

Automated Wildlife Recognition

Josef Pavlíček¹, Jan Jarolímek¹, Jiří Jarolímek², Petra Pavlíčková¹, Stanislav Dvořák³, Jan Pavlík¹, Petr Hanzlík¹

¹ Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

² Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Czech Republic

³ Military Forests and Farms, state-owned enterprise, division Karlovy Vary, Czech Republic

Abstract

The estimation of wildlife populations is an issue currently being solved at workplaces on many levels. Knowledge of wildlife population and localization is not only very important for reducing damage to agricultural and forest growth, which arises from the local overgrowth of certain animal species, but also for the protection of endangered species of animals and plants.

The article presents the results of a research carried out during 2017 as the first partial objective of a complex automated wildlife estimation project, namely the recognition of game in a free landscape without vegetation cover from an UAV (unmanned aerial vehicle). The paper describes a method of finding game animals in a selected area and identifies problems with the recognition of the animals hiding in the vegetation. These results play an important role in solving the overall complex problem of automated game recognition.

Keywords

Wildlife, automated recognition, picture segmentation, neural networks, mimicry, false positive recognition.

Pavlíček, J., Jarolímek, J., Jarolímek, J., Pavlíčková, P., Dvořák, S., Pavlík, J. and Hanzlík, P. (2018) "Automated Wildlife Recognition", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 1, pp. 51-60. ISSN 1804-1930. DOI 10.7160/aol.2018.100105.

Introduction

Various efforts to accurately estimate the population of game are centuries old. Nowadays, the census results are the basis for determining the amount of hunting needed to ensure stable game numbers. In most European countries, including the Czech Republic, the population of hooved animals has increased in recent decades, causing more and more damage to forest and field cultures (Bartoš et al., 2010). In order for the population estimates to properly function as a control method, its results should roughly correspond with reality. At present day, most commonly used methods in the Czech Republic account only for 10-33% of the actual population. The accuracy of estimates of game conditions is eloquently evidenced by the comparisons of the spring population with the number of hunted animals (Bartoš et al., 2005). According to statistics from some countries, sometimes more animals were caught than the amount estimated by the census for the whole population, which is also the case in the Czech Republic. That is why new alternative and more efficient methods are being sought.

Better results can be usually achieved using more powerful equipment, but its use alone does not guarantee the quality of outputs. Various methods, such as telemetry monitoring, are used to track game, which are particularly useful for research on game ethology, but are not well suited for population estimation (Jarolímek et al., 2014; Masner et al., 2014). The first findings of aviation census have been published more than forty years ago (Graves et al., 1972). Estimating game population from aircraft or helicopter is practiced, for example, in the Scandinavian countries (Liberg et al., 2010). Thermo-vision is also used frequently (Gill et al., 1997, Focardi et al., 2001 and others), but mostly only for ground imagery. In contrast, mainly in the US and Canada, the use of thermo-vision is expanding not only in land censuses, but also in aerial imaging. The results of monitoring of various animal species in various environments (Wyatt et al., 1980; Bayliss and Yeomans 1989, Wiggers and Beckerman 1993, Focardi et al. 2001, Garel et al., 2010, Fuentes et al., 2015) were published.

The current development of the use of autonomous flying devices and Artificial Intelligence tools for image evaluation brings a new dimension for the use of "aerial" sensing and game monitoring methods.

Unmanned vehicles are nowadays, mainly due to the massive expansion of the so-called multi-copters (multiple motor helicopters), known mainly by the term "drones". Officially, however, the term UAV (Unmanned Aerial Vehicle) is used. Unmanned vehicles offer a variety of new applications by combining air and ground imagery. At the same time, it brings new methods of retrieving data from selected areas in real time. Some types of UAVs are capable of covering an area of several square kilometres, making them cheaper and more affordable alternative to conventional aircraft. Due to the lower scanning height, one can also get very detailed images from the unmanned vehicle (Eisenbeiss, 2011). In addition to capturing images, monitoring can also be performed „on the fly“, when video is transmitted to the operator screen but not recorded.

Automated image recognition is one of the most important technologies at the moment. It is generally referred to as machine vision. Although machine vision is already very elaborate (Gonzales and Woods, 2002), its practical application provides a number of technical problems. These must always be addressed specifically for a given task. Image processing of a recognized object consists of a series of steps. First, it is needed to capture and digitize the image and then use the image pre-processing method to improve the image, which is especially focused on grayscale, brightness and contrast adjustment, histogram equalization, image sharpening, and various filtration methods. Another important step is to use segmentation methods to distinguish a recognized object from the background. It is primarily segmentation by thresholding, image dyeing algorithms, edge detection and linking methods and various algorithms for object filling. Usually, the gradient change of pixel brightness is used. After the image processing is completed, the next phase, the description of the object, follows. The most well-known methods of object description are the moment method, Fourier descriptors and chain codes, which can also be used for so-called structural description of objects. The final stage of the image processing is the object's classification (recognition). The task of classification is to include objects found in the image in a group of previously known classes

(Parker, 2011). The object recognition itself can be realized using artificial intelligence methods or statistical analysis (or by combining them). Typically, the acquired description of the object of a so-called classifier will be presented, which, with a certain degree of accuracy, can determine what the object is. The classifier is familiar with the objects that can be submitted to it. This process is called learning.

A particular example is the SIFT (Scale-Invariant Feature Transform) method, which was first used to detect objects in a picture scene. According to Noviyanto and Arymurthy (2013), the identification of cattle shows that the SIFT method has the best results. From the training set of images, object vectors were calculated, which were subsequently searched in test pictures. If the vectors obtained during training and testing were sufficiently matched, the object was detected and recognized at the same time. However, this principle can be equally well used in classification. From the training sets (one for each class), the flag vectors are obtained by the algorithm and are then compared with the vectors counted for the test set. In the next step, using the selected classification method, it is decided to divide the elements of the test set into individual classes. SIFT consists of four main steps: (1) detection of extremes within scale-space; (2) refining the location of significant points; (3) assigning orientation to significant points; (4) compiling a descriptor of significant points (Lowe 1999 and Lowe 2004).

Yu, Wang and Kays (2013) have published an analysis that shows that the combination of SIFT and cLBP (compound local binary pattern) can serve as a useful technique for recognizing animals in real complex situations. They use sparse coding spatial pyramid matching (ScSPM), which extracts dense SIFT descriptors and mobile-structured LBP (cLBP) as a local function that generates global functions via weighted sparse encoding and max pooling using the multi-scale kernel pyramid and sorts images according to the linear support vector machine algorithm.

The photo (or digital stream) is actually a two dimensional array of points that have a defined light intensity (Gonzales, 2002). These points are called "pixels" and their value is most often given in RGB. The captured object is then recorded as an array of points with a defined intensity due to lighting (Gonzales, 2002). The linking of points then creates the final image in the human brain (Russ, 2008). This image can be a real object

display, or non-existing object, meaning the brain is hallucinating. The hallucination state is the state in which the displayed object is interpreted in a human brain differently than the image producer should be. Hallucinatory states are unwanted deviations in the interpretation of the reality model display. However, they are useful tools for people (and probably for living organisms) just to filter the image and perceive reality. Hallucination is the misinterpretation of reality given by the property of the brain to generalize the presented pattern, place it in the context of a known (and logically valid reality) and anchor it. Thanks to this generalization feature, one can drive the car and concentrate on the road but not see its structure (asphalt stones, colour differences, etc.). Unnecessary information is filtered out. In terms of image processing and object recognition the equivalent to such hallucinations is the false positive recognition of an object that actually isn't there.

Therefore, for the purpose of the article, the fundamental differences between the vision of humans and machines must be taken into account. A person is able to generalize and anchor the image in a given reality and link it to the context. This means that human brain not only analyzes the shape, size and colour of an object and compares it to an abstract image of given object type (while taking into account possible colour variants, usual size ranges and general shape), but also its link to the environment (objects outside / inside, which room is the object in, is it on the ground or on a table and so on). The image is therefore interpreted as a real instance of an abstract object in the context

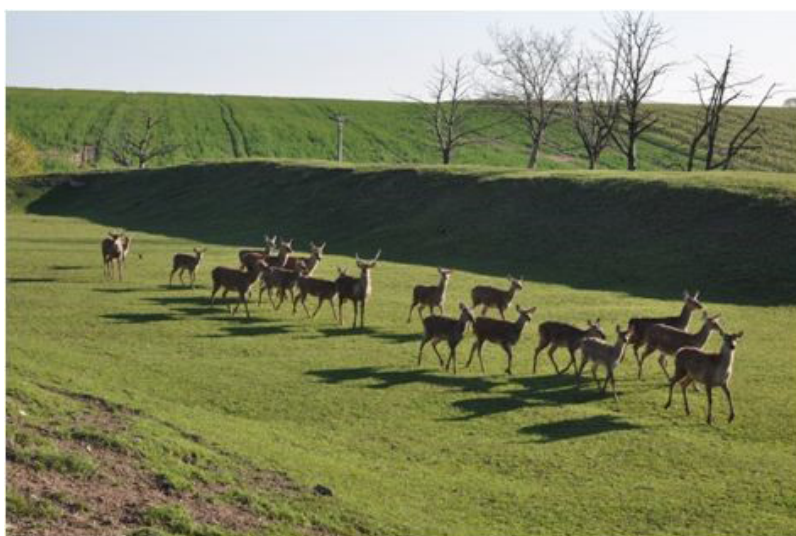
of the environment in which the observer lives. While a machine can be taught what an object looks like by supplying large sample of images, it is extremely difficult for the machine to learn how to use the context portion of vision properly.

According to (Gonzales, 2002), (Russ 2008) the machine does not see the same way humans do. The machine sees a matrix of pixels of varying intensity. For the machine the circle is not "round" but it is a vector of pixel coordinates of similar intensity which are the same distance from the defined centre. The reality the machine sees is diametrically different from humans. It is, in fact, similar to the thinking of an engineer who uses CAD tools to construct devices in "curves". Such a person transforms from a real image into a vector space in order to achieve their goal. In order for the machine to see what needs to be seen by it, it is necessary to teach it and give it the context of the problem. This is the most basic problem of computer vision and is it being tackled with a variety of successes by a number of research teams.

Material a methods

The first partial goal of a complex automated game estimation project is to recognize game in an open area without a vegetation cover from a UAV. To obtain the baseline images, the stag and fallow deer animals were optically scanned in the farm (Figure 1).

To fulfil this objective requires a synchronization



Source: authors' own processing

Figure 1: Environment and animals used to retrieve the baseline images.

of a whole range of tasks that the machine must perform. The basic steps are shown in the diagram (Figure 2).

Stream cutting

It serves to divide the data stream into the sequentially following images (Figure 3). This division is an easy task for machine image processing. Of course, the image recognition algorithm can also be applied to a data stream (which is a sequence of images). However, due to the algorithm's tuning and other practical reasons, it is better to work with static photos.

Image restoration

At this stage, cleaning the photo from "normal noises", such as changes in brightness, takes place. These are due to light conditions changing during recording (change in light intensity). By balancing the brightness and colour layers, it is ensured that the algorithm processing the photograph has the necessary properties to perform segmentation.

Image segmentation

Image segmentation is a basic task of image processing. It separates unnecessary "noise" objects from a photograph and creates a favourite image to be processed.

Assuming the task is to search for a herd of animals

in the open air, the following knowledge base serves as a starting point:

- animals are in a free space, without growing vegetation
- the target are hooved animals (deer)
- the target is a herd - not an individual
- animals are scanned from a flying autonomous machine from a height of about 50 m

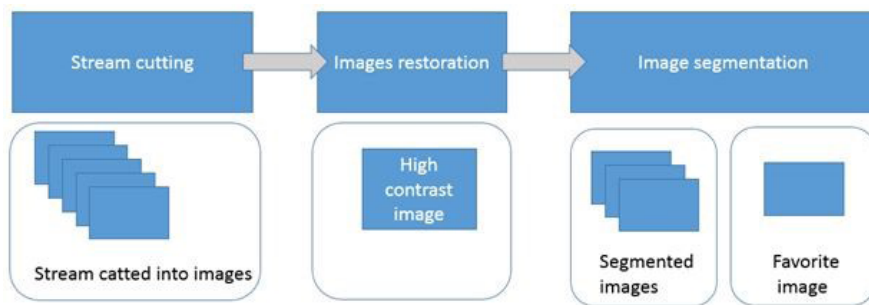
The task is to localize a herd on a pasture based on the data stream of images, see Figure 4.

To be able to locate the object, it was first needed:

- Select a favourite area (based on the knowledge base)
- Search the favourite area to for objects that resemble the animals (or herd)

Selecting the favourite area

Selecting the favourite area will be done by means of thresholding. First, it is necessary to find a suitable threshold (or some threshold function) that will allow to separate (de-segment) "noise objects". Such a noise object is probably a road, a building, or the area with corn and fruit trees. Favored areas are green surfaces with a smooth terrain, see Figure 5.



Source: authors' own processing

Figure 2: Flow diagram of data stream processing.



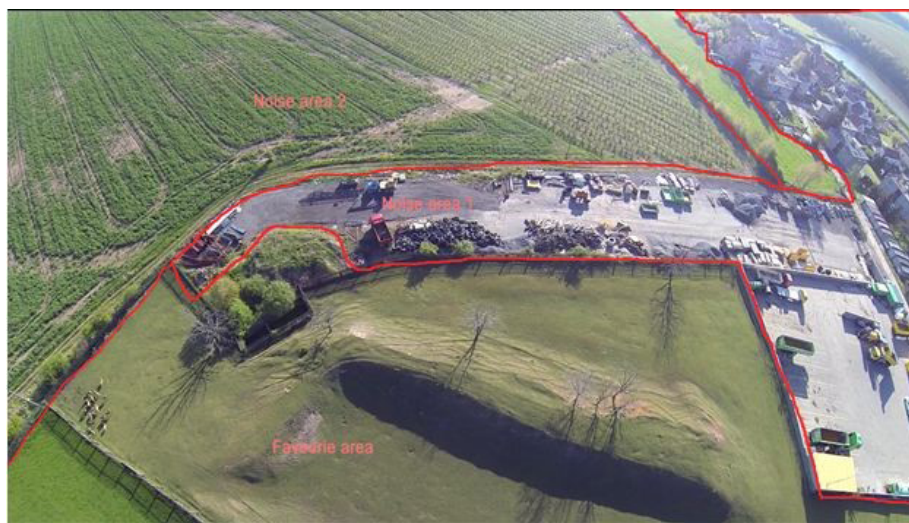
Source: authors' own processing

Figure 3: Video stream split into separate images.



Source: authors' own processing

Figure 4: An example of target object - herd.



Source: authors' own processing

Figure 5: Picture of areas that need to be machine-separated (the areas have been marked manually for the purpose of the article)

So how to choose the right method of selecting a given area? In our case, one of the segmentation methods that takes into account the structure of the search area had to be used. This „structural filter“ is then applied to the entire image. The chosen method (for simplicity and applicability) will be:

- Creating mask sized $A \times B$ pixels (e.g. $A = 10$, $B = 10$, resulting in 10×10 pixels)
- This mask will be moved through the image and its average RGB value will be evaluated. In the case of the model submitted, it will be the average RGB value in the interval:
- $R \in \langle 100, 140 \rangle$, $G \in \langle 110, 125 \rangle$, $B \in \langle 80, 90 \rangle$
- The area selected by this method will then be the favourite selected object. Inside this favourite object (pasture) there will be

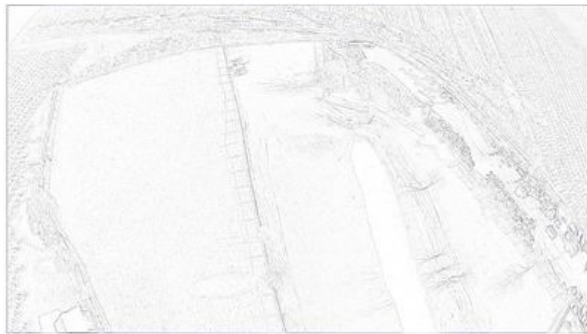
objects of interest (animals, herd). It should be remembered that the size of the mask influences how precisely the machine selects the area. At the same time, the filter must not be too strict. Otherwise, it would discard the area with the animals as well, which would render the whole task pointless.

- For better image processing, the input image is processed by the Difference of Gaussians edge detector.
- Resulting images are merged.



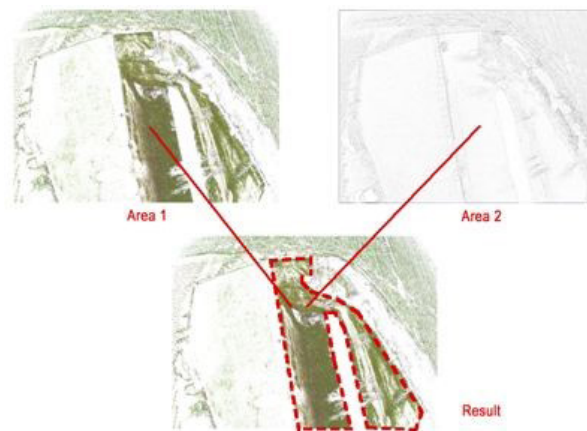
Source: authors' own processing

Figure 6: Areas selected by the RGB mask.



Source: authors' own processing

Figure 7: Area processed by an edge detector.



Source: authors' own processing

Figure 8: Resulting favourite object.

Finding game in the selected area

Finding game in the selected area is a key task. It is important to realize, that the animals try to mask optically and are therefore very difficult to be seen for a number of predators. The human eye is an excellent perceptual organ, but the interpretation of the seen image is performed by the brain. In computer vision, the intelligence of Homo Sapiens and his excellent eyesight (in the case of dog breeds probably an excellent sense of smell) is missing. It is therefore necessary to provide

the machine with the necessary context of the problem and to work within the limitations of computer vision (with limited resolution quality).

What is the context:

- a favourite area is where the game should reside (may or may not be there)
- in the favourite area, the game may be split into groups (creating herds, meaning there will be a neighbouring of objects of a similar type)
- the herd will probably move (partially due to the noise and presence of the UAV)



Source: authors' own processing

Figure 9: Detail of animals in the open.

Locating game can be done in several ways:

- According to shape properties of found objects
- By detecting objects using Artificial Intelligence tools
- According to colour profile

In our case, the simplest option by working with the colour profile of the object was used. Each animal is a light object on a darker surface. If this object moves (if the animal is in motion) then its localization is relatively easy. The goal is to recognize the favourite objects on a video stream and the change of the centre point of favourite objects (possible herd movement). Each favourite object found has its own area, which must be larger than the minimum size (to distinguish, for example, the stones in the pasture that may have a similar colour when viewed from a high angle in direct sunlight) and at the same time smaller than the maximum size. The animals on the pasture will then have a more pronounced colour intensity than the average pasture area (area). Such a moving object is probably a sought-after animal, or an object of similar size and colour (a herding

dog for example). This method prevents counting of other objects, such as a maintenance car, wooden shed, etc., resulting in reduced amount of false positive recognitions.

Results and discussion

The problem of game recognition can be divided into two basic types of tasks. Recognizing an object (animal) in an area where it is likely to occur and to distinguish the actual animal species. The first type of task simplifies the problem of just finding an object in a favourite area. This step, however, is necessary prerequisite for the second task, i.e. the recognition of the species. Due to the limits in available technology, it was decided to focus on the first task. Similar problems were also investigated by different teams (Hanzlik et al., 2014), (Pavlickova et al., 2017) who use two different approaches. Algorithmic approach, when machines work virtually based on the brightness of the image and on its threshold. Here, research teams have to successfully solve image segmentation, i.e. separating uninteresting areas from interesting (favourite). Such a task is fairly simple in laboratory conditions when the machine scans an object ideally lit by artificial light in a clean environment. These applications are currently being used in industry and, although not without a complications, their applications in general are more or less successful (e.g. face recognition, smile recognition by Apple, Congitech etc.).

Other applications are in the automotive industry and in microelectronics (smart cars track obstacles at different angles around the car - VW, Škoda). Another approach is based on artificial intelligence methods. To a large extent, Google is currently contributing with TensorFlow convolutional neural network technology. It is designed for the task of finding typical images in the data stream, and its application for agriculture has not yet been explored. Hanzlik and Pavlíček research (Hanzlik et al., 2014) is a successful attempt within the area of agriculture. The authors worked with perceptron neural networks and convolutional neural networks. The disadvantage of this approach is the need to carry out a series of learning cycles (not to mention the need for a huge amount of learning data - and these are often absent), which are very slow and good result at the end of the learning cycle is not guaranteed. Currently, there is no satisfactory method to teach the networks to achieve a certain result. Therefore, their use is still experimental and is being developed in areas where these properties do not matter.

Recognizing game is very complicated. Unless one builds on facts such as the animal size and the ideal free space, but rather try to recognize the game through a computer vision in its natural environment, there are a number of struggles. The primary problem is, of course, the mimicry of the animals. They are very successful in hiding from predators by their camouflage, which is of course especially effective in their natural



Source: authors' own processing

Figure 10: Animal mimicry and its machine recognition.

habitat. Finding such an object with computer vision is possible on experimental images (and with help from humans), but automated search is very complicated. Most machines lack important knowledge of the search context and are producing high amount of false positives. In the case of the presented image, the apparent animal is not revealed at first sight because the machine is unable to distinguish it from the surroundings thanks to the mimicry. Edge or colour-based segmentation methods fail. Although for every image you will eventually find a suitable filter and more or less recognize the animals, in real conditions, the changes in light intensity and therefore the change of colour shade will cause the machine to be very inaccurate and virtually useless. A lying stone may be misinterpreted as the typical white spot on the animals behind thus resulting in false positive identification. This is not unusual in nature itself. Dog-like predators have a solid sight, which makes it possible for them to locate moving objects (and the machine can successfully detect moving objects due to a change in their relative coordinates), but to find lying animals they use smell. Since machine is relying purely on vision, successful recognition is much more complicated.

Although Artificial Intelligence tools promise to eliminate the context issue and are more or less trying to simulate the functioning of the human brain (similarly to how one actually sees - or interprets the seen objects as a human), this research path is still lengthy. The paper showcases really solid results in game location based on image segmentation and combination of motion (using statistical methods) and the knowledge context which was used to look for the animals. The experiment proved that with an autonomous machine, a herd can be found with a high degree of accuracy. The issue is, that if a herd is in its natural environment, it is complicated to find it, if it is not in motion. This problem is the logical consequence of animal mimicry and without the proper olfactory equipment of the machine, the task of proper recognition using only sight is challenging. It does not mean, however, that the herd cannot be found at all. If it is not physically covered by any object (trees, rocks etc.) it has been proven that the herd location in the open air is possible and relatively successful. However, it is not possible to count the number of individual animals with a high degree of accuracy. Conversely, according to 16 record streams, a relatively accurate approximate count can be established for a herd in motion. Thanks to the movement of the flying machine, it is possible to create a series of pictures of the herd. These

images can be grouped together using triangulation methods and be appropriated to measuring points (or surfaces). These points can be, for example, a tree, a corner of fencing, a roof of a building, a road, a river, or a parking area (see Figure 5). A suitable triangulation point is a point that is lonely (its surroundings are visually monotonous) and its appearance is unique enough. For this reason, it is advisable to select three distinct objects (such as tree in a pasture, a pond and a field fence) so that their location is clearly visible in multiple frames. Moving animals are then possible to be recognised based on machine comparison of these pictures. As part of the experiment, a unique identifier was assigned to each discrete cluster of pixels with the appropriate brightness and colour. If this cluster was in motion, it was an animal. If the cluster divided, it was an animal that was covered by another animal. The reverse (animal entering cover) corresponds with two clusters merging. This method proved to be suitable, but only in the experimental conditions, it suffers from quite lengthy data preparation and complicated creation of a suitable cluster recognition algorithm. Now the task of our research team is to convert this method from laboratory models into readily applied solutions.

Conclusion

The described method was tested during the summer of 2017. Based on the calibration of 16 video streams by the operator (finding objects - animals on the pasture), the designed algorithm has proven to be very satisfactory. Machine hallucination occurs approximately on one image in 10 minutes, which is negligible due to the large number of images in the overall video stream. Using a simple statistical method, it is possible to distinguish the deviations in the number of animals found and eliminate those errors. There is, of course, a big problem if the game is hidden by a bush, a tree, or it is hidden behind one another. Another problem are the "long suns" when the shade of the bodies cover other animals and the machine cannot distinguish between them. Either it counts a small amount, or it links them to a single and then discards it, thanks to the maximum object size limit. If the herd is in motion, the ability of the machine to count precisely is almost 100 percent, assuming all animals are moving and their bodies are not overlapping. This problem can be mitigated by the use of statistical method. However, if the animals are stationary or lying, then the recognition becomes very problematic. Tackling

this issue is the subject of another planned stage of our research.

Acknowledgements

The results and knowledge included herein have been obtained owing to support from the following

institutional grants. Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences in Prague, grant no. 20171016, „Možnosti využití metod rozpoznávání obrazu z výškového snímkování pro monitoring zvěře“ (Utilization of image recognition methods for aerial wildlife monitoring).

Corresponding authors:

Ing. Josef Pavlíček, Ph.D.

Faculty of Economics and Management, Czech University of Life Sciences Prague

Kamýcká 129, 165 00 Prague - Suchbát, Czech Republic

Phone: 00420 224 385 475, E-mail: pavlicek@pef.czu.cz

References

- [1] Bartoš, L., Kotrba, R. and Pintíř, J. (2010) „Ungulates and their management in the Czech Republic“, in: Apollonio, M., Andersen, R. and Putman, R. (eds.) „*European Ungulates and their Management in the 21st century*“, Cambridge University Press, London, UK, pp. 243 - 261. ISBN 978-0-521-76061-4.
- [2] Bartoš, L., Kotrba, R., Pluháček, J. and Dušek, A. (2005) „Sčítání zvěře termovizní metodou srovnávací studie publikovaných zkušeností“ (in Czech), Research study for Lesy ČR, Prague.
- [3] Bayliss, P. and Yeomans, K. M. (1989) „Correcting bias in aerial survey population estimates of feral livestock in northern Australia using the double-count technique“, *Journal of Applied Ecology*, Vol. 26, pp. 925-933. E-ISSN 1365-2664. DOI 10.2307/2403702.
- [4] Eisenbeiss, H. (2011) „*The Potential of Unmanned Aerial Vehicles for Mapping*“. [Online]. Available: <http://www.ifp.uni-stuttgart.de/publications/phowo11/140Eisenbeiss.pdf>. [Accessed: 20 Jan. 2018].
- [5] Focardi, S., De Marinis, A. M., Rizzotto, M. and Pucci, A. (2001) „Comparative evaluation of thermal infrared imaging and spotlighting to survey wildlife“, *Wildlife Society Bulletin*, Vol. 29, No. 1, pp. 133-139. E-ISSN 1938-5463.
- [6] Fuentes, M. M. P. B., Bell, I., Hagihara, R., Hamann, M., Hazel, J., Huth, A., Seminoff, J. A., Sobtchik, S. and Marsh, H. (2015) „Improving in-water estimates of marine turtle abundance by adjusting aerial survey counts for perception and availability biases“, *Journal of Experimental Marine Biology and Ecology*, Vol. 471, pp. 77-83. ISSN 0022-0981. DOI 10.1016/j.jembe.2015.05.003.
- [7] Garel, M., Bonenfant, C., Hamann, J. L., Klein, F. and Gaillard, J. M. (2010) „Are abundance indices derived from spotlight counts reliable to monitor red deer *Cervus elaphus* populations?“, *Wildlife Biology*, Vol. 16, No. 1, pp. 77-84. ISSN 0909-6396. DOI 10.2981/09-022.
- [8] Gill, R. M. A., Thomas, M. L. and Stocker, D. (1997) „The use of portable thermal imaging for estimating deer population density in forest habitats“, *Journal of Applied Ecology*. Vol. 34, pp. 1273-1286. E-ISSN 1365-2664. DOI 10.2307/2405237.
- [9] Gonzales, R. C. and Woods, R. E. (2002) „Digital image processing“, 2nd edition. ISBN 9780201180756.
- [10] Graves, H. B., Bellis, E. D. and Knuth, W. N. (1972) „Censusing white-tailed deer by airborne thermal infrared imagery“, *Journal of Wildlife Management*, Vol. 36, No 2, pp. 875-884. E-ISSN 1937-2817. DOI 10.2307/3799443.
- [11] Hanzlik, P., Kozisek, F. and Pavlicek, J. (2006) „Decision Support Systems, Intelligent Agent, Knowledge engine, Agricultural DSS“ [Online]. Available: <http://www.naun.org/main/NAUN/mcs/2015/a322002-233.pdf> [Accessed: 15 Jan. 2018].

-
- [12] Jarolimek, J., Vanek, J., Jezek, M., Masner, J. and Stoces, M. (2014) „The telemetric tracking of wild boar as a tool for field crops damage limitation“, *Plant, Soil and Environment*, Vol. 60, No. 9, pp. 418-425. ISSN 1214-1178.
 - [13] Liberg, O., Bergström, R., Kindberg, J. and Von Essen, H. (2010) „Ungulates and their management in Sweden“, in: Apollonio, M., Andersen, R. and Putmann, R. eds.: „*European Ungulates and their Management in the 21st century*“, Cambridge University Press, London, UK, pp. 37-70. ISBN: 978-0-521-76061-4.
 - [14] Lowe, D. G. (2004) „Distinctive image features from scale-invariant keypoints“, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110. E-ISSN 1573-1405, ISSN 0920-5691. DOI 10.1023/B:VISI.0000029664.99615.94.
 - [15] Lowe, D. G. (1999) „Object recognition from local scale-invariant features“, *International Conference on Computer Vision*, Corfu, Greece, pp. 1150-1157. ISBN 0-7695-0164-8.
 - [16] Masner, J., Vanek, J. and Stoces, M. (2014) „Spatial Data Monitoring and Mobile Applications – Comparison of Methods for Parsing JSON in Android Operating System“, *AGRIS on-line Papers in Economics and Informatics*, Vol. 6, No. 1, pp. 37-46. ISSN 1804-1930.
 - [17] Noviyanto, A. and Arymurthy, A. M. (2013) „Beef cattle identification based on muzzle pattern using a matching refinement technique in the SIFT method“, *Computers and Electronics in Agriculture*, Vol. 99, pp. 77-84. ISSN 0168-1699, DOI 10.1016/j.compag.2013.09.002.
 - [18] Parker, J. R. (2011) „*Algorithms for Image Processing and Computer Vision*“, Wiley Publishing, Indianapolis. 504 p. ISBN 978-0-470-64385-3.
 - [19] Pavlickova, P., Hanzlik, P. and Pavlicek, P. (2017) „Feasibility study KZ 13 Baby weeds“, Faculty of Economics and Management, Czech University of Agriculture Prague.
 - [20] Russ, J. C. (2008) „*Introduction to Image Processing and Analysis*“, 6th Edition. ISBN 978-084937073.
 - [21] Wiggers, E. P. and Beckerman, S. F. (1993) „Use of thermal infrared sensing to survey white-tailed deer populations“, *Wildlife Society Bulletin*, Vol. 21, pp. 263-268. ISSN 00917648.
 - [22] Wyatt, C. L., Trivedi, M. and Anderson, D. R. (1980) „Statistical evaluation of remotely sensed thermal data for deer census“, *Journal of Wildlife Management*, Vol. 44, pp. 397-402. E-ISSN1937-2817. DOI 10.2307/3807970.
 - [23] Yu, X., Wang, J., Kays, R., Jansen, P. A., Wang, T. and Huang, T. (2013) „Automated identification of animal species in camera trap images“, *Eurasip Journal on Image and Video Processing*, Vol. 52. E-ISSN 1687-5281. DOI 10.1186/1687-5281-2013-52.