Research on Key Technology for Processing Images from Infrared Trigger Cameras

LI Jian[1, a], ZHANG Lei[1, b], YAN BaoPing[1, c]

[1]Computer Network Information Center, Chinese Academy of Sciences, Beijing, 100190, China

[a]email: lijian@cnic.cn, [b]email: zhanglei@cnic.cn, [c]email: ybp@cnic.cn

Keywords: image similarity; infrared trigger; image feature analysis; ROI; color histogram

Abstract. In recent years, infrared trigger cameras are widely employed in the fields of wildlife monitoring and protection as they can collect massive image data. It is of great realistic significance about how to conduct automatic processing of these image data and extract key information. This paper has a detailed analysis about features of images in this type by automatically grouping them based on time intervals in order to eliminate the impact of light intensity on images in different time frames, automatically eliminating the background of those images within the groups and extracting the image area of individual animals (namely Region of Interest, short for ROI) and identifying the image features of these ROI areas in terms of species. The above-mentioned key technique has been tested for many times and achieved quite good results, but still has the potential of further optimization and improvement.

Introduction

In recent years, infrared trigger cameras are vigorously promoted and widely applied, in particular deployed extensively in the wild. They can produce massive image data for the study, monitoring and protection of wildlife. As is introduced in the references [1][2][3], successful application cases have achieved effective results in observation. However, with the growing number of infrared trigger cameras and accumulating shooting time, the amount of image data is getting more and more massive, thus it becomes a very arduous and heavy task to process these image data only by hand. Also, the time delay will affect the progress of scientific research and its timeliness. In that case, how to analyze and process these massive image data in a fast and accurate way and extract effective information is a key issue of urgency, which is of extreme significance to liberating biologists from piles of low-level and repetitive labor. However, so far, the research related to automatic analysis and processing towards this type of images is quite limited and it is also pretty difficult to find associative research material.

Analysis of Image Data Features

Based on different working principles, the infrared trigger cameras can be divided into two categories which are by fixed-interval shooting and animal-trigger shooting. The former focuses on behaviors of individual or group animal for detailed study and then fit for frequent appearance of study objects, such as bird nestling and brooding, while the latter applies to the unpredictable situation with low appearance frequency of target objects, in particular the monitoring activities about covert large or medium-sized rare beasts and birds. We only focus on the second category, which is image shooting triggered by target animals (so are the images mentioned in the follow-up of this paper). The images of this type have the following features:

1.   The camera is in a fixed position so that the shooting area will not change frequently, therefore the image background will not often alter, which means the images shot by the same camera have quite similar backgrounds.

2.   The images often are attached with shooting time. The time interval between images in continuous capture triggered by the same individual creature is quite short while the interval triggered by different animal objects is much longer.

3.   The shooting in different time frames and affected by the light intensity will lead to large difference with images shot by the same camera in different time periods.

The above-mentioned features indicate that different images actually have big similarities, which is clearly a different advantage from common image sets and of great help to improving image analysis and processing capabilities.

Image Processing Procedure

Our main purpose is to find out the region of interest among massive images (i.e. the imaging part within the picture of animal individuals, or hereinafter referred to as ROI) so that species can be automatically identified, and information about time, location and number of species showing up can automatically go through automatic statistical analysis. To this end, numbers of algorithms with regards to image analysis and processing will be involved. In practice, the related interfaces of OpenCV are mainly called. Open Source Computer Vision Library (short for OpenCV) [4] is a cross-platform computer vision library based on open source distribution, which can support Linux, Windows, Mac OS, IOS and Android. Being light and efficient, it is composed of a series of C functions and a few C++ types. It can offer Python, Ruby and Java language interfaces at the same time, thus realizing many common algorithms in terms of image processing and computer vision [5][6]. With the support of OpenCV, the development progress has been noticeably accelerated.

In combination with image correlation, this paper intends to propose the general processing flow as indicated in the Figure1, which includes three major models, and also three important processing steps–image data pre-processing model, species identification model and data statistic analytical model. All these models involve core algorithm related to image processing, which will be introduced in this chapter in detail.

Fig.1. General Flow Chart

3.1 Pre-processing of Image Data

With an increase of infrared trigger cameras numbers and gradual growth of shooting time, massive image data will be obtained. However, these original image data are varied in quality, some with too much exposure and some being vague and lacking fidelity. Also, images in different shooting time vary widely. Therefore, before these data are analyzed and processed, it is very necessary to work on pre-processing. In combination with features of images in this type, the pre-processing work involves the following steps:

3.1.1 Automatic Grouping of Images

As the shooting time of infrared trigger cameras is not fixed, mainly triggered when the animals pass or are active, so the shooting can happen in both daytime and at night and the image quality is significantly influenced by the light intensity. Hence, the parameters like image colors or exposure rate in different times are also different, leading to large divergence in images, even if the image backgrounds are significantly different. However, as the series shooting in a short time usually have

the same light intensity, so their backgrounds have much similarity. Thus, the images need to be grouped automatically according to the time first and go through the following pre-processing as indicated by Figure2:

    1.   All the infrared trigger cameras are assigned with a device number. Those images shot by different cameras will be divided into different groups for the later separate processing;

    2.   The images from the same camera will be put in sequence based on the shooting time;

    3.   Identify the shooting interval between two adjacent images after sequencing one after another. If it is larger than the designated time (such as 100 seconds, then an ideal value can be got through statistics), and the two images will be divided into different groups, otherwise will be deemed as in the same group.
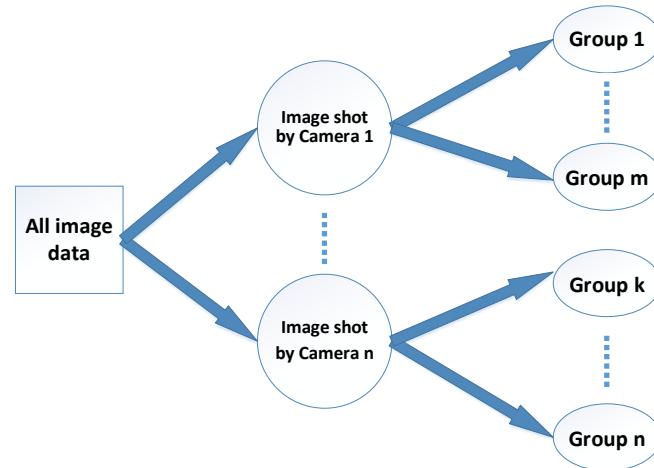
Fig.2. Diagram of Image Data Grouping Steps

After the experiment and tests, it is discovered that every group has three to six images on average. Generally speaking, we can see each group of images as shot by the camera triggered by the same one or more animal individuals. Different groups will involve different animal individuals.

It is a quite simple process in practice, because among the image data is hidden EXIF additional information, which can be used for automatic grouping. EXIF (Exchangeable image file format) [7], can be attached to JPEG, TIFF and RIFF files in order to add contents and relevant property information about camera shooting. These information can be edited randomly by starting with OXFFE1 as a mark, followed by two bytes to indicate the information length of EXIF, hence the maximum volume of EXIF information is 64KB. There is a variety of metadata information recorded by EXIF, mainly including the following kinds: shooting date, shooting device, shooting parameter, image processing parameter, image description and copyright information, GPS positioning data and thumbnail, etc. These information usually can achieve significant effect if being used in detail.

After being divided into groups successfully, the image groups need to be analyzed and processed. The images in each group have relatively fixed background, which usually can be seen as static and unchanged, due to the same shooting angle and short time interval. What is different with images is about the part of animal activity (i.e. ROI). Therefore, our priority is on how to automatically find and analyze the ROI.

3.1.2 Eliminate Invalid Images

When testing the actual images, it is discovered that some original images lack fidelity or become vague. This might be owed to jitter by wind, or the animal too close to the camera. So, this kind of images are regarded as invalid, which need to be eliminated, otherwise they will affect the follow-up processing result and increase the error. There are several ways of processing, the images can be directly identified or similarity can be compared as for the images in the same group.

3.1.3 Select Region of Interest

The images in the same group have very similar background. What are different between them is that as the animal is moving, its position in the image often changes. By overlap comparison of images, it is easy to find out the image area of animal activity, i.e. region of interest, which is the focus of our research. This is the method and thinking behind this most direct analysis: a whole-background image is shot beforehand as the benchmark image data. Then ROI area is obtained by comparing all the images pending for processing with the benchmark one. However, this method is not desirable in practice, because in different periods of time, the light intensity has a very big impact on the image. And, the single whole-background image obviously cannot offset the impact of light intensity. If the whole-background images with various light intensity are provided, there will be much more calculations. Therefore, this thinking only has certain value as auxiliary reference, and will not be adopted in this paper.

Because of short time interval between images in the same group, so the impact of light intensity is very limited and it will get ideal results by comparing every two images according to the time sequence. In practice, suppose there are three images in this group and they are marked as picture A, picture B and picture C according to time sequence. After directly comparing picture A and picture B in content, two ROI as indicated by Figure3 (a) are obtained. Similarly, after picture B and picture C are compared in terms of content, the two ROI as indicated by Figure3 (b) are obtained. Then in combination with time factor, we can preliminarily determine that the ROI_1 in Figure3 (a) is the ROI area in picture A, the ROI_2 in Figure3 (a) and Figure3 (b) is the ROI area in picture B and the ROI_3 in Figure3 (b) is the ROI area in picture C. In the actual testing, this analysis result basically conforms to reality.
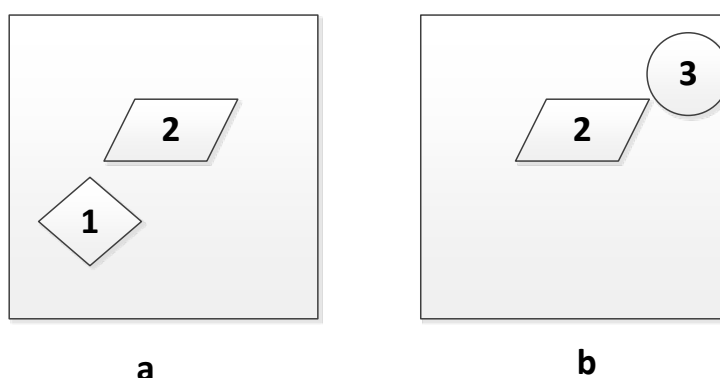

Fig.3. Diagram of selecting ROI Within the Group

3.1.4 Select Representative Images

As the images within each group get overlap comparison, the ROI can be obtained. In order to improve calculation efficiency, it is unnecessary to conduct follow-up analysis and processing on all the images and every group can select one or two pictures to meet the follow-up requirements. The principle behind the selection of representative images are set as: ROI should be as large as possible; meanwhile, this region should locate in the center of image, if possible, rather than at the edge of image so as to avoid the incomplete shooting caused by the animal body being unable to completely enter the shooting area of camera.

By selecting representative images and ensuring no loss of key data as well, the image data scale will be diminished to a large degree so that the workload for users to process images by hand will decrease dramatically. Right now, the analysis and processing of images has been realized to some extent. However, we can still continue to extract features from images for the processing method of species identification and statistical analysis so as to further improve automation degree.

3.2 Key algorithm for species identification

There are many ways of species identification, such as gene bar-coding, sound, morphological description and image identifications. The main research object of this paper is images, so it gets

very difficult to identify the species mainly based on the image features. It is of critical meanings to adopt feature type and comparison algorithm in the identification procedure, as it not only determines the accuracy of whole identification algorithm, but also exerts great impact on the real-time performance. The image features include feature of many types, like color histogram and texture, etc. Now, this algorithm mainly uses color histogram for fast comparison.

As the scatter diagram of colors in images, color histogram can demonstrate the probability of different colors seen in the image. As the impact of image changes (such as translation, blur, rotation and zoom) on color histogram is limited, so color histogram is suitable for retrieval of similarity in image colors.

Currently, the most common color histogram is gray histogram, which is to select statistic values of H (hue) and S (saturation) channels from the HSV hue space. In this paper, RGB three-channel color histogram is adopted, which means that the R, G and B channel histograms will be counted respectively and then a 3D histogram will come out of three histograms by recombination of them for comparison.

There are many ways of histogram comparison. Different ways offer different results concerning different histogram data. In order to reduce error, here we use four ways to compare in this paper, including correl, chi-square, intersection and Bhattacharyya [8]. The final image similarity is calculated based on the comparison result out of four methods with the same weight.

In practice, color histogram comparison algorithm has certain degrees of limitation. For example, the analyzed result is the histogram features of whole ROI without regard to distributed location of color features. This limitation will exert certain influence on the final identification result. Therefore, this algorithm is fitted to species with pretty unique color. However, the similarity of this algorithm might be affected concerning the species with very closely similar colors.

The above-mentioned procedure mentions a key factor, i.e. the feature database of species images, where the feature type and number in this database will directly affect the accuracy of the comparison algorithm. In the extreme case, for those new species not covered by this feature database, it is very hard to get correct analysis result no matter how the comparison algorithm is optimized. Therefore, this feature database needs to be supplemented in the process of system application, which means that the result and features of species identification need to be stored in this database as a new feature record after the experts review them.

3.3 Data Analysis Statistics

In the use of species identification model, we can preliminarily determine the specific type of this species. In combination with the EXIF information of original image, we can get the time, location longitude and latitude and volume of this species showing up and store the relevant information into the database. Then aggregation statistics and analysis can be carried out based on the scientific demand. For example,

1) add up the species type and number that appeared in a certain time period and in a certain location;

2) Time frame, activity area and species volume of certain species;

3) Analyze and study the living habits and activity rules of certain species;

4) If there are quite a number of cameras, the distribution areas of some species can be roughly outlined.

Due to some errors in species identification, so the result of automatic identification cannot be guaranteed as totally accurate. Those inaccurate results of identification will lead to the errors in data in this model. Hence, while trying every means to improve accuracy of identification model, the man-made check mechanism is offered. The experts are allowed to check whether the species identification result is correct or not and then alter or supplement, which will enhance the general accuracy in the whole system and ensure the scientific property of data. At the same time, the data checked by experts will also put into the feature database for follow-up species identification, which will help further improve the accuracy of species identification by itself.

Conclusion

The method of image processing adopted in this paper can reduce the number of image data to be processed by a large margin in actual application and cut the work load of users to check by hand. Meanwhile, certain effects can be achieved by using image features of color histogram to identify species, so that rough calculation about species distribution, active time and numbers can be carried out. However, besides these effects achieved above, there are still large room for improvement. For example, as the histogram features have some shortcomings, the follow-up work needs more image features (such as image texture, etc.) to improve the accuracy of species identification. With the gradually growing data scale in the species image database and increase of image features, the algorithm will run much longer. So in the follow-up work, the OpenMP[8] and JOMP[9] method can be taken into consideration to get the key algorithm into parallel optimization, which is believed to obtain better results.

References

[1] Wang Liang, Zhao Wei, Yang Zengwu, Tao Jinding, Bao Xinkang The Application of Infrared Camera-Trapping in Wildlife Survey in Mountainous Regions. Chinese Journal of Wildlife, 2014 35 (4) 381-387.

[2] MIAO Baohe, HANG Wenbo, YANG Ke, SUN Nan, FU Fang, WU Jing, LU Xinyan, DENG Qi, XIA Canwei. Using Auto-Trigger camera system to document wildlife species in beijing wangping coal mining areas. JOURNAL OF BEIJING NORMAL UNIVERSITY (NATURAL SCIENCE). 2011 47 (3).

[3] MA Ming, XU Feng, R.S.CHUNDAWAT, Kubanych JUMABAY, WU YiQun, Aizezi, ZHU MaHong. Camera trapping of snow leopards for the photo capture rate and population size in the Muzat Valley of Tianshan Mountains. Acta Zoologica Sinica, 2006 52 (4) 788-793.

[4] OpenCV, open source computer vision, http://opencv.org/.

[5] Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts [C]//ACM Transactions on Graphics (TOG). ACM, 2004 23 (3) 309-314.

[6] Bradski G, Kaehler A. Learning OpenCV: Computer vision with the OpenCV library [M]. O'reilly. 2008.

[7] Japan Electronics and Information Technology Industries Association. Exchangeable image file format for digital still cameras: Exif Version 2.2[J]. JEITA CP-3451, April, 2002.

[8] Dagum L, Menon R. OpenMP: an industry standard API for shared-memory programming [J]. Computational Science & Engineering, IEEE, 1998 5 (1) 46-55.

[9] Bull J M, Kambites M E. JOMP--an OpenMP-like Interface for Java [C]. Proceedings of the ACM 2000 conference on Java Grande. ACM, 2000 44-53.