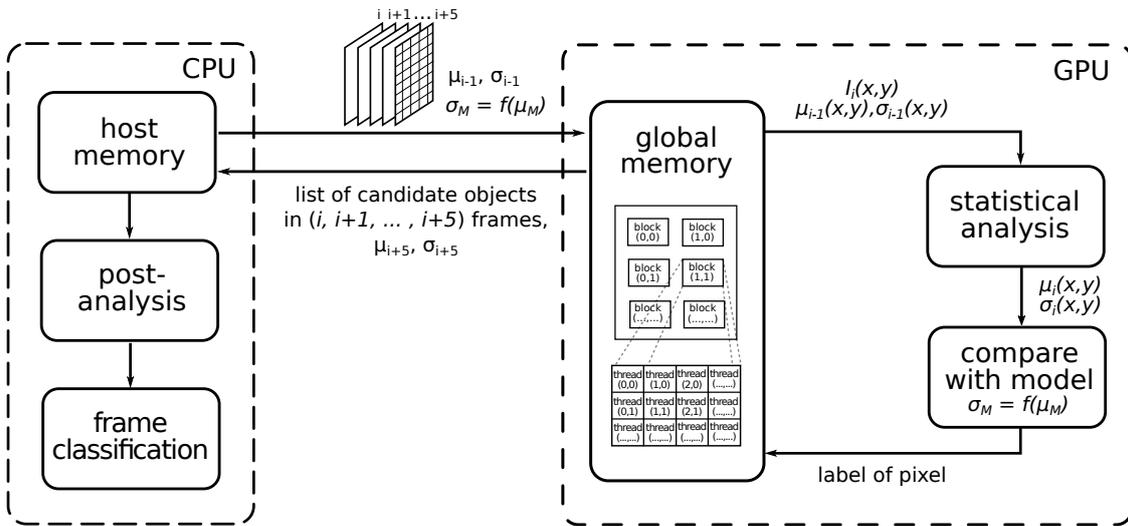


Figure 14. Block diagram of the GPU-accelerated pipeline.

5. Verification of Algorithm Performance

During the test, we focused on the two main features of the algorithms: (a) the ability to detect meteors in the single frame and (b) the ability to detect an event as such. The frame classifier has four possible outcomes: true positive, shown in Figure 15 (TP, the case when the meteor is present in the frame and it is correctly detected by the algorithm), false positive (FP, the case when the meteor is not present in the frame, but it is falsely detected by the algorithm), true negative (TN, the case when the meteor is not present in the frame and the algorithm is not producing any alert) and false negative, shown in Figure 16 (FN, the case when the meteor is present in the frame and the algorithm is not producing any alert).

Performance of the detection algorithm depends on the geocentric velocity of the meteoroids and the geometry of the meteor appearance. When a meteoroid enters the top layers of the Earth’s atmosphere, its movement is not followed by any significant change in brightness. Thus, it is hard to distinguish the beginning of the event and fluctuation caused by speckle noise, and frames capturing the beginning of the meteor trail are the main source of the FN classification. Another problem for the processing algorithm is faint meteors, for example meteoroids entering the atmosphere at a small angle, so the overall duration of the event is short, and changes in the brightness are weak. It is therefore difficult to track the brightness changes of the neighboring pixels, and the false detection rate is higher for frames capturing shorter events.

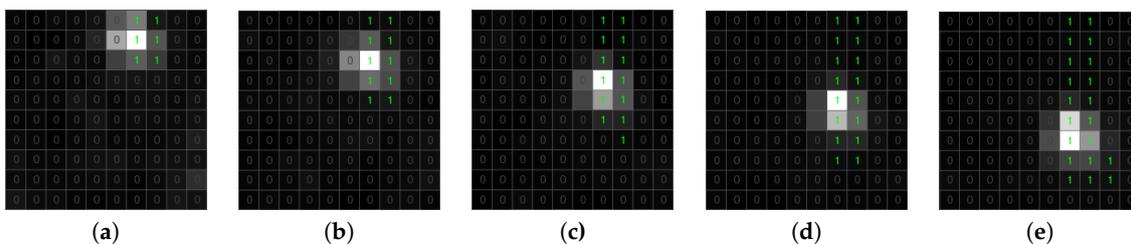


Figure 15. True positive example. (a) Time T. (b) Time T + 15 ms. (c) Time T + 30 ms. (d) Time T + 45 ms. (e) Time T + 60 ms.

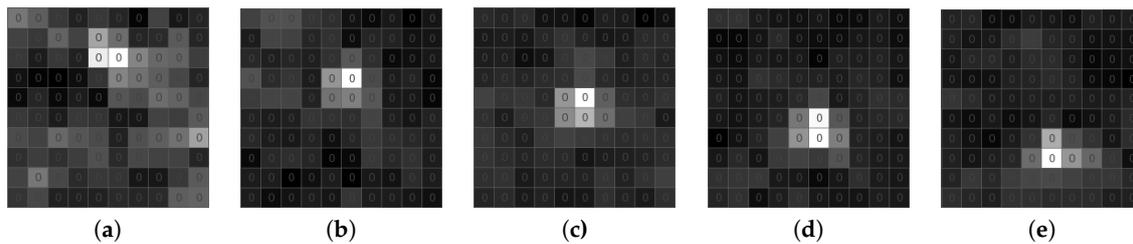


Figure 16. False negative example. (a) Time T. (b) Time T + 15 ms. (c) Time T + 30 ms. (d) Time T + 45 ms. (e) Time T + 60 ms.

5.1. Experimental Setup

For the purpose of this paper, testing data were acquired with the system MAIA (Meteor Automatic Imager and Analyzer) [45]. This system uses image intensifier XX1332 and GigE (Gigabit Ethernet) progressive scanning camera JAI CM-040GE, running at a frame rate of 61 fps and a bit depth of 10 bits. The spatial resolution of the camera is 776×582 pixels (approximately 6 arcmin/pixel), corresponding to a field-of-view of 52° . The limiting stellar magnitude is +8. The astrometric precision of the system is quite good: the standard deviation is better than 0.04° both for naked and intensified systems. MAIA works in double-station configuration, and camera systems are deployed in two places: Ondřejov and Kunžak, the distance between both stations being 92.5 km.

To evaluate the performance of the proposed algorithm, we processed 30 video sequences with a total number of 2419 frames, acquired during different nights by the use of the MAIA system. All video sequences contain a meteor, and we manually labeled all 1169 frames on which meteors are recorded. Frames at the beginning of each video sequence contain only static objects, so it is possible to build the model. We compared our algorithm with three other methods: the first one is an algorithm that is currently in use within the MAIA project; the second one is the widely-used UFOCapture; and the third one is our reimplementation of a meteor detector used within CMN [46]. Originally, the Python-based software RPi Meteor Station (RMS [47]) was running on the Raspberry Pi platform.

The algorithm currently in use within the MAIA project takes into account the high temporal resolution of video sequences. It creates a list of static objects and detects new objects in the next frames. Each new object is placed in the list of temporary objects as an object for the next investigation. To find a meteor, the trajectories of these temporary objects are followed. The algorithm is implemented in the pipeline known as dMAIA. The goal of the pipeline is obtaining the sequential measurement of the meteor and its apparent coordinates in comparison with real stars in the background. Detected meteors are the subject of further measurements, particularly the measurement of brightness, the measurement of range of height (especially the beginning heights) and the determination of the atmospheric trajectory. Details about the measurement may be found for example in [48,49].

The most common methods of meteor detection discover meteor tracks in video sequences with low temporal resolution. In this case, the meteor track presents a line in each frame. The RMS algorithm is based on this frame feature. Its basic concept is line detection by kernel-based Hough transform in a reconstruction image from 64 frames. In our implementation of this algorithm, we reconstructed images from 15 frames, which was enough to detect a meteor.

5.2. Results

To compare the execution time of a tested algorithm, we used a personal computer with Intel Core i5-3210M 2.5 GHz x4, 16 GB of DDR3/1600 MHz memory and NVIDIA GeForce GT 635M 2 GB GDDR5 graphics card. To include UFOCapture in the test, we developed a custom virtual DirectShow camera. As we can see in Table 1, the implementation of the proposed algorithm significantly reduces the time needed to process one frame of the video sequence. Note that the time needed to build a model is 1.19 s, so while we are updating the model once per 36,600 frames (i.e., ten minutes of recording), there is an overhead of 0.03 ms per processed frame.

Table 1. Time for processing a single frame.

	CPU (ms)	GPU (ms)
UFOCapture	62.6	-
RMS	47.1	-
dMAIA	27.2	16.1
Proposed	12.2	10.3

Table 2 summarizes the results of particular algorithms. Following our hypothesis of the more difficult detection of shorter events, we performed a test on the subset of video sequences containing events longer than 25 frames. The results of those tests are summarized in Table 2b, and one can see a significantly lower number of FN detections for events longer than 25 frames.

Furthermore, we evaluated the ability of the algorithms to find a meteor event (i.e., a streak of light in consecutive frames) in the video sequence. Our algorithm was able to detect all meteors in the video sequences, and the currently used algorithm missed two meteors, while the algorithm based on RMS missed three meteors. UFOCapture missed only one meteor, but also produced a high number of false positives.

Table 2. Confusion matrix based on frame classification. (a) Analysis of all available video sequences. (b) Analysis of the video sequences capturing events longer than 25 frames.

	(a)				(b)			
	TP	FP	FN	TN	TP	FP	FN	TN
UFOCapture	920 78.69%	320 25.19%	249 21.31%	950 74.81%	425 83.99%	93 25.34%	81 16.01%	274 74.66%
RMS	860 73.57%	14 1.1%	309 26.43%	1256 98.9%	391 77.27%	0 0%	115 22.73%	367 100%
dMAIA	902 77.15%	22 1.73%	267 22.84%	1248 98.26%	420 83%	10 2.72%	86 17%	357 97.18%
proposed	994 85.03%	12 0.94%	175 14.97%	1258 99.06%	462 91.03%	5 1.36%	44 8.07%	362 98.64%

It is worth noting that we also had the possibility to investigate the usability of the tracking algorithm incorporated in a University of Western Ontario processing pipeline for high temporal resolution of video sequences. This algorithm is an evolution of the Astrobiology Instrumentation for Meteor Imaging and Tracking system [50]. It has an advantage in time processing compared with the proposed algorithm. The time of a single frame processing is 6.3 ms. However, this algorithm requires the accurate setting of input parameters for each video sequence, which has a significant effect on the precision of meteor detection. The algorithm proposed in this paper removes this disadvantage.

6. Conclusions

This paper focuses on methods of meteor detection in video sequences with high frame rates. We proposed the algorithm of frame classification based on the comparison between temporal statistical characteristics of a pixel and the model built on the relation between the mean and the standard deviation of the pixel.

The results showed high performance in accuracy and speed. The precision of the proposed algorithm is 0.9881, and the recall is 0.8503. The proposed algorithm is significantly faster compared to state-of-the-art algorithms. The implementation using computing on a GPU reduced the processing time of a single frame and had a duration of 10.3 ms per frame, which means that it is possible to process single frame in a time shorter than the exposure time (16.4 ms for a frame rate of 61 fps). The parameters of this implementation need further investigation to obtain a trade-off between accuracy and speed.

Acknowledgments: This work was supported by Grant No. 17-05840S Multicriteria Optimization of Shift-Variant Imaging System Models of the Czech Science Foundation and by the Grant Agency of the Czech Technical University in Prague, Grant No. SGS17/099/OHK3/1T/13 Real time classification of optical transients. We also would like to thank Peter Gural, who kindly provided us his tracking/clustering processing pipeline.

Author Contributions: S.V. conceived the idea and designed experiments. M.N. performed the experiments and collected data. Both authors analyzed the data and contributed to writing the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hey, J.S.; Stewart, G. Radar observations of meteors. *Proc. Phys. Soc.* **1947**, *59*, 858.
- Popowicz, A.; Malcher, A.; Bernacki, K.; Fietkiewicz, K. A Passive FPAA-Based RF Scatter Meteor Detector. *Publ. Astron. Soc. Pac.* **2015**, *127*, 152.
- Spurný, P.; Borovicka, J. The autonomous all-sky photographic camera for meteor observation. In Proceedings of the Asteroids, Comets, and Meteors (ACM 2002), Berlin, Germany, 29 July–2 August 2002; Volume 500, pp. 257–259.
- Castro-Tirado, A.J.; Jelínek, M.; Vítek, S.; Kubánek, P.; Trigo-Rodríguez, J.M.; de Ugarte Postigo, A.; Sanguino, T.M.; Gomboš, I. A very sensitive all-sky CCD camera for continuous recording of the night sky. *Adv. Softw. Control Astronomy II Proc. SPIE* **2008**, *7019*, 70191V.
- Trigo-Rodríguez, J.M.; Lyytinen, E.; Jones, D.C.; Madiedo, J.M.; Castro-Tirado, A.J.; Williams, I.; Llorca, J.; Vítek, S.; Jelínek, M.; Troughton, B.; et al. Asteroid 2002NY40 as a source of meteorite-dropping bolides. *Mon. Not. R. Astron. Soc.* **2007**, *382*, 1933–1939.
- Trigo-Rodríguez, J.M.; Madiedo, J.M.; Williams, I.P.; Castro-Tirado, A.J. The outburst of the κ Cygnids in 2007: Clues about the catastrophic break up of a comet to produce an Earth-crossing meteoroid stream. *Mon. Not. R. Astron. Soc.* **2009**, *392*, 367–375.
- Porubčan, V.; Williams, I.; Kornoš, L. Associations between asteroids and meteoroid streams. *Earth Moon Planets* **2004**, *95*, 697–712.
- Ocana, F.; Zamorano, J.; de Miguel, A.S.; Izquierdo, J.; Gallego, J.; Palos, M.; Muñoz-Ibáñez, B.; Rodríguez-Coira, G.; García, R.; Santamaría, A.; et al. First scientific results of the Fireball Detection Station at UCM Observatory. In Proceedings of the International Meteor Conference (31st IMC), La Palma, Spain, 20–23 September 2012; Volume 1, pp. 178–181.
- Gural, P. The American Meteor Society’s filter bank spectroscopy project. In Proceedings of the International Meteor Conference, Mistelbach, Austria, 27–30 August 2015; pp. 16–23.
- Weryk, R.J.; Brown, P.G. Simultaneous radar and video meteors—I: Metric comparisons. *Planet. Space Sci.* **2012**, *62*, 132–152.
- Christou, A.; Margonis, A.; Oberst, J. A time-domain deconvolution technique for meteor photometry applied to double-station CCD image data. *Astron. Astrophys.* **2015**, *581*, A19.
- Madiedo, J.M.; Trigo-Rodríguez, J.M. Multi-station video orbits of minor meteor showers. *Earth Moon Planets* **2008**, *102*, 133–139.
- Jenniskens, P.; Gural, P.; Dynneson, L.; Grigsby, B.; Newman, K.; Borden, M.; Koop, M.; Holman, D. CAMS: Cameras for Allsky Meteor Surveillance to establish minor meteor showers. *Icarus* **2011**, *216*, 40–61.
- Roggemans, P.; Johannink, C.; Breukers, M. Status of the CAMS-BeNeLux network. In Proceedings of the International Meteor Conference, Egmond, The Netherlands, 2–5 June 2016; pp. 254–260.
- Vida, D.; Šegon, D.; Gural, P.S.; Martinović, G.; Skokić, I. CMN_ADAPT and CMN_binViewer software. In Proceedings of the International Meteor Conference, Giron, France, 18–21 September 2014; pp. 18–21.
- Wiśniewski, M.; Żołądek, P.; Olech, A.; Tyminski, Z.; Maciejewski, M.; Fietkiewicz, K.; Rudawska, R.; Gozdalski, M.; Gawroński, M.; Suchodolski, T.; et al. Current status of Polish Fireball Network. *Planet. Space Sci.* **2017**, *143*, 12–20.
- Campbell-Brown, M. Results from the Canadian Automated Meteor Observatory. In Proceedings of the International Meteor Conference, Egmond, The Netherlands, 2–5 June 2016; pp. 50–52.
- Howie, R.M.; Paxman, J.; Bland, P.A.; Towner, M.C.; Cupak, M.; Sansom, E.K.; Devillepoix, H.A. How to build a continental scale fireball camera network. *Exp. Astron.* **2017**, *43*, 237–266.

19. Colas, F.; Zanda, B.; Vaubaillon, J.; Bouley, S.; Marmo, C.; Audureau, Y.; Kwon, M.K.; Rault, J.L.; Caminade, S.; Vernazza, P.; et al. French fireball network FRIPON. In Proceedings of the International Meteor Conference, Mistelbach, Austria, 27–30 August 2015; pp. 37–40.
20. Gardiol, D.; Cellino, A.; Di Martino, M. PRISMA, Italian network for meteors and atmospheric studies. In Proceedings of the International Meteor Conference, Egmond, The Netherlands, 2–5 June 2016; pp. 76–79.
21. Hawkes, R.; Jones, J. Television Observations of Faint Meteors—I: Mass Distribution and Diurnal Rate Variation. *Mon. Not. R. Astron. Soc.* **1975**, *170*, 363–377.
22. Hawkes, R.; Jones, J. An inexpensive meteor observing system. *Observatory* **1973**, *93*, 233–235.
23. Hawkes, R. Video Techniques for Observation of the Leonid Storms. *Highlights Astron.* **1998**, *11*, 1017–1019.
24. Jordan, P.; Jerram, P.; Fryer, M.; Stefanov, K. e2v CMOS and CCD sensors and systems for astronomy. *J. Instrum.* **2017**, *12*, C07008.
25. Trigo-Rodriguez, J.M.; Madiedo, J.M.; Gural, P.S.; Castro-Tirado, A.J.; Llorca, J.; Fabregat, J.; Vítek, S.; Pujols, P. Determination of meteoroid orbits and spatial fluxes by using high-resolution all-sky CCD cameras. *Earth Moon Planets* **2008**, *102*, 231–240.
26. Gural, P.S. Algorithms and software for meteor detection. *Earth Moon Planets* **2008**, *102*, 269–275.
27. Molau, S. The meteor detection software MetRec. In Proceedings of the International Meteor Conference (17th IMC), Stara Lesna, Slovakia, 20–23 August 1998; pp. 9–16.
28. Gural, P.S. An operational autonomous meteor detector: Development issues and early results. *WGN J. Int. Meteor Organ.* **1997**, *25*, 136–140.
29. Gural, P.S.; Larsen, J.A.; Gleason, A.E. Matched filter processing for asteroid detection. *Astron. J.* **2005**, *130*, 1951.
30. Kozak, P. “Falling Star”: Software for Processing of Double-Station TV Meteor Observations. *Earth Moon Planets* **2008**, *102*, 277–283.
31. Brown, P.; Weryk, R.; Kohut, S.; Edwards, W.; Krzeminski, Z. Development of an all-sky video meteor network in Southern Ontario, Canada The ASgard System. *WGN J. Int. Meteor Organ.* **2010**, *38*, 25–30.
32. Gural, P.; Segon, D. A new meteor detection processing approach for observations collected by the Croatian Meteor Network (CMN). *WGN J. Int. Meteor Organ.* **2009**, *37*, 28–32.
33. Silađi, E.; Vida, D.; Nyarko, E.K. Video meteor detection filtering using soft computing methods. In Proceedings of the International Meteor Conference, Mistelbach, Austria, 27–30 August 2015; pp. 27–30.
34. Molau, S.; Gural, P. A review of video meteor detection and analysis software. *WGN J. Int. Meteor Organ.* **2005**, *33*, 15–20.
35. *Photography—Electronic Still-Picture Cameras—Methods for Measuring Opto-Electronic Conversion Functions (OECFs)*; International Organization for Standardization: Geneva, Switzerland, 1999.
36. Fliegel, K.; Švihlík, J.; Páta, P.; Vítek, S.; Koteň, P. Meteor Automatic Imager and Analyzer: Current Status and Preprocessing of Image Data. In *SPIE Optical Engineering + Applications*; International Society for Optics and Photonics: Bellingham, WA, USA, 2011; p. 81351S.
37. Welford. Note on a Method for Calculating Corrected Sums of Squares and Products. *Technometrics* **1962**, *4*, 419–420.
38. Starck, J.L.; Murtagh, F. *Astronomical Image and Data Analysis*; Springer: Secaucus, NJ, USA, 2006.
39. Starck, J.L.; Murtagh, F. Automatic noise estimation from the multiresolution support. *Publ. Astron. Soc. Pac.* **1998**, *110*, 193.
40. Popowicz, A.; Smolka, B. A method of complex background estimation in astronomical images. *Mon. Not. R. Astron. Soc.* **2015**, *452*, 809–823.
41. Masias, M.; Freixenet, J.; Lladó, X.; Peracaula, M. A review of source detection approaches in astronomical images. *Mon. Not. R. Astron. Soc.* **2012**, *422*, 1674–1689.
42. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis, and Machine Vision*; An International Thomson Publishing Company: London, UK, 1996.
43. Cook, S. *CUDA Programming: A Developer’s Guide to Parallel Computing with GPUs*; Elsevier: Newnes, Australia, 2012.
44. Vítek, S.; Švihlík, J.; Krasula, L.; Fliegel, K.; Páta, P. GPU accelerated processing of astronomical high frame rate videosequences. *Appl. Digit. Image Process.* XXXVIII **2015**, 9599, 95992L.
45. Vítek, S.; Páta, P.; Koteň, P.; Fliegel, K. Long-Term Continuous Double Station Observation of Faint Meteor Showers. *Sensors* **2016**, *16*, 1493.

46. Vida, D.; Zubović, D.; Šegon, D.; Gural, P.; Cupec, R. Open-source meteor detection software for low-cost single-board computers. In Proceedings of the International Meteor Conference (IMC2016), Egmond, The Netherlands, 2–5 June 2016.
47. RPi Meteor Station. Available online: <https://github.com/CroatianMeteorNetwork/RMS> (accessed on 28 December 2017).
48. Koten, P.; Spurný, P.; Borovička, J.; Evans, S.; Elliott, A.; Betlem, H.; Štork, R.; Jobse, K. The beginning heights and light curves of high-altitude meteors. *Meteorit. Planet. Sci.* **2006**, *41*, 1305–1320.
49. Koten, P.; Stork, R.; Páta, P.; Fliegel, K.; Vitek, S. Simultaneous analogue and digital observations and comparison of results. In Proceedings of the International Meteor Conference, Egmond, The Netherlands, 2–5 June 2016; pp. 133–136.
50. Gural, P.S.; Jenniskens, P.M.; Varros, G. Results from the AIM-IT meteor tracking system. In *Modern Meteor Science an Interdisciplinary View*; Springer: Dordrecht, The Netherlands, 2005; pp. 541–552.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

C Appendix C

VITEK, Stanislav, et al. Influence of Camera Setting on Vehicle-to-Vehicle VLC Employing Undersampled Phase Shift On-Off Keying. *Radioengineering*, 2017, 26.4: 947.

Author's contribution: image processing, data analysis.

Influence of Camera Setting on Vehicle-to-Vehicle VLC Employing Undersampled Phase Shift On-Off Keying

Stanislav VITEK¹, Jiri LIBICH¹, Pengfei LUO², Stanislav ZVANOVEC¹, Zabih GHASSEMLOOY³,
Navid Bani HASSAN³

¹ Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, Prague, 166 28, Czech Republic

² Research Department of HiSilicon, Huawei Technologies Co., Ltd, Beijing, 100085, China

³ Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom

vitek@fel.cvut.cz, oliver.luo@hisilicon.com, xzvanove@fel.cvut.cz, z.ghassemlooy@northumbria.ac.uk,
navid.hassan@northumbria.ac.uk

Submitted February 16, 2017 / Accepted August 4, 2017

Abstract. *This paper focuses on the performance analysis of a camera-based vehicle-to-vehicle visible light communication system employing undersampled phase shift on-off keying modulation under interference scenario. Two Nissan Qashqai front lights with daylight running light emitting diodes based lamps are used for communications. The bit error rate (BER) performance of the proposed system is experimentally measured for a transmission span up to 24 m focusing mostly on the side interference due to reflections. Based on experimental data we demonstrate reduction of the system performance due to the side reflection and illumination of the detector by other light sources which have to be taken into account during further data processing. We provide with further statistics for particular shutter speed and transmitter power setting and discuss BER improvement especially to meet FEC via the method of the adaptive region of interest.*

Keywords

Visible Light Communication, Vehicle-to-Vehicle communication, undersampled phase shift on-off keying, low bitrate, camera

1. Introduction

Today's dense road traffics in combination with relatively fast cars require additional improvements in the car and road safety. Transportation has become an essential part of everyday life. New safety systems have been developed thanks to newly emerged technologies such as anti-collision and intelligent transport systems (ITS) [1]. The widespread deployment of ITS will achieve a high level of continuous traffic flow, safety, and a lower carbon footprint at a global level. Vehicular networking (VN) is an essential component of ITS, which offers vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. The information

gathered via V2V and V2I provides the drivers with real-time information on the traffic and road conditions including collisions, congestion, surface condition, traffic signal violations, emergency brakes, etc. [2], [3] Furthermore, such information can be used by the local transport departments to develop efficient traffic applications (e.g., speed advisory, platooning) or by automotive manufacturers to develop safety applications (like collision avoidance). To achieve this, ITS have taken advantages of the information and communication technologies (ICT) to assist the vehicle users. The most widely adopted ICT are based on the short-range radio frequency wireless technologies (i.e., ultra-high frequency and very high-frequency) employing IEEE 802.11 protocols. However, longer range communications have been proposed based on worldwide interoperability for microwave access (WiMAX) IEEE 802.16, global system for mobile Communications (3G, 4G) requiring extensive and very costly infrastructure deployment compared to the shorter range system [4], [5].

ITS has also been considering the use of visible light communications (VLC) [6] by way of exploiting the potential of light-emitting diode (LED) based lighting fixtures in both vehicles and roadsides (i.e., street lamps, traffic lights, car lights). This approach will lead in the long run to reduced cost, greener environment and the release of highly need RF frequency spectrum for other applications. VLC (in the spectrum range of 380 to 780 nm) has attracted the attention of many researchers in the last decade with dedicated IEEE 802.15.7r1 standard [7], which defines network architectures, physical and medium access control (MAC) layers. The LED light sources adapted are characterized mostly by their lower power consumption, relatively higher energy efficiency, switching capability and much wider bandwidth compared to the incandescent and fluorescent lights. Hence LEDs are not only suitable for illumination but also for data transmission, and sensing as well as for very precise localization (e.g. in [8] reported precision <1.7 cm for indoor navigation), which cannot be achieved in the RF domain.

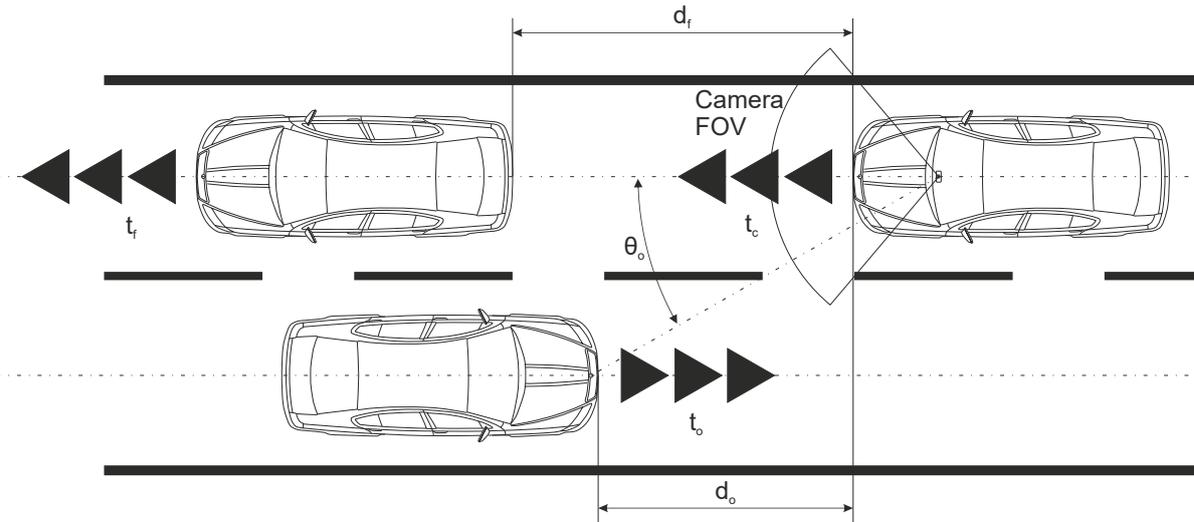


Fig. 1. Typical road traffic situation for a two-lane road.

VLC systems employ LEDs as the transmitting source and a photodiode (PD) or a camera sensor (CS) as the receiving module. Almost all modern vehicles have already LED-based head and brake lights, and indicator lights, thus being possible for the concept of vehicular VLC (VVLC) as a new cost-effective way to implement V2V and V2I communications [9–13]. Furthermore, LED-based roadside units can also be used for both signaling and broadcasting safety-related information to vehicles on the road.

Several test use-cases and experimental results have been published for VVLC networks consisting of onboard units, vehicles, and road side units (i.e., traffic lights, street lamps, digital signage, etc.) [15], [14]. Recent studies reported on V2V communications use either PDs or a CS to detect oncoming vehicles and subsequently control the illumination pattern of the head lights to avoid the glaring. Furthermore, the camera concept can also be used as a receiver (Rx) as part of the VLC based ITS [16], [17], known as the optical camera communications (OCC) [18]. VLC systems based on the LED transmitter (Tx) and a camera based Rx were proposed for automotive applications in [19], where a signal reception experiments were performed under laboratory conditions with fixed positions and then within outdoor environment with changed distance of the receiving camera Rxs with up to a 15 Mb/pixel/s error-free performance. In [20], it was shown that under the driving condition the Rx could detect and accurately track an LED Tx array with an error-free communication over a transmission range of 25 – 80 m. Contrary to a typical VLC communications scenario (with a data source, driver, Tx, PD-based Rx, and processing units), the camera-based VLC system can receive and separate multiple light sources within its field of view (FOV) and extract the information using image sequence processing. This detection technique also offers a unique feature, for example, utilization of the multiple-input and multiple-output (MIMO) capabilities supporting parallel data transmission [21].

Nowadays, many new cars are equipped with a built-in camera for data recording and detection of road condition for safety reason and headlight illumination control. Such a camera can also be used for data communications as well as sensing as part of VLC based ITS. Because of the low frame rate of standard low-cost cameras, OCC only offers a low data rate, typically tens of bits per second, which is sufficient for traffic related information transmission as part of ITS. With respect to such a low data, communication will have a form of periodic broadcast of the short messages, such as hazardous location warning, traffic signal violation warning, etc. This paper is focused on the experimental verification of the camera setting on V2V VLC employing the undersampled phase shift on-off keying modulation scheme, which uses two square wave carriers at the same frequency, with the same amplitude and opposite phases to represent bits 0 and 1.

The paper is organized as follows. At first, the main concept of the V2V system model is introduced in Sec. 2 taking into consideration typical V2V VLC scenario and the main drawbacks that have to be considered in the experimental part. We have focused especially on the side interference reflection. Section 3 outlines setup of the experimental campaign. The results from the OCC VLC system measurement including the bit error rate (BER) performance for different camera parameter settings are given in Sec. 4. Finally, the conclusion is discussed in Sec. 5.

2. System Scenario

Figure 1 shows a typical two line road case where vehicles are communicating with each other using OCC VLC [22]. The cars have two cameras to detect the cars at the front as well as the oncoming vehicles. In detecting vehicles, the beam direction and shape can be adjusted to

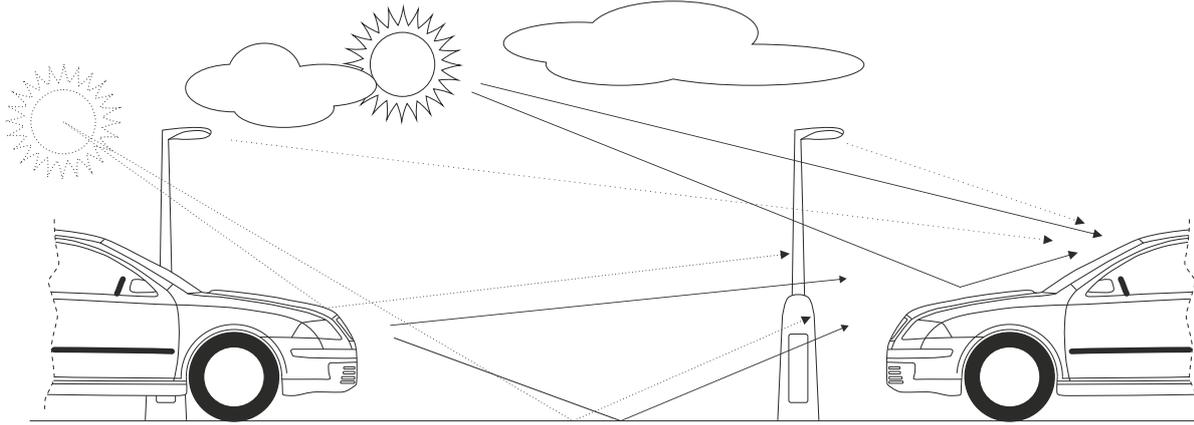


Fig. 2. A C2C VLC link with possible noise sources.

avoid dazzling of other drivers. The VLC communication link is used to transmit traffic related information between the cars.

Figure 2 depicts the both line of sight (LOS) and NLOS (non-LOS) paths including as well interference due to lights from the Sun, other vehicles, street light, etc. which can reach the camera based Rx positioned within the cars. When a light beam illuminates a surface, it results in reflection, transmission, and absorption from the road surface [23]. Contrary to the PD-based Rx, which combines the light rays from LOS and NLOS paths, the camera-based Rx can spatially distinguish each reflected and non-reflected light sources, and process them separately. The reflection from the road surface not only depends on the tarmac road materials but also on the weather conditions. A dry surface road has a diffuse characteristic where an incident light ray is reflected at many angles with some reaching the Rx, whereas under the wet road surface condition the reflection is specular similar to a mirror (including polarization properties) [24].

The received optical power from a single head light can be expressed in the case of the LOS propagation by [25]:

$$P_{\text{LOS}} = \begin{cases} \frac{I_R \cdot A_r \cdot \cos(\Theta_{\text{BLOS}})}{LER \cdot d_{\text{LOS}}^2}, & 0 \leq \Theta_{\text{BLOS}} \leq \Psi, \\ 0, & \Theta_{\text{BLOS}} > \Psi \end{cases} \quad (1)$$

where I_R is luminous intensity of the car light in a particular direction, Θ_{BLOS} denotes the angle between the PD surface normal and the incident Line-of-sight direction, d_{LOS} is the direct distance between the source and detector, A_r and Ψ are active surface area and field of view of the PD, respectively, and LER stands for the luminous efficiency of radiation.

For the NLOS propagation path between the source and detector the received optical power is given as

$$P_{\text{NLOS}} = \begin{cases} \iint dP_{\text{NLOS}} dS, & 0 \leq \Theta_{\text{BNLOS}} \leq \Psi, \\ 0, & \Theta_{\text{BNLOS}} > \Psi \end{cases} \quad (2)$$

$$dP_{\text{NLOS}} = \frac{dI_R \cdot A_r \cdot \cos(\Theta_{\text{BLOS}})}{LER \cdot d_{\text{NLOS}}^2} \quad (3)$$

where d_{NLOS} is the NLOS path length, Θ_{NLOS} denotes the angle between the road surface or surrounding reflecting object and the PD surface normal, and S is the entire area of the surface that has been illuminated by the source.

2.1 Undersampled Phase Shift On-Off Keying

VLC links based on camera Rxs typically use undersampled frequency shift on-off keying (UFSOOK) [26], employing two square wave based carrier signals at frequencies of f_{c0}, f_{c1} to represent bits 0 and 1, respectively. Additionally, there is a square wave signal with higher frequency

$$f_{\text{cFH}} > \frac{1}{\text{shutter speed}}, \quad (4)$$

which represents the intermediate state or also known as the frame header (FH).

In UFSOOK, aliasing is intentionally introduced, and the original data can be decoded from the sampled results. Note that at the Rx sampling is done at the camera frame rate f_c . With a camera based Rx, a number of frames per second (fps) with a fixed or variable exposure time can be captured. Normally, for commercial cameras, $f_c \leq 60$ fps to support data transmission with no flickering. Therefore, the baseband data should be modulated at higher frequencies, and the modulated signal must conform to the flicker-free illumination requirement.

Undersampled phase shift on-off keying (UPSOOK) [27], [28] uses opposite phases to represent bits 0 and 1 and scheme supports flicker-free OCC. Note that the critical flicker-fusion frequency (CFF) of an average human eye f_{eye} is about 100 Hz [29] and CFF of the camera f_{ca} is given by

$$f_{\text{ca}} = \frac{1}{\text{shutter speed}}. \quad (5)$$

In UPSOOK, a square wave subcarrier with a frequency

$$f_s = n \cdot f_c, \quad f_{\text{eye}} < f_s < f_{\text{ca}} \quad (6)$$

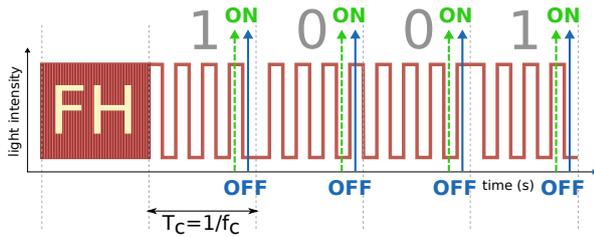


Fig. 3. UPSOOK waveform and two possible undersampled cases for the same transmitted signal.

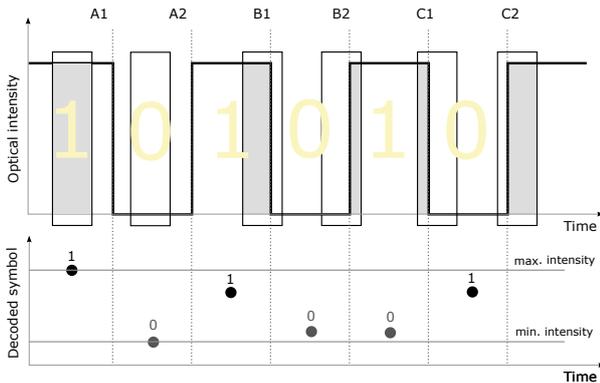


Fig. 4. Illustration of sampling in UPSOOK based OCC VLC.

where n is an integer with a typical value higher than two. Figure 3 shows an example of the UPSOOK modulated signal with a FH. The FH is a 50% duty cycle square wave with a frequency of 10 kHz, which is much higher than f_{ca} . For the modulated bits 1 (mark) and 0 (space) the symbol duration $T_c = 1/f_c$ to ensure that the symbol rate $R_s = 1/T_c$ and f_c are the same. Therefore, the camera can correctly undersample the bit stream at each sampling strobe.

The UPSOOK signal can be expressed as

$$s(t) = \lceil \cos(2\pi f_c t + \theta_n) \rceil, \quad 0 < t \leq T_c \quad (7)$$

where $\lceil \cdot \rceil$ denotes a square wave with the phase θ_n modulated by an input data a_n given by

$$\theta_n = \begin{cases} 0^\circ, & a_n = 1, \\ 180^\circ, & a_n = 0. \end{cases} \quad (8)$$

Since the sampling phase of the camera at the Rx is generally not controlled, there will be a random phase difference between Tx and the camera. Figure 3 illustrates two possible undersampled scenarios obtained by two groups of sampling strobes: green (dashed lines) and blue (solid lines) strobes. Ideally, the sampling points should be at the center of each symbol. Nevertheless, owing to the asynchronous transmission, the sampling point is randomly shifted within the symbol period.

Note, the sampling points of A1 and A2, and B1 and B2 results in normal and shorter sampling intervals, respectively as shown in Fig. 4. The sampling points of C1 and C2 illustrate the extreme case where samplings are right at the transition edges of the pulse, thus may result in wrong decision making between the logic levels of 0 and 1. Therefore,

this effect together with the integration time of the camera sensor will result in reduced signal to noise ratio (SNR). At higher shutter speed the integration time is shorter and therefore the probability that the integration window will be between two symbols is rather low. However, the SNR can be improved by increasing the integration time but at the cost of the integration window spreading over two symbols. In the extreme case, when the center of the integration window is at the edge of a symbol, then the probability of 1 being read as 0 and vice versa increases, thus leading to higher BER.

3. Experimental Measurement

We have carried out experimental verification of the proposed OCC VLC system composed of two LED-based daylight running light (DRL) modules and a camera-based Rx in an indoor environment. The focus of our investigation was mostly on the non-LOS reflected paths to mimic the outdoor condition. A pseudorandom binary sequence of length of 100 symbols was generated and subsequently converted into the UPSOOK format [27] (i.e. total length of the frame is 102 symbols including FH and one symbol representing symbol 1 following FH). A Tektronix AFG3252C arbitrary waveform generator was used to generate the UPSOOK signal for intensity modulation (IM) of two LED headlights for Nissan Qashqai vehicle, see Fig. 5.

Based on the results published in [30], which outlines that the typical duration of V2V VLC link is in the order of seconds, we set the transmission rate to 50 fps. It means that transmission of a single frame takes approximately 2 second. An 18-megapixel Canon EOS 100D camera with a shutter speed ranging from 1/4000 to 30 seconds was used to capture and record the IM LED light sources, see Fig. 6. The captured video stream was stored in the camera as MOV files with a resolution of 720p. The focal length of the camera was set to 18 mm. The transmission link length was set within the range of 2 m to 26 m, with the step of 6 m. Note, at 30 mph the thinking distance, the braking distance and the average distance to stop are around 9, 14 and 23 m, respectively. Two different shutter speeds were used, 1/800 and 1/4000.



Fig. 5. Experimental setup: the Tx.



Fig. 6. Experimental setup: the Rx.

The frequency of the square wave carrier was chosen 200 Hz (i.e. $n = 4$ in (6)). The data stream was packaged into data frames. Each frame began with a header consisted of a symbol formed by a 10 kHz signal (camera sees this signal as a half of the peak to peak optical power) followed by a single bit of logical 1 to correct the potential unsynchronized phase between the Tx and the Rx.

In [27], a system utilizing a mapping of bits via two LED Tx's to increase the data rate was reported. In our experimental work, we have used both DRLs to transmit the same data. During tests, we analyzed OCC VLC result when the front light low beam (LB, in parallel used only for illumination, not carrying VLC) turned on or off in order to influence the VLC link performance. We modulated the light intensity by applying two sets of current, 200 mA and 600 mA.

4. Results

We considered two types of detection with the camera based Rx:

- (i) camera emulates PD Rx – averaging entire camera FOV (see Fig. 7a) into one value. Such a detector does not need any further image processing, but all reflections and noises are integrated in the reception and affect the quality of transmission significantly.
- (ii) The adaptive region of interest (ROI) – adaptively capturing always one Tx area (ROI for a distance of 2 m see Fig. 7b) – when a detector needs a tracking algorithm in case of moving cars and real-time image-processing.

In both cases as mentioned above, the off-line data processing provides the simple sum of pixel values in the selected ROI (or whole frame). The frames are then converted from RGB color space to grayscale. Note that FH were manually identified and also thresholds to distinguish between symbol 0 and symbol 1 were founded separately for particular video sequence.

Figure 8 then illustrated the example of decoded UP-SOOK waveform with cropped captured frames. Although the quality of this link is perfect, one may see the effect of the

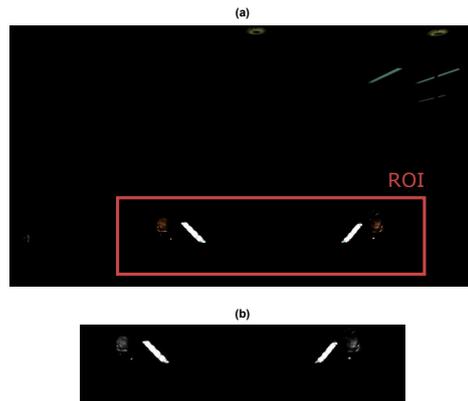


Fig. 7. Example of image data. (a) Whole camera FOV. (a) ROI covers both left and right LED headlights.

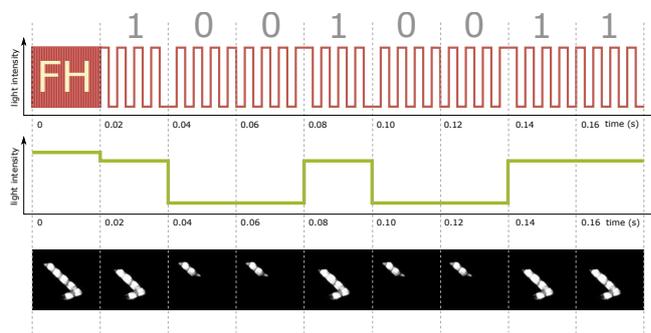


Fig. 8. Example of decoded data. Shutter-speed 1/4000, LB off. LED current 600 mA.

phase shift between the sampling strobe and the waveform, which are labeled as B1 and B2, respectively, in Fig. 4.

From data processing regarding BER statistics, it was observed that most of the errors were due to a phase shift of sampling time and symbol position. Contrary to the emulating PD (whole picture averaging), the link length was not a limiting factor for data transmission in the case of adaptive ROI as can be demonstrated in Fig. 9 and 10. Because of a very low-speed data rate, it is not uncommon to obtain relatively high values of BER in contrast to conventional telecommunication links, especially when a link connection lasts only a few seconds. The primary source of the increased BER due to the reflection of the low beam especially for the case of shutter speed of 1/800 where the BER is above the forward error correction (FEC) 7% BER limit of $3.8 \cdot 10^3$ [31]. This can be clearly seen from Fig. 9a, where values of BER in such case do not meet FEC limit for distances longer than (i) 8 m for the shutter-speed of 1/800 and the LED current of 200 mA with LB is off; (ii) 14 m for the shutter speed of 1/800 and the LED current of 600 mA; (iii) 8 m for the shutter-speed of 1/800 and the LED current of 200 mA; (iv) 8 m for the shutter-speed of 1/4000 and the LED current of 200 mA when LB is off. BER does not meet the FEC limit for the shutter-speed of 1/800 and the LED current of 200 mA when LB is on. This has to be taken into account in further preprocessing of the received data. It is also evident, that quality of the link is significantly worsen when lower LED current

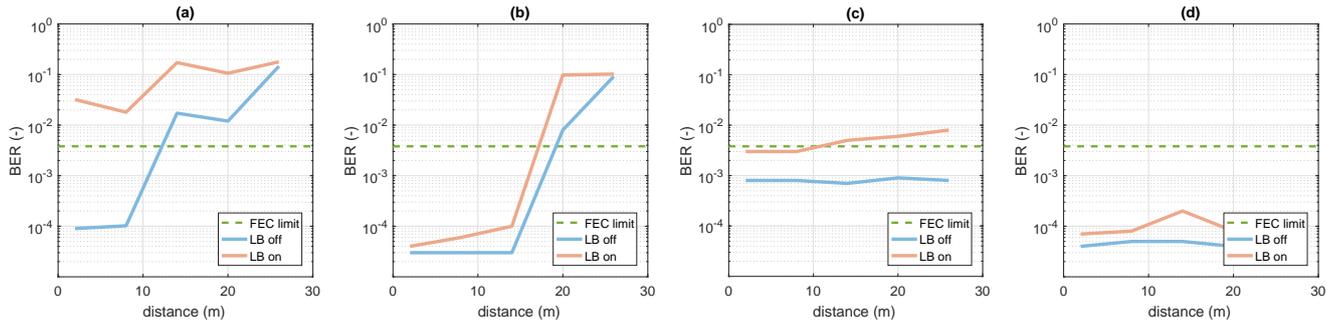


Fig. 9. V2V OCC BER as a function of the distance from Tx for camera Rx emulates PD. a) Shutter-speed 1/800. LED current 200 mA. b) Shutter-speed 1/800. LED current 600 mA. c) Shutter-speed 1/4000. LED current 200 mA. d) Shutter-speed 1/4000. LED current 600 mA.

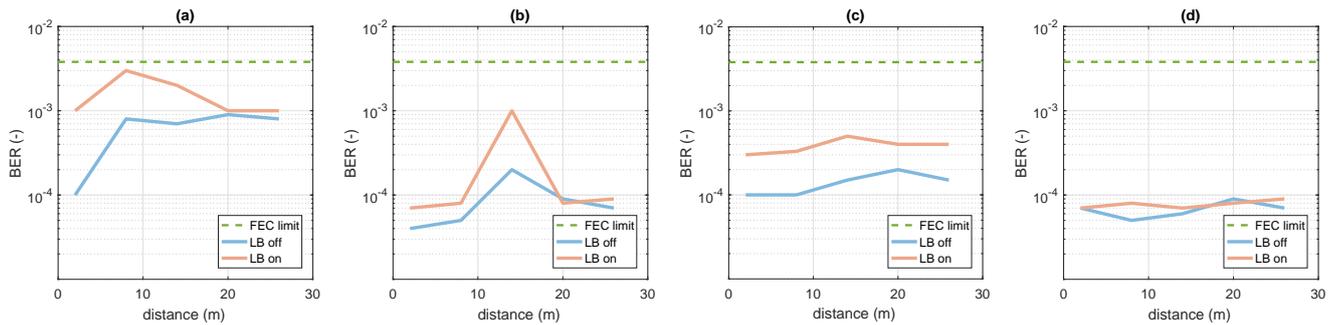


Fig. 10. V2V OCC BER as a function of the distance from Tx for camera Rx uses adaptive ROI. a) Shutter-speed 1/800. LED current 200 mA. b) Shutter-speed 1/800. LED current 600 mA. c) Shutter-speed 1/4000. LED current 200 mA. d) Shutter-speed 1/4000. LED current 600 mA.

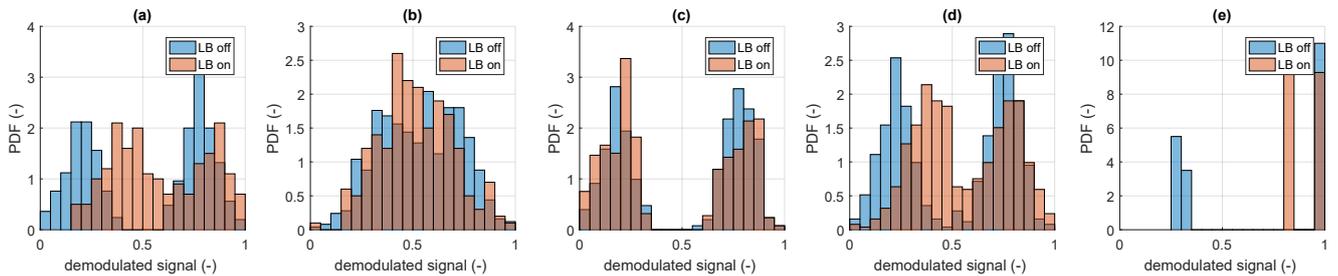


Fig. 11. Examples of received data PDF. (a) Entire FOV. LED current 200 mA. Shutter-speed 1/800. Distance 2 m. (b) Entire FOV. LED current 200 mA. Shutter-speed 1/800. Distance 26 m. (c) Entire FOV. LED current 600 mA. Shutter-speed 1/800. Distance 2 m. (d) Entire FOV. LED current 600 mA. Shutter-speed 1/800. Distance 26 m. (e) Selected ROI. LED current 600 mA. Shutter-speed 1/800. Distance 26 m.

is used. Faster shutter speed reduces BER as can be seen in Fig. 9b. Considerable improvement is achieved when ROI implemented – as shown in Fig. 10 BER has been reduced below the FEC limit for both shutter speeds thus providing seamless communications over the entire transmission span.

The camera based Rx needs to consider several issues that influence link performance: (i) when recording a video stream, the entire picture may not be captured at the same time (i.e., rolling shutter effect) (ii) asynchronous data transmission. The latter is when the camera captures the image exactly at the transition time of the symbol; it is not possible to distinguish the symbols 0 and 1 because of the integrator functionality of the image sensor as described above. For the sake of completeness, Fig. 11 displays examples of the probability density functions (PDF) of the decoded data, which clearly demonstrating how particular effect results in overlapping of detected 0 and 1 symbols. The histograms show PDF calculated from ten consecutive frames, i.e. 1020 symbols.

The values obtained as a sum of all pixels in the grayscale image were normalized relative to the highest value in that sequence. While Figures 11a and 11b illustrates how the distance between the Tx and the Rx affects distinctiveness of symbols, Figures 11c and 11d show the role of the LED current. Figure 11e then shows the impact of the ROI.

5. Conclusion

Even though the car to car communications have been considered as new ways to improve road safety and several papers have been published recently a very limited number of papers have been devoted to the influence of the camera setting on the optical camera communications. Based on the experimental campaign employing common camera as part of the OCC system we tested adaptable low beam and surrounding reflection impact on the system performance. We have realized a VLC link to experimentally evaluate perfor-

mance over a distance up to 24 m under several scenarios. It is evident that proper setting of the shutter speed and detected picture processing can result in BER being above the FEC limit in case of the low beam interference. These aspects have to be taken into account since they could considerably impact transmission of the messages between two moving vehicles when there are only short timeslots available.

Acknowledgments

The research is supported by the European social fund within the framework of the project CZ.1.07/2.3.00/30.0034. The joint research was supported by the EU COST ICT Action IC 1101 and by the Grant Agency of the CTU in Prague, grant no. SGS14/190/OHK3/3T/13.

References

- [1] NASIM, R., KASSLER, A. Distributed architectures for intelligent transport systems: A survey. In *Proceedings of the Second Symposium on Network Cloud Computing and Applications (NCCA)*. 2012, p. 130–136. DOI: 10.1109/NCCA.2012.15
- [2] CAVENEY, D. Cooperative vehicular safety applications. *IEEE Control Systems*, 2010, vol. 30, no. 4, p. 38–53. DOI: 10.1109/MCS.2010.937003
- [3] WANG, Q., FAN, P., LETAIEF, K. B. On the joint V2I and V2V scheduling for cooperative VANETs with network coding. *IEEE Transactions on Vehicular Technology*, 2012, vol. 61, no. 1, p. 62–73. DOI: 10.1109/TVT.2011.2167249
- [4] MORGAN, Y. L. Notes on DSRC & WAVE standards suite: Its architecture, design, and characteristics. *IEEE Communications Surveys & Tutorials*, 2010, vol. 12, no. 4, p. 504–518. DOI: 10.1109/SURV.2010.033010.00024
- [5] LIBICH, J., ZVANOVEC, S., MUDROCH, M. Mitigation of time-spatial influence in free-space optical networks utilizing route diversity. In *Proceedings of the SPIE LASE International Society for Optics and Photonics*. 2012, vol. 8246, p. 1–6. DOI: 10.1117/12.907928
- [6] GHASSEMLOOY, Z., ALVES, L. N., ZVANOVEC, S., et al. *Visible Light Communications Theory and Applications*. CRC Press, 2017. ISBN: 9781498767538
- [7] JUNGnickel, V., et al. A European view on the next generation optical wireless communication standard. In *Proceedings of the IEEE Conference on Standards for Communications and Networking (CSCN)*. 2015, p. 106–111. DOI: 10.1109/CSCN.2015.7390429
- [8] LUO, P., et al. Experimental demonstration of an indoor visible light communication positioning system using dual-tone multi-frequency technique. In *Proceedings of the 3rd International Workshop in Optical Wireless Communications (IWOW)*. 2014, p. 55–59. DOI: 10.1109/IWOW.2014.6950776
- [9] LEE, S. J., et al. Evaluation of visible light communication channel delay profiles for automotive applications. *EURASIP Journal on Wireless Communications and Networking*, 2012, vol. 2012, no. 1. DOI: 10.1186/1687-1499-2012-370
- [10] YU, S. H., et al. Smart automotive lighting for vehicle safety. *IEEE Communications Magazine*, 2013, vol. 51, no. 12, p. 50–59. DOI: 10.1109/MCOM.2013.6685757
- [11] YOU, S. H., et al. Demo: Visible light communications for scooter safety. In *Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services (ACM)*. 2013, p. 509–510. DOI: 10.1145/2462456.2465707
- [12] LIU, C. B., SADEGHI, B., KNIGHTLY, E. W. Enabling vehicular visible light communication (V2LC) networks. In *Proceedings of the Eighth ACM International Workshop on Vehicular Inter-Networking*. 2011, p. 41–50. DOI: 10.1145/2030698.2030705
- [13] CAILEAN, A., et al. Visible light communications: Application to cooperation between vehicles and road infrastructures. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*. 2012, p. 1055–1059. DOI: 10.1109/IVS.2012.6232225
- [14] YAMAZATO, T., et al. Image-sensor-based visible light communication for automotive applications. *IEEE Communications Magazine*, 2014, vol. 52, no. 7, p. 88–97. DOI: 10.1109/MCOM.2014.6852088
- [15] YAMAZATO, T., et al. Vehicle Motion and Pixel Illumination Modeling for Image Sensor Based Visible Light Communication. *IEEE Journal on Selected Areas in Communications*, 2015, vol. 33, no. 9, p. 1793–1805. DOI: 10.1109/JSAC.2015.2432511
- [16] BELLE, A., et al. Development of IEEE 802.15.7 based ITS services using low cost embedded systems. In *Proceedings of the 13th International Conference on ITS Telecommunications (ITST)*. 2013, p. 419–425. DOI: 10.1109/ITST.2013.6685583
- [17] SAHA, N., et al. Survey on optical camera communications: Challenges and opportunities. *IET Optoelectronics*, 2015, vol. 9, no. 5, p. 172–183. DOI: 10.1049/iet-opt.2014.0151
- [18] LUO, P., et al. Experimental demonstration of a 1024-QAM optical camera communication system. *IEEE Photonics Technology Letters*, 2016, vol. 28, no. 2, p. 139–142. DOI: 10.1109/LPT.2015.2487544
- [19] TAKAI, I., et al. LED and CMOS image sensor based optical wireless communication system for automotive applications. *IEEE Photonics Journal*, 2013, vol. 5, no. 5, p. 6801418–6801418. DOI: 10.1109/JPHOT.2013.2277881
- [20] NAGURA, T., et al. Tracking an LED array transmitter for visible light communications in the driving situation. In *Proceedings of the 7th International Symposium on Wireless Communication Systems (ISWCS)*. 2010, p. 765–769. DOI: 10.1109/ISWCS.2010.5624361
- [21] NAGURA, T., et al. Improved decoding methods of visible light communication system for ITS using LED array and high-speed camera. In *Proceedings of the 71st IEEE Vehicular Technology Conference (VTC 2010-Spring)*. 2010, p. 1–5. DOI: 10.1109/VETECS.2010.5493958
- [22] KINOSHITA, M., et al. Motion modeling of mobile transmitter for image sensor based I2V-VLC, V2I-VLC, and V2V-VLC. In *Proceedings of the IEEE Globecom Workshops (GC Wkshps)*. 2014, p. 450–455. DOI: 10.1109/GLOCOMW.2014.7063473
- [23] KOMINE, T., NAKAGAWA, M. Fundamental analysis for visible-light communication system using LED lights. *IEEE transactions on Consumer Electronics*, 2004, vol. 50, no. 1, p. 100–107. DOI: 10.1109/TCE.2004.1277847
- [24] CAHYADI, W. A., et al. Efficient road surface detection using visible light communication. In *Proceedings of the Seventh International Conference on Ubiquitous and Future Networks*. 2015, p. 61–63. DOI: 10.1109/ICUFN.2015.7182498
- [25] LUO, P., et al. Performance analysis of a car-to-car visible light communication system. *Applied Optics*, 2015, vol. 54, no. 7, p. 1696–1706. DOI: 10.1364/AO.54.001696
- [26] ROBERTS, R. D. Undersampled frequency shift ON-OFF keying (UFSOOK) for camera communications (CamCom). In *Proceedings of the 22nd Wireless and Optical Communication Conference (WOCC)*. 2013, p. 645–648. DOI: 10.1109/WOCC.2013.6676454

- [27] LUO, P., et al. Undersampled phase shift ON-OFF keying for camera communication. In *Proceedings of the Sixth International Conference on Wireless Communications and Signal Processing (WCSP)*. 2014, p. 1–6. DOI: 10.1109/WCSP.2014.6992043
- [28] UYSAL, M., et al. *Optical Wireless Communications: An Emerging Technology*. Springer, 2016. ISBN: 9783319302010
- [29] WILKINS, A., VEITCH, J., LEHMAN, B. LED lighting flicker and potential health concerns: IEEE standard PAR1789 update. In *Proceedings of the IEEE Energy Conversion Congress and Exposition (ECCE)*. 2010, p. 171–178. DOI: 10.1109/ECCE.2010.5618050
- [30] WU, L. C., TSAI, H. M. Modeling vehicle-to-vehicle visible light communication link duration with empirical data. In *Proceedings of the IEEE Globecom Workshops (GC Wkshps)*. 2013, p. 1103–1109. DOI: 10.1109/GLOCOMW.2013.6825140
- [31] PROAKIS, J. G. *Digital Communications*. 4th ed. McGraw-Hill, 2000. ISBN: 978-0072321111

About the Authors ...

Stanislav VÍTEK graduated at the Czech Technical University in Prague in 2002 and received Ph.D. degree from the Czech Technical University in Prague in 2008. Recently he is an assistant professor with Dept. of Radioelectronics at the Faculty of Electrical Engineering at the Czech Technical University in Prague. Member of numerous international teams focused to networks of robotic telescopes. His main research interests are signal and image processing, embedded systems, machine learning and database systems.

Jiří LIBICH received his M.Sc. from the Czech Technical University in Prague in 2009 and Ph.D. degree in Radioelectronics in 2014. Since then, he has worked as research assistant in the CTU in areas of free space optics and visible light communications.

Pengfei LUO received his Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT), China

in 2013. He was a research fellow of Department of Physics and Electrical Engineering, Northumbria University, UK (12/2013–10/2014), and a project assistant in BUPT (11/2014–03/2016). Currently, he is a research engineer at Research Department of HiSilicon, Huawei Technologies Co., Ltd.. His research interests include optical camera communication, indoor positioning technology.

Stanislav ZVÁNOVEC a full professor and deputy-head of the Department of Electromagnetic Field, Faculty of Electrical Engineering, CTU. He leads Free-space and fiber optics team at CTU, was vice-chair of WG1 of EU COST project IC1101 OPTICWISE, holder of several projects. Published more than 170 papers, several books and book chapters, he is editor of book Visible Light Communications: Theory and Applications. His main research interests are wireless optical communications, visible light communications and optical fiber technologies.

Zabih GHASSEMLOOY A Research Fellow, City University, U.K (87). Sheffield Hallam University, UK (88–2004), University of Northumbria Newcastle, UK (2004), Chinese Academy of Science President's International Fellowship (2016). Published 695 papers (259 journals and 6 books). Research interests optical wireless communications, free space optics, and visible light communications. He is the fellow of the IET, a senior member of IEEE and a member of OSA.

Navid Bani HASSAN received the B.Sc. and the M.Sc. degree in electrical engineering from the Ferdowsi University of Mashhad in 2012 and Tarbiat Modares University, Tehran, in 2015, respectively. He is currently pursuing the Ph.D. degree at Northumbria University at Newcastle, Newcastle upon Tyne, UK. His research interests include car-to-car optical communications, free space optics, underwater optical communications, and digital signal processing.

D Appendix D

CHVOJKA, Petr, et al. Analysis of nonline-of-sight visible light communications. *Optical Engineering*, 2017, 56.11: 116116.

Author's contribution: image processing, data analysis.

Analysis of nonline-of-sight visible light communications

Petr Chvojka,^{a,*} Stanislav Vitek,^a Stanislav Zvanovec,^a Zabih Ghassemlooy,^{b,c} and Sujan Rajbhandari^d

^aCzech Technical University in Prague, Faculty of Electrical Engineering, Prague, Czech Republic

^bNorthumbria University, Faculty of Engineering and Environment, Newcastle Upon Tyne, United Kingdom

^cChinese Academy of Sciences, QIEM, Haixi Institutes, Quanzhou, China

^dCoventry University, Centre for Mobility and Transport and School of Computing, Engineering and Mathematics, Coventry, United Kingdom

Abstract. We analyze the channel properties of a nonline-of-sight (NLOS) ceiling-to-device and device-to-device visible light communication systems by considering various receivers' orientation and variable fields of view (FOVs). Analyses based on the recursive indoor channel model show that for a particular transmitter configuration, the pure NLOS path can offer higher 3-dB channel bandwidth (up to 14 MHz) compared with the link with LOS and NLOS components. We also show how the receiver rotation (orientation) influences the probability of receiving signals via the NLOS path compared with the LOS and NLOS paths. Moreover, based on the experimental campaign, we demonstrate that shadowing observed at the receiver due to people's movement results in decreased received power level (up to 1.8 dB), thus resulting in reshaping of the probability density function of received power. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.OE.56.11.116116]

Keywords: channel modeling; nonline-of-sight link; shadowing; visible light communications.

Paper 170996 received Jun. 28, 2017; accepted for publication Nov. 13, 2017; published online Nov. 30, 2017.

1 Introduction

Recent development in highly energy-efficient solid-state lighting devices has opened new possibilities for their use in visible light communications (VLC).¹ VLC based on visible light-emitting diodes (LEDs) has multiple features including illumination, data communications, localization, and sensing.² VLC can be adopted in a multitude of applications, including indoor high-speed wireless connectivity with speeds of hundreds of Mb/s to Gb/s over a typical transmission span of 1 to 2 m.^{1,3} Most research works reported on VLC have focused on the proof-of-concept demonstration at higher data rates over a very short transmission range (i.e., <0.5 m at a few Gb/s). However, in typical indoor environments, both the LEDs and the channel introduce bandwidth bottleneck, which limits the achievable data rates. Most commercial white LEDs have a typical modulation bandwidth B_{LED} of a few MHz. However, gallium nitride (GaN) micro-LEDs with a dimension of <100 μm can potentially offer higher B_{LED} (i.e., hundreds of MHz).⁴ However, these small micro-LEDs offer a trade-off between B_{LED} and the transmit optical power P_t ,⁵ and hence limited applications due to the low level of P_t (i.e., orders of mW). However, to be able to fully exploit the channel bandwidth B_{ch} of an indoor environment particularly for the nonline-of-sight (NLOS) configuration, channel modeling and characterization remain an open issue and therefore needs a fresh look.⁶

In VLC, the computationally based channel models such as the recursive method, Monte Carlo ray tracing, or the ceiling bounce model have been commonly adopted from infrared (IR) communications. Contrary to the visible band, in the IR band, the reflections from walls are independent of the wavelength.^{7–10} The extensive comparison between IR and VLC channel parameters was published in Ref. 10 confirming the improvement in the coherence bandwidth and the

root-mean-square (RMS) delay spread in VLC, while diffuse, specular, and mixed specular–diffuse wavelength-dependent reflections were considered to attain a realistic channel impulse response. Moreover, in most VLC channel modeling, the receiver (Rx) is considered to be pointing toward the transmitter (Tx) mounted on the ceiling. To improve the channel DC gain and the signal-to-noise ratio, Rxs with tilting angles of 45 deg and 90 deg with respect to the Tx were investigated in Ref. 10. In Ref. 11, a detailed analysis of the Rx orientation and its impact on the received optical power level as well as the intersymbol interference was reported. However, LOS VLC links can temporarily experience blocking due to shadowing, thus resulting in the link failure and performance degradation.¹² The undesirable impact of the NLOS environment on the accuracy of VLC-based indoor localization, with the highest errors near the edge of the room, was reported in Ref. 13. A new analytical method with low computational complexity to determine the NLOS channel impulse response was presented in Ref. 14. In Ref. 15, the effect of multipath reflections on the signal-to-interference-plus-noise ratio (SINR) was investigated, showing that higher order reflections have a stronger adverse impact on the SINR than the first-order reflection when utilizing narrow field of view (FOV) Rxs.

The NLOS (diffuse) indoor optical wireless systems have been extensively investigated both theoretically and experimentally.^{16–18} In Ref. 19, a simple analytical model for LOS and diffuse IR channels for indoor environments for approximate prediction of the path loss and bandwidth was investigated showing that moderate directivity would be sufficient at data rates >100 Mbit/s. The shadowing effects on LOS and diffuse channels were analyzed in Ref. 20, where it was illustrated that the shadowed LOS configuration exhibited higher path loss, RMS delay spread, and

*Address all correspondence to: Petr Chvojka, E-mail: chvojpe8@fel.cvut.cz

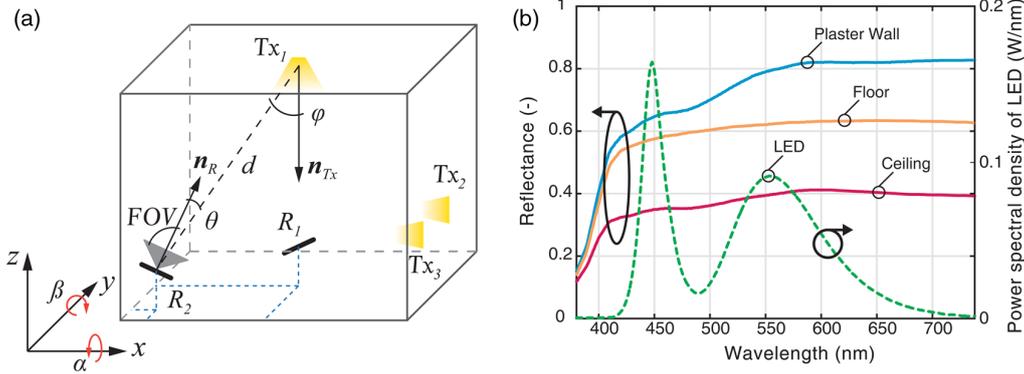


Fig. 1 (a) A geometry of the Tx and Rx in a room for the LOS scenario including the vector rotation and (b) optical spectrum of a white LED (dashed line) and reflectance of the indoor materials, adapted from Ref. 7 (solid lines).

power penalty compared with the unshadowed LOS. To overcome shadowing and to improve the link availability and performance, an NLOS scheme based on the multispot diffuse configuration and Rx's with multiple narrow FOV receiving elements employing angle diversity was introduced in Refs. 21 and 22, respectively.

In this paper, we investigate the NLOS VLC channel considering a range of Rx orientations with arbitrary positions both in azimuth and elevation, and provide the essential channel statistics. In contrast to Ref. 11, in this paper, we (i) outline detailed analysis with the emphasis being on B_{ch} , (ii) consider Rx's with different FOVs and show how the VLC link performance is prone to the Rx's rotation, (iii) further extend our work to the device-to-device (D2D) communications and show that NLOS VLC can in specific scenarios outperform LOS VLC in terms of B_{ch} by up to 14 MHz, and finally (iv) extend our work to provide statistics on shadowing due to people's movement by means of an extensive measurement campaign using a fish-eye camera-based Rx. The paper is organized as follows: in Sec. 2, the system model and experimental setup are briefly outlined, whereas in Sec. 3, the results from simulations and camera-based measurements are shown. Finally, the conclusions are given in Sec. 4.

2 System Description

2.1 Visible Light Communications Channel Model

In NLOS VLC, the channel impulse response can be obtained using the recursive method, which will be described here only briefly due to space constraint. Here, both the Tx (i.e., point source emitting a unit pulse of optical intensity) and the Rx elements are given as

$$\text{Tx} = \{\mathbf{r}_{\text{Tx}}, \mathbf{n}_{\text{Tx}}, m\} \quad (1)$$

and

$$R = \{\mathbf{r}_R, \mathbf{n}_R, A_R, \text{FOV}\}, \quad (2)$$

where \mathbf{r}_{Tx} represents a source position, \mathbf{n}_{Tx} and \mathbf{n}_R are the Tx and Rx orientation vectors normal to the surfaces of Tx and R, respectively, \mathbf{r}_R is the Rx position, and A_R is photodetector surface area [see Fig. 1(a)]. Lambertian order of the source radiation pattern is given as $m = -1/\ln(\vartheta_{1/2})$, where $\vartheta_{1/2}$ is

the semiangle at half power of an LED. To capture a rotation of the Rx, \mathbf{n}_R is calculated according to²³

$$\mathbf{n}_R = \mathbf{M}_z(\gamma)\mathbf{M}_x(\alpha)\mathbf{M}_y(\beta)\mathbf{n}_n, \quad (3)$$

where $\mathbf{n}_n = (0, 0, 1)$ is the orientation vector for the Rx facing up toward the ceiling. \mathbf{M}_z , \mathbf{M}_x , and \mathbf{M}_y are the rotation matrices of \mathbf{n}_n rotated around z -, x -, and y -axes by angles of γ , α , and β , respectively, in a counterclockwise direction using a right-handed coordinate system, see Fig. 1(a). Thus, for the Rx facing up toward the ceiling (i.e., $\alpha = \beta = 0$) the rotation matrices are given as²³

$$\mathbf{M}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \quad (4)$$

$$\mathbf{M}_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \quad (5)$$

$$\mathbf{M}_z(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

To determine channel impulse response for the k 'th reflection between Tx and R, a wall surface was discretized into N reflecting elements each having Lambertian intensity profiles. Based on the recursive algorithm, which is the sum of power contributions from each element, the channel impulse response is given as⁷

$$h^{(k)}(t; \text{Tx}, R) = \frac{m+1}{2\pi} \sum_{i=1}^N \frac{\rho_{\text{avg},i} \cos^m(\varphi_i) \cos(\theta)}{d_i^2} \cdot \text{rect}(2\theta/\pi) h^{(k-1)}(t - d_i/c; \{\mathbf{r}_{\text{Tx}}, \mathbf{n}_{\text{Tx}}, 1\}, R) dA, \quad (7)$$

where $\rho_{\text{avg},i}$ is the average reflectance of the i 'th element, θ is the angle between \mathbf{n}_R and $(\mathbf{r}_{\text{Tx}} - \mathbf{r}_R)$, φ_i is the angle between \mathbf{n}_{Tx} and $(\mathbf{r}_R - \mathbf{r}_{\text{Tx}})$, d_i represents the distance between the Tx and R, $h^{(k-1)}$ represents the impulse response of order

$(k - 1)$ between i 'th element and R , dA is the reflecting element area, c is the speed of light, and $\text{rect}(x)$ is the rectangular function equal to 1 for $|x| \leq 1$ and 0 for $|x| > 1$. The average reflectance approximation was used to simulate the wavelength dependency of the reflecting surfaces, which is given as⁷

$$\rho_{\text{avg},i} = \frac{1}{P_t} \int_{\lambda} \theta(\lambda) \rho(\lambda) d\lambda, \quad (8)$$

where $P_t = \int_{\lambda} \theta(\lambda) d\lambda$ represents the total radiant power from the LED, $\theta(\lambda)$ is the LED power spectral density (PSD), $\rho(\lambda)$ is the spectral reflectance, and λ is the wavelength. Figure 1(b) shows a typical LED PSD and the wavelength-dependent reflectance from plaster, floor, and ceiling adopted from Ref. 7. The Fourier transform on the received channel impulse response was applied and B_{ch} was determined according to $|H(B_{ch})|^2 = 1/2|H(0)|^2$. Note that the effect of Rx's spectral responsivity is not included in the channel model on assumption that white light LEDs with a broad optical spectrum and silicon photodiode are used, and hence will have a limited influence on the results presented here. Since the majority of research on the VLC channel modeling has not included the wavelength response of the Rx, we would like to ensure a like-to-like comparison with already reported data.⁷⁻¹⁵

2.2 Experimental Setup

The NLOS VLC link performance may be affected by shadowing due to the user's mobility or objects within a room, but not as much as the LOS VLC systems. Here, we experimentally investigate shadowing using a fish-eye camera-based (Moonglow ASC-N1B) Rx with a semiangle FOV ~ 90 deg. The Rx was located in the center of a room with dimensions of $\sim 5 \text{ m} \times \sim 5 \text{ m} \times \sim 3 \text{ m}$ (width \times length \times height) at the location $(2.5, 2.5, 0.8)$ facing up toward the ceiling (no rotation was considered). Here, we considered three people randomly moving around within the room, see Fig. 2(a). A 5-min-long video was recorded using the fish-eye camera and 5 samples per second were consequently extracted for further data processing to determine the

shadowing probability P_{sh} of the individual elevation angles. A sample from the recorded video is shown in Fig. 2(b). Note that, we considered only samples with at least one person captured in the video to determine the statistics. Therefore, samples without people were not considered in the image processing.

The individual samples were processed as follows: the distortion introduced by the lens was corrected by adopting the calibration procedure as in Ref. 24, which allowed us to back-project any pixel point into the 3-D space and then estimated the elevation of the shadowed region. For the foreground object segmentation (i.e., people induced shadowing at the Rx), adaptive Gaussian mixture model the background subtraction method was employed.²⁵ Finally, the processed images, see Fig. 2(c), were used to determine P_{sh} of the individual elevation angles ϕ measured from the horizontal plane as the ratio between white N_w and black N_b pixels, which is given by

$$P_{sh} = N_w/N_b. \quad (9)$$

3 Results and Discussions

For numerical simulations in MATLAB, we have considered a rectangular room with the dimensions of $5 \text{ m} \times 5 \text{ m} \times 3 \text{ m}$ (width \times length \times height) as shown in Fig. 1(a). An LED light source with P_t of 1 W was utilized as in line with the recent literature.^{7,10} We have considered three locations for the Tx and two for the Rx: Tx₁ is located at the center of the ceiling with $\mathbf{r}_{\text{Tx}1} = (2.5, 2.5, 3)$ with $m = 1$ (i.e., $\vartheta_{1/2} = 60$ deg), Tx₂ and Tx₃ are mounted on the wall with $\mathbf{r}_{\text{Tx}2} = (5, 2.5, 1)$ and $\mathbf{r}_{\text{Tx}3} = (5, 1, 1)$, respectively, for D2D communications with $m = 20$ (i.e., $\vartheta_{1/2} \sim 15$ deg) with a narrow radiation pattern, and R_1 and R_2 positioned at the center (as in the experimental part) and in the corner of the room with $\mathbf{r}_{R1} = (2.5, 2.5, 0.8)$ and $\mathbf{r}_{R2} = (0.5, 1, 0.8)$, respectively, for the worst-case scenario. Note that, we have adopted the following parameters of $A_R = 1 \text{ cm}^2$, and $\text{FOV} = \{60 \text{ deg}, 70 \text{ deg}, 80 \text{ deg}, \text{ and } 90 \text{ deg}\}$. As in Ref. 19, we considered $k = 3$ reflections modeled as purely diffuse, which can be approximated by the interior of an integrating sphere.

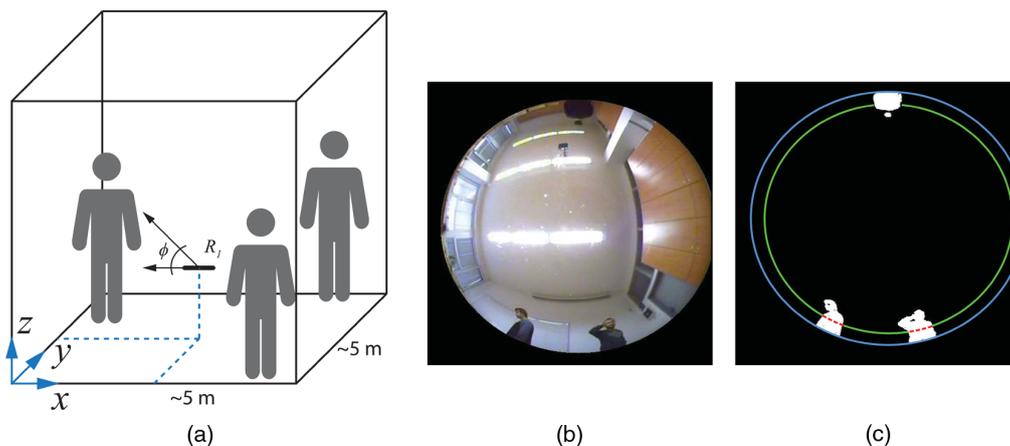


Fig. 2 (a) The experimental setup for VLC with shadowing, (b) a captured video sample, and (c) a processed image with a highlighted elevation angle of 7 deg. The calculated P_{sh} is $\sim 10\%$. Note that the elevation angle and its shadowed part are illustrated by green and red (dashed) colors, respectively.

The focus of the paper was to determine the system bandwidth, which is independent of modulation schemes.^{7,10} Adopting multicarrier schemes such as orthogonal frequency division multiplexing (OFDM) and multiband carrier-less amplitude and phase (*m*-CAP) modulation, equalization or error control code techniques can significantly enhance the system performance by eliminating the ripples in the frequency response or by adapting the transmitted signal to the system frequency response. However, knowledge about the system performance (without any additional advanced technique) in terms of bandwidth, received power, and other parameters is essential for an efficient VLC system design. Thus, we focused on the channel characterization without any additional technique that can improve the system performance, which makes the results of the paper more general and allows the results to be compared with most of the existing research.

3.1 Channel Bandwidth

Figures 3(a) and 3(b) show B_{ch} against the Rx's rotation angle β (i.e., rotation around the *y*-axis) for a range of FOVs for Tx₁ placed at the center of the ceiling and for R₁ and R₂ located in the middle and corner of the room,

respectively. This provides an NLOS configuration for the typical case of indoor VLC. The color regions in each figure denote the range of rotation angles for FOV values, where the received optical signal also includes the LOS component. For example, for a detector with an FOV of 80 deg for the link Tx₁ – R₂, see Fig. 3(b), the rotation angle β is within the range of –120 deg to 40 deg. Note that, the Rx is rotated around the *y*-axis in a counterclockwise direction as shown in Fig. 1(a). We can see from Fig. 3(a) that B_{ch} is almost the same over the wide range of β independent of the FOV. However, for the Rx with an FOV of 60 deg (gray curve) and for $|\beta| > \sim 165$ deg (i.e., the Rx is facing toward the floor), we notice a significant improvement in B_{ch} by up to 14 MHz when only the NLOS path is considered. This is because the received power from the first reflection from the floor is considerably higher than signal power from higher order reflections from other surfaces within the room.

The unlimited B_{ch} is achieved for the LOS case, i.e., for $|\beta| < 50$ deg for an FOV of 60 deg. Note that, this range of β decreases with an increase in the values of FOV, see Fig. 3(a), thus contributing to the limitation of the user's mobility. However, the pure NLOS link can still support adequate B_{ch} of >7 MHz. Note, in the NLOS scenario higher $|\beta|$ leads to higher B_{ch} . Figure 3(b) shows B_{ch} for the worst

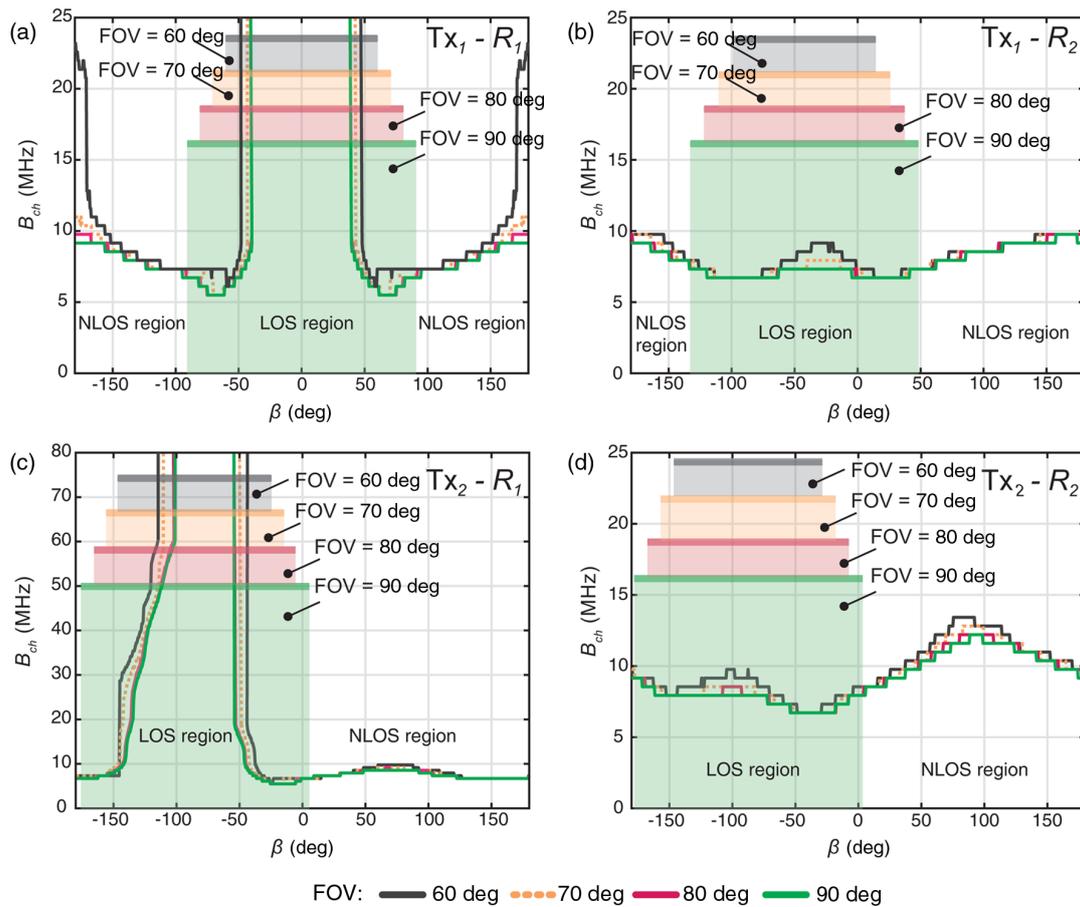


Fig. 3 The channel bandwidth as a function of the rotation angle β around *y*-axis for Tx placed (i) on the ceiling (upper row) and Rx's positions: (a) in the middle (R₁), and (b) in the corner of the room (R₂); and (ii) on the wall for D2D communications (bottom row) and Rx's position: (c) R₁, and (d) R₂. Note that color regions denote the presence of LOS path for particular FOVs.

scenario where the source is located on the ceiling. With R_2 placed in the corner of the room, we observe similar profiles for both LOS and NLOS scenarios for a range of FOVs. However, the pure NLOS offers slightly higher (up to ~ 3 MHz) B_{ch} compared with the LOS for the FOV = {70 deg, 80 deg, and 90 deg} because of lower received power levels of the LOS path.

Figures 3(c) and 3(d) show B_{ch} for both the Rx's positions and for the Tx_2 placed on the wall for D2D communications. With the Rx tilted toward the Tx (i.e., see the LOS regions of individual Rxs), we observe a significant improvement in B_{ch} compared with the case where the source is placed on the ceiling, see Fig. 3(a). This is because of the presence of a strong LOS component. In the best case scenario, B_{ch} is ~ 62 MHz for all values of FOV. Note that the NLOS case does not outperform the LOS link in D2D as

was observed in the previous cases. For the Rx placed in the corner of the room (i.e., R_2) as in Fig. 3(d), the measured B_{ch} of the signal with the LOS component is lower (up to ~ 8 MHz) for all values of FOV for $\beta > \sim 25$ deg compared with the NLOS case. This is attributed to two main reasons: (i) lower power contributions from the LOS path compared with the higher order reflections, since the Tx and the Rx are not pointing to each other as is the case in most D2D communications, and (ii) higher power contributions from the first reflection than from the higher order reflections, thus resulting in slightly increased B_{ch} compared with LOS. Moreover, compared with Fig. 3(c), we observe an improvement up to 4 MHz in B_{ch} for $\beta > \sim 50$ deg for the pure NLOS channel, which is caused by stronger NLOS components received by the detector placed in the corner of the room.

To further illustrate the importance of the research findings for the D2D communications, we have evaluated the system performance for the Tx_3 placed on the central position of the wall and R_1 rotated around the y -axis. The measured B_{ch} against the rotation angle β is shown in Fig. 4. Compared with the results shown in Figs. 3(a) and 3(c), the system with NLOS path outperforms the LOS in B_{ch} up to ~ 4 MHz. Note that, for the LOS path B_{ch} is only ~ 6 MHz for the range of $-180 \text{ deg} < \beta < 5 \text{ deg}$. This is significantly lower compared with Figs. 3(a) and 3(c) due to the lower intensity of the LOS component arrived at the photodetector.

Next, we considered the effect of the rotation of the Rx by the angle α around the x -axis on B_{ch} for NLOS D2D communications. The results for both R_1 and R_2 are shown in Figs. 5(a) and 5(b), respectively. B_{ch} is almost flat for all values of FOV and α , i.e., independent of the user's mobility. However, for $\alpha < -10$ deg the B_{ch} of NLOS is slightly (by ~ 2.5 MHz) higher compared with the LOS as shown in Fig. 5(b). Note that, for detectors with FOV = {60 deg, 70 deg, and 80 deg} and FOV = {60 deg and 70 deg} no LOS component was received for any value of α as shown for R_1 and R_2 in Figs. 5(a) and 5(b), respectively.

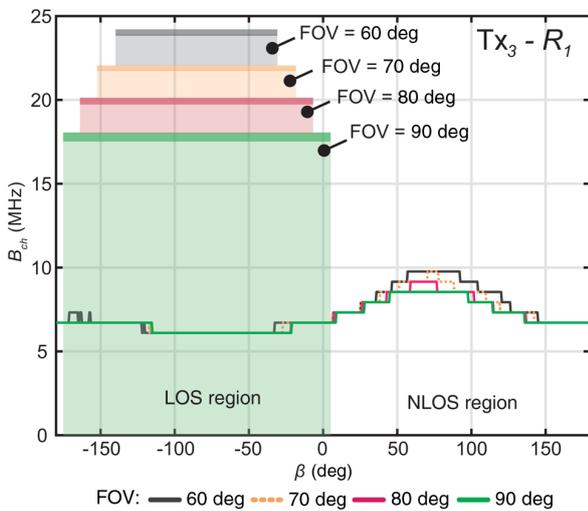


Fig. 4 The channel bandwidth as a function of the rotation angle β around y -axis for the Tx_3 and the R_1 for D2D communications. Note that color regions denote the presence of LOS path for particular FOVs.

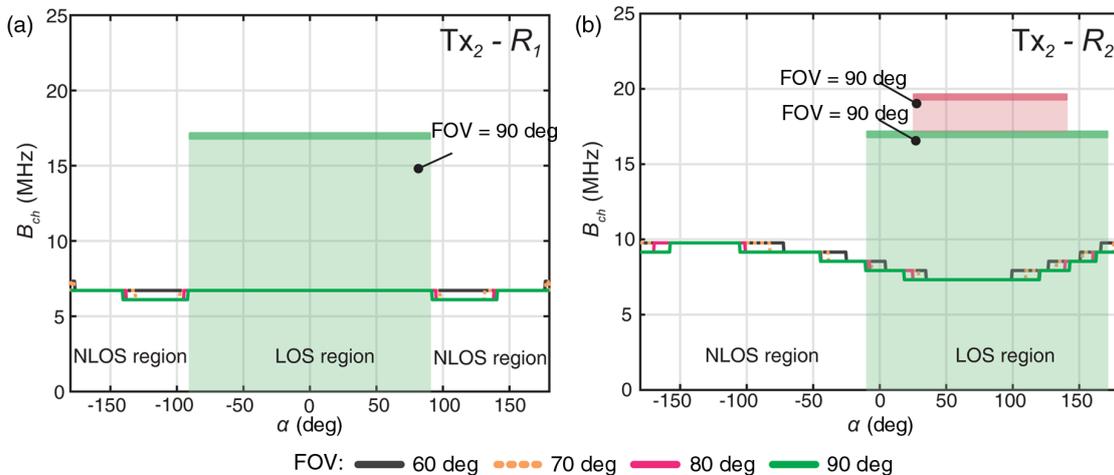


Fig. 5 The channel bandwidth against the rotation angle α around the x -axis for the Rx's position: (a) in the center (R_1) and (b) in the corner of the room (R_2) for D2D communications. Note that the color regions denote the presence of LOS component for particular Rxs' FOVs (otherwise NLOS is observed).

3.2 LOS and NLOS Probability

Considering the rotation of the Rx, the numerical simulations show that for a Rx with a narrow FOV (i.e., 60 deg) the probability density function of receiving a signal including the LOS path is approximately half that of the pure NLOS. As shown in Fig. 6(a) for the case of $Tx_1 - R_1$ link and

an Rx with an FOV of 60 deg, ~69% of transmission is established via the pure NLOS paths at the cost of reduced optical power levels compared with the LOS. However, the user's mobility is enhanced. An increase in the Rx's FOV to 90 deg resulted in equal transmission via LOS and NLOS paths, see Fig. 6(b).

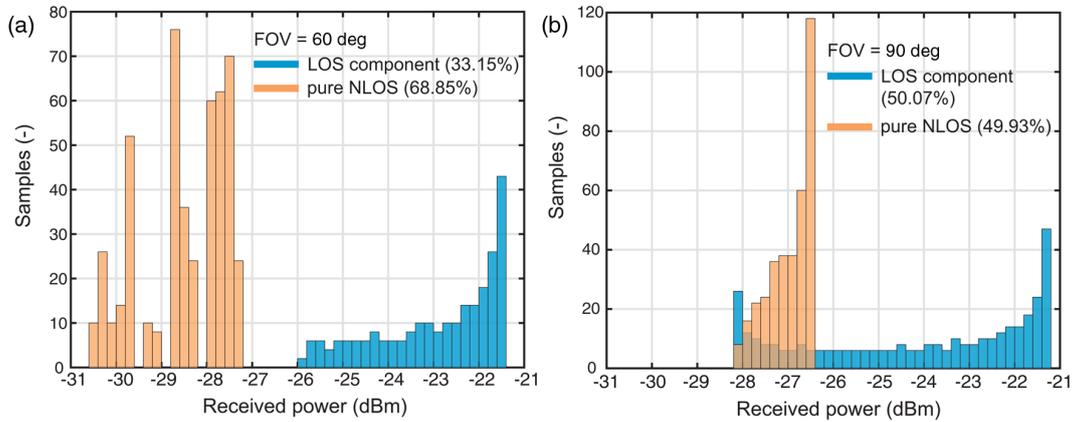


Fig. 6 The histograms of the received power for the configuration $Tx_1 - R_1$ and full rotation around y -axis for a Rx: (a) FOV = 60 deg and (b) FOV = 90 deg with separated signals with LOS and pure NLOS received power components.

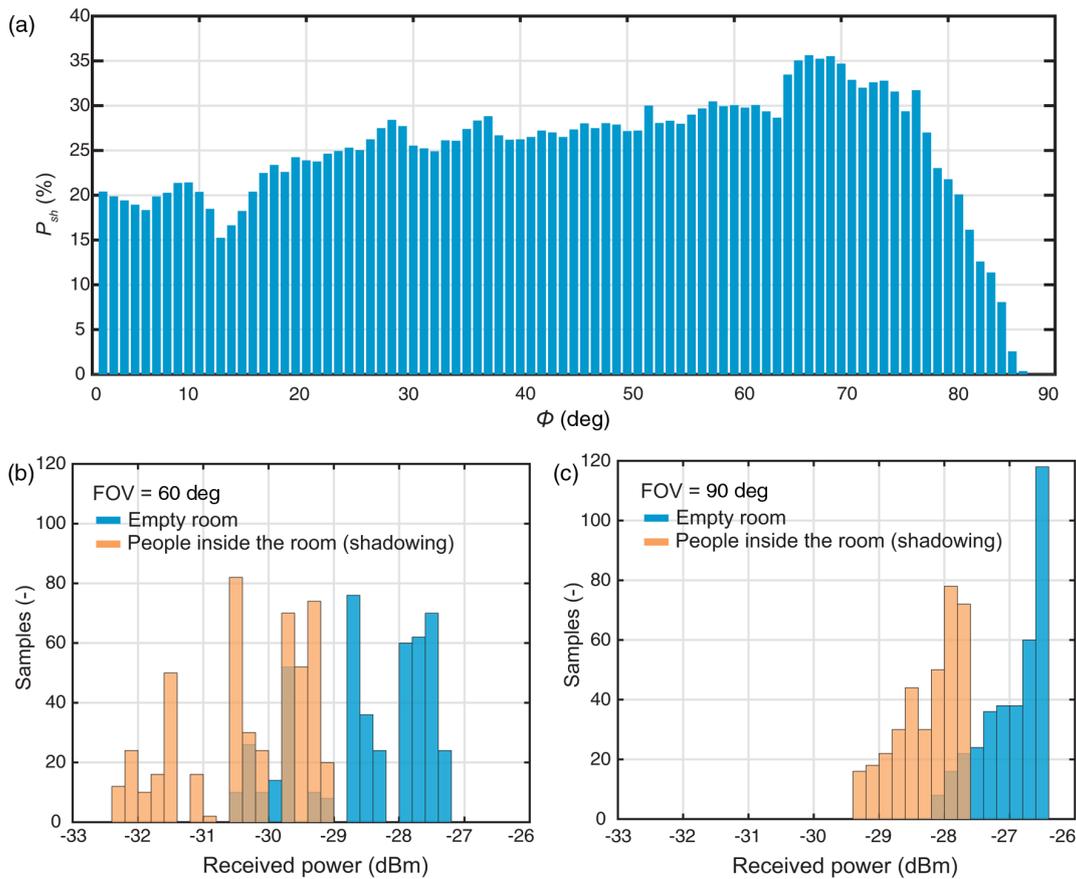


Fig. 7 Configuration $Tx_1 - R_1$ with shadowing caused by people: (a) the measured maximum shadowing probabilities for the range of elevation angles; and histograms of the pure NLOS received power for, and (b) FOV = 60 deg and (c) FOV = 90 deg.

3.3 Shadowing

Using the setup shown in Fig. 2(a), we carried out measurements for the maximum shadowing probabilities P_{sh} for a range of elevation angles ϕ as well as histograms of the received power for the pure NLOS for the FOVs of 60 deg and 90 deg, see Fig. 7. Note that, the highest measured P_{sh} of $\sim 35.6\%$ is observed at ϕ of 67 deg. Clearly, higher ϕ is affected more by people moving around the Rx. However, shadowing at these elevation angles occurs rarely in comparison with low elevation angles. Based on the measured data, the received power level for the Tx₁ – R₁ configuration was consequently calculated. The same system model as in Sec. 2 was adopted in the simulation and only pure NLOS and shadowing [i.e., the measured P_{sh} values from Fig. 7(a)] due to people's movement around the Rx were considered. The total optical received power is dropped by 1.8 dB for the Rx with an FOV of 60 deg. For the Rx with a wider FOV of 90 deg, the received power is decreased by only 1.3 dB (i.e., such an Rx is less sensitive to shadowing, since signals are captured over a much wider range of elevation angles). Note that, shadowing also affects the histograms of received optical power for pure NLOS as shown in Figs. 7(b) and 7(c) for the FOVs of 60 deg and 90 deg, respectively.

4 Conclusion

In this work, we outlined the analysis of the NLOS VLC based on numerical simulations by including the receiver's orientation and variable FOVs. We considered a typical ceiling-mounted transmitter for the case of D2D communications and showed that the pure NLOS channel in specific cases offered superior performance compared with the link with both LOS and NLOS paths resulting in the increased channel bandwidth up to 14 MHz. Moreover, the receiver's rotation increased the probability of transmission via the pure NLOS channel compared with the case of the LOS path for narrow FOV receiver. Based on the experimental campaign, we showed that higher elevation angles and receivers with a narrow FOV are more prone to the shadowing caused by people's movement within a room resulting in the received power level drop of 1.8 and 1.3 dB for the receiver with the FOV of 60 deg and 90 deg, respectively.

Acknowledgments

The work was supported by GACR 17-17538S.

References

- Z. Ghassemlooy et al., *Visible Light Communications Theory and Applications*, CRC Press, Boca Raton (2017).
- H. Chun et al., "LED based wavelength division multiplexed 10 Gb/s visible light communications," *J. Lightwave Technol.* **34**(13), 3047–3052 (2016).
- K.-D. Langer et al., "Rate-adaptive visible light communication at 500 Mb/s arrives at plug and play," *SPIE Newsroom* (14 November 2013).
- R. X. G. Ferreira et al., "High bandwidth GaN-based micro-LEDs for multi-Gb/s visible light communications," *IEEE Photonics Technol. Lett.* **28**(19), 2023–2026 (2016).
- S. Rajbhandari et al., "High-speed integrated visible light communication system: device constraints and design considerations," *IEEE J. Sel. Areas Commun.* **33**(9), 1750–1757 (2015).
- Z. Ghassemlooy et al., "Emerging optical wireless communications—advances and challenges," *IEEE J. Sel. Areas Commun.* **33**(9), 1738–1749 (2015).
- K. Lee, H. Park, and J. R. Barry, "Indoor channel characteristics for visible light communications," *IEEE Commun. Lett.* **15**(2), 217–219 (2011).
- S. Long et al., "Investigating channel frequency selectivity in indoor visible-light communication systems," *IET Optoelectron.* **10**(3), 80–88 (2016).
- Y. Qiu, H.-H. Chen, and W.-X. Meng, "Channel modeling for visible light communications: a survey," *Wireless Commun. Mob. Comput.* **16**, 2016–2034 (2016).
- F. Miramirkhani and M. Uysal, "Channel modeling and characterization for visible light communications," *IEEE Photonics J.* **7**(6), 1–16 (2015).
- C. Le Bas et al., "Impact of receiver orientation and position on visible light communication link performance," in *Proc. Int. Workshop on Optical Wireless Communications*, pp. 1–5 (2015).
- P. Chvojka et al., "Channel characteristics of visible light communications within dynamic indoor environment," *J. Lightwave Technol.* **33**(9), 1719–1725 (2015).
- W. Gu et al., "Impact of multipath reflections on the performance of indoor visible light positioning systems," *J. Lightwave Technol.* **34**(10), 2578–2587 (2016).
- C. Chen, D. Basnayaka, and H. Haas, "Non-line-of-sight channel impulse response characterisation in visible light communications," in *IEEE Int. Conf. on Communications (ICC)*, Kuala Lumpur, pp. 1–6 (2016).
- Z. Chen, D. Tsonev, and H. Haas, "Improving SINR in indoor cellular visible light communication networks," in *IEEE Int. Conf. on Communications (ICC)*, Sydney, pp. 3383–3388 (2014).
- J. B. Carruthers and J. M. Kahn, "Modeling of nondirected wireless infrared channels," *IEEE Trans. Commun.* **45**(10), 1260–1268 (1997).
- H. Hashemi et al., "Indoor propagation measurements at infrared frequencies for wireless local area networks applications," *IEEE Trans. Veh. Technol.* **43**(3), 562–576 (1994).
- F. R. Gfeller and U. Bapst, "Wireless in-house data communication via diffuse infrared radiation," *Proc. IEEE* **67**(11), 1474–1486 (1979).
- V. Jungnickel et al., "A physical model of the wireless infrared communication channel," *IEEE J. Sel. Areas Commun.* **20**(3), 631–640 (2002).
- J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental characterization of non-directed indoor infrared channels," *IEEE Trans. Commun.* **43**(2/3/4), 1613–1623 (1995).
- S. T. Jovkova and M. Kavehard, "Multispot diffusing configuration for wireless infrared access," *IEEE Trans. Commun.* **48**(6), 970–978 (2000).
- J. B. Carruthers and J. M. Kahn, "Angle diversity for nondirected wireless infrared communication," *IEEE Trans. Commun.* **48**(6), 960–969 (2000).
- H. Goldstein, C. Poole, and J. Saffko, *Classical Mechanics*, Addison Wesley, Essex (2002).
- D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Beijing, pp. 5695–5701 (2006).
- Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proc. 17th Int. Conf. on Pattern Recognition*, Cambridge, pp. 28–31 (2004).

Biographies for the authors are not available.

E Appendix E

CHAVEZ-BURBANO, P., et al. Optical camera communication system for Internet of Things based on organic light emitting diodes. *Electronics Letters*, 2019.

Author's contribution: experimental work, setup of the measurements.

Optical camera communication system for Internet of Things based on organic light emitting diodes

P. Chavez-Burbano[✉], S. Vitek, S.R. Teli, V. Guerra, J. Rabadan, R. Perez-Jimenez and S. Zvanovec

The inclusion of organic light emitting diodes (OLEDs) in high-end devices, such as TVs and smart-phones, along with the insertion of cameras embedded in daily use devices, provides the opportunity to establish optical camera communication (OCC) systems based on OLED emitters for Internet of Things (IoT). This Letter presents an experimental demonstration of this OLED-based OCC system for IoT. The results suggest that despite the low emitted power of OLED devices, long range links can be established based on OCC. One outdoor and two indoor scenarios are tested, validating BER below 10^{-6} for short range and give just 3.56×10^{-3} for long range links.

Introduction: Nowadays, organic light emitting diodes (OLEDs) have been massively introduced within both the solid state lighting and the high-resolution displays markets due to their low power consumption and cost-efficient manufacturing process. Worldwide companies such as Samsung, LG, Apple, and Sony, have been integrating OLEDs in their high-end smartphones (i.e. Samsung Galaxy S7, iPhone X), high-end TVs (i.e. Sony A1, LG OLED65B7V) and flexible decorative lamps.

OLEDs have been proposed as emitters for implementing visible-light communication (VLC) systems and for future 5G systems [1]. Haigh *et al.* presented an OLED-based implementation using on-off keying (OOK) at 250 kbps with a bit error rate (BER) of 10^{-6} [2]. This system was improved by using pre-equalisation and multiple-input multiple-output (two emitters and nine receiver sections), increasing the data rate up to 1.8 Mbps [3]. In [4], a 54.9 Mbps OLED-based system was introduced. The authors used a high performance artificial neural network equaliser for classifying the input signal based on a training sequence for an RGB polymer LED. In the same way, orthogonal frequency division multiplexing has been proposed for increasing the data rate of OLED-based VLC systems. In [5], this modulation was combined with power allocation in order to achieve 10 Mbps using a 350 kHz bandwidth emitter. Chen *et al.* demonstrated a 51.6 Mbps transmission with a BER of 3.8×10^{-3} by implementing offset quadrature amplitude modulation [6]. However, these high-rate solutions were reached at extremely short distances (below 15 cm), and further experiments only achieved 1 m span limit. To the authors' knowledge, currently there is no literature addressing tests with OLED based VLC systems for medium and longer distances, outdoor cases, neither the use of cameras as receivers.

In this work, the idea of using optical camera communication (OCC) systems for IoT based on OLEDs, instead of other traditional VLC schemes already proposed for data transportation in IoT networks [7], is presented and experimentally validated. In this way, longer distances can be easily reached for systems that require relatively low data transmission rates (up to 1 kbps). Since OCC systems can be affected by the blooming effect, different exposure times were tested in this work. High exposure times affect several pixels rows due to the overlapped scanning of rolling shutter cameras, creating visible white halos around the illuminated portions and therefore image distortion between the bits' bands.

System description: The proposed OCC system is based on the utilisation of OLED devices as transmitters and a rolling-shutter camera as a receiver. There are several types of OLED emitters that can be used. Bendable panels are implemented in rooms, on clothing or embedded into decoration appliances. OLED monitors are usually implemented as advertisement devices. Finally, OLED screens are implemented as part of electronic devices such as smartphones and tablets. The signal is modulated using OOK during the experiments. Nonetheless, its under-sampled variants in frequency, amplitude or phase can be used for avoiding possible flickering. These modulations are part of the communication modes proposed for OCC by the IEEE 802.15.7m Optical Wireless Communications Task Group [8]. At the beginning of the transmission, a calibration signal (clock-like signal) is sent in order to adjust the system's thresholds. The OLED-device transmits the modulated signal by switching between the on and off states accordingly at frequency f_s .

On the receiver side, two types of cameras can be implemented: smart-device cameras such as those embedded into watches, smartphones, tablets or glasses, and closed circuit television (CCTV)

cameras such as the ones used for security and surveillance purposes. As it is shown in Fig. 1, each frame is captured by the camera at predetermined frames per second (fps) using N_y pixels of vertical resolution, and stored for its corresponding processing. To speed up the processing performance, a smaller image with only the emitter's signal information is required, so the first step is the detection of the region of interest (ROI), which can be multi-target (several simultaneous emitters). The first frame is binarised using Otsu's algorithm and dilated for finding the sources as image regions, which extrema points can be extracted. The obtained coordinates that define the boundaries of the ROI are used for the image cropping. In the multi-target case, the frame is cropped in several images, one per target. Since the first frame contains the calibration signal, the histogram information of its three channels (red, green and blue) is used for calculating the corresponding threshold values by assuring that the 50% of the data belong to each bit's level. Then, the remaining frames are binarised applying these threshold values, and the obtained B/W image (binary matrix) is projected on a 1D vector by averaging each row. Finally, this vector is decoded using (1) which calculate the corresponding number of pixels per bit (n_{bit}) in order to acquire the transmitted data. The number of consecutive ones or zeros are divided by the estimated n_{bit} . Nevertheless, since there may be sampling frequency offset, the final output takes into account the decimal part of the division, adding one bit when needed.

$$n_{bit} = \frac{N_y}{f_s} \cdot fps \quad (1)$$

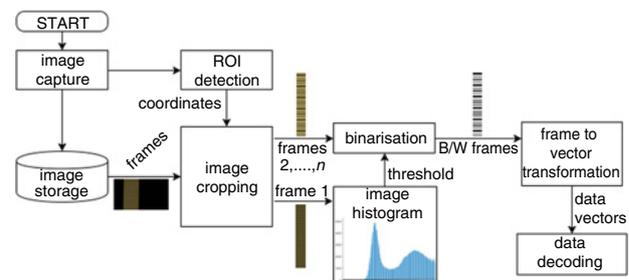


Fig. 1 OCC data processing flowchart

To validate the proposed system, three test cases were performed at the Faculty of Electrical Engineering building of the Czech Technical University in Prague: an indoor case up to 4 m, a corridor up to 120 m, and finally an outdoor test up to 120 m. In all cases, a 751 m bendable bar type OLED panel was used as the emitter, with an area of 187.84 mm × 37.84 mm and 85% of spatial luminous uniformity. This panel was switched by a MOSFET-based driver which implements an OOK non-return to zero modulation. An USB rolling-shutter camera with variable exposure time (from 100 μs to 30 s) was selected as the receiver. This camera has a video resolution of 720 × 540 pixels, manual white balance of 0 dB, gain of 3 dB, selectable frame rate (fps), brightness and contrast of 25/100, zero Hue, and saturation of 70. For distances up to 40 m, the camera worked with a lens of 12.5 – 75 mm and F/1.8. For longer indoor test outdoor tests, the camera was mounted on a 10/1000 macro telephoto lens with focal length of 1000 mm, geometrical relative aperture 1 : 10 and angular field-of-view (FOV) of 2.5°.

At the beginning of the transmission, a calibration sequence was sent, then the data were encapsulated within 2-bytes long packets that had specific start and end delimiters (0xA sequence). In total, 1200 frames were recorded and stored per each trial. Finally, the videos were off-line processed using MATLAB routines following the procedure described in Fig. 1.

Indoor experiment: The indoor experiment was performed within a 6 m × 6 m room. The distance between the emitter and the receiver was set to 4 m, while the switching frequency varied from 500 to 5000 Hz with steps of 500 Hz. The camera's frame rate was changed among five values: 17.5, 19.01, 25, 30, and 50 fps; while its exposure time was set to two different values: 400 and 500 μs.

Fig. 2 depicts measured BER for different switching frequencies. It can be seen that the trials of 30 and 50 fps reached the best BER results, less than 10^{-2} for 5 kHz. Taking into account the packet encoding and the fact that each 2-byte packet was repeated twice in order to ensure its detection, the achieved data rates were approximately 120

and 200 bps, respectively. In general, the system worked properly up to 2.5 kHz, for all the frame rates, with less than 10^{-5} BER.

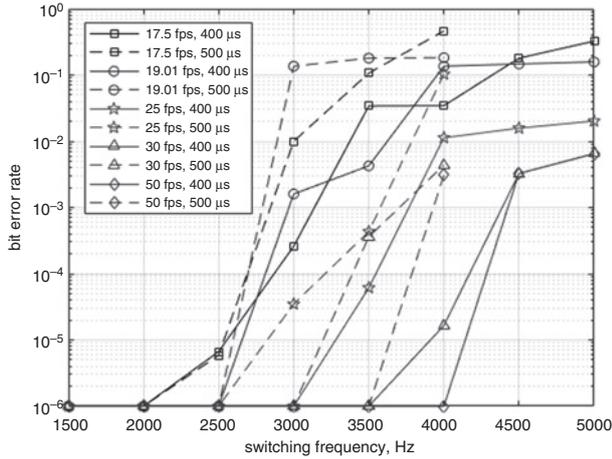


Fig. 2 BER versus switching frequency of the 4 m-long indoor scenario

Indoor corridor measurements: For longer indoor distances, two sets of trials were carried out in a 200 m corridor. In both experiments, the switching frequency was set to 1 kHz, the frame rate was set to 25 and 50 fps, and the exposure time was 400 μ s. These values were selected using (1) and considering that the target link range was longer than in the previous experiment.

In the first case, the distance ranged from 10 to 40 m in steps of 5 m and the receiver used the 75 mm focal-length lens. The system worked as expected compared to the previous experiment, for both camera frame rates, the measured BER value remained less than $< 10^{-6}$ for all cases.

In the second case, the distance varied from 20 to 120 m in steps of 20 m. Furthermore, the camera was mounted on the telescope, and the captured images were vertically compressed 2 : 1 to simplify the location of the emitter at the beginning of the trials. As shown in Fig. 3, the BER increased from 9.68×10^{-3} at 40 m to 5.29×10^{-2} at the limit test distance (120 m), demonstrating that the OLED-based OCC system can achieve longer distances. However, the use of the telescope may introduce misalignment errors and presented some problems related to the stability of the receiver.

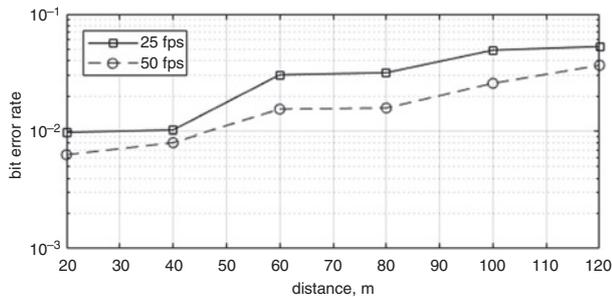


Fig. 3 Dependence of BER on distance in the case of OCC system tested within an indoor corridor using a telescope at the receiver side

Outdoor measurements: Finally, the experimental tests were performed on the roof of the faculty building (approximately 30 m above ground). The measurements were taken in June 2018 during a scattered cloudy afternoon with the following atmospheric conditions measured by the meteorological station placed at one side of the link: temperature of 18°C, wind speed of 9 km/h and 60% humidity. The small wind speed and mostly the reduced link range limited the possible harmful effect of turbulence. In this experiment, two link distances were tested (60 and 120 m), and the switching frequency was established to 1 kHz. The camera was mounted on the telescope with a binning of 2 : 1. In addition, the tested frame rates were 25 and 50 fps, and the exposure time was 400 μ s.

In this case, as shown in Table 1, using a 50 fps frame rate for a distance of 60 m, the 1 kHz transmission achieved less than 3×10^{-5} BER, and BER of 3.56×10^{-3} was experienced in the case of 120 m link.

Table 1: BER measured in the case of the outdoor scenario.

Distance (m)	25 fps	50 fps
60	2.77×10^{-3}	$< 3 \times 10^{-5}$
120	6.54×10^{-3}	3.56×10^{-3}

Conclusion: Based on the results of the different tested indoor and outdoor cases, the feasibility of an OCC system for IoT based on OLED devices was demonstrated. The indoor transmissions showed that the system can be implemented with a standard USB camera of 30 fps, obtaining BERs below 1.7×10^{-5} for measurements inside a standard room (4 m) with switching frequency up to 4 kHz, and less than $< 10^{-6}$ for distances shorter than 40 m with 1 kHz.

These experiments also proved that the camera's exposure time influenced the decoding accuracy due to the blooming effect. Using smaller exposure time limited the overlapping read time between the rows, minimising the white halos. The exposure time of 400 μ s allowed an easier calibration due to the better distribution of the bands' width and improved the threshold levels' calculation and the decoding process. The laboratory tests showed that for each selected camera's frame rate, the BER's measurements are better with an exposure time of 400 μ s rather than 500 μ s.

The results shown in Table 1, verified that an outdoor OLED-based OCC system is feasible. In the same way, the number of projected pixels from the panel in the frame in the outdoor tests assured that longer distances can be easily achieved so this system can even be applied for Smart Cities. Implementations for environmental conditions monitoring, disaster detection, or smart traffic management are viable using spatial division multiple access allowing several emitters spatially separated within the FOV of the camera.

Acknowledgments: This work was supported in part by Escuela Superior Politecnica del Litoral, Ecuador; by Horizon 2020 MSCA ITN project no.764461 (VISION); and by project GACR 17-17538S.

© The Institution of Engineering and Technology 2019

Submitted: 10 December 2018 E-first: 5 February 2019

doi: 10.1049/el.2018.8037

One or more of the Figures in this Letter are available in colour online.

P. Chavez-Burbano (*Facultad de Ingenieria en Electricidad y Computacion, Escuela Superior Politecnica del Litoral, Guayaquil, Guayas, Ecuador*)

✉ E-mail: paxichav@espol.edu.ec

S. Vitek, S.R. Teli and S. Zvanovec (*Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic*)

V. Guerra, J. Rabadan and R. Perez-Jimenez (*IDeTIC, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain*)

References

- Zvanovec, S., Chvojka, P., Haigh, P., *et al.*: 'Visible light communications towards 5G', *Radioengineering*, 2015, **24**, (1), pp. 1–9
- Haigh, P., Ghassemlooy, Z., Rajbhandari, S., *et al.*: 'Visible light communications using organic light emitting diodes', *Commun. Mag.*, 2013, **51**, (8), pp. 148–154
- Haigh, P., Ghassemlooy, Z., Papakonstantinou, I., *et al.*: 'A MIMO-ANN system for increasing data rates in organic visible light communications systems'. IEEE Proc. Int. Conf. on Commun., Budapest, Hungary, June 2013, pp. 5322–5327
- Haigh, P., Bausi, F., Le Minh, H., *et al.*: 'Wavelength-multiplexed polymer LEDs: towards 55 mb/s organic visible light communications', *J. Sel. Areas Commun.*, 2015, **33**, (9), pp. 1819–1828
- Le, S., Kanesan, T., Bausi, F., *et al.*: 'A 10 Mb/s visible light transmission system using a polymer light-emitting diode with orthogonal frequency division multiplexing', *Opt. Lett.*, 2014, **39**, (13), pp. 3876–3879
- Chen, H., Xu, Z., Gao, Q., *et al.*: 'A 51.6 Mb/s experimental VLC system using a monochromatic organic LED', *Photon. J.*, 2018, **10**, (2), pp. 1–12
- Kim, C., and Koh, S.: 'Device management and data transport in IoT networks based on visible light communication', *Sensors*, 2018, **18**, (8), p. 2741
- Nguyen, T., Islam, A., Yamazato, T., *et al.*: 'Technical issues on IEEE 802.15.7 m image sensor communication standardization', *Commun. Mag.*, 2018, **56**, (2), pp. 213–218

F Appendix F

VÍTEK, Stanislav; MELNIČUK, Petr. A Distributed Wireless Camera System for the Management of Parking Spaces. *Sensors*, 2018, 18.1: 69.

Author's contribution: main idea, design of the system, data analysis.

Article

A Distributed Wireless Camera System for the Management of Parking Spaces

Stanislav Vítek *  and Petr Melničuk

Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2,
166 27 Prague, Czech Republic; melnicukpetr@gmail.com

* Correspondence: viteks@fel.cvut.cz; Tel.: +420-22435-2232

Received: 22 September 2017; Accepted: 22 December 2017; Published: 28 December 2017

Abstract: The importance of detection of parking space availability is still growing, particularly in major cities. This paper deals with the design of a distributed wireless camera system for the management of parking spaces, which can determine occupancy of the parking space based on the information from multiple cameras. The proposed system uses small camera modules based on Raspberry Pi Zero and computationally efficient algorithm for the occupancy detection based on the histogram of oriented gradients (HOG) feature descriptor and support vector machine (SVM) classifier. We have included information about the orientation of the vehicle as a supporting feature, which has enabled us to achieve better accuracy. The described solution can deliver occupancy information at the rate of 10 parking spaces per second with more than 90% accuracy in a wide range of conditions. Reliability of the implemented algorithm is evaluated with three different test sets which altogether contain over 700,000 samples of parking spaces.

Keywords: parking space; occupancy; smart city; internet of things; raspberry pi; low-power consumption; histogram of oriented gradients; support vector machine

1. Introduction

Recent population growth in urban areas negatively affects vehicle traffic in city centers. In addition to the negative impact on the environment, finding a vacant space in parking lots during peak hours may be almost impossible. Shoup reports that, based on the sample of 22 studies performed in 11 cities (Detroit, Washington, New Haven, London, Freiburg, Jerusalem, Cambridge, Cape Town, New York, San Francisco, and Sydney) between 1927 and 2001, drivers spend on average eight minutes finding a vacant space [1]. Naturally, cruising time may vary, numerous studies performed in European cities (Netherlands, Belgium) report average cruise time for on-street parking of 30 s and off-street parking 50 s [2].

Systems able to manage this problem can be categorized into counter-based, sensor-based, and image or video based. The first two categories have a couple of drawbacks: counter-based systems could help only with information about a total number of vacant spaces, a sensor-based system costs a lot because of the number of sensors required to cover the entire parking lot. However, the third category is usually considered as quite expensive and producing a significant amount of data, which are unable to transmit over the wireless network, the growth in low-cost, low-power sensing and communication technologies enables a wide range of physical objects and environments to be monitored in fine spatial and temporal detail. A network of dedicated low-power devices connected to the cloud then could be part of the Internet of Things (IoT) platform for smart cities [3].

The oldest video systems for automatic detection of vacant parking spaces were based on a simple calculation of the differences between consecutive frames [4]. Motion detection is not cost-effective because of the need to process large amounts of redundant data. The system requires continuous

image acquisition, which does not provide room for any energy savings. Although systems using temporal characteristics are not efficient, their accuracy may be relatively high, as shown by recent works [5,6]. All the below-mentioned approaches do not consider the temporal link between the consecutive frames and solve each image individually.

The system described in [7] is computationally less demanding. The system determines decision about occupancy from the ratio of the number of edges calculated by Laplace's operator in the defined areas. The authors report an accuracy of over 95% for one test sequence of a parking lot with 53 parking spaces. The robustness of the classification algorithm is, however, quite low when conditions are worse. Delibatov et al., [8] model a parking slot as a volume in the 3D space. This approach eliminates problems with overlaps. Parking spaces are semi-automatically marked and then modeled by the probability density that a pixel is part of the parking space. For detection, a rotationally invariant local binary pattern (LBP) descriptor with a support vector machine (SVM) classifier with a nonlinear core is used. The LBP-SVM classifier works on small blocks that can be classified as one of six classes: cars, land, trees, skies, buildings and the rest. The final result of the occupancy of the parking space is obtained by the weighted sum of the results of the block classification where the scales are given by the models of parking spaces.

A number of works are focused on testing of the algorithms' robustness. De Almeida et al., [9] verify the usability of texture descriptors based on LBP and local phase quantization (LPQ) descriptors using the SVM classifier. The authors showed an accuracy of 99% when the classification algorithm was tested on the same view of the car park as it was trained. If a different view of the car park was selected for training, they reached an accuracy of around 89%. Algorithms were tested on the PKlot [10] database of almost 700,000 samples.

In addition to the descriptors already mentioned, there are works employing wavelet transform [11], scale-invariant feature transform (SIFT) [12], histogram of oriented gradients (HOG) [13,14], integral channel features [15], or Hue histogram descriptor [16]. Fusek et al., used AdaBoost based algorithm and also compared such approach with HOG descriptor [17].

From the point of view of reliability, it is worth mentioning the works focused on Convolutional Neural Networks (CNN). Amato et al., [18] describes the system using deep learning with accuracy better than 90%. CNN was trained to directly decide about the occupancy status of the individual parking spaces seen by the video cameras. The authors used the CNRPark dataset [19] consisting of roughly 12.000 labeled images and the CNRPark-EXT dataset including 150.000 labeled samples. Classification of one parking space takes around 300 ms Raspberry Pi model B.

The paper is organized as follows. Section 2 is dedicated to some aspects of camera systems design and Section 3 introduces proposed system. Section 4 discussed results including comparisons with some already published papers and Section 5 describes real-world scenarios. Section 6 concludes the paper.

2. Design of Camera System

In this paper, we will focus on wireless camera systems. A typical arrangement of such a system can be seen in Figure 1. The first camera covers parking spaces 01, 02, 03, 09, and 10; the second camera covers parking spaces 03, 04, 05, 06, 07, and 08. Information on the occupancy of parking spaces is transferred to the server, which then evaluates the data and transmits it to the end user. In this case, one parking space (03) is visible from both cameras, so there should be an algorithm to solve this overlap. The server also has a database to keep all messages for further analysis.

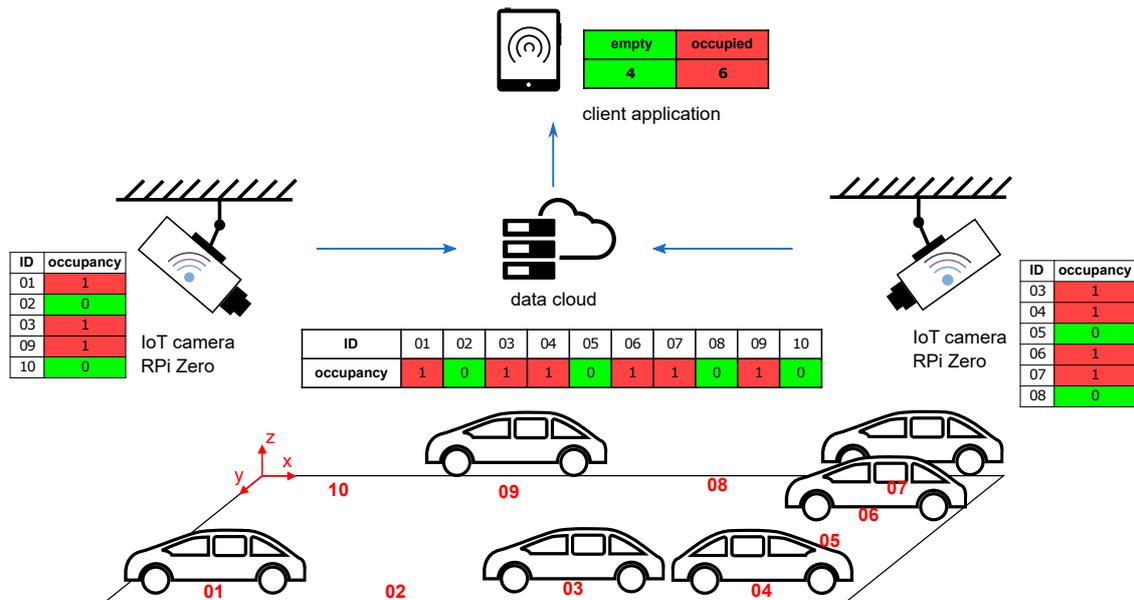


Figure 1. Multicamera parking slot management.

It is sufficient to describe the occupancy of the single parking space with only 1 bit (0 = vacant/1 = occupied). The transfer of such information (including some error-correcting code) to the central server can be conveniently implemented via any of IoT wireless networks [20]. For example, the LPWAN (Low-Power Wide-Area Network) Sigfox platform allows users to transmit 96-bit messages 140 times a day, which approximately corresponds to an update interval of 10 min when sending regular messages. Since irregular intervals could better capture the peak situation, the time between messages can be adaptively modified according to the significance of changes in occupancy. Information about the occupancy of parking slots including navigation to the vacant space could be distributed via the Internet, for example using Vehicle-to-vehicle communication [21].

For wireless cameras, it is also necessary to solve their power supply. Cameras usually consume more energy than simpler sensors, and the design is therefore more complicated. The basic trick to reduce power consumption is the use of sleeping mode with a periodic wake-up. It is not necessary to update the parking capacity at a high sampling frequency, as slower changes are expected. When in sleeping mode, the power consumption of single unit can be minimized to the order of mW, thus significantly prolonging the battery life. Another way to optimize power consumption can be solar panels, which can achieve complete energy self-sufficiency of the camera [22].

An essential part of the design of camera systems is the consideration of the relative position of the cameras and the sensed objects. It is desirable to find the most efficient solution to maximize the coverage area with the minimum number of cameras. However, the design must take into account further consequences associated with placing the camera in a certain position [23]. The illustration image (Figure 2) indicates three different camera positions relative to the plane of the car park. The size of the area and the number of parking spaces to be covered by the camera depends on its location and its field of view. The large FOV usually associated with wide-angle lenses, intruding unwanted geometrical distortions. Camera A can easily cover all the parking spaces but at the cost of overlaps, which make occupancy harder to distinguish. Camera C can capture parking spaces without overlapping, but all of them come into view only when the camera is high above the ground or when using a large FOV lens. There may be overlaps in camera B but not as distinctive as in camera A. Option B seems to be the most advantageous as it is a compromise between the FOV and the overlaps. Moreover, the cameras can be placed on public lighting pillars or nearby buildings.

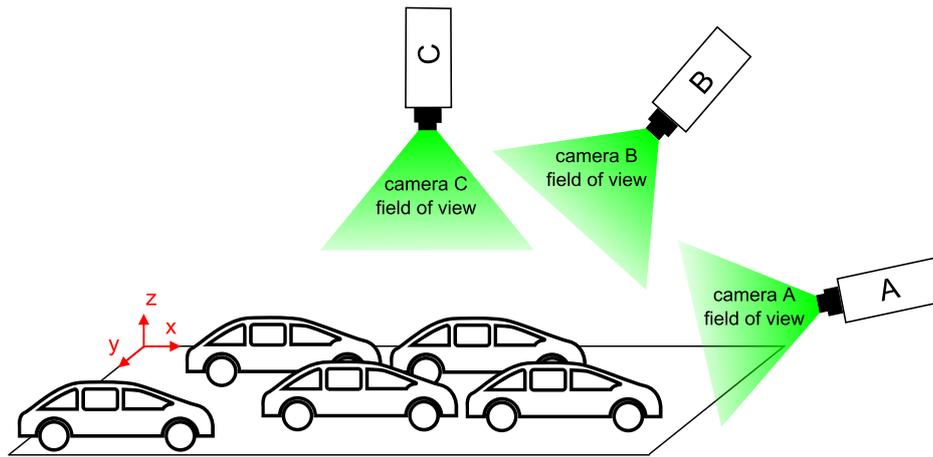


Figure 2. Possible camera locations.

3. Implementation

Prototypes are based on Raspberry Pi Zero with camera module connected to the CSI-2 interface. Camera module uses CMOS sensors OmniVision OV5647 with a spatial resolution of 2592×1944 pixels and lens with focal length of 3.6 mm (FOV is 54×41 degrees). As a power supply we used two Li-Io batteries of 3.7 V ($2 \times 3.7 \text{ V} \times 1800 \text{ mAh}$) and DC-DC step-up converter 5 V/2 A.

The module is equipped with Wi-Fi adapter 802.11 b/g/n and LPWAN SigFox Node UART Modem working at 868 Mhz. While the Wi-Fi interface serves to configure the module, the SigFox modem serves to transmit information about the availability of parking spaces to the central server. Naturally, choice of this IoT communication platform limits the number of parking space that can be managed to approximately 90 (assuming that 6 bits will be used to identify camera module), which is enough for experimental purposes. All image processing tasks are done on Raspberry Pi, extracted information of vacant/occupied parking spaces is then transmitted through a central server to a client application.

3.1. Training Set and Descriptor

Training data are images displaying occupied and free parking spaces. Images were obtained with a camera mounted on a telescopic rod from a height of about 4 m. Overall, about 1000 images were collected in different daytimes and weather conditions. 1120 positive samples and 1378 negative were extracted from these data. After preliminary experiments we considered HOG descriptor as a most suitable and robust solution for our dataset. Test showed that HOG descriptor is highly invariant to light conditions. Extraction of feature vectors using the HOG descriptor requires input data of constant size, so the samples were resampled to the same size, aspect ratio 3:2.

According to our experiments, the standard HOG feature descriptor is not able to adequately describe differently rotated vehicles and produces a high number of false positives. Our findings are supported by work of Fusek et al., [17], where authors report less than 44% accuracy of the HOG based detector. Thus, reasonably use the HOG descriptor requires describing the relative position of the car and the camera. For the purpose, we divided the positive samples into several categories according to the orientation of the car in the picture. The orientation is described by a single value corresponding to the angle between the wheels of the vehicle and the horizontal line. Figure 3a displays example of the training sample with manually labeled wheels (points A and B).

Although this model is very simple, it allows us to very easily distinguish the samples according to the orientation. The angle of the vehicle rotation can yield values from -180° – 180° , but dimensionality can be reduced by adoption of the following simplification:

- it is possible to park a vehicle in two ways (for example -90° and 90° , or -135° and 45°), therefore, the samples corresponding to these angles can be merged;
- a typical vehicle is horizontally symmetric, so it is possible to group corresponding samples (for example 135° and 45°).

Figure 3b demonstrates the similarity of differently rotated vehicles: quadrants A and B are horizontally symmetric, quadrants A and C represents opposite direction of parking and quadrants A and D show the opposite orientation of the cars in the parking space relative to the horizontal symmetry. Based on this consideration, we have merged positive samples into a single quadrant. After experiments, we resampled samples to the size of 72×48 pixels (i.e., feature vector has the length of 1440). This size of the sample was chosen as an optimal size according to the typical size of the car in the field of view of the camera. We divided samples into ten groups, representing vehicle rotations 0° – 90° with the step of 10° . For examples see Figure 4.

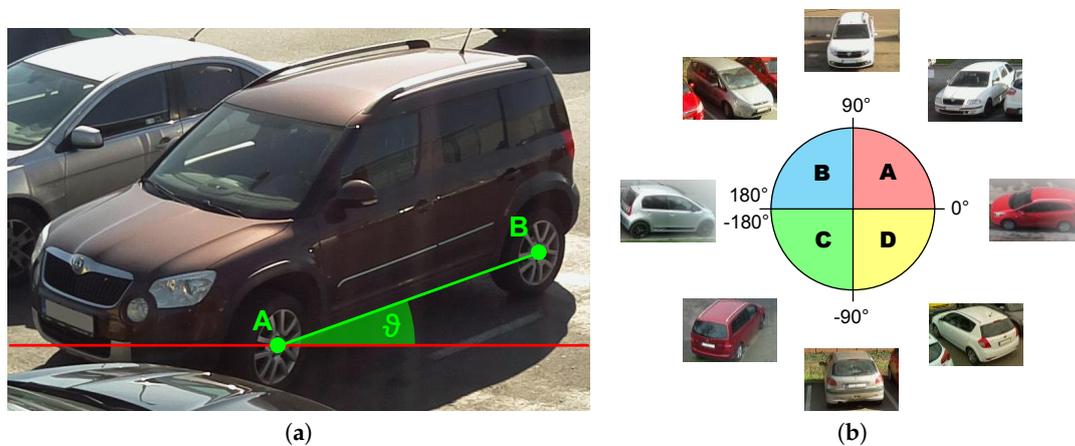


Figure 3. Orientation of the vehicle in the sample. (a) Example of the training sample with manually labeled wheels. (b) Demonstration of the similarity of differently rotated vehicles.



Figure 4. Groups of training samples.

3.2. Learning Algorithm

Since our target platform is Raspberry Pi Zero or possibly any other similar low-power platform, the main focus of the implementation is on efficiency and low complexity of used algorithms (i.e., lower usage of computer resources and thus lower power consumption). For this reason, we performed a comparison of popular learning algorithms performance: SVM with linear kernel, Logistic Regression (LR) and Random Forrest (RF). Tests were performed on our dataset described in Section 4.2, accuracy ACC (ratio between a number of correct classification and a total number of samples), area under ROC (Receiver operating characteristic) and also the average time needed to the classification of one parking space have been investigated. Results are summarized in the Table 1. The best learning algorithm regarding speed is LR, but this algorithm exhibits bad performance with our dataset. As the best trade-off between computational difficulty and accuracy, we selected SVM. Due to the relatively large

size of the feature vector and a small number of samples, the training set data is linearly separable. Thus, the linear kernel is a natural choice.

Table 1. Comparison of learning algorithms.

	ACC	ROC	Time
RF	0.910	0.946	110 ms
SVM	0.931	0.955	100 ms
LR	0.638	0.782	30 ms

3.3. Classification of the Parking Space

Unlike the training phase, it is not possible to specify the position of the vehicle during classification, since its exact position and size are not known. Thus, the algorithm works with a previously known approximate position of a parking space, which is naturally larger than the vehicle itself. The classification is then carried out using a sliding window in variously resampled copies of the section of the parking area (see Figure 5). From each copy, a sliding window of size 72×48 produces several segments for which the HOG feature vectors are calculated. SVM classification model then predicts the class. If the model classifies the given feature vector as a vehicle, the prediction is equal to one. Otherwise, the prediction value is zero. Summing up the results for all segments produced by sliding window, the algorithm calculates a classification score that is thresholded to obtain the final classification result. The value of the threshold is determined from the statistical analysis, for details see Section 4. The classification of a single parking space is relatively computationally inexpensive; with Raspberry Pi Zero it takes about 100 ms. Thus, the system can classify about ten parking spaces per second.

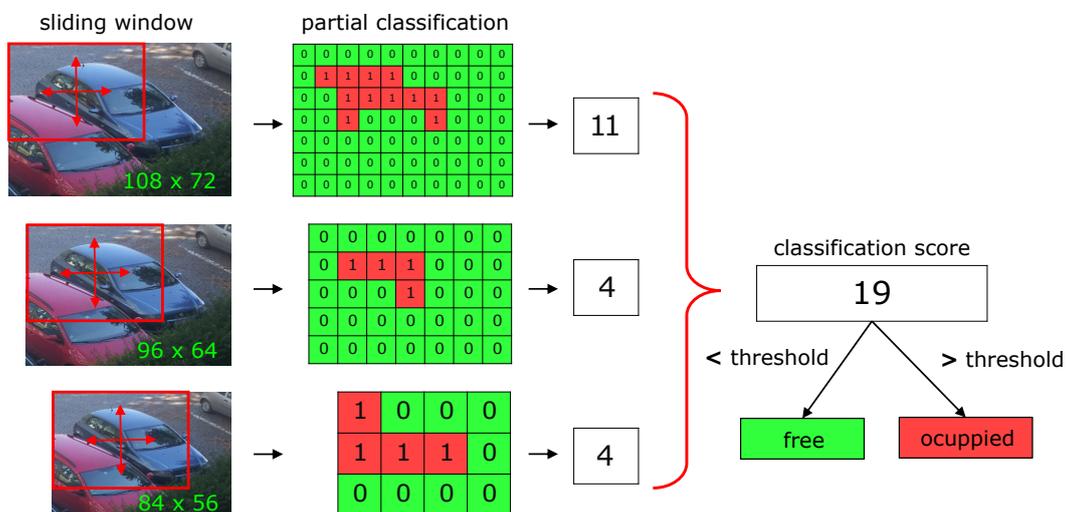


Figure 5. Classification of the single parking space.

If the system uses more cameras with an overlapping field of view, the parking space in the overlay is classified by all cameras. The decisive classification score is the value with the greatest difference from the chosen threshold.

3.4. Power Consumption

The module works with the voltage of 5V. We measured the current consumption of the module during typical actions: idle, dense scanning (classification of single parking space), Wi-Fi connected

and SigFox data transmission. For details about current consumption see Table 2. Note, that the Raspberry Pi Camera Modules require 250 mA to an operation, and the SigFox module has current consumption 65 mA during transmit and 15 mA during receive operation, typical sleep mode current is 2 μ A. The timing of the actions is determined by the frequency of data transmissions and a number of parking spaces. Assuming that the module is working during the daytime, with the limitation of the SigFox network it is possible to send one message per five minutes. A Wi-Fi interface is switched on during normal operation, and it may be switched on with a command sent over the SigFox network. Peak current occurs when the camera is in action. This is typically 10 s during the above-mentioned five minutes interval (to classify up to 100 parking spaces). The average power consumption of the single module is about 1.5 W during daytime (6 am to 8 pm during the summer).

Table 2. Current consumption of the module.

Activity	Current
idle	80 mA
dense scanning	830 mA
Wi-Fi connected	250 mA
data transmission	145 mA

4. Validation of the Results

For the purpose of testing the entire system, validation of the results was carried out on the databases PKLot and two smaller databases collected by the authors of the article—PKSlots described in Section 4.2 and FELSlot described in Section 4.3.

4.1. PKLot

PKLot database was created from 12,417 images of size 1280 \times 720 pixels captured with low-cost HD camera on two different parking lots in sunny, cloudy and rainy days. The first parking lot was captured in two different view angles, so there are three different datasets (see Figure 6, images taken from [10]). The whole dataset contains 695,899 samples of parking spaces, but since some of the samples have missing info about occupancy, we used only 667,076 of them (348,125 negative and 318,951 positive samples).

The format of data definitions used in the PKLot database is different from the data format we have designed for our system. The PKLot defines the parking space using a contour close to the assumed position of the car and automatically calculates the parameters of the rotated rectangle from contour coordinates, see Figure 7. However, the parameters of the rectangle were not satisfactory for the separation of the parking spaces. The main problem was that the angle parameter often does not correspond to the actual orientation of the parking space. We solved this problem by using the original contours, from which we constructed a straight line (green dashed line in Figure 7) whose angle we considered to be the orientation of the place.

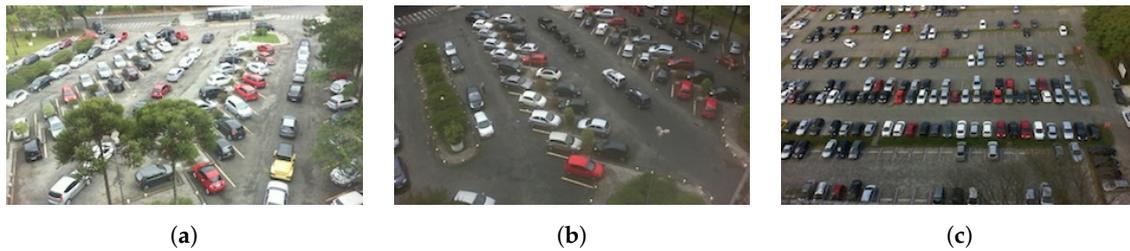


Figure 6. Parking lots used in PKLot database (a) Federal University of Parana; (a) Federal University of Parana, captured from 4th floor of the building; (b) Federal University of Parana, captured from 5th floor of the building; (c) Pontifical Catholic University of Parana, captured from 10th floor of the building.

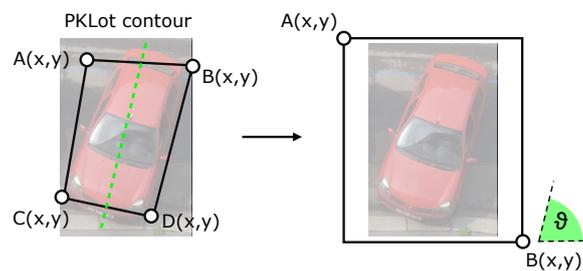


Figure 7. Transformation of PKLot.

Figure 8a shows the dependence of the false-rate of classification on the threshold setting, with the distinction whether a vacant or occupied place is recognized. Figure 8b displays accuracy ACC, the maximum value of ACC is 91.8% for the threshold equal to 15. Maximum accuracy, however, may not be the most important criterion when setting the threshold for a classification score. Efforts can be made to reduce false negatives (FN, i.e., algorithm did not recognize occupied place) even at the cost of precision, as lower FN is suited when submitting information to the user. For example, let have the threshold set to value 8, the accuracy of ACC reduces to 90.7%, but the number of situations where the system denotes the occupied parking space as vacant decreases to 6%. Table 3 shows confusion matrix for PKLot dataset and two values of the threshold. For the sake of completeness, note, that true positive (TP) means that algorithm correctly recognized car occupying parking space, true negative (TN) means that algorithm correctly recognized vacant space and false positive (FP) means that algorithm did not recognize vacant place. Since authors of the PKLot database labeled samples also in terms of weather, it is possible to study the impact of weather conditions to algorithm performance. Table 4 shows the confusion matrix corresponding to PKLot samples obtained in cloudy, rainy or sunny weather.

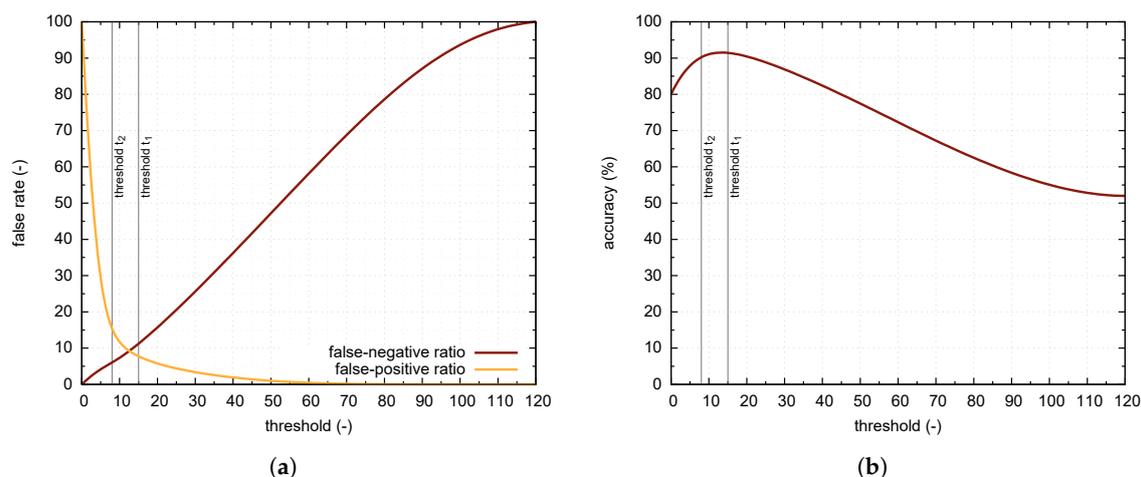


Figure 8. Statistical evaluation of algorithm performance (a) Threshold influence to the classification false rate; (b) Threshold influence to accuracy (ACC) of the classification.

Table 3. PKLot dataset: confusion matrix for two different thresholds.

	$t_1 = 15$				$t_2 = 8$			
TP	285,542	FN	33,409	TP	299,812	FN	19,139	
	89.5%		10.5%		94.0%		6.0%	
FP	21,562	TN	326,563	FP	42,566	TN	305,559	
	6.2%		93.8%		12.2%		87.7%	
	ACC = 91.8%				ACC = 90.7%			

Table 4. PKLot dataset: confusion matrix for different weather conditions, threshold $t_2 = 15$.

	Cloudy				Rainy				Sunny			
TP	77,663	FN	8426	TP	55,803	FN	3177	TP	152,076	FN	21,806	
	90.2%		9.8%		94.6%		5.4%		87.5%		12.5%	
FP	9358	TN	130,677	FP	2110	TN	31,478	FP	10,094	TN	164,408	
	6.7%		93.3%		6.3%		93.7%		5.8%		94.2%	
	ACC = 92.1%				ACC = 94.3%				ACC = 90.8%			

4.2. PKSlots

PKSlots is a custom dataset collected by the authors of this paper. The dataset consists of 774 samples of parking slots that were obtained from a total of 142 very different shots (different angles, different weather conditions, etc.) taken by the use of different cameras under different conditions. See Figure 9 for examples of both positive and negative samples.

Using ROC analyses we determined the optimal threshold ($t = 8$). Table 5 displays a confusion matrix for this value of threshold. Of the 774 samples, only 53 (25 cars and 28 vacancies) were incorrectly classified, and accuracy reached a value of 93.2%. Due to the significant variability of the images, this is a very satisfactory result, and it is also good that the FN value is lower than FP.



Figure 9. PKSlots dataset (a) Positive samples; (b) Negative samples.

Table 5. Confusion matrix for PKSlots dataset, $t = 8$.

TP	423 94.4%	FN	25 5.6%
FP	28 8.6%	PN	298 91.4%
ACC = 93.2%			

4.3. FELSlot

Testing dataset FELSlot is another dataset collected by authors of this paper. It consists of 5.549 images of the parking slot at the Faculty of Electrical Engineering of Czech Technical University in Prague. Cameras (IoT RaspberryPi modules described in this paper) were placed in four different positions, 71.084 samples were obtained in total during two consecutive days. For examples see Figure 10.

Table 6 shows confusion matrix for the threshold $t = 15$. Accuracy is then $ACC = 92.2\%$.

Table 6. Confusion matrix for FELSlot dataset, $t = 15$.

TP	27.639 85.2%	FN	4784 14.8%
FP	773 2.0%	PN	37,888 98.0%

4.4. Comparison with Other Algorithms

To evaluate the performance of our method, we tried to compare the results we achieved with other known algorithms. Because the authors of papers typically present the results in different forms, the following text is a simple comparison without a deeper analysis. Baroffio et al., tested algorithm based on Hue histogram and linear SVM [16] and published accuracy reached on PKLot database (UFPR04—96%, UFPR05—93%, PUCPR—87%). Amato et al., [18,19] refer accuracy of their CNN based method up to 99%, but their results strongly depend on the choice of the learning dataset. Almeida et al., [9], authors of PKLot database, used SVMs trained on histograms of textural features. Their paper refers to following accuracy: UFPR04—99%, UFPR05—84%, PUCPR—84%. Fusek et al., [17] refer accuracy of HOG based detector 44% and accuracy of the detector based on AdaBoost 94%. Our method, trained on our custom dataset, reached an accuracy of 96% on UFPR04, 83% on UFPR05 and 94% on PUCPR.

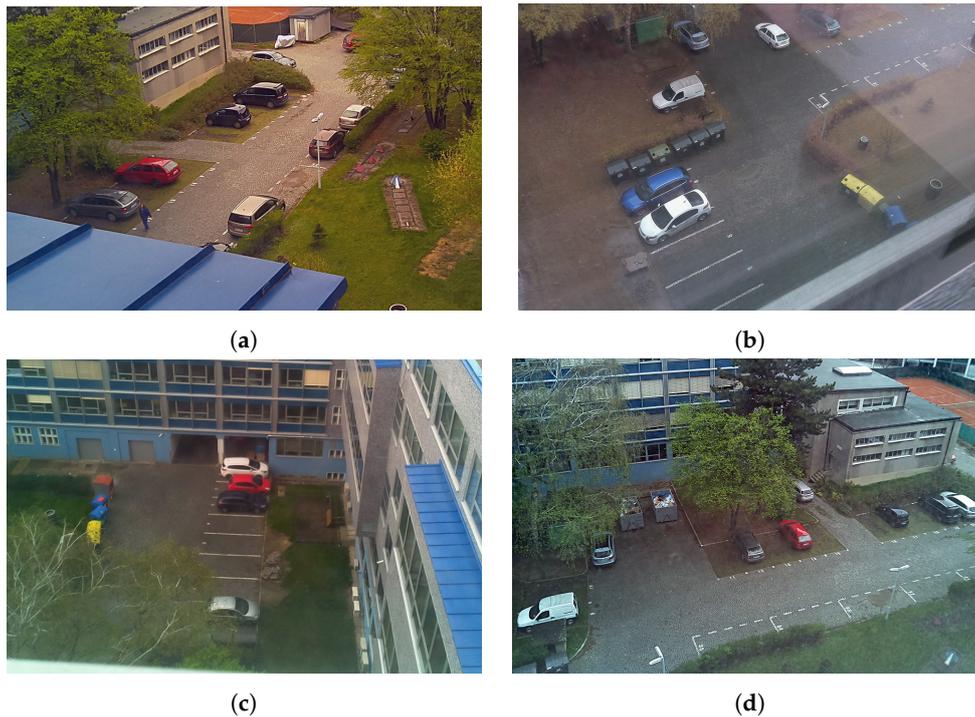


Figure 10. FELSlot dataset (a) Position A1–14 parking slots, 1615 images; (b) Position A1–14 parking slots, 695 images; (c) Position B1–8 parking slots, 1635 images; (d) Position B2–16 parking slots, 1604 images.

5. Working Scenarios

The system described in this paper is able to work in two basic modes: (a) mode with predefined parking places and (b) mode with parking places determined automatically.

5.1. Predefined Parking Spaces

In this mode, an operator has to select and label expected parking spaces. For this purpose, the operator can connect the module using Wi-Fi interface and use GUI (graphical user interface) to define parking places—see Figure 11. In case of cameras with overlapping fields of view, an operator has to label parking spaces in overlapped areas; the occupancy is then evaluated as described in Section 3.3. Evaluation of parking spaces in overlapped areas is done in the central server. In this scenario, users receive information about the occupancy of a particular parking space.

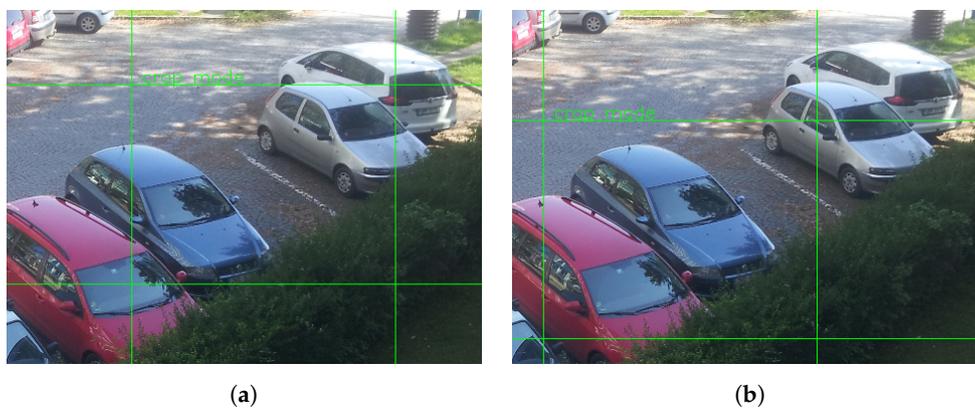


Figure 11. Manual definition of two neighboring parking spaces. (a) Example of vacant parking space. (b) Example of occupied parking space.

5.2. Determination of Parking Space Position

The previous working scenario, described in Section 5.1, depends on the input of the operator. In the more challenging scenarios, like on-street parking, a position of parking spaces may change in time. Thus, we propose to use a simple procedure to determine and update a position of parking places. For this purpose, the procedure employs the single-pass version of classification algorithm described in Section 3.3, for details see Algorithm 1.

Algorithm 1: Algorithm of determination of the vehicles position

Data: single frame

Result: positions of vehicles within the FOV

for blocks in the frame **do**

- | calculate HOG features;
- | classify the block;

end

for blocks classified as a vehicle **do**

- | **if** block is not 4-connected with blocks classified as a part of vehicle **then**
- | | the block is determined as a boundary of a blob;
- | **else**
- | | block is determined as a part of a blob;
- | **end**

end

identify blobs;

for blobs **do**

- | find the center of the mass;
- | find the largest possible rectangle with center point closest as possible to the center of the blob and corners on the boundary of the blob;
- | consider the center point of the given rectangle as a center of the vehicle;

end

Figure 12a display an example, how the frame is processed. Set of blocks classified as a vehicle take a form of a blob (red color). Blocks belonging to the blob with less than 4-connectivity with other blocks in the blob are considered as a blob boundary (yellow color). Determination of the blob boundaries allows us to divide the blob into a few smaller ones that fit the position of the vehicles. The algorithm then estimates centers of mass of a new blob (or blobs, if there are more than one car—see Figure 12b). Rectangles estimates around the blobs then define a position of parking spaces.

This procedure can be run for a couple of days after installation of the module to train positions of the available parking spaces and on the regular basis to update them. When about ten cars are present in the FOV, the procedure takes about 90 s; current consumption is the same as for classification.

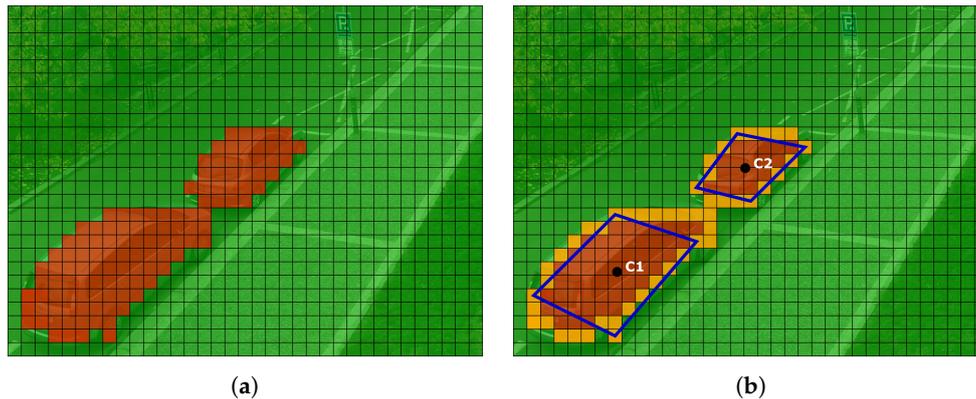


Figure 12. Determination of parking spaces position. (a) Blocks classified as a vehicle. (b) Boundaries of the blobs and determination of the vehicle position.

6. Conclusions

A fast and reliable algorithm designed to run on low-power consumption embedded systems is used to classify images of parking space to recognize if they are occupied or vacant. The extracted information can be used in any application aiming to help drivers to find free parking space. In order to increase the accuracy of the detector based on HOG descriptors and SVM, we have provided a descriptor with additional information about rotation of the vehicle. We also employed dense scanning in three different scales to improve detector performance.

Performance of our algorithm was determined on the publicly available dataset for parking occupancy detection PKLot. We also collected two new datasets based on the different approach (more views and weather conditions) and then evaluated them. Average time to classify one parking space is about 100 ms on Raspberry Pi Zero. For the real-world applications, we prepared two different scenarios that are able to manage off-street as well as on-street parking, with a self learning procedure. Clearly, there is room for further improvement, for example, the design of predicting algorithms. Tests in the real environment showed that the system works well for a wide range of light conditions. We would also like to solve possible security issues—our plan is to implement our system with the use of IBM Bluemix, which is employing secured MQTT (Message Queue Telemetry Transport) protocol.

Acknowledgments: This work was supported by the Grant No. 17-05840S Multicriteria Optimization of Shift-Variant Imaging System Models of the Czech Science Foundation and by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/099/OHK3/1T/13 Real time classification of optical transients.

Author Contributions: Stanislav Vitek and Petr Melničuk conceived and designed the experiments; Petr Melničuk performed the experiments; Stanislav Vitek and Petr Melničuk analyzed the data; Stanislav Vitek wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shoup, D.C. Cruising for parking. *Transp. Policy* **2006**, *13*, 479–486.
2. Van Ommeren, J.N.; Wentink, D.; Rietveld, P. Empirical evidence on cruising for parking. *Transp. Res. Part A* **2012**, *46*, 123–130.
3. Ji, Z.; Ganchev, I.; O'Droma, M.; Zhao, L.; Zhang, X. A cloud-based car parking middleware for IoT-based smart cities: Design and implementation. *Sensors* **2014**, *14*, 22372–22393.
4. Hasegawa, T.; Nohsoh, K.; Ozawa, S. Counting cars by tracking moving objects in the outdoor parking lot. In Proceedings of the IEEE Vehicle Navigation and Information Systems Conference, Yokohama, Japan, 31 August–2 September 1994; pp. 63–68.
5. Del Postigo, C.G.; Torres, J.; Menéndez, J.M. Vacant parking area estimation through background subtraction and transience map analysis. *IET Intell. Transp. Syst.* **2015**, *9*, 835–841.

6. Tschentscher, M.; Neuhausen, M. Video-based parking space detection. In Proceedings of the Forum Bauinformatik, Bochum, Germany, 26–26 September 2012; pp. 159–166.
7. Hsu, C.; Tanaka, T.; Sugie, N.; Ueda, K. Locating vehicles in a parking lot by image processing. In Proceedings of the IAPR workshop on machine vision applications, Nara, Japan, 11–13 December 2002; pp. 475–478.
8. Delibaltov, D.; Wu, W.; Loce, R.P.; Bernal, E.A. Parking lot occupancy determination from lamp-post camera images. In Proceedings of the 2013 16th International Conference on IEEE Intelligent Transportation Systems-(ITSC), The Hague, The Netherlands, 6–9 October 2013; pp. 2387–2392.
9. De Almeida, P.R.; Oliveira, L.S.; Britto, A.S.; Silva, E.J.; Koerich, A.L. PKLot—A robust dataset for parking lot classification. *Expert Syst. Appl.* **2015**, *42*, 4937–4949.
10. Parking Lot Database. Available online: <http://web.inf.ufpr.br/vri/news/parking-lot-database> (accessed on 25 December 2017).
11. Kłosowski, M.; Wójcikowski, M.; Czyżewski, A. Vision-based parking lot occupancy evaluation system using 2D separable discrete wavelet transform. *Bull. Pol. Acad. Sci. Tech. Sci.* **2015**, *63*, 569–573.
12. Bhaskar, H.; Werghi, N.; Al-Mansoori, S. Rectangular Empty Parking Space Detection using SIFT based Classification. In Proceedings of the International Conference on Computer Vision Theory and Application VISAPP, Algarve, Portugal, 5–7 March 2011; pp. 214–220.
13. Bulan, O.; Loce, R.P.; Wu, W.; Wang, Y.; Bernal, E.A.; Fan, Z. Video-based real-time on-street parking occupancy detection system. *J. Electron. Imaging* **2013**, *22*, 041109.
14. Rybski, P.E.; Huber, D.; Morris, D.D.; Hoffman, R. Visual classification of coarse vehicle orientation using histogram of oriented gradients features. In Proceedings of the 2010 IEEE Intelligent Vehicles Symposium (IV), San Diego, CA, USA, 21–24 June 2010; pp. 921–928.
15. Ahrnbom, M.; Astrom, K.; Nilsson, M. Fast classification of empty and occupied parking spaces using integral channel features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 9–15.
16. Baroffio, L.; Bondi, L.; Cesana, M.; Redondi, A.E.; Tagliasacchi, M. A visual sensor network for parking lot occupancy detection in smart cities. In Proceedings of the 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), Milan, Italy, 14–16 December 2015; pp. 745–750.
17. Fusek, R.; Mozdřeň, K.; Šurkala, M.; Sojka, E. AdaBoost for parking lot occupation detection. In *Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013*; Springer: Berlin, Germany, 2013; pp. 681–690.
18. Amato, G.; Carrara, F.; Falchi, F.; Gennaro, C.; Meghini, C.; Vairo, C. Deep learning for decentralized parking lot occupancy detection. *Expert Syst. Appl.* **2017**, *72*, 327–334.
19. Amato, G.; Carrara, F.; Falchi, F.; Gennaro, C.; Vairo, C. Car parking occupancy detection using smart camera networks and deep learning. In Proceedings of the 2016 IEEE Symposium on Computers and Communication (ISCC), Messina, Italy, 27–30 June 2016; pp. 1212–1217.
20. Abas, K.; Porto, C.; Obraczka, K. Wireless smart camera networks for the surveillance of public spaces. *Computer* **2014**, *47*, 37–44.
21. Torrent-Moreno, M.; Mittag, J.; Santi, P.; Hartenstein, H. Vehicle-to-vehicle communication: Fair transmit power control for safety-critical information. *IEEE Trans. Veh. Technol.* **2009**, *58*, 3684–3703.
22. Abbas, Z.; Yoon, W. A survey on energy conserving mechanisms for the internet of things: Wireless networking aspects. *Sensors* **2015**, *15*, 24818–24847.
23. Aghajan, H.; Cavallaro, A. *Multi-Camera Networks: Principles and Applications*; Academic Press: Cambridge, MA, USA, 2009.



