Local image processing in distributed monitoring system

Ya A Bekeneva¹, V D Petukhov¹ and O Yu Frantsisko²

¹Saint-Petersburg Electrotechnical University “LETI”, 5, Professor Popov str., Saint Petersburg, 197376, Russian Federation
²Kuban State Agrarian University named after I.T. Trubilin, 13, Kalinina str., Krasnodar, 350044, Russian Federation

E-mail: yana.barc@mail.ru

Abstract. In this paper an approach to local image processing in distributed monitoring systems is considered. The proposed approach makes possible to optimize traffic transmitted from monitoring devices generating images or video streams. Image processing, object recognition and identification are performed locally. Based on information received from all monitoring devices, including images, event logs are generated locally for their further transmission to the central data collection, processing and analysis unit. The method of local image processing was tested on the example of face recognition of people who took part in the experiment. Under various experimental conditions, it was found that the proposed method has a sufficiently high accuracy.

1. Introduction

Nowadays distributed monitoring systems have a complicated hierarchical structure and often consist of a large number of heterogeneous devices. One of the important part of modern monitoring systems is photo and video surveillance equipment. The multimedia flows received from such devices are transmitted to the central device for processing and further analysis. The continuous transmission of a large amount of video data creates a heavy load on communication channels and increases the requirements for their bandwidth at the design stage of a distributed monitoring system. At the same time, a significant part of video usually does not contain any useful information; therefore, the transfer of large amounts of data is not expedient.

In [1] an approach to distributed event logs formation was proposed. The event logs are based on all data received from all monitoring devices located at each place where an event can be recorded. The approach is based on the fact that each monitoring area is equipped with a certain set of monitoring devices, while such devices register the same event simultaneously or with an acceptable delay, which can be known or determined in advance. For the local event log formation using data from photo or video recording tools, it is necessary to extract the required information from multimedia data: images or video stream, and the observed object must not only be recognized, but also identified.

In this paper a method of local face recognition for monitoring systems is proposed. This method is designed for systems where people are the objects of observation but it can also be used for other objects of observation, for example, cars, with the introduction of appropriate adjustments.
2. Related works
In recent years, the problem of pattern recognition has become widespread in many areas. The problem of determining the presence of an object of the desired type in the image is often solved: a person [2], an animal [3], a vehicle [4], a road sign [5], etc. can be both simple enough, for example, to determine the type of animal in the image [6], and more complex, for example, determining the age group to which the object in the image belongs [7].

Frequently solved tasks are determining the shape of an object [8], determining the trajectory of an object's movement based on the analysis of several images [9], etc.

In tasks related to monitoring the movements of objects of observation in industrial areas, employees, as well as vehicles carrying out the transportation of goods, are most often the objects of observation. Thus, in monitoring systems, the main task of image processing is not only for determining the presence of an object of observation in the image and object classification, but also its identification. In case of vehicle identification, text recognition methods can be used as technologies used to obtain unique vehicle registration numbers. In cases where the objects of observation are employees, it is necessary not only to determine the presence of a person in the frame, but also to identify him. The face of the same person may differ in the image depending on his / her facial expressions, as well as the accessories used (hats, glasses, protective masks, etc.). One of the conditions necessary for the above-mentioned approach to the formation of event logs is the local identification of the monitored object.

3. Approach description
The main feature of the designed solution is local data processing. To do this, all operations for collecting, storing and processing data must be performed directly on the local device.

The current goal of the approach is to develop the functionality of face recognition, based on this, three main work algorithms can be distinguished: gathering information, training the model, recognition process.

The module for collecting photographs for training, named as face finder, performs the role of collecting information about the observed object, namely, during operation, the id of the observed object and its name are entered into the program, which will be saved on the device. The module is a cycle that completes after successfully collecting the n-th number of photos of the observation object for recognition. As soon as the module receives n photos (in the current implementation, the number of photos is taken equal to thirty) it completes its work, saves the compressed images to the device for their further processing.

Thus, we collect data to train the model to find a specific person. Each new object of the system monitoring must go through such a data collection procedure, since the data is stored on the local device and external computers are not involved in their processing, and the data as a whole does not go beyond the local machine, it was decided not to encrypt the data.

After collecting data for analysis, the next step is to train the model to search for specific persons, according to the parameters. This stage involves the collection of all user data collected in the previous stage of work and data "instructor" in the OpenCV Recognizer (recognizer). The training is performed using a specific function of OpenCV. LBPH Face Recognizer, which is included in the OpenCV package, is used as a recognizer.

After collecting the data and training the model, we move on to the third block of performing face recognition, namely the recognition process itself.

After face recognition and determination of person’s identifier used in the monitoring system (personal ID or surname), a message containing the identified identifier and other necessary characteristics of the event (time, place, direction of movement, etc.) is generated (figure 1). The message is transmitted to the node, which performs the function of forming a single event log in a specific monitoring area. Thus, instead of continuously transmitting images or a video signal, a local transmission of a small message occurs, containing the necessary attributes of the event and only when an object of observation is detected in the image. At other times, when there are no objects of observation in the frame, the transmission of multimedia data does not occur.
This algorithm implies that if the face in the frame of our camera was previously photographed and trained in the system, then the "recognizer" will make processing that will return its identifier and index, with a percentage indicator of how confident the recognition system is in this feature.

![Algorithm of one image processing](image)

**Figure 1.** Algorithm of one image processing.

4. Experiments
To test the operation of the face recognition module, it was required to collect a database of subjects’ faces in N numbers. The test involved 10 subjects. For this, each subject was entered into the device database. Then the model was trained on this data.

The next step was a sequential “scan” of the objects, with a check whether the model was able to find out the object, the number of errors and the number of successful results.
The result of the first test is presented in table 1 in the second column, in this column the result is presented for the database, which includes 30 source photographs for each subject, and 10 test approaches for each subject, the percentage shows successful recognition in these 10 attempts.

Next, to test the possibility of improving facial recognition, by increasing the amount of initial data for training, a second test was conducted, where the number of source photographs for each subject was doubled, 10 approaches were also performed for each subject, the results are shown in the table in column 3.

| People ID | Successful 30 | Successful 60 | Difference |
|-----------|---------------|---------------|------------|
| 0         | 80%           | 100%          | +20%       |
| 1         | 70%           | 80%           | +10%       |
| 2         | 70%           | 70%           | 0%         |
| 3         | 90%           | 80%           | -10%       |
| 4         | 60%           | 70%           | +10%       |
| 5         | 70%           | 70%           | 0%         |
| 6         | 60%           | 70%           | +10%       |
| 7         | 90%           | 100%          | +10%       |
| 8         | 80%           | 90%           | +10%       |
| 9         | 80%           | 90%           | +10%       |

The overall best percentage of module operation for this test can be estimated at 82%.

Tests were carried out both under perfect conditions and with additional difficulties (clothing items, glasses, lighting problems).

5. Conclusion
The approach to data processing proposed in this paper is aimed at optimizing the data transmission process in distributed monitoring systems [10, 11], reducing the amount of transmitted data and reducing the load on the central node. The approach also allows to reduce the requirements for communication channels, but increases the requirements for the equipment used, since it involves the usage of their computing power to implement image processing. The experiments showed that the accuracy of recognizing human faces is quite high, including in conditions when additional elements of clothing are used that partially hide or change the face.

As the next stages of research and development, it is planned to investigate the accuracy of face recognition at different distances from cameras, as well as to investigate the accuracy of identifying other objects of observation in the image, for example, cars or trucks.

The study showed that the proposed approach is promising for its use in various monitoring systems [12].

Acknowledgments
This work was supported by the Russian Federation President Award SP-2581.2019.5.

References
[1] Bekeneva Y A 2020 An Approach to the Distributed Generation of Event Logs Based on Data from Heterogeneous Monitoring Devices 9th Mediterranean Conf. on Embedded Computing 1-4
[2] Nguyen D T, Li W and Ogunbona P O 2016 Human detection from images and videos: A survey Pattern Recognition 51 148-75
[3] Tack J L P, West B S, McGowan C P, Ditchkoff S S, Reeves S J, Keever A C and Grand J B 2016 AnimalFinder: A semi-automated system for animal detection in time-lapse camera trap images Ecological Informatics 36 145-51
[4] Choudhury S, Chattopadhyay S P and Hazra T K 2017 Vehicle detection and counting using haar feature-based classifier 8th annual industrial automation and electromechanical engineering conf. 106-9

[5] Le T T, Tran S T, Mita S and Nguyen T D 2010 Real time traffic sign detection using color and shape-based features Asian Conf. on Intelligent Information and Database Systems 268-78

[6] Shalika A U and Seneviratne L 2016 Animal classification system based on image processing & support vector machine Journal of Computer and Communications 4(1) 12-21

[7] Zamansky A, Sinitca A M, Kaplun D I, Dutra L M and Young R J 2019 Automatic Estimation of Dog Age: The DogAge Dataset and Challenge Int. Conf. on Artificial Neural Networks 421-6

[8] Poda X and Qirici O 2018 Shape Detection and Classification Using OpenCV and Arduino Uno RTA-CSIT 2280 128-36

[9] Nowosielski A, Frejlichowski D, Forczmanski P, Gosciewska K and Hofman R 2016 Automatic analysis of vehicle trajectory applied to visual surveillance Image Processing and Communications Challenges 7 89-96

[10] Atanasov I, Pencheva E, Nametkov A and Trifonov V 2019 On Functionality of Policy Control at the Network Edge International Journal on Information Technologies and Security 3(11) 3-24

[11] Lvovich Ya E, Tishukov B N, Preobrazhenskiy A P, Pitolin A V and Kravets O Ja 2019 Complex-Structured Objects Optimization During Modeling on the Population Algorithms Adaptation Basis International Journal on Information Technologies and Security 3(11) 41-50

[12] Tsaregorodtsev A V, Kravets O Ja, Choporov O N and Zelenina A N 2018 Information Security Risk Estimation for Cloud Infrastructure International Journal on Information Technologies and Security 4(10) 67-76