Data Container for Autonomous Cars

Janusz Bobulski[1], Mariusz Kubanek[1], Joanna Kulawik[1] and Sabina Szymoniak[1]*

[1] *Czestochowa University of Technology, Computer Science Department,*

73 Dabrowskiego Str., Czestochowa, Poland

ABSTRACT

. In the future, autonomous cars will become commonplace. However, they re-quire the development of complex software. Therefore, our research focused on developing a multimedia container structure containing three types of images: RGB, Lidar and infrared, calibrated adequately against each other. An additional goal is to establish libraries of programs for creating, saving and saving these types of files. It will also be necessary to develop a method for synchronizing data from Lidar, RGB and infrared cameras. This type of file could be used in autonomous vehicles, facilitating data processing by the intelligent autonomous vehicle management system and providing the driver with valuable information.

Keywords: autonomous cars, multimedia container, image processing

INTRODUCTION

Autonomous cars are increasingly breaking into our consciousness. No one seems to have any doubts that self-driving cars are the future of motoring. Manufacturers promise that moving the first of them to showrooms is the prospect of the next few years. Many experts believe that creating a network of communicating autonomous cars will be able to eliminate accidents. However, it is necessary to develop effective detection methods of objects around the moving vehicle to make this possible. In bad weather conditions, this task is complicated based on the RGB image. Therefore, in such situations, you should be supported by information from other sources, such as Lidar or infrared cameras. The problem is the different data formats that individual types of devices return. In addition to these differences, there is a problem with synchronizing these data and the formatting. The project's goal is to develop a file structure that could be containing different types of data. This type of file is called the multimedia container (Jangblad, 2018).

A multimedia container is a container that contains many data streams, which al-lows you to store complete multimedia material in one file. The data streams in such a container should be indicated streams of images, films, sounds, subtitles, and additional information, i.e. metadata. This type of file could be used in autonomous vehicles, facilitating data processing by the intelligent autonomous vehicle management system (Wahib and Abdessalem, 2017; Ze-Niam et al., 2014; Troncy, 2011).

STATE OF ART

There is a wide range of multimedia containers. They differ in their use and the functions they provide. These containers are classified according to the type of data for which they have been specialised. So you can distinguish containers that store sound, audio-video materials, and containers for images. The WAV (wave form audio for-mat) is the most important among the containers that store sound. It allows you to store uncompressed audio files. In addition to the audio stream, it contains information describing the stream (codec, sampling frequency, number of channels).

The most popular containers store audio-visual materials are AVI, MOV, MP4, MKV, and ogg. AVI (Audio Video Interleave) stores data in two or three parts. The first is the header, which contains data about the video file (screen resolution, number of streams, number of frames displayed in one second). The second part of the AVI format is data. In turn, the third, optional, part contains information about the location of the other elements in the file. AVI uses loss compression.

The MOV format allows you to save audio, video and text tracks in one file. MOV also uses loss video compression. The MP4 container primarily stores audio-visual data. It also allows you to save other types of data (still images, signatures). This format is commonly used to store digital streams on websites. This container assumes an object structure. Contains objects called boxes. Each such part consists of three elements: size, type and data.

The MKV container, i.e. the popular Matroska, has been designed to store images, sound and other information such as subtitles, chapters, menus. The main task of this container is to strive for data flexibility to enable the joining of different data streams.

The ogg container also allows you to store various multimedia data (audio, video, graphics, subtitles). Lossy sound compression is done here using the Vorbis codec. It is responsible for filtering those frequency ranges that are not audible to the human ear.

The most crucial image container is the TIFF (Tagged Image File Format) container, which allows you to save raster graphics. This container was developed in 1986 by Aldus Corporation for DTP and postscript printing applications and then extended by Microsoft and Hewlett-Packard. The TIFF container is widely used, including photo processing, advanced text composition, text scanning and recognition support, faxing and medical imaging.

The TIFF file starts with a header. Thanks to so-called tags in the file header, this format has become flexible because several images and text data can be placed in one file. Tags describe the entire file in TIFF format, including dimensions, compression, description, colour depth, an indication of actual image data (IFD). There may be several IFD segments in the file. TIFF uses lossless compression in the form of LZW (Lemple-Zif-Welch) or CCITT Group 4 algorithms. This container also allows you to save images with a more extensive colour range than other graphic formats, determined by the number of bits per channel. A typical application of the TIFF container is to describe data from scanners, from devices processing an image from an analogue source into a digital form (frame grabbers) and image retouching programs. TIFF allows you to describe two-level image data and grayscale, palette, and full-colour in several colour spaces (EAIFF, 2020; Ho and Shujun, 2016).

The TIFF file has an 8-byte header that indicates image data, or IFD (Image File Directory). Its first two bytes contain information about the order of the bytes that were used in the file. If the value 4949h is entered there, it means that the order little endian was used. If the value 4D4Dh is entered there, it means that the big-endian order has been used. The following two bytes of the header contain the value 42, which is the file signature that defines that the file is a TIFF file (Bobulski, 2016).

The last four bytes of the header contain the offset, i.e. the offset in bytes, to the first IFD. The IFD entry contains image information and a pointer to the actual image data. Each TIFF file must contain at least one IFD having at least one access. The first two bytes of IFD contain the number of entries in the image file directory. A sequence of 12-byte directory entries follows them. The last four bytes contain offset to the next IFD, if it exists. Otherwise, this part of IFD is filled with zeroes. It is impossible to save data from the Lidar next to the RGB and infrared data in the above multimedia containers, so we decided to address this problem and develop a new standard.

RESEARCHES

The research comprises five main parts. The first stage will be obtaining data showing the image in different formats of the same scene. For this purpose, a stand will be made where the RGB camera, IR camera and Lidar will be mounted in the same direction. In front of them, there will be a so-called scene on which objects simulating vehicle surroundings and obstacles will appear. Then, for various scenarios, data will be downloaded in three RGB, IR and Lidar formats. The goal is to collect 100 sample datasets.

The second stage of the project is the Normalisation of data from Lidar. Each Lidar model has an individual, well-defined range of testable space expressed in degrees. Laser beams sent at different angles from a central point towards the analysed scene return coordinates of points for found objects in the space bent along the arc. Distances are expressed in meters. Images recorded with a traditional camera flatten the Cartesian space of the recorded scene to 2D dimension linearly, the unit of measurement is pixels Bobulski, 2018; Ashok, 2009). It is, therefore, necessary to develop an algorithm whose purpose will be to normalise data from Lidar to enable the data obtained using this technique to be referenced in the analysis of traditional digital images and then to implement it.

Developing the structure of the data container will be a crucial stage of research. A file structure will contain normalised data from Lidar, an RGB bitmap, and an IR image. In addition, the file will contain so-called metadata describing the contents of the container (Bobulski, 2017).

The next stage will be the development of a program library to save the developed data structure to a file. It will be necessary to create appropriate algorithms and then to implement them.

The last stage will be the development of a program library enabling access to data contained in the developed container and saved in a file. It will be necessary to create appropriate algorithms and then implement them, which will provide data extraction and visualisation tools.

Lidar

Light scattering is as much used as absorption. Scattering of light, namely electro-magnetic waves, is the phenomenon of the interaction of light with matter. The direction of light propagation changes, except for the phenomena described by the reflection and refraction of light. It causes the illusion of the so-called glow centre. Scattering is a phenomenon based on which it is possible to determine the properties of the substrate on which light was directed, e.g. from a laser.

A laser is a commonly known device that emits electromagnetic radiation in the visible, ultraviolet or infrared range, using the phenomenon of forced emission. The laser radiation is coherent, usually polarised, and has a beam with a minimal divergence and a small line width
Lidar technology is an active, remote sensor system that generates light in a laser measuring the exact distance. Devices based on this method of operation resemble radar, but the fundamental difference in their case is that laser light is used instead of microwaves.
The laser sends very short, measured, but intense light pulses of a specific wavelength defined by the optical system. Along the way, the light is scattered, which is observed with a telescope in the same device. Then it is recorded with a detector - a photodiode or a photomultiplier, as well as CCD and CMOS cameras - in this way, the intensity of the observed, scattered light is examined and finally, data is obtained. Additionally, differences in laser return time and wavelengths create precise, three-dimensional representations and characteristics of the substrate - map-visualisation.

The precise representation offered by the technology allows clear and accurate road structure data and the identification of obstacles to avoid collisions. In addition to the radar above, it can be compared through the sonar prism, which works on the principle of sound waves - it is more widely known from submarines. Both radar and sonar do not offer the same accuracy as Lidar technology because they can only determine the location of an object, and Lidar creates a 3D representation of it. This makes it valuable, among others, in autonomous vehicles.

Lidars are divided into two types depending on the function:

• Airborne Lidar - installed on-board Lidar aircraft, is used to collect data and create 3D models of the observed landscape. Their precision allows for models to take into account the shape of land, seabed and rivers. It is divided into two sub-types: topographic and bathymetric.

• Terrestrial Lidar (TLS) laser scanning - these systems are often stationary as opposed to aerial. However, they can be installed on moving vehicles or tripods. They are used, among others in conventional topography, researching and documenting the community's cultural heritage of a given area, road observation or infrastructure analysis. Therefore, they are divided according to their purpose into stationary and mobile.

Lidar parameters

In our experiments, we used Lidar Benewake model CE 30-D, whose detection angle is 60o horizontally and 4o vertically, while distance detection achieves 90% accuracy in the range of 0.4m to 28m. Its use allows you to register the scene in the form of a point cloud.

In the first stage of our experiments, we calibrated images RGB \ IR with Lidar data. The results of this operation show figure 1.

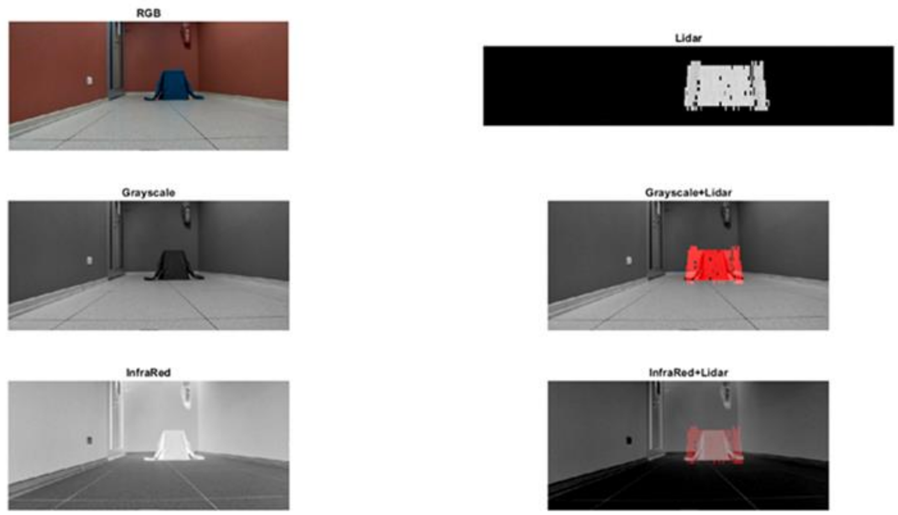

Fig. 1. Results of calibration.

Calibration

The calibration of the images consisted of adjusting them to each other so that the Lidar data was displayed correctly. The easiest way to match the data is to find the middle of both sets and superimpose them.

The resulting images, combining RGB, IR and Lidar data, are saved to one stand-ard file. After calibrating the distance, it is possible to add distances to objects in numerical values. Ultimately, a data stream similar to the movie but containing additional data will be created.

CONCLUSION

As shown by preliminary studies, combining RGB and InfraRed images with Lidar data allows for more straightforward data analysis. Thanks to this application, it will be possible to display the distance to the object in a colour photo. Such information can be beneficial for drivers and systems in autonomous cars.

The integration of various types of data allows for easier management and simultaneous access to them. The use of this type of container reduces data fragmentation on data carriers. In addition, it facilitates data transfer as you only need to upload one file, not several.

ACKNOWLEDGEMENTS

This research was funded by the Polish Minister of Science and Higher Education under the name "Regional Initiative of Excellence" in the years 2019 - 2022 project number 020/RID/2018/19, the amount of financing 12,000,000 PLN.

REFERENCES

Advogato: Microsoft patents ASF media file format, stops reverse engineering, www.advogato.org (last access 2020)

Ashok B., (2009) Multimedia Technologies, McGraw Hill Education.

Bobulski, J. (2016) 2DHMM-based face recognition method, Advances in Intelligent Systems and Computing. Vol. 389. doi:10.1007/978-3-319-23814-2_2.

Bobulski, J. (2017)Multimodal face recognition method with two-dimensional hidden Markov model. Bulletin of the Polish Academy of Sciences: Technical Sciences, 65(1), 121-128. (2017). doi:10.1515/bpasts-2017-0015

Bobulski, J. (2018) Iris identification method using only a section of the pattern. Optica Applicata, 48(1), 149-160. doi:10.5277/oa180114

EA IFF 85 - Standard for Interchange Format Files, http://www.martinreddy.net/gfx/2d/IFF.txt (last access 2020)

Ho, A. T. S., Shujun L. (2016) Handbook of Digital Forensics of Multimedia Data and Devices, Enhanced E-Book. John Wiley & Sons.

Jangblad M. (2018) Object Detection in Infrared, Images using Deep Convolutional Neural Networks, Master of science thesis, Uppsala Universitet

Troncy R. at al. (2011) Multimedia Semantics: Metadata, Analysis and Interaction, Wiley-Blackwell

Wahiba B. Abdessalem K., Dey,N. (2017) Mining Multimedia Documents, Chapman and Hall CRC

Ze-Nian L., Drew M.S., Jiangchuan L. (2014) Fundamentals of Multimedia, Springer