Software and algorithmic complex for evaluating the efficiency of an automated optical-electronic system

Dmitry Lovtsov
Lebedev Institute of Precise Mechanics and Computer Engineering (IPMCE)
Moscow, Russia
dal-1206@mail.ru

Dmitry Gavrilov
Lebedev Institute of Precise Mechanics and Computer Engineering (IPMCE)
Moscow, Russia
gavrilov.da@mipt.ru

***Abstract* — This paper presents a set of operation quality assessments of systems intended for detecting and localising objects of interest depending on the content of input video. The development of an input video registration generator has been carried out based on object and camera movement trajectories and on image distortions in a frame. A description of the main modules of the complex has been presented and experimental studies of localisation detection algorithms have been conducted.**

Keywords — modelling optoelectronic systems, visual information processing, efficiency, primary video processing algorithms, detection and localisation algorithms

I. Introduction

At present, satellite aerospace imagery is widely used to solve a multitude of different cross-sector applications [1, 2]. The central problem in the field of image processing is the creation of effective automated optical-electronic ground-space monitoring systems (AOES GSM), which provide a solution to specific problems with high accuracy, have resistance to various aggressive factors and a clearly defined and investigated range of applicability [3, 4, 5].

The architecture of an efficient AOES GSM [3], which is an efficient automated optoelectronic system for special (interbranch, etc.) purposes providing processing of diverse visual information, has been substantiated. Quality evaluation of such a system should be carried out taking into account the parameters characterizing the solution of certain particular tasks of AOES GSM using a method of consistent approach to the effective rational solution of a common task.

Automatic target tracking (ATT) is an AOES GSM operation mode, in which without the direct participation of human operator, but only under his/her control, change in the orientation of the axis or direction of movement of some elements of the complex following changes in trajectories and speeds of the observed objects [1-4] is provided. A universal way of evaluation and functional diagnosis is to compare the error probabilities of algorithms used in information and mathematical support of AOES GSM.

The aim of the work is to model the visual information organization processes allowing to automate the functional check of AOES GSM. The result is achieved by providing an opportunity to evaluate the correspondence between input video parameters and numerical coefficients used for evaluation of the quality of tracking an object of interest, as well as an opportunity to modify the video signal to create unique parameters of input video, which makes it possible to improve the quality of evaluation of target detection and tracking systems.

II. METHODS FOR EVALUATION OF INFORMATION EFFICIENCY AND FUNCTIONAL DIAGNOSIS OF AN AUTOMATED OPTOELECTRONIC SYSTEM

In order to achieve this goal, a programme and algorithmic complex for assessing the quality of AOES GSM has been developed, with the help of which incoming video sequences are recorded using a set of specific parameters. The parameters may be trajectories of objects or camera movement, and distortions of the frame images.

The methodology developed to evaluate the information efficiency and functional diagnosis of AOES GSM includes the following steps:

1. Formation of incoming video, recording raw data on camera and target movement
2. Determination of objective parameters of the information flow affecting the quality of tracking, including configuration of object and camera trajectories, configuration of dynamic parameters, input information flow parameters, flow distortions, camera shake telemetry.
3. Mathematical modelling of the input information video stream.
4. Generation of reference data.
5. Calculation of numerical parameters to evaluate the object tracking quality.
6. Comparison of calculated and reference values.
7. AOES GSM efficiency evaluation based on comparison of calculated and reference values.

III. DESCRIPTION OF THE HARDWARE-SOFTWARE COMPLEX FOR ASSESSING INFORMATION EFFICIENCY AND FUNCTIONAL DIAGNOSIS OF AN AUTOMATED OPTOELECTRONIC SYSTEM

A software-algorithmic complex for assessing the quality of AOES GSM performance depending on the content of the input visual signal has been developed. It consists of the following independent modules:

- Simulator – generates the incoming video and records the original data on camera and target movement (Ground Truth as to foreign literature), i.e. information on the location of blips in each frame, i.e. the position of their centres and mask;
- Visualiser – generates video with markers of the blips, detected by AOES GSM;
- Evaluator – provides quantitative indicators of the teleautomatic system operation quality; the

978-1-6654-3941-1 @2021 IEEE

evaluator accepts the Ground Truth data and the AOES GSM's operation result to its entry, then checks how close the initial values are to those provided by the AOES GSM.

A set of objective parameters affecting the quality of the tracking is formed at the stage of preparing the original video image. The result of the work of the software and algorithmic complex is a correspondence between input video parameters and numerical coefficients for evaluation of the object of interest tracking quality.

A. Simulator

The input video simulator consists of three separate modules, including: a trajectory interpolator; a dynamic parameter interpolator; and a 2.5-dimensional scene generator.

The trajectory interpolator converts the input set of key points of movement trajectories, as well as target and camera velocities, into smooth trajectories broken down into frames with a given video frequency.

The dynamic parameter interpolator converts a set of key points in a time diagram of dynamic parameters into sets of parameters for each frame. Interpolation between key points is linear.

The video generator uses a set of discrete trajectories to generate an image for each frame. In addition, the generator records complete scene information, which is then used to evaluate the quality of tracking.

B. 2.5D scene model

2.5D scene model is a scene composed of two-dimensional layers ordered by distance from an observer. Most of the scene elements are sprites, i.e. they are not volumetric models, but animated 2D images that undergo projective transformations. In general, scene elements can be substituted depending on the position (angle) of the object relative to the observer.

In the implementation of a 2.5-dimensional scene model, the model is described by the following characteristics:

- The camera is fixed at a single point; the camera trajectory is a set of turns in the direction of its optical axis.
- The background image is "attached" to an infinitely distant sphere.
- Target images do not change perspective; target sprites can only rotate around the camera's optical axis, approaching and moving away. However, the distance of a target from an observer can be replaced with a sprite scale. This degree of freedom is 2.5-dimensional space.

C. Generator

The video generator is a module implemented in the Matlab environment and generating a frame image for each moment in time.

The main functions of the generator:

1. Determining the background area falling into the frame - to carry out this function, the image is scaled to take into account the distance of the object from the camera and the focal length of the lens, then a translation vector and rotation are applied to it.

2. The following actions are performed for each target trajectory:

- A target sprite image is scaled and rotated according to the set scale and angle values.
- The sprite image is superimposed on the background image of the current frame at specified coordinates.

D. Trajectory interpolator

The trajectory interpolator is designed to convert a target trajectory into its 2.5*D* model.

Input data for the simulator is prepared by a specially developed utility that solves the problem of creating a smooth target trajectory. Figure 1 shows an example of how a smooth trajectory can be constructed by reference points.

Fig. 1. Construction of a smooth trajectory along reference points

The configuration also specifies the type of trajectory - camera or target - and the video frame rate. Cartesian coordinates and velocities are specified in metres and metres per second, respectively.

The main task of the video generator is to create a video sequence along the trajectories of objects and the camera. The objects are presented as sprites and move against a panoramic image background. The camera trajectory in this case describes the relationship between the frame coordinate system and the real reference system. In addition, the video generator adds distortions to each frame in the form of interference, the parameters of which are specified in the incoming task.

Vibrations loaded from telemetry files are superimposed on the camera trajectory. The telemetry file provides information on both the movement of the machine relative to the ground and the camera's vibrations relative to the machine. In this way, the total vibration of the camera relative to the ground is calculated, which is superimposed on the trajectory.

The simulator output is fed with video sequences of the scene broken down into frames, as well as a set of parameters and Ground truth (recording of camera and targets movements).

IV. MATHEMATICAL MODELLING OF INTERFERENCES AND EFFECTS

A. Mathematical modelling of interferences and effects

- Blur (decrease in contrast) – creates a halftone image by contrasting the original halftone image.
- Exposure error - adds an exposure correction value to the intensity value of each point in the original image, which can take values within the range of [-1...1].
- Heterogeneous backlighting – simulates a point light source located in the upper left corner of the frame. The distribution of intensity from the

source as a function of distance is described by the expression:

$$I = A \cdot e^{-\frac{x^2}{2\sigma^2}},\ \sigma = \text{мax}(w,h) \cdot S$$

where *S* is backlight scale factor taking values within the range of [0...1]; the unit corresponds to the backlight of the whole frame; *A* is backlight amplitude taking values within the range of [0...1]; *w*, *h* are frame image dimensions.

- Focus error is simulated by blurring the boundaries of objects.
- A blurred image caused by camera movement is simulated by superimposing an averaging two-dimensional filter parameterised by length (amplitude) and tilt angle.

Some types of interference may lead to improvements in the values of randomly selected parameters, such as creating additional texture. In order to highlight the features created by distortions, it is necessary to evaluate the value of each parameter before and after the distortions are superimposed. The norm of difference of these parameters is accepted as a component of the distortion value metric. Thus, "artefact" key points are taken into account as a noise component.

B. Mathematical modelling of noise

A physically correct mathematical model of camera noise has been built. The process of generating and reading the signal has the following features. In the monochrome image, the pixels are located in the nodes of the square grid. Upon total absence of light, a dark current $i_{i,j}^{t}$ flows into the cell. At the end of the exposure, the amount of charge accumulated in each pixel $q_{i,j}$ is read out. The main sources of matrix noise are photon number fluctuations, dark noise, and reading noise. The number of photons n carried in a unit of time is a random value with the Poisson distribution $P(\lambda,n)$. This type of noise becomes noticeable with low light intensity. Noise caused by fluctuations in the dark signal is defined using the following form:

$$n^t = \sqrt{N_t},$$

Where N_t is the dark signal value expressed in electrons.

The reading noise is due to the fact that the electronic circuitry that reads the signal from the photosensitive matrix also generates noise.

The input parameters of the developed model are:

- Dark frame matrix $q_{i,j}^{d}$ is a record of a dark signal accumulated over a fixed time in absence of light.
- Matrix pixel sensitivity calibration matrix (flat field) - pixel sensitivity $s_{i,j}$ is fairly homogeneous and contains units of broken pixels r_b, which intensity is generated by the simulator at the beginning of the next run and is not measured during the run.
- The conversion factor of the number of accumulated electrons into matrix counts *k*.
- Source frame $I_{i,j}$ is treated as a record of the light intensities, falling per image pixel and expressed in camera counts for a single exposure time.
- Exposure value – the number of electrons from the source frame accumulated in each cell of the matrix has been determined for application of the model:

$$N_{i,j}^{e,l} = k \cdot \tau \cdot I_{i,j}.$$

Average number of dark electrons have been calculated

$$N_{i,j}^{e,t} = k \cdot q_{i,j}^{d} \cdot \tau,$$

where τ is the exposure time.

Vibrations loaded from telemetry files are superimposed on the camera trajectory. The telemetry file provides information on both the movement of the machine relative to the ground and the camera's shaking relative to the machine. This calculates the total vibration of the camera relative to the ground, which is superimposed on the trajectory.

The output of the simulator is fed to the scene images collected in a video sequence broken down into frames.
In addition, , the simulator generates at the output a set of parameters Ground truth (recording camera movements and targets).

V. AOEC GSM INFORMATION EFFICIENCY EVALUATION

The evaluation result is the foreground mask or area of the image occupied by the blips. The AOEC's task in this case is to create a foreground mask segmented by markers for each frame of the input video.

The main components of the evaluation result are: segmentation of objects, re-capture, deviation of the centre position, probability of detecting an object.

As a result of segmentation, labels are assigned to a particular blip. The blip number is the label of most visible pixels of a given blip.

The following rating characteristics are calculated:

- The share of pixels incorrectly assigned to this blip (*false positive*);
- The share of pixels mistakenly not assigned to this blip (*false negative*).

Segmentation evaluation factors are numerically described as follows:

- The true detection factor of blip pixels:

$$TP = \frac{Fs \cap Ft}{As}.$$

- The false-positive factor of blip pixels:

$$TP = \frac{Ft - Fs \cap Ft}{As}.$$

- The false-negative factor of detected mark pixels:

$$FN = \frac{-Fs \cap Ft}{As}.$$

where *Fs* is a matrix of size $N \times M$ containing the original foreground binary mask resulting from the simulation; *Ft* is a matrix of size $N \times M$ containing the foreground binary mask resulting from the tracking; *As* is number of *Fs* points with a value of 1; *At* is number of *Ft* points with a value of 1.

Overcapture is a situation where the blip number changes. The average number of captures per frame is defined as a quantitative characteristic.

To determine the deviation of the centre position, the distance between its true centre and the centre calculated by the tracker is calculated for each blip. The average characteristic of the number of blips and frames is then calculated.

For each frame, the success, q, of object detection depending on the size of the surrounding area, ε, is evaluated. In this case, a successful detection is an event where there is at least one detected object for the true blip, the centre of which lies within the radius of the specified surroundings.

VI. EXPERIMENTAL RESEARCH

Experimental studies of algorithms for primary video signal processing were conducted using the developed software and algorithmic complex.:

- An algorithm for selecting unique fragments (*A*-1);
- An algorithm to tie together moving parts of athen object (*A*-2);
- A pyramidal algorithm for selecting and calculating properties of objects that differ from the average intensity of the local background [6] (*A*-3);
- An algorithm and computation of properties of objects that differ from the local background by the grey level [7] (*A*-4).

The main metrics for evaluating the algorithms are the probability of detecting a target and the probability of "false alarm". In the course of the experiment, the algorithms were tested using a video signal simulator using a synthetic video and real video recordings of the background/target environment, and the video signal bandwidth was investigated. As a result, the results of each of the tested algorithms were visualised and quantified. The quality of the algorithm's work is monitored visually by the operator using a video file generated by the Visualizer module. In the course of the test, quantitative characteristics of the algorithm's performance and statistical data calculated by the "Evaluator" module are recorded. The test results in the form of averaged data are presented in Table 1.

Table 1 shows the average probability of detection and the average probability of "false alarm" for each of the algorithms being tested.

In general, a false alarm is a situation where a target is displayed while in reality there is no target. The probability of a false alarm is estimated by the value of the false alarm frequency. The developed software and algorithmic complex provides statistical analysis and estimation of information efficiency of processing visual information in AOES GSM, as well as comparison with the specified probability of correct detection of the target and false alarm. The obtained quantitative characteristics of the algorithms' quality make it possible to adjust their functional characteristics and to select the optimal parameters depending on the purpose of interpretation.

TABLE I. ALGORITHM TEST RESULTS

Initial processing method	**Average probability of detection**	**Average probability of "false alarm".**
Algorithm A-1	0,72	0,31
Algorithm A-2	0,79	0,28
Algorithm A-3	0,71	0,35
Algorithm A-4	0,68	0,38

VII. CONCLUSION

A software and algorithmic complex has been developed to evaluate the information efficiency and quality of the multilevel automated opto-electronic ground-space monitoring system based on information and mathematical methods of input signal simulation and visualization of the information signal, including methods for mathematical simulation of dynamic parameters, trajectory and noise configurations of the camera and object; quantitative indicators of quality and numerical control of AOES GSM results have been determined and justified, including those which include an integrated structure for solving the evaluation task.

REFERENCES

[1] Gavrilov D.A., Lovtsov D.A. “Information theory fundamentals of ground-space monitoring automated special-purpose optical-electronic system,” *Materials of the International scientific-practical conference* ”*Information Innovative Technologies*”, MIREA, 2020, pp. 155 – 159.

[2] Gavrilov D.A., Melerzanov A.V., Shchelkunov N.N., Zakirov E.I. “Use of Neural Network Based Deep Learning Techniquesfor the Diagnostics of Skin Diseases”, *Biomedical Engineering*, 2019, vol. 52, No. 5, pp. 348-352.

[3] Lovtsov D.A., Gavrilov D.A. “Automated special purpose optical electronic system’s functional diagnosis”, *Proceedings of International Seminar on Electron Devices Design and Production, SED-2019* (23 – 24 April 2019), Prague, IEEE, 2019, pp. 70 – 73.

[4] Buryi A.C. “Information and mathematical support for computer software quality control”, *Pravovaya informatika*, 2019, № 2, pp. 15 – 25. (in Russian)

[5] Gavrilov D.A, Mestetskiy L.M., Semenov A.B. “A method for aircraft labeling in remote sensing images basedon continuous morphological models”, *Programming and Computer Software*, 2019, vol. 45, No. 6, pp. 303–310.

[6] Gavrilov D.A. “Quality assessment of objects detection and localization in a video stream”, *Vestnik MGTU im. N.E. Baumana. Seriya* “*Priborostroenie*”, BMSTU, 2019, №2, pp. 40-56.

[7] Gavrilov D.A. “A neural network algorithm for automatic detection and tracking of an object in a video signal”, *Tr. 16th Nat. conf. on artificial intelligence* (24–27 sept. 2018 г.), vol. 2, 2018, pp. 188 – 190.