

Target Identification Using Sensors of Different Nature

Anete Vagale¹, Agris Nīkitenko², Eduards Slava³, Ottar L. Osen⁴
¹⁻³ Riga Technical University, Latvia
⁴ Norwegian University of Science and Technology, Norway

Abstract – In the present paper, a new and improved visual sensor data fusion method is proposed that uses visible and far-infrared light sensors. Additionally, lux meter data are used for decision level fusion of beliefs of recognised target classes. The database consisting of 4 ambient light condition images is created using Canon and FLIR cameras.

The developed approach has been tested using database images, neural network training and classification, particularly for low light level conditions. Enhancements of target identification precision are proved by practical implementation and testing of the proposed method.

Keywords – Hyperspectral sensor data processing, object recognition in various light conditions, sensor data fusion, target identification.

I. INTRODUCTION

For visual object identification, different types of sensors might be used. The most used visual sensors are visible and infrared (IR) light sensors, because of their affordability and ease of use. Nevertheless, even these visual sensors fail to work in ambient conditions, such as a lack of visible light or too low/high temperatures [1]. The erroneous target identification may have negative consequences depending on the field of application. The solution to this problem might be to use a combination of sensors of a different nature.

The rest of the paper is organised as follows. Section II introduces the target identification task, gives a review about visual sensors, sensor data fusion techniques and related studies. In Section III, the proposed approach of target visual identification is described. Section IV describes the practical implementation and testing of the proposed approach. Conclusions are given at the end of the paper.

II. SENSORS FOR TARGET IDENTIFICATION

In the image analysis, objects can be separated into *target* objects and the *background*. It is not difficult to detect the existence of targets, but identifying the targets is more challenging [2].

A. Definition of the Target Identification

Automatic target recognition (ATR) is the ability of an algorithm or a device to recognise targets or objects, based on data obtained from sensors [3].

The advantage of using ATR in some applications is the elimination of the human support for target recognition. The major drawbacks of using humans are the slow reaction time and the dependency on ambient light conditions, although they

are good at some target recognition tasks. The architectures of ATR systems can be divided into the three high-level functional blocks: *a) data acquisition*, *b) signal processing* and *c) target recognition and tracking* [4]. The ATR algorithm initially processes the image, then detects target object, performs segmentation, calculates features, performs selection and classification, makes priorities, performs tracking and chooses target points [3]. The functions of ATR in mobile robot vision systems or military target identification systems need to be used in real time and they should adapt to dynamic situations.

Two possible erroneous detection cases are *blindness* (false negative) and *ghost object* detection (false positive). In these cases, the cost of mobile robot trajectory can increase, the collision can happen or in the military field the opponent can remain invisible for the system.

B. The Diversity of Sensors Used in Target Identification

Visual sensors usually employed in target identification use emitted and reflected light from the target object. Image processing algorithms together with visual sensors make up a computer vision system.

Visual sensors used in vision systems may be divided into 4 main groups, based on the area of the electromagnetic spectrum that they perceive from the environment: *a) infrared (IR)* light sensors, *b) visible* light sensors, *c) ultraviolet (UV)* light sensors and *d) multispectral* and *hyperspectral* sensors.

a) IR Light Sensor

In *ISO 20473:2007* [5] standard, the wavelength of IR light spectrum is divided into the three subcategories:

- *near-IR* with wavelengths of 0.75–3.00 μm ;
- *mid-IR* with wavelengths of 3–50 μm ;
- *far-IR* with wavelengths of 50–1000 μm .

The mid-IR and far-IR spectrum sensors are known as the *thermal* spectrum sensors that perceive heat radiation from the environment. The major advantage of employing sensors in the IR spectrum for visual target identification is the independence of visible light conditions – targets can be “seen” even in complete darkness that is outside the human visual range.

b) Visible Light Sensor

Visible light sensor is the most often used visual sensor that perceives electromagnetic waves with the wavelength of 380–750 nm. These sensor data represent the colour parameters of the target. The drawback of these sensors is the dependency on visible light conditions, they cannot be used in complete darkness [6].

c) UV Light Sensor

The wavelengths of UV light spectrum according to the *ISO-21348* standard are 10–400 nm [5]. UV sensor output images give more information on details of the object surface that is not so sensitive information in this case. For this reason, the technology is not suitable in terms of target identification.

d) Multispectral and Hyperspectral Sensors

Both sensors perceive multiple narrow spectra. *Hyperspectral Imaging (HSI)* data are stored as 3D data. Both sensors may include spectrums from UV to IR spectrum. Hyperspectral sensors perceive many more but narrower spectra, compared to multispectral sensors. The advantage of these sensors is the high level of details; however, the drawback is the large amount of data about the environment that the sensor perceives.

e) Visual Sensor Technologies and Their Possible Combinations

Multiple visual sensor technologies and their combinations might be employed using the previously mentioned visual sensors. The most often used technologies are the *Time-of-Flight (ToF)* camera and the *Light Detection and Ranging (LiDAR)* [7]. Both technologies employ active sensors to detect the reflection time of the emitted signal. Thereby the so-called point cloud is acquired providing additional information to the observer – depth, which enables one to analyse a three-dimensional space [8].

f) Visual Sensor Network

Visual sensor network (VSN) consists of multiple camera nodes and its major advantage is the redundancy of the system. VSN also has low cost and a wide field of applications [9]. The major drawback of VSN is relatively high energy consumption rates in order to perform all the tasks – image perception, processing and data transmission.

C. Sensor Data Fusion

When multiple sensors are used in a common system, the main effort is related to fusion of the data provided by the sensor system [10]. It is possible to fuse not only multiple visual sensor images, but also to combine visual and non-visual sensor data simultaneously.

Data fusion of visual sensors can occur at one of the three abstraction levels [11] – *pixel*, *feature* and *decision fusion*. In *pixel*-level fusion, pixels of both images are fused with certain probability. For *feature*-level fusion at first the image features are extracted and only then fused with certain probability. In *decision*-level fusion, the objects in the image are classified at first and then decisions are fused.

In order to choose an appropriate visual sensor combination for sensor data fusion, the following criteria must be considered (Table I).

Table I shows that the visible light camera is available in terms of the low cost and high sharpness level of the outlines. Particular combination of sensors depends on application environment specifics as well as on the necessity to compensate drawbacks of particular sensors by advantages of the others.

TABLE I
VISUAL SENSOR COMPARISON BY DIFFERENT CRITERIA

VISUAL SENSOR	SPECTRUM, NM	PARAMETER	OUTLINE SHARPNESS	DEPENDENCY ON LIGHT	COST
IR light	8000–12000	IR spectrum	Low	No	Fairly high
Visible light	380–750	Colour (visible light)	High	Yes	Low
UV light	10–400	UV spectrum	High	No	Fairly high
Hyperspectral multispectral	at least 400–2400	Visible light, IR and UV spectrum	High	No	Very high

D. Related Studies

The most frequently used sensor combinations for target visual identification are the following.

a) Existing Technologies of Sensor Combination

Visual sensors are usually combined with different distance measurement sensors. In [12], a mobile luggage carriage robot uses a *multi-directional camera* that obtains images with 360° field of view, together with a *laser range finder*. The drawback of this technology is that it operates properly only in good light conditions.

In [13], a micro aerial vehicle vision system that uses a visible light camera together with a sonar is used for navigation purposes and visual target tracking.

The pedestrian classification task, in [11], is performed using *visible* and *far-IR light cameras*. This technology gives good precision in different light and weather conditions and the paper compares the results of all three sensor data fusion abstraction levels.

The path and obstacle detection system of *John Deere Gator* mobile robot in [14] consists of a *monocular camera* and *2D lasers*. The drawback of this project is the fact that the robot is effective only when moving on a previously known path.

A system made of a *visible light camera* and a *2D LIDAR sensor* is used in [15], to obtain environmental data for object detection and obstacle trajectory identification. The developed project may be used in car parking assistance or collision warning systems.

In [16], the authors have created an automatic 3D thermal model representation system by combining *far-IR light camera Jenoptik IR-TCM 640* with *depth-sensing camera Microsoft Kinect*. In this case, the thermal information about the object is combined with the spatial information.

The stereo vision system, described in [17], comprises *IR light* and *visible light cameras*. The field of application of this system is nuclear power stations for detection of stainless steel plate deformation.

To sum up, different types of sensor combinations are used for specific applications, based on the individual sensor and the overall system advantages and drawbacks. Out of the reviewed technologies, the most often used visual sensor is the visible light camera because it provides a high detailing level. The second most frequently used sensors are IR sensors of different spectra because they provide extended data of the environment.

b) Acquisition of the Uncertainty Estimate

In computer vision systems that combine multiple visual sensors, it is necessary to pay attention to different visual sensor data fusion techniques. The paper [11] focuses on *decision abstraction level* fusion that uses *uncertainty estimates*.

The uncertainty of data can be defined in multiple ways. The paper [18] reviews the most often applied methods:

1. *Bayesian probability theory*;
2. *Dempster-Shafer theory (DST)*, also known as the *theory of belief functions*.

The Bayesian theory uses the “*probability*” concept as the strongly defined number that determines, how often the event will occur if the experiment is repeated multiple times. Belief values of DST represent the interpretation, where the statement can be true with a certain belief. In this context, the belief is subjective, that is why it is necessary to find a way to obtain the belief value automatically, while the model remains mathematically simple.

In multiple studies, the DST is used for sensor data fusion. In the beginning, it is necessary to obtain *belief* values of the fact that the target class was identified in the image [18]. In [19], the authors analyse DST and its drawbacks. The first drawback of DST is the belief modelling problem, i.e., to determine how to create the probability allocation function. There are some specific application methods, but there is no general method that could be used regardless of the problem considered. The second important issue is to determine, which method should be used to combine belief values.

To sum up, the DST is appropriate for belief value acquisition because in this case the Bayesian probability theory does not characterise the uncertainty of data in general.

c) Fusion of the Uncertainty Estimate

Data fusion at the decision level can be performed using different fusion rules. One of the common rules is the *Dempster Combination Rule (DCR)* [19].

In [20], the authors offer the method for target identification, using algorithm of Dempster belief theory function to fuse the uncertainties of IR and visible light images. The developed algorithm consists of the three stages:

1. determining the *Frame of Discernment (FoD)*;
2. obtaining *Basic Probability Assignment (BPA)* values;
3. applying the *DCR*.

At the first stage, all proposition values that are parameters of FoD, should be determined. *Basic Belief Assignment (BBA)* values are the uncertainties obtained from the proposition values. The second stage is to assign basic probability values, where the sum of all FoD elements should be equal to 1. The third stage is applying DCR, which fuses two belief values into a common probability estimate.

Fusion equation (1) [20] is used for target position detection, but the principle can also be used for target identification situation:

$$I = \frac{P(T)}{P(T)+P(B)} \cdot I_1 + \frac{P(B)}{P(B)+P(T)} \cdot I_2 \quad (1)$$

Parameters I_1 and I_2 in the formula are the target positions in the IR and visible light images. $P(T)$ and $P(B)$ are probability values for the IR and visible light images, respectively.

In [19], the authors offer an improved fusion method for the visible and far-IR light sensor data that is based on the *Conditional Update Rule (CUR)*. This method is compared with other most often used belief fusion rules – *DCR*, *Jeffrey-like Evidence Update Rule (JUR)*, *Linear Condition Update Rule (LUR)* [19]. The improved CUR method shows the improved results in the situations when the probability value has been changed fast. In these experiments, the drawback of the DCR method is the increase of decision uncertainty when the belief values are changing, but the JUR is connected only to the latest belief.

In [18], the authors propose a new method for data fusion that is based on DST – implementation of the consensus operator that is also using belief values. This consensus operator is based on the new parameter – “*opinion*”. Experiments show better results than with other methods, such as *DCR* or *Yager’s Combination Rule*.

Pixel-level fusion can also be used to combine visual sensor data. In [21], pixel-level fusion for visible and far-IR sensor data is performed, using the *Poisson fusion algorithm* to combine two image gradients. *Markov Random Field (MRF)* based model is used for the optimisation of gradient fusion. In this method, a weight map is used. The results of the method are good, but it focuses on the lowest – pixel-level fusion [21].

The analysed methods can be compared by their abstraction level, used sensors, existence of fusion weights and applied combination rule (Table II).

TABLE II
COMPARISON OF UNCERTAINTY COMBINATION RULES

ARTICLE	ABSTRACTION LEVEL	COMBINATION RULE	SENSORS	FUSION WEIGHTS
[33]	Decision level	DCR	IR and visible light	Yes
[32]	Decision level	CUR	Far-IR and visible light	No
[31]	Decision level	New combination rule	Heterogenous sensor system	No
[34]	Pixel level	Poisson fusion algorithm	Far-IR and visible light	Yes

The comparison of methods shows that most methods use decision-level fusion, which allows focusing on the qualitative fusion approach, rather than low-level image recognition. In the referenced sources, at least one of the sensors is the visible light sensor because in the vision systems it is important to have a qualitative and detailed image. Some of the methods use sensor belief weights or image weight maps that open more opportunities for sensor data fusion techniques.

Since there is no universal and generally applicable method, the next section focuses on a new and improved, practically-implemented and tested method for applications with different light conditions.

III. APPROACH OF TARGET VISUAL IDENTIFICATION

This section reviews the technical solution to the implementation of the proposed method. The technical solution comprises two visual sensors and an additional non-visual sensor. The selection of the abstraction level of data fusion is also reviewed.

Two visual sensors chosen are *visible light* and *far-IR light sensors*. The same visual sensor combination is used in the approach described in [11]. The advantage of this combination is the fusion of different types of parameters and the redundancy in different light conditions.

For both visual sensor image fusion, the highest, *decision abstraction level* is chosen. The *sensor weight* concept for image fusion is introduced. The weights are used to define which one of the sensors is more reliable in the current situation. Therefore, a sensor with the highest weight value has higher credibility and impact on the output result. The visible light sensor performance depends directly on ambient light conditions; therefore, the weight size is based on additional sensor data, for example, a *lux meter*.

To summarise, the proposed target identification approach is split into 5 sections:

- sensor data acquisition;
- illumination value determination;
- image recognition;
- weight determination;
- decision fusion.

This approach focuses on the development of the 4th and 5th sections. Figure 1 provides visual representation of the method.

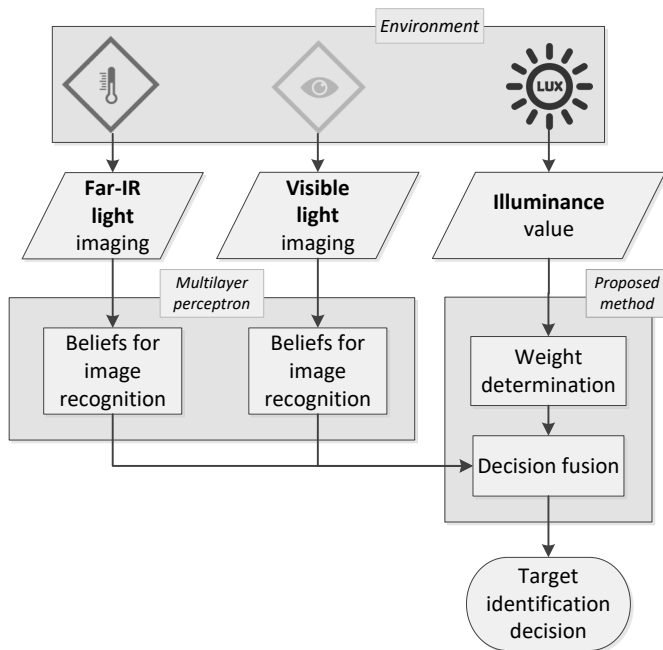


Fig. 1. The method structure for target identification.

A. The Acquisition of Fusion Weights

At first, for decision-level belief fusion, the *sensor belief weights* are obtained. The weights are determined using a function that possesses an illumination parameter from the environment.

In the proposed method, the illumination parameter is used and it is obtained using a lux meter. This additional sensor helps “switch over” between visual sensors. This is how the *credibility* of sensors is changed. For example, in the case when the lux meter detects a low illumination level, higher credibility is given to the far-IR sensor and lower – to the visible light sensor.

The Engineering Toolbox site [22] represents light level correlation and corresponding illumination values. The “sunlight” value is assumed to be the maximum possible illumination value: $LuxMax = 107\,527$ lx. The visible light sensor weight w_v is calculated using (2) by considering a maximum illumination value:

$$w_v = \frac{Lux}{LuxMax}, \quad (2)$$

where Lux is the lux meter illumination level.

The sum of both weights is always equal to 1 (3):

$$\sum w = w_t + w_v = 1. \quad (3)$$

The far-IR light sensor credibility weight w_t is calculated using (4):

$$w_t = 1 - w_v. \quad (4)$$

As a result, both credibility weight calculation equations (2) and (4) are used for visual sensor belief fusion.

B. The Fusion of Sensor Beliefs

When both visual sensor credibility weights are obtained, they need to be used in sensor belief fusion. In the proposed method, image classification is done using an artificial neural network – *multilayer perceptron (MLP)*. In the system, 2 images are perceived from different visual sensors simultaneously. The lux meter perceives illumination values at the same time. Both images are classified separately using different MLPs. As a result, two belief vectors v_v and v_t are obtained as seen in (5).

$$\begin{aligned} v_v &= \{p_{v1}, p_{v2}, p_{v3}, \dots, p_{vn}\}; \\ v_t &= \{p_{t1}, p_{t2}, p_{t3}, \dots, p_{tn}\}, \end{aligned} \quad (5)$$

where p_{vn} and p_{tn} values represent the values of belief that the image in the input belongs to the classification set n , respectively for visible and far-IR light images.

When both the sensor credibility weights and classification belief values are obtained, the resulting decision Res_n about target identified in the image is made using (6).

$$Res_n = p_{vn} \cdot w_v + p_{tn} \cdot w_t. \quad (6)$$

In contradiction to the previously mentioned systems [14], [20] and [21], the proposed approach uses additional sensors to obtain the weight factors for the used image detection sensors.

The situation when belief values are fused with the constant and equal weights ($w_t = 0.5$ and $w_v = 0.5$) is called *equal fusion*. In this case, an additional parameter from environment is not considered at all.

The field of application of this system is night vision and low visibility conditions for military scope, object recognition in the real time or localisation for mobile robotics. The proposed system is especially effective in low visibility situations, when it is necessary to identify living beings that emit thermal radiation, such as humans or animals.

IV. PRACTICAL IMPLEMENTATION

The implementation process of the system is divided into the following steps:

- reading of the sensor data;
- artificial neural network implementation;
- credibility weight calculation;
- belief fusion.

For image processing and object recognition, the *OpenCV* library is used [23]. This library is compatible with C++ programming language.

A. Implementation of Multilayer Perceptron Using OpenCV Library

One of the most often used machine learning methodologies for image recognition is an artificial neural network. That is why the developed approach of image recognition is performed using multilayer perceptron.

For the practical implementation, the already developed algorithm is taken from *Kaggle* website [24]. An additional library that is used for the solution is the *Boost v1.64.0* library set [25] for C++ language that supports linear algebra functions, random number generation, multithreading, image processing, and unit testing.

The structure of software implementation is the following:

- reading of the training set;
- extraction of image features;
- training of Bag-of-Words (BoW) [26];
- training of neural network;
- identification of target;
- fusion of belief vectors;
- evaluation of a network.

The four input parameters of the software are the following:

- visible light image directory;
- far-IR light image directory;
- neural network input layer size;
- training set proportion in relation to the testing set.

The feature extraction is performed using the *KAZE* algorithm described in [27], and feature compilation is implemented using *Bag-of-Words* strategy [24].

When features are extracted, neural network training starts. A separate neural network is trained for each visual sensor. After finishing the training of both neural networks, they are tested on test sets. The output of testing is two belief vectors, one for each visual sensor, where each element of the vector represents a belief value within the range [0; 1]. Belief values represent which target class has been recognised in the image. The sum of these beliefs in one vector equals one. The vector element with the largest belief value is chosen as the most probably recognised target class.

The next step, after the acquisition of both visual sensor belief vectors, is the fusion of these belief values to make the final decision about the recognised target. Fusion of belief values is performed after the credibility weights are obtained. In the simplest way, the weights are constant values that depend on a light condition identifier. In the testing example, the following constant weight values are taken for different light conditions:

- 1 (sunlight) – $w_v = 0.9$, $w_t = 0.1$;
- 2 (overcast day) – $w_v = 0.7$, $w_t = 0.3$;
- 3 (twilight) – $w_v = 0.3$, $w_t = 0.7$;
- 4 (overcast night) – $w_v = 0.1$, $w_t = 0.9$.

The other weight calculation equations are (2) and (4). In this case, the weight calculation is more precise and adaptation is done proportionally to the illumination value.

The next step of software is the fusion of belief vectors using credibility weight values. This is implemented based on (6). When the fusion is performed, the most probable belief value for each fused vector is chosen to be the resulting target class.

Results are combined and compared using a *confusion matrix* (Fig. 2) on the basis of the real and predicted classification values to determine the correct and incorrect classification cases.

		Predicted	
		1. class	2. class
Real	1. class	TT	FT
	2. class	FT	TT

Fig. 2. Confusion matrix for evaluation of 2 target class identification.

The number of matrix columns and rows in the table is equal to the number of target classes. The elements on the matrix diagonal represent the number of correctly identified images (*True True (TT)*), but the rest represents the number of incorrectly identified classes (*False True (FT)*).

Matrix elements are used to calculate the overall precision of testing using (7):

$$Precision = \frac{TT}{FT + TT}. \quad (7)$$

The precision evaluation result is employed to compare different fusion techniques and the truthfulness of target identification.

B. Hardware of the Solution

The used sensors are *Canon EOS 6D digital camera* and *FLIR VUE Pro far-IR light camera* [28]. *Canon EOS 6D* camera has a 20.2 megapixel sensor for visible light image acquisition, and *FLIR VUE Pro* is a thermal (far-IR) camera of professional use that is often employed for unmanned aerial vehicle (UAV) systems.

The *FLIR VUE Pro* camera is controlled by user application for *Android* or *iOS* operating systems, and it needs a *Bluetooth* connection.

Simultaneously when both visual sensors perceive images from the environment, an additional sensor – lux meter *LX-1010B* – perceives the light and returns an illumination value.

C. Image Database for Testing

An image database both for visual light and far-IR light images together with corresponding illumination values was generated for the experiment. The database has the following features:

- database contains images of multiple objects;
- each object has images in different light conditions;
- both sensors perceive images from the same angle and at the same light conditions simultaneously.

The database that contains 44 visible light and 44 far-IR light has been created according to the previously listed requirements. These images have 4 different light conditions:

1. *direct sunlight* – illumination value of 53 000 lx;
2. *overcast day* – illumination value of 420 lx;
3. *twilight* – illumination value of 0 lx;
4. *overcast night* – illumination value of 0 lx.

These illumination values are obtained using the lux meter while creating the database of images. The obtained illumination data show that real values are twice lower or even 0 lx for very low light conditions than those defined in the table [22]. These imprecisions of the illumination value data can occur due to the limitations of the lux meter device and non-linear dependency of light conditions and illumination value.

The four objects that are represented in the database are:

1. *a construction dryer* heating for 10 minutes;
2. *a hot mug*;
3. *a facial model*;
4. *a heated electric kettle* with hot water inside.

In order to test different physical parameters, such as light and temperature, the database contains objects that have both the artificially increased temperature and the natural temperature.

D. Testing of the Developed Method

Testing is done using the created database as input data. Each light condition is tested separately. Perceptron of each sensor neural network is trained to recognise 4 target classes. When the neural network training is done, 4 different target class images in the same light conditions are passed as input for the testing and the result values are combined in the belief vector. When two belief vectors of the same timestamp images from different visual sensors are obtained, they can be combined using credibility weights to create the resulting belief vector. Adapted weights that depend on illumination

values are used for testing and the case of equal fusion is also tested. These two fusion results can be compared to show the advantage of adapted fusion based on additional non-visual sensor data.

As an example, the 4th light condition – overcast night testing – is considered. When two “construction dryer” class images are passed as input, the perceptron outputs have two belief vectors – Flir (for a thermal image) and Foto (for a visible light image):

Foto 0: [0.12347877, 0.40169856, 0, 0.47482267].

Sequence of belief vector values corresponds to the target class sequence mentioned before. The highest belief value 0.47482267 in this vector is the 4th one that corresponds to the “electric kettle” class. In this case, the target recognition is false.

The belief vector of the same timestamp, but thermal image: Flir 0: [0.45739141, 0.29959834, 0, 0.24301022].

The highest belief value 0.45739141 is for the 1st class that is a “construction dryer” class. That means the recognition is performed correctly.

Both belief vectors are fused using equal fusion weights:

Output50: [0.29043508, 0.35064846, 0, 0.35891646].

In this case, the highest belief value 0.35891646 corresponds to the “electric kettle” class, and it is a false recognition again.

The second fusion is performed using adapted fusion weights:

Output: [0.45738512, 0.29960027, 0, 0.2430146].

In this case, the highest belief value 0.45738512 corresponds to the correctly recognised “construction dryer” class.

In the example, the visible light sensor did not identify the target object correctly in the low light conditions. It was identified correctly in the thermal image because even in low light conditions the object had some temperature features that made it easier to identify. Therefore, a thermal imaging sensor is a good addition to the computer vision system. The equal fusion did not identify the target object correctly, but the proposed method fusion did.

The next test is done for the “electric kettle” object class.

The results are the following:

Foto 1: [0.12347877, 0.40169856, 0, 0.47482267];

Flir 1: [0.062196929, 0, 0.055165347, 0.88263768]

Output50: [0.092837848, 0.20084928, 0.027582673, 0.67873019];

Output: [0.062198084, 7.5792182e-06, 0.055164307, 0.88262999].

The output belief vectors show that in this case the object was identified correctly in all cases, because in all vectors the highest belief value was the 4th one.

The results of the 3rd testing – a “hot mug” object class are the following:

Foto 2: [0.12347877, 0.40169856, 0, 0.47482267];

Flir 2: [0.1419034, 0.7709766, 0, 0.087120041];

Output50: [0.13269109, 0.58633757, 0, 0.28097135];

Output: [0.14190306, 0.77096963, 0, 0.087127358].

Similar to the first test, the object was not identified correctly in the visible light image, but in the thermal image it was identified correctly. Identification was also done correctly for both types of fusion.

The last test is done using “facial model” object class images:

Foto 3: [0.12347877, 0.40169856, 0, 0.47482267];
 Flir 3: [0.0092320433, 0.43331987, 0.55744815, 0];
 Output50: [0.066355407, 0.4175092, 0.27872407, 0.23741134];
 Output: [0.0092341993, 0.43331927, 0.55743766, 8.9589184e-06].

Here the correct identification appears only in the thermal image and both types of fusion.

Results of the identification can be combined in Table III where the values correspond to the real target class identification case. The coloured cells mean that the real target class belief value was not the highest one in the belief vector; therefore, the target was not identified correctly.

TABLE III
 BELIEF VALUES OF TARGET IDENTIFICATION IN THE 4TH LIGHT CONDITIONS

No.	TARGET CLASS	FOTO	FLIR	$w_1 = w_2 = 0.5$	$w_1 = 0.000019, w_2 = 0.999981$
1	Construction dryer	0.12347877	0.45739141	0.29043508	0.45738512
2	Hot mug	0.40169856	0.7709766	0.58633757	0.77096963
3	Facial model	0.00000000	0.55744815	0.27872407	0.55743766
4	Electric kettle	0.47482267	0.88263768	0.67873019	0.88262999

The resulting data are represented in the confusion matrices that allow calculating a precision value of the method.

The confusion matrix for equal fusion weights $w_1 = w_2 = 0.5$:

$$\begin{pmatrix} \text{dryer} & \text{mug} & \text{face} & \text{kettle} \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The precision value for equal fusion is 0.5 because 2 out of 4 matrix values are on the diagonal.

The confusion matrix for adapted weights $w_1 = 0.000019, w_2 = 0.999981$:

$$\begin{pmatrix} \text{dryer} & \text{mug} & \text{face} & \text{kettle} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The precision value in this case is 1 because all elements are on the diagonal of matrix.

This test is only performed for one light condition. The same tests are also done for the rest light conditions. The result precision is calculated and represented in Table IV.

TABLE IV
 THE SUMMARY OF PRECISION VALUES

LIGHT CONDITIONS	FOTO	FLIR	EQUAL FUSION	THE DEVELOPED METHOD
1	0.75	0.75	0.75	0.75
2	1.00	1.00	1.00	1.00
3	0.75	1.00	1.00	1.00
4	0.25	1.00	0.50	1.00

The table of resulting precision values shows that the precision of target identification of the developed method is always equal or higher in comparison with an equal weight fusion method. The improvement is observed in the low light conditions.

The parameters that could be changed in the testing are neural network input layer size, size of training and testing sets and multiple different light conditions could be viewed. In the alternative situation, the perceptron can be trained only by using images of good quality and then testing can be done to images of any light conditions.

V. CONCLUSION

In the paper, the existing methods of decision fusion and their drawbacks in different light conditions have been analysed. A new and improved method for target identification has been developed based on the current situation analysis. The proposed method has been implemented and its results presented using a generated image database comprising images of different objects in different spectra.

The case of target identification without data fusion has been compared to the equal and weighted fusion results. The weighted fusion shows equal or higher precision in comparison with the equal and no fusion cases. Improvement has been observed especially in the low light level conditions that proves the enhancement of the target identification precision.

ACKNOWLEDGMENTS

The present research has been supported by the European Regional Development Fund project “The Competency Centre of the Latvian Electric and Optical Equipment Productive Industry”.

REFERENCES

- [1] K. Lenac, I. Maurovic, and I. Petrovic, “Moving Objects Detection Using a Thermal Camera and IMU on a Vehicle,” in *2015 International Conference on Electrical Drives and Power Electronics (EDPE)*, pp. 212–219, 2015. <https://doi.org/10.1109/edpe.2015.7325296>
- [2] S. Ghorashi, “Spatial Selection and Target Identification are Separable Processes in Visual Search,” *Journal of Vision.*, vol. 10, no. 3, pp. 1–12, 2010. <https://doi.org/10.1167/10.3.7>
- [3] B. Bhanu, “Automatic Target Recognition: State of the Art Survey,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, no. 4, pp. 364–379, 1986. <https://doi.org/10.1109/taes.1986.310772>
- [4] R. E. Bethel, B. Shapo, and C. M. Kreucher, *PDF Target Detection and Tracking, Signal Processing*, vol. 90, no. 7, pp. 2164–2176, 2010.

- [5] "Optics and photonics. Spectral bands," 2007. <https://doi.org/10.3403/30112984>
- [6] I. N. Perevozčikov, *Multisensornaja Tehnologija Izmerenija Geometričeskikh Parametrov Tipovyh Detalej*, 2009.
- [7] H. Grünbergs, *3D Punktu Mākoņa Apstrāde Objektu Izsekošanai Telpā*. Rīgas Tehniskā Universitāte, 2014.
- [8] S. Rodríguez-Jiménez, N. Burrus, and M. Abderrahim, "A-Contrario Detection of Aerial Target Using a Time-of-Flight Camera," *Sensor Signal Processing for Defence (SSPD 2012)*, 2012. <https://doi.org/10.1049/ic.2012.0110>
- [9] A. Canclini, L. Baroffio, M. Cesana, A. Redondi, and M. Tagliasacchi "Object recognition in visual sensor networks based on compression and transmission of binary local features," [Online]. Available: <http://www.icassp2014.org/files/ST/1569921811%20-%20Object%20recognition%20in%20visual%20sensor%20networks.pdf>
- [10] F. Castanedo, "A Review of Data Fusion Techniques," *The Scientific World Journal*, vol. 2013, pp. 1–19, 2013. <https://doi.org/10.1155/2013/704504>
- [11] J. Thomanek, M. Ritter, H. Lietz, and G. Wanielik, "Comparing Visual Data Fusion Techniques Using FIR and Visible Light Sensors to Improve Pedestrian Detection," *2011 International Conference on Digital Image Computing: Techniques and Applications*, pp. 119–125, Dec. 2011. <https://doi.org/10.1109/dicta.2011.27>
- [12] M. Kristou, A. Ohya, and S. Yuta, "Target Person Identification and Following Based on Omnidirectional Camera and LRF Data Fusion," *2011 RO-MAN*, pp. 419–424, Jul. 2011. <https://doi.org/10.1109/roman.2011.6005248>
- [13] T. Nakamura, S. Haviland, D. Bershadsky, D. Magree, and E. N. Johnson, "Vision-Based Closed-Loop Tracking Using Micro Air Vehicles," *2016 IEEE Aerospace Conference*, pp. 1–12, Mar. 2016. <https://doi.org/10.1109/aero.2016.7500873>
- [14] A. Rechy Romero, P. V. Koerich Borges, A. Elfes, and A. Pfrunder, "Environment-Aware Sensor Fusion for Obstacle Detection," *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 19–21, Sep. 2016. <https://doi.org/10.1109/mfi.2016.7849476>
- [15] P.-M. Hsu, M.-H. Li, and Y.-F. Su, "Object detection and recognition by using sensor fusion," *11th IEEE International Conference on Control & Automation (ICCA)*, pp. 56–60, Jun. 2014. <https://doi.org/10.1109/icca.2014.6870895>
- [16] S. Soldan, J. Rangel, and A. Kroll, "3D Thermal Imaging: Fusion of Thermography and Depth Cameras," 2014 [Online]. Available: www.ndt.net/?id=17665
- [17] A. Charbal, J. E. Dufour, F. Hild, M. Poncelet, L. Vincent, and S. Roux, "Hybrid Stereocorrelation Using Infrared and Visible Light Cameras," *Experimental Mechanics*, vol. 56, no. 5, pp. 845–860, Jan. 2016. <https://doi.org/10.1007/s11340-016-0127-4>
- [18] S. Acharya and M. Kam, "Evidence Combination for Hard and Soft Sensor Data Fusion," *14th International Conference on Information Fusion Chicago, Illinois, USA, July 5–8, 2011* [Online]. Available: http://fusion.isif.org/proceedings/Fusion_2011/data/papers/133.pdf
- [19] X. Fan, P. Shi, J. Ni, and M. Li, "A Thermal Infrared and Visible Images Fusion Based Approach for Multitarget Detection under Complex Environment," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–11, 2015. <https://doi.org/10.1155/2015/750708>
- [20] A. Wang, J. Jiang, and H. Zhang, "Multi-sensor Image Decision Level Fusion Detection Algorithm Based on D-S Evidence Theory," *2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control*, pp. 620–623, Sep. 2014. <https://doi.org/10.1109/imccc.2014.132>
- [21] J. Sun, H. Zhu, Z. Xu, and C. Han, "Poisson image fusion based on Markov random field fusion model," *Information Fusion*, vol. 14, no. 3, pp. 241–254, Jul. 2013. <https://doi.org/10.1016/j.inffus.2012.07.003>
- [22] Engineering Toolbox team, "Illuminance – Recommended Light Level," 2017 [Online]. Available: www.engineeringtoolbox.com
- [23] OpenCV team, "About OpenCV," 2017 [Online]. Available: <http://opencv.org/about.html>
- [24] A. Mathews, "Is It a Cat or Dog? A Neural Network Application in OpenCV," 2017 [Online]. Available: <https://picodelimiao.github.io/blog/2016/01/31/is-it-a-cat-or-dog-a-neural-network-application-in-opencv>
- [25] R. Rivera, "Boost C++ Libraries," 2014 [Online]. Available: <http://www.boost.org>
- [26] A. Hota, "Comparison of some Bag-of-Words models for image recognition," *2014 X International Symposium on Telecommunications (BIHTEL)*, pp. 1–5, Oct. 2014. <https://doi.org/10.1109/bihitel.2014.6987648>
- [27] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE Features," *Lecture Notes in Computer Science*, pp. 214–227, 2012. https://doi.org/10.1007/978-3-642-33783-3_16
- [28] I. FLIR Systems, "FLIR Vue Pro and Vue Pro R User Guide," 2016 [Online]. Available: <http://www.flir.com/uploadedFiles/sUAS/Products/Vue-Pro/FLIR-Vue-Pro-and-R-User-Guide.pdf>



E-mail: anete.vagale@edu.rtu.lv

Anete Vagale received the Master's degree in Intelligent Robotic Systems from Riga Technical University, Latvia, in 2017.

She is a Doctoral student of the study programme "Computer Systems" and works as a Research Assistant at the Department of Artificial Intelligence and System Engineering. She is developing her Doctoral Thesis related to visual sensor data fusion and complicated decision making for object identification purpose. Her research interests include human-robot interaction, artificial intelligence and robot vision.



for his Doctoral Thesis in 2006. At the moment, he is a member of IEEE and ACM as well as represents Latvia in NATO STO AVT panel.

E-mail: agris.nikitenko@rtu.lv

Agris Nīkitenko has received his *Dr. sc. ing.* degree in 2006 from Riga Technical University, as well as *B. sc. ing.* and *M. sc. ing.* degrees focusing on artificial intelligence in his theses.

At present, he is a Professor at the Department of Artificial Intelligence and Systems Engineering of Riga Technical University and Vice Dean of Academic Affairs at the Faculty of Computer Science and Information Technology. His scientific interests cover artificial intelligence and autonomy as well as their application in robotics.

He has received an award of Verner fon Siemens



Eduards Slava received the Master's degree in Intelligent Robotic Systems from Riga Technical University, Latvia, in 2017.

He is a Doctoral student of the study programme "Computer Systems" and works as a Robotics Engineer. His Master's Thesis is related to object recognition and semantics using deep neural networks.

E-mail: eduards.slava@gmail.com



Ottar L. Osen graduated with MSc in Cybernetics from the Norwegian Institute of Technology in 1991 (NTH, now the Norwegian University of Science and Technology (NTNU)).

He is the Head of R&D at ICD Software AS and an Associate Professor at NTNU in Ålesund.

In addition to 12 years in academia, he has worked for 15 years as a Consultant in the maritime and offshore industry both in Norway and internationally.

He has taken part in the engineering phase on several large offshore projects on the Norwegian continental shelf. He has broad experience in subsea control systems and SCADA systems in general.

His research interests belong to the fields of artificial intelligence, cybernetics, instrumentation, real-time systems, industrial control systems, mechatronics, modelling and simulation.

E-mail: ottar.l.osen@ntnu.no