

# REVIEW OF METHODS FOR THE SURVEILLANCE AND ACCESS CONTROL USING THE THERMAL IMAGING SYSTEM

---

**Mate Krišto**

**Mate Krišto**

PhD student at Department of Informatics,  
University of Rijeka

---

**Article info**

Paper category: Review paper

Received: 21.12.2016.

Accepted: 27.12. 2016.

JEL classification: F5, O31

---

## ABSTRACT

*This paper presents methods for human detection for application in the field of national security in the context of state border surveillance. Except in the context of state border security, the presented methods can be applied to monitor other protected object and infrastructure such as ports and airports, power plants, water supply systems, oil pipelines, etc. Presented methods are based on use of thermal imaging systems for the human detection, recognition and identification. In addition to methods for the detection of persons, are presented and methods for face recognition and identification of the person. The use of such systems has special significance in the context of national security in the domain of timely detection of illegal crossing of state border or illegal movement near buildings, which are of special importance for national security such as traffic infrastructure facilities, power plants, military bases, especially in mountain or forests areas. In this context, thermal imaging has significant advantages over the optical camera surveillance systems because thermal imaging is robust to weather conditions and due to such an infrared thermal system can successfully applied in any weather conditions, or the periods of the day. Featured are procedures that has human detection results as well as a brief survey of specific implementation in terms of the use of infrared thermal imagers mounted on autonomous vehicles (AV) and unmanned aerial vehicles (UAV). In addition to the above in this paper are described techniques and methods of face detection and human identification based on thermal image (thermogram).*

### **Keywords:**

national security, negative effects on the national economy, infrared thermal imaging, area surveillance and access control, human detection and identification

---

## 1. INTRODUCTION

Safety and security are perhaps the most important issues in the modern world in almost every aspect of human work and life, but the most significant is in the area of national security in terms of protection of state border and other objects that represent a significant national interest. The importance of preserving security in the domain of national security, of particular importance in preventing and cutting of terrorist activities, as well as other illegal activities such as smuggling of narcotics across the border and human trafficking. In this regard, special emphasis is placed on the timely detection of the persons who illegally cross the border with the objective of carry out terrorist attacks or committing other illegal activities as a direct attack on national security, and the freedom and security of citizens and private or state property, which in the end has a direct negative impact on the entire economic system of each state which may be manifested in the form of a reduced number of tourist arrivals, increasing business risk, etc. Therefore, the continuously evolving technology, techniques and methods for the implementation of security measures especially when it comes to the security of the state level, or the level of each economic subject. Each country seeks the application of different security measures and techniques to increase the level of security to all citizens, and limits and protect areas and facilities that are of significant importance for the state and its citizens. In this context, developing methods and techniques (and ways of their implementation) that enable easy and timely detection of potential threats, especially in context of illegal border crossing, as well as access to certain facilities that are of national importance. Each country, as well as enterprises seek to maintain the cooperation and coordination of security policies and practical actions in the fight against the so-called, asymmetric security challenges, risks and threats - terrorism, organized crime, illegal migration and human trafficking, especially women and children.

Given the above, this paper presents methods of using thermal imaging systems, as well as computer vision, in border control, as well as other protected areas. The paper is structured so that in second part are described the fundamentals of infrared thermal imaging, in third an overview of research in the domain of application of thermal imaging in surveillance and human detection, and fourth part describes methods of face recognition and identification of persons based on thermogram.

## 2. FUNDAMENTALS OF INFRARED THERMOGRAPHY

The use of infrared (IR) thermal imaging systems is not new, since the thermal imaging camera is already widely applied in different areas, security and surveillance, medical diagnostics and detection of various anomalies in architecture, civil engineering, building construction, mechanical engineering, etc. Infrared (IR) thermal sensors are special types of sensors designed to detect the infrared radia-

tion of different wavelengths. IR offers the capability of imaging scenes based upon either the IR light reflectance or upon the IR radiation they are emitting. IR radiation is emitted in proportion to heat generated/reflected by an object and, thus, IR imaging is commonly referred to as thermal imaging. As IR light is invisible to the human eye, IR illumination and imaging systems offer a useful method for covert surveillance, as such; IR imaging commonly forms the basis for night-vision systems (Solomon and Breckon, 2011). The importance of the application of thermal imaging systems in border control and the fact that modern thermal imaging cameras have extremely high reliability and scope in the context of human and vehicles detection. Both individual systems have the ability to detect people at distances up to 20 km, and vehicles even through 20 km, regardless of atmospheric conditions. The application is particularly suitable for the colder part of the year when the temperature difference between the environment and the observed object is even more pronounced. In connection with that, it should be emphasized that the possibility of detection at long distances solely related to the detection of thermal silhouettes of people, while the possibility of applying the recognition of faces and possible identification of the person is reduced to a distance of several dozen meters (Systems, Inc. FLIR, 2016).

The infrared spectrum (IR) of electromagnetic radiation is not visible to the human eye as it has a greater wavelength, and therefore a lower frequency. The infrared spectrum covers the wavelength range from 0.7 to 300 microns, which corresponds to the frequency range from 1 to 430 THz, which is larger than the wavelength of the spectrum of visible light, but shorter than the wavelength of terahertz microwaves.

Infrared is primarily divided into three sub-areas:

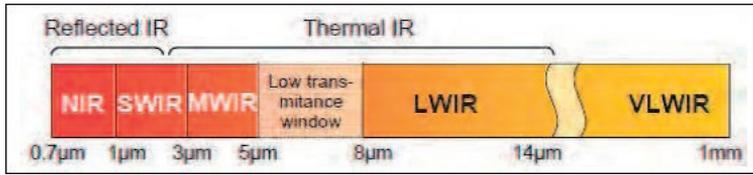
- IR-A in the range of 0.7 to 1.4 microns;
- IR-B in the range from 1.4 to 3 microns;
- IR-C in the range from 3 to 1000 m.

In addition, the spectrum of infrared radiation can be divided according to the following scale:

- Near-Infrared - NIR, ranges from 0.7 to 1 microns ( $\mu\text{m}$ );
- Short-Wave infrared - SWIR, ranges from 1 to 3  $\mu\text{m}$ ;
- Mid-Wave Infrared - MWIR, ranges of 3 to 5  $\mu\text{m}$ ;
- Long-Wave Infrared - LWIR, ranges from 8 to 14  $\mu\text{m}$ ;
- Very Long-Wave infrared - VLWIR, in a range greater than 14  $\mu\text{m}$  (Figure 1).

In this context, the authors, point out that the NWIR and SWIR bands are sometimes referred to as "the reflected infrared radiation" while MWIR and LWIR bands called "thermal infrared radiation", which do not require an additional light source and heat, considering that the emitted thermal radiation sensors can detect completely passive and thus get an image of the environment solely by reading the emission of thermal energy of the observed object (Bhowmik, M. K., Saha K., Majumder S., et al., 2011).

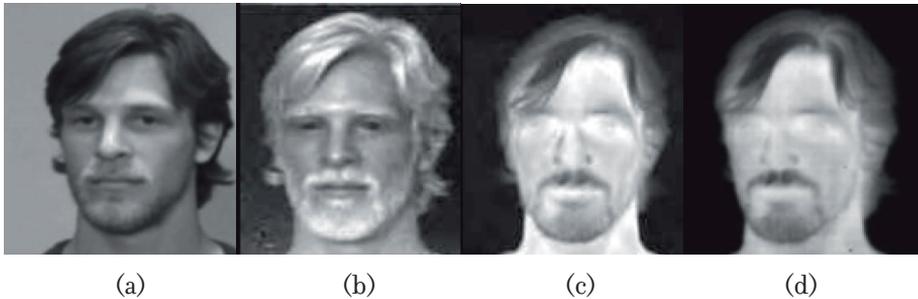
**Figure 1.:** Infrared band of the electromagnetic spectrum showing the different IR sub-bands



Source: Bhowmik, M. K., Saha K., Majumder, S., et al., 2011

The most commonly used systems are FLIR (Forward Looking Infrared) systems that are using MWIR and LWIR wavelength infrared radiation (Figure 2.), show the visibility of faces in different band of electromagnetic and IR radiation. It should be emphasized that, as a rule, comes to video cameras system for signal processing and human detection sent to a continuous video signal which requires extra processing steps of the input signal and in this case it's called real-time detection, or online processing and detection. Such results have the highest value since it enables timely reaction, especially when it comes to national security in terms of movement detection in the monitored area. In addition to the online system, there are also offline systems, which recordings are saved and after that, starts image processing in order to detect people or vehicles. Although the results of such systems have less information value in border control and monitored regions, its application in the different areas where it is not necessarily reporting real-time human detection (Bhowmik, M. K., Saha K., Majumder, S., et al., 2011).

**Figure 2.:** A face simultaneously imaged in the (a) visible spectrum, 0.4 - 0.7  $\mu\text{m}$ ; (b) short-wave infrared, 1.0 - 3.0  $\mu\text{m}$ ; (c) mid-wave infrared, 3.0 - 5.0  $\mu\text{m}$ ; (d) long-wave infrared, 8.0-14.0  $\mu\text{m}$ .

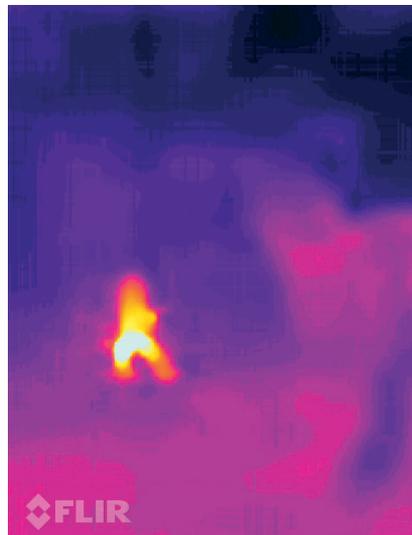


Source: Bhowmik, M. K., Saha K., Majumder, S., et al., 2011

### 3. HUMAN DETECTION USING INFRARED THERMOGRAPHY

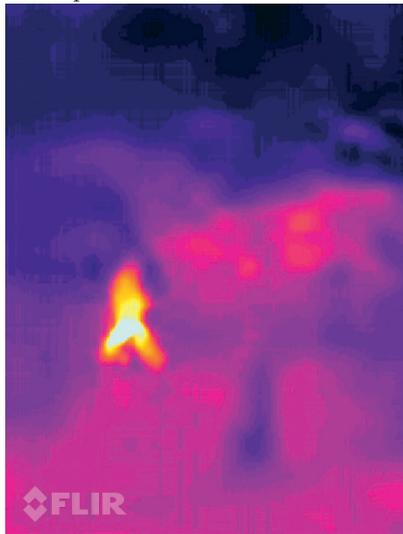
Human detection primarily relates to the detection of movement of persons, without attempting identification. In the context of state border surveillance and access control, motion detection is sufficient for an adequate and timely response, especially when it comes to illegal border crossings. Figure 3 shows the thermal silhouette of a person in motion. The person was recorded at night, normal and harsh weather conditions (rain) at a distance of about 50 meters. The objective is to demonstrate the robustness of thermal imaging cameras to weather conditions, which highlights the main advantage of this system compared to standard optical cameras that do not have the ability to record in these conditions. Recording was conducted using thermal camera implemented in smartphone CAT S60, and the recording resolution is 640x480 pixels, 25 FPS. Implementation of thermal cameras in smartphones further demonstrates the possibilities of application of infrared thermal imaging in the field of control of persons and access control. Although in this case a relatively short distance, surname of this type of camera allows timely reaction in certain segments, such as close supervision of supervised areas, and mobility devices is definitely an added advantage since it is extremely mobile and most unobtrusive, given that a small sensor that can be hidden in a variety of items that will raise doubts in people who are planning to make unauthorized access to a controlled area. A visual analysis shows that the thermal silhouette of a person clearly visible, and further research and development of algorithmic support can contribute to the development of a fully autonomous system control area.

**Figure 3.:** Thermal silhouette of a person - rain



Source: Author

Figure 4.: Thermal silhouette of a person normal weather conditions -



Source: Author

In the context of the application of thermal imaging systems in the control of the protected areas Dulski et al., (Dulski, R., Kastek M., Trzaskawka, P., et al., 2011), describe the use of systems based on a combination of different sensors. Specifically, these authors present a system that uses a combination of radar system (using radar millimeter wavelengths, which is detecting objects at a distance of 780 m), passive infrared thermal imaging camera (shooting from a distance of 480 m) and standard optical cameras that covers the visual part of the electromagnetic spectrum (recorded from a distance of 760 m), as input sensors for image acquisition. The authors point out that such systems increase the probability of detection of an intruder in the protected area, but our focus is on thermal imaging systems.

In the context of thermal imaging cameras, as one of the input sensor, it was pointed out that the input image necessarily need preprocessing using standard techniques of digital image processing such as detection of objects, edge detection, detection and recognition of colors, textures, etc. (Solomon and Breckon, 2011; Zhou, H., Wu, Y., Zhang, J., 2010). In addition, processing of thermogram has some specifics in relation to the processing of standard optical images taken in the visible spectrum. Dulski, et al., (Dulski, R., Kastek M., Trzaskawka, P., et al., 2011), state that the infrared thermal imaging must be binarised, based on detection thresholds set in advance, so that the application of the mask can define different levels of objects emission of the infrared radiation. Masks are defined by the level of brightness (or white level) of individual segments of the image that have a value above the set threshold, while all other segments are assigned a lower level of illumination (to black) and have a value below the set threshold, and are marked as background.

Breckon, et al., (Breckon, T. P., Ji, W. H., Richardson, J., 2012), also describe a combined autonomous system for surveillance and human detection, which also can be applied for vehicles detection, using optical and thermal images. The authors propose a two-stage approach to automatically detect people and vehicles with a time filtering. This approach primarily to detect the initial segments within the scene that might contain object, followed by feature extraction of isolated segments based on which carried out a secondary classification of persons/vehicle. In order to achieve effective detection of real-time segments within the initial scene, the isolated sub-regions which can contain objects, and then through the extraction of features and classification of selected sub-regions it is possible to achieve real-time processing. Further improvements can be achieved using adaptive background model based on the Mixture of Gaussian (MOG). This approach allows each pixel to be processed as a set of Gaussian distribution, and enables the simultaneous operation of changes in movement or the noise of each pixel of each image affected by a certain period of time. Ultimately, this model makes it possible to capture a new image every time situation is changed in a way that evaluates each pixel in order to evaluate whether it is a front (object pixels) or background pixels that are discarded. The whole process of modeling is taking place "in the background system" which enables automatic identification of the foreground (object) segments each image that comes from the input video signal making it possible to conduct effective extraction of features and classification of each new object that appears on the scene, but also objects appear on the scene and remain motionless after appearing (e.g. a hiding person). Furthermore, the authors state that the use of this model and applying the selected classifier, such as SVM, Bayes classifiers, Artificial Neural Network, KNN (Witten, I. H., Eibe, F., Hall, M., A., 2011; Harrington, P., 2012), can classify objects based on offline classifiers training and direct online application to input images. Breckon, et al., (Breckon, T., P., Ji, W. H., Richardson, J., 2012), as a classifier in the last phase of identification, describe the use of Support Vector machine - SVM (Burges, Christopher J.C., 1998), and Decision Forest - DF, (Breiman, L., 2001), algorithms.

Breckon, et al., (Breckon, T., P., Ji, W. H., Richardson, J., 2012), conducted an independent learning system for examples of identification of cars and people, as well as for each type of image (optical and thermal), with the application of additional classifiers for the classification of sub-categories (civilian/soldier). Furthermore, authors put a significant emphasis on the temporal filtering in terms of high frequency input images into the system in relation to the possible occurrence of a person on the recordings. The experiment used a camera that captures 25 frames per second (FPS), and therefore it is clear that a certain number of the images do not contain the object and should be rejected. By rejection of background pixels, the processing system retains only images where objects are detected based on pixels with high intensity lighting, or only images containing a person and which are suitable for further processing and classification. Extracted images, which contain an object,

during processing, are grouped into time groups that actually makes only a small part of the total recordings and a smaller part of the time slot in relation to the entire period of training supervision (recording). In this sense, for moving objects (the same is for static objects) changes the scale, perspective in observing the object as well as the level of the presence of noise, can significantly affect the distribution of the extracted features of any segment of the image that can contain the object. These variations can largely be systemic, in terms of changes with respect to the scale and perspective of observation, or may be random with respect to the level of noise present, and thus may have a certain impact on the detection and identification of the selected object images. In conclusion, this paper pointed out that the simple time filtering effect for reporting the detection of the object, with respect to the sensors used, is very important. Namely, it is emphasized that in this way reduces the number of duplicate reports detection of objects, this approach ultimately leads and to reduce the number of false positive detection. It was also noted that this approach largely sets the framework for further research and integration of multi-modal system for object detection and reporting of users in the post-classification stage. Further development opportunities are open especially in the context of the robustness of the sensors as well as in the field of combining different types of sensors to the smooth and continuous operation of the system to be functional even in the event that one of the sensors fails.

In the context of monitoring the open areas with the goal of detecting people and vehicles, Breckon, et al., (Breckon, T., P., Gaszczak, A., Han J., et al., 2013), also described a combined monitoring system based on optical and thermal images, and further combined with the radar signal. In the mentioned paper is described the use of combined systems installed on unmanned ground vehicle (UGV), and unmanned aerial vehicle (UAV). The emphasis of the paper is on automatic processing, fusion of input signals, and reporting the results of the surveillance in the classes of persons/vehicle. The authors describe in detail used platforms (UGV and UAV) with descriptions of the technological capabilities of vehicles and equipment installed on these platforms. The focus was put on automatic processing, detection and interpretation of recordings. Regarding the use of UGV, the authors state that the automated detection based on a two-phase system which is the first stage relies on the rapid localization of regions that may contain object using Haar cascade classifier (Viola, P., and Jones M., 2001), while the second stage is to confirm the presence of the object by using SVM (Burges, Christopher J.C., 1998). Presented system, during the first phase, using Haar cascade classifier for identifying objects in certain segments of the input image. The concept is designed in a way that combine more related weak classifiers to create a strong classifier, and the authors cite the example of the use of boosted classifiers using Haar features (Viola, P., and Jones M., 2001). Haar features are obtained from the spatial response of Haar basic function and its derivation, but in practice is calculated as the sum of the difference between varying rectangular localized sub-regions, which, although limited to the individual features, can very effec-

tively calculate (relying only on integer math). Individually, these features are weak discriminative classifiers, however, when connected in cascade yields a strong discriminative classifier based on a calculation of the difference in brightness and color. Breckon, et al., further point out that the resulting classifier is trained using the Ada-Boost algorithm (Dollár, P., Tu, Z., Tao H., Belongie, S., 2007), in order to choose the maximum discriminant subset of Haar features from around the extensive feature set. Final Haar classifier cascade consists of a series of key weak classifiers. Mention simple classifiers actually represent degenerative decision tree that Haar features taken as input for the weak classifiers and returns a response aisle/rejection. Selected segment images must give a positive answer to all classifiers in cascade in order to be classified as an instance of an object on which to learn a strong classifier. The classifier can be tested in the input image based on a number of levels and the location of use of a sliding window of credit. Nature cascades, arranged according to the most discriminative features, provides early rejection of most of the window based on the results in searching the minimum features subset to be used for evaluation of the cascade. Haar cascade classifier successfully combine multiple classifiers and initially reject the negative image regions and focuses on segments that have the potential, which certainly saves resources and time (Breckon, T., P., Gaszczak, A., Han J., et al., 2013). The same authors state that the system in the second phase requires the confirmation of the object, which is based on the use of SVM classifiers. Classifier generation is performed, following the aforementioned methodology, for both people and vehicle by performing training over a set of manually labeled positive and negative examples. For people detection, they use a data set of approximately 2000 positive examples (people) and approximately 11,000 negative (non-people) examples randomly selected from the same source imagery. The negative set is again randomly sub-sampled to generate 2000 negative (non-people) examples for each stage of the cascaded Haar classifier training whilst it is used in full for training the secondary SVM classifier with the addition of another 2400 negative examples. This training procedure is performed for both visible-band and thermal-band imagery to generate a trained classifier for detection that operates independently in each modality. For thermal imagery, an additional two-stage classifier pair is generated using only the upper-torso of the human body to train detection in some partial occlusion cases. The tests were conducted in the optical and thermal band of the spectrum, but in this paper, focus is only on the results of testing the thermal spectrum. The authors created additional pair of two-stage classifiers only for the person upper torso in order to train detection in cases where the body is partially hidden. On thermal images, people have shown in various environmental and weather conditions as well as in different positions of the body. Human detection is carried out in relation to the buildings and openings (windows, doors) on the building. Before the process of detection of persons, is performed detection of openings on the building using detection of edges, and then the straight lines extracted by using PCA method. Detec-

tion of human thermal traces on windows or doors of the building are obtained using the average pixel value that is set as the threshold, with a predetermined sub-region of the thermal image, which corresponds to the dimensions of the any openings in a building and scaled to the dimensions of 50 x 50 pixels. From this, the rated areas extracted eigenvector that covers the form of distribution of thermal trace of the opening of the building and the ratio of openings before scaling. Eigenvector is designed chaining vertical and horizontal histograms projection scaled images with the set threshold and ratio for each opening. The resulting eigenvector generates input set for four class neural network (full body, half body, empty hole, non-person) and carried out the detection of whether the visible whole person, half the person and ignore the holes in which no person or non-person. In the context of detection of persons using thermal imaging cameras mounted on unmanned aircraft (UAV), authors also used the procedure described above, but they in the second phase used multivariate Gaussian detector forms to confirm the detection of human thermal silhouette. The authors also described the process of determining the location and measured the distance from the sensors. Results showed that the presented system, and proposed techniques for the construction of an autonomous system for human detection give good and reliable results ranging from 90 to 100% with error in estimating the distance of  $\pm 5$  m. Although the present system gives a good result, in some situations it is less than 100%, which is not enough for absolute reliability for application in the field of national security. For absolute reliability and use in national security, the result always must be 100%; due to a slight deviation can lead to misinterpretation and wrong decision. Wrong measuring distance can also lead to making the wrong assessment, given that the deviation of 5 m, may mean that the person is not detected crossed the state border and did not enter the monitored and protected area, and to conduct the police to such a person, if it comes to the state border, could lead to violations of the border areas, and even an international incident, especially when it comes to crisis areas between the countries. In addition, a negative effect may arise in the case of using such a system on the UAV for offensive action, when a deviation of 5 m can result in unintended victims, or to miss targets.

Gilmore, et al., (Gilmore, E., T., Ugbome, C., Kim C., 2011), described the thermal imaging human detection system in urban environments implemented with a Matlab on a laptop platform interfaced with a low-end microcontroller. This system primarily focuses on digital, infrared thermal imagery and pedestrian candidate selection and detection using available detection algorithms, and on simple, economical implementation, which is possible through the direct interfacing between the algorithm in Matlab and low-end external hardware for indication of the detection result. The objective was to achieve the human detection almost in real time. The authors described a system that is based on a three steps. The first step, preprocessing is done to obtain hotspots from the image using the methods of histogram equalization and segmentation by a threshold approach. A threshold approach is a technique

often used to separate foreground objects from background environment based on their differences in image intensity. In a simple threshold approach, a single intensity threshold is used to generate a binary image from the original image. A simpler approach is based on determining a single binary threshold intensity of the original image. The second step involves the implementation of morphological operations to remove background noise from the image, in this case, techniques of erosion and dilation are successfully removed most of the noise from the image, and the application of these techniques have been successfully detected regions of interest (ROI) in the images. The third stage is the selection of the appropriate characteristics of the human body using different metric measures such as the ratio restrictions, local histogram filters and/or matching morphological human model. Gilmmore et al., for human detection used height and width and are thus automatically remove all non-human objects in the image. The results of the tests given by the accuracy of detection of 90% with 10% false positives, which the authors characterize as a good result given the equipment used in the testing and algorithmic support. As main advantages, the authors pointed out the possibility of application platforms developed in Matlab and use simple microcontroller, which enables the successful implementation of all steps in terms of online, and offline data processing, formatting, image detection and calculates the difference and simplicity in use, flexibility and efficiency. As the shortcomings pointed out the weakness of their algorithm to perform analysis of complex image and the inability to simultaneously, detect more people with regard to the threshold set by only one person. A major drawback of the system is the inability to detect people in different positions (lying, stoop, squat, etc.).

Despite some disadvantages presented system is a good indication that it is possible to develop a cost-effective and efficient system for human detection and therefore economic viability given that shows how thermal imaging system for detecting people as possible built using very simple techniques and inexpensive components. It is also extremely important that the system despite its simplicity has a high detection rate. The disadvantage is, as the authors themselves noted, the inability to simultaneously detect more than one person, what can be achieved by applying more advanced classification algorithms, and the question remains how the system is based on components of this type can be a reliable long-term, due to the control of the state border and other protected area requires high reliability for continuous (24/7) long-term use.

Mieziako and Pokrajac (Mieziako, R., and Pokrajac D., 2008), presented a method for detecting people in low-resolution infrared videos. Presented system use feature set based on a recursive model patches. Feature extraction based on the histogram edges orientation without creating a dense grid. The basic premise of the feature set is to extract information on the gradient and display them as an integral image to quickly compute the difference between a "patches". One set of "patches" is defined as the relative position in the search window, while the other set of "patches"

is used to yield more local descriptors. Details of the detection and the human presence are obtained by calculating the histogram ratio between different "patches", followed by a linear SVM trained on the individual thermograms. In the first stage, the image is converted to grey scale image (if the input image is not black and white), followed by computing the edge gradient of every pixel that is composed of the gradient scale and orientation. Followed by construction of 3D images in a way that every part of the 3D image is one orientation bin highlighted as an integral image, and the sum of the pixels. This representation allows rapid extraction of features histogram's gradient of the sub region of the input image of any size using permanent access to the integral image. For each search window, is calculated expensive models location "patch", a corresponding histogram vector is calculated from the resulting 3D image. The ratio between the histogram of each set of patch models becomes an eigenvector. Linked eigenvectors all sets of patch model in the search window, linear SVM classified it as a human/non-human. In the same paper, the authors pointed out that the calculation of the eigenvectors must take several steps. Initially, the edge detection is performed by convolving vertical and horizontal Sobel masks used in object detection (Levi, K. and Weiss, Y., 2004). Except of Sobel masks, proposed by the authors, other methods for edge detection probably can achieve equally good results, such as Canny edge detector (Nadernejad, E., Sharifzadeh, S., Hassanpour, H., 2008), which is one of the most commonly used edge detectors of which should implement appropriate tests and comparisons with the proposed Sobel masks. In conclusion, the authors (Mieziako, R., and Pokrajac D., 2008), pointed out that the application of the described model - using histogram gradient for the construction of 3D images and models histogram ratios, recursive model for the calculation of eigenvectors and use linear SVM classifier gives as good results in the human detection in infrared thermal imaging video of low resolution.

#### 4. FACE DETECTION AND IDENTIFICATION

In addition to the previously described human detection methods, another promising area of application of infrared thermal imaging in the field of national security, and surveillance of protected buildings is face recognition on thermal imaging footage and possible identification of a person based on face thermograms.

With regard to detection and face recognition to thermal imaging, footage reveals a certain advantage over standard optical image. Thermal imaging is insensitive to the level of the lighting conditions. The standard photo and video sensors record image based on the reflection of the visible part of the spectrum of the subject. On the other hand, infrared thermography uses infrared band of the spectrum, and depending on the amount of radiation of the detected objects, in this case the human body or face. In this regard, it should be noted that the infrared thermal imaging sensors have the ability to record the structure of veins structure and tissue on the face, which can

be further utilized in the field of face detection and identification of a person (Jafri, R., Arabnia H., R., 2009).

Cutler (Cutler, 1996), noted that the structure of veins and facial tissue is unique for each person, such as a fingerprint or papillary lines. Based on that, he spent experiment using eigenfaces and infrared images to do automated face recognition. 24 subjects were used in a database of 288 images. A recognition accuracy of 96% was achieved for frontal views, 96% for 45° views, and 100% for profile views. This compares favorably with visible light techniques. However, to the structure of blood, vessels and tissue were visible on the thermogram; the images must be high resolution. Cutler experiment used an image resolution of 160 x 120 pixels on which the structure was only moderately visible. Jafri and Arabnia (Jafri, R., Arabnia H., R., 2009), pointed out the disadvantages in the use of thermograms for face recognition. Although the infrared thermal imaging system is robust to a wide range of impacts on the record because the focus is on capturing thermal radiation of the observed person, significant gaps lie in the economic viability due to the usually high cost of thermal imaging systems, then as other types of problems. Jafri and Arabnia (Jafri, R., Arabnia H., R., 2009), as one of disadvantages, also mentioned low resolution of thermal sensors and high noise in the captured images (although the previously described specific techniques that successfully solved the problem of noise), and an additional problem is the high sensitivity thermal imaging camera on the glass (or glasses) as the glass is an obstacle to recording thermograms high resolution, or any other resolution because thermal radiation can't penetrate through glass.

Bebis, et al., (Bebis, G., Gyaourova, A., Singh, S., Pavlidis, I., 2006), emphasized one of the major problems in face detection and possible identification of person, the high impact of emotional and health status of people on results. In fact, this is perhaps one of the most significant problems of thermal imaging to identify a person on the basis of the thermogram as the different emotional and medical conditions greatly affect the body temperature which is why solving these problems are very difficult and demanding. Solutions of the mentioned disadvantages can be found in combining techniques of infrared thermal imaging with optical recording of persons with regard to the techniques recorded completely different facial characteristics. Bebis et al., have proposed an approach that is based on the simultaneous recording of optical and thermal imagery and their subsequent merger in order to obtain maximum characteristics with both images. The authors describe two approaches, one based on analyzing pixels (pixel-based) and the second to feature extraction (feature-based). To calculate the optimal border mergers obtained recordings authors have proposed the use of genetic algorithms (GA). The authors also conducted extensive research on the available data, and the results showed that the approach based on analyzing pixels gave better results than an approach based on the calculation of the intrinsic characteristics, although this approach is more computer demanding.

Socolinsky et al., (Socolinsky, D., A., Wolff, L., B., Neuheisel, J., D., Eveland, C., K., 2001), described face recognition and identification using thermograms and they conducted parallel analysis of the thermal and optical images with a singularity. The authors, for the purposes of the study, developed a special sensor capable of simultaneous recording of optical and thermal images (CCD and LWIR micro bolometer). The resulting recordings have special significance since it shows an identical recording in two different areas of the spectrum, which is not the case with the previously described techniques, which are normally used footage captured separately. Face detection on the obtained images were tested using two algorithms - eigenfaces and ARENA algorithm that is somewhat simpler and describe it as an algorithm with a nearest neighbor (1-NN). ARENA algorithm first reduces the dimensionality of the image pixel at a very coarse resolution by replacing each pixel corresponding mean values of gray values on the square adjacent pixels. The test images were taken under different lighting conditions except that the brightness setting for each person changed several facial expressions. The results show that both the tested algorithm provides differently on optical images when changing the brightness (illumination while the thermogram had no impact), however the thermograms were encountered problems when a person is wearing glasses, although this problem successfully solved using the eigenfaces with respect to the reduced impact of eye region on the overall score. Ultimately, better results are achieved by using thermograms than optical images and the use of eigenfaces methods than the ARENA algorithm.

Seal et al., (Seal, A., Ganguly, S., Bhattacharjee, D., Nasipuri, M., and., Basu D., K., 2013), described a method for automated person identification based on face thermograms respectively using image of facial veins structure. The authors present a new approach to the identification of a person based on the structure of facial veins structure based on the assumption that this structure is unique for each person. The application of the three methods for obtaining the image of the facial veins face thermograms. The first is bit-plane slicing, and transformation around the medial axis, the second is morphological erosion gray values with a transformation around the medial axis and the third is the application of the Sobel operators for the extraction of data on the blood flow. Method of bit-plane slicing is a method wherein each pixel is displayed in black and white in 8 bits, which means that the image is made up of eight one-bit planes (bit-1) starting from the 0-plane bit, which is the least significant (LSB plane) to the bit-plane 7, which is the most important plane (MSB plane). The advantage of this method is that the yield relative importance of each bit in the pixel. The unique aspect only four bit-planes of higher order (4 to 7) contain most of the visually significant information, while the bit-plane of the lower visual order contain fewer important information as they are obtained from a lower contrast. This is the method used for extraction of thermal physiological features of the face of grey scale images and create folders contour areas with constant temperature. The transformation of the medial axis is used for the extraction of lines connecting all points of

equal temperature in order to obtain the image structure of the blood vessels (similar to the application of isothermal lines connecting areas of equal temperature on maps weather forecasts). Another commonly used method is the morphological erosion of grey values, and is also used in order to extract the thermal physiological features of the face and create a region of constant and equal temperature, after which, as with previous methods used by the transformation of the medial axis in order to obtain images of structure blood vessels of the face. The third proposed method is the application of Sobel operators to obtain images of facial veins structures, and is based on the detection of discontinuities in the image. In this case, as the dashed changes in pixel intensity, and the operator is composed of a pair of 3 x 3 mask. The masks are designed to correspond to the maximum discontinuities in intensity values, and move vertically and horizontally with respect to the grid of pixels provided that is used by one mask for each direction (one vertical, one horizontal movement). Masks are combined in order to find the maximum gradient magnitude at each orientation point gradient. After applying some of the methods presented extracted detailed (key) points or minutiae that are commonly used in identifying person's fingerprint and that is the main similarity identification of people using the structure of the blood vessels of the face. The facial veins structures structure arranges similar to the fingerprint ridges (with fingerprint analysis of the structure of the ridge on the print) on the basis of which it receives a single thermal face print. Additional processing and extraction of eigenvectors of the structure of minutiae that connect areas of equal temperature of the resulting vectors will have different lengths for each person and for the classification of eigenvectors they proposed multilayer perceptron which has achieved recognition accuracy of 95.24%. Regarding the above, it should be noted that the person is recorded with thermal imaging camera at close range (from a distance of about 60 cm) in order to obtain maximum distinguishable facial thermogram. As the main advantage of this approach the authors pointed out the ease of implementation provided that it is not necessary knowledge of geometry or the specific features of the face, but the disadvantage is the necessity of recording the face at close range, the applicability of the system to a single face with a constant background, considering that with the changing background can get bad results. Although this system, because of its limitations, is not applicable for the identification of people with longer distances, according to the authors, is likely to be successfully implemented in the context of authentication access protected buildings and other areas of national security.

Bourlai and Čukić (Bourlai, T., and Čukić, B., 2012) presented the results of research regarding the problems related to the intra-spectral and inter-spectral face recognition in different conditions and with different recording distance. The authors examine the advantages and disadvantages of face recognition in images taken in different band of the infrared spectrum compared to images taken in visible light in controlled and uncontrolled conditions. Infrared images are captured outdoors at night from a distance of 30 to 120 m. It is also performed to test the recognition

between different distances from which the recording picture, as well as relations between the images taken in different spectra and different distances. Comparison of the performance of facial recognition are performed between the images recorded in the shortwave infrared (SWIR) in relation to the images from the visible part of the spectrum, and from images taken in medium infrared spectrum (MWIR vs. MWIR), as well as between the images taken in the visible spectrum and MWIR band of the spectrum in a controlled environment, and between the images captured in the spectrum near infrared region (NIR) as well as between the NIR images and images from the visible part of the spectrum with the change of the shooting distance. Test results showed that the inter-spectral face recognition level of 100% between images from the visible part of the spectrum and SWIR image when all the images from the database recorded in controlled conditions from a distance of 30 m. In the case of half-controlled conditions and shooting distance of 50 m, the recognition rate is reduced to 90%, and when the distance is further increased (to 106 m), the recognition rate is reduced to 80%. In the most demanding scenario, or in a completely uncontrolled, the results were unsatisfactory at the border, and mainly because they were non-cooperative target person and are recorded under different conditions. Decreased percentages of recognizing the different scenarios have been shown in testing MWIR of the infrared spectrum compared to other parts of the spectrum. However, promising results are obtained when it comes to comparison between MWIR images at reasonable terms when they are comparable with the results of the comparison images from the visible part of the spectrum, and the added benefit is that these images can capture in complete darkness. Testing conducted in the NIR part of the infrared spectrum, which was conducted at different distances from 30 to 120 m, as authors expected, showed worse results, especially when the recording distance is increased to more than 90 m when the facial features significantly less prominent due to weather conditions. The same trend was also evident when testing between different parts of the spectrum and between different shooting distances when comparing them with images from the NIR spectrum recorded with a 30 m. Nevertheless, when compared to images from the visible part of the spectrum and NIR regions that were taken from a distance of 120 m, using Bayes ML algorithm, the recognition accuracy is achieved of 92%. The described experiment is just one in a series of experiments and research that can be carried out regardless of facial recognition in demanding circumstances and recording conditions.

## 5. DISCUSSION

The use of thermal imaging cameras in the area of national security shows promising results. However, there are rare cases in which achieved 100% accuracy, which in some segments of national security is not sufficiently reliable. One area of the state border surveillance and protected facilities, where necessary absolute re-

liability of the control system in the context of detection and control of movement of persons in controlled areas. On the other hand, thermal imaging systems show absolute reliability when it comes to just about detecting movement, which in some cases may be good enough. Such examples are the use of infrared thermal imaging in the semi-autonomous systems where large investors human operator who makes the final decision whether a person or an animal, and the system performs the localization of regions of interest in the video to the human crew could quickly has made a decision on further action.

## 6. FUTURE WORK

For future research opens up a completely new field for improving the efficiency and reliability of thermal imaging systems. In the context of identifying persons to thermal silhouette of the whole body, opens the possibility of developing the system for identification by way of walk, considering that human gait unique to each person and people can identify known people by gait at greater distances. In this context, possible research in terms of identifying a person and distinguishing authorized/unauthorized person so that the autonomic system can detect unauthorized persons and to activate the alarm.

The study of the thermal imaging system, except in the area of national security, it is possible in other areas such as the expertise of traffic accidents in the context of detecting invisible traces of braking due to thermal footprint remains on the tarmac for some time after leaving the car, however, remains to investigate how long the thermal footprint remains and what is the reliability of such a sign for calculation of braking distance and vehicle speed. Further research in this direction could be the creation of a database of thermal traces of well-known manufacturers of tires and rubber based on the mark to determine the type of vehicle.

## 7. CONCLUSION

Based on the review of methods described above, it is clear that the application of infrared thermal imaging in the context of detection of people in open areas gives good results, both in urban areas and in areas outside of urban areas. In this regard it should be noted that in the context of the control of open areas of the application of reduced solely to the detection of persons without any possibility of identification with respect to thermal imaging cameras when shooting from long distances have the option of recording only the thermal silhouette of a person. On the other hand, to infrared thermal imaging system successfully applied to identify people based on facial thermogram recording faces must be carried out at close range and only one face, because only in this way can get a picture of the entire structure of the blood vessels of the face.

The described fields of application certainly open wide spectrum regarding future research in this area, especially in the context of identifying people based on facial thermograms. Furthermore, in the context of state border described techniques and methods it is possible, with some adjustment used at border crossings to detect hidden objects, or when you view the baggage or cargo space. In this context, it should be noted that thermal imaging has certain advantages over other methods that are currently being applied. It is certainly the most important application in order to detect hidden persons (illegal immigrants) within the cargo, which is a common method of attempting to illegally cross the state border.

The above models and ways to detect people in the infrared thermal imaging recordings are not the only, or in certain segments closest to the ideas related to the development of future dissertation. The dissertation topic will be primarily focused on the description of automatic unattended system detection of persons using infrared thermal imaging camera. Furthermore, the idea is also to be part of such research conduct analysis of images obtained on the basis of which it will be possible to calculate the eigenvalues of the whole person and the most common postures when illegal attempts to cross the border or unauthorized access to the protected areas (upright, hunched, crawling, etc.). It will also further attention be focused on the detection of persons on the basis of partial visibility, or will seek to define the characteristic vectors by which it will be possible to establish unequivocally that it is a person.

## REFERENCES

- Solomon, C., and Breckon, T., "Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab"; John Wiley & Sons, Ltd. (2011)
- Systems, Inc. FLIR., "FLIR HRC-Series Thermal Imaging Cameras - FLIR Infrared Cameras", Accessed December 21, 2016. <http://www.flir.com/cs/emea/en/view/?id=42061>
- Bhowmik, M. K., Saha K., Majumder S., et al., "Thermal Infrared Face Recognition - A Biometric Identification Technique for Robust Security system", Reviews, Refinements and New Ideas in Face Recognition, (2011)
- Dulski, R., Kastek M., Trzaskawka, P., et al., "Concept of data processing in multisensor system for perimeter protection", SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, (2011)
- Zhou, H., Wu, Y., Zhang, J., "Digital Image Processing - Part one", Ventus Publishing, (2010)
- Zhou, H., Wu, Y., Zhang, J., "Digital Image Processing - Part two", Ventus Publishing, (2010)
- Breckon, T., P., Ji, W. H., Richardson, J., "Consistency in multi-modal automated target detection using temporally filtered reporting"; Proc. SPIE 8542, Electro-Optical Remote Sensing, Photonic Technologies, and Applications VI, 85420L (November 19, 2012); <http://dx.doi.org/10.1117/12.974559>
- Witten, I. H., Eibe, F., Hall, M., A., "Data Mining: Practical Machine Learning Tools and Techniques"; Morgan Kaufmann Publishers, (2011)
- Harrington, P., "Machine Learning in Action"; Manning; (2012)
- Burges, Christopher J.C. "A tutorial on support vector machines for pattern recognition", Data mining and knowledge discovery 2.2 (1998): 121-167
- Breiman, L., "Random forests", Machine learning 45.1 (2001): 5-32
- Breckon, T., P., Gaszczak, A., Han J., et al., "Multi-modal target detection for autonomous wide area search and surveillance", Proc. SPIE Security and Defence: Unmanned/Unattended Sensors and Sensor Networks; (2013)
- Viola, P., and Jones M., "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition, Vol. 1., (2001)
- Dollár, P., Tu, Z., Tao H., Belongie, S., "Feature mining for image classification", IEEE Conference on Computer Vision and Pattern Recognition, (2007)
- Gilmore, E., T., Ugbome, C., Kim C., "An IR-based Pedestrian Detection System Implemented with Matlab-Equipped Laptop and Low-Cost Microcontroller", International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 5, (2011)
- Mieziako, R., and Pokrajac, D., "People detection in low resolution infrared videos", Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, IEEE, (2008): 1- 6
- Levi, K., and Weiss, Y., "Learning object detection from a small number of examples: the importance of good features", Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, Vol. 2., IEEE, 2004
- Nadernejad, E., Sharifzadeh, S., and Hassanpour, H., "Edge detection techniques: Evaluations and comparison" Applied Mathematical Sciences, 2.31 (2008): 1507-1520

- Jafri, R., and Arabnia H. R., "A survey of face recognition techniques" *Jips* 5.2 (2009): 41-68
- Cutler, R. G., "Face recognition using infrared images and eigenfaces", University of Maryland, (1996)
- Bebis, G., et al. "Face recognition by fusing thermal infrared and visible imagery." *Image and Vision Computing* 24.7 (2006): 727-742
- Socolinsky, D. A., et al. "Illumination invariant face recognition using thermal infrared imagery", *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001
- Seal, Ayan, et al. "Automated thermal face recognition based on minutiae extraction" *International Journal of Computational Intelligence Studies* 2.2 (2013): 133-156
- Bourlai, T., and Cukic B., "Multi-spectral face recognition: identification of people in difficult environments", *Intelligence and Security Informatics (ISI)*, 2012 IEEE International Conference on. IEEE, 2012