Intelligent emotion recognition system in neural network basis

K.D. Fathutdinova[1], A.M. Vulfin[1], V.I. Vasilyev[1], A.D. Kirillova[1], A.V. Nikonov[1]

[1]Ufa State Aviation Technical University, K. Marks St. 12, Ufa, Russia, 450008

Abstract. The human factor plays a significant role in ensuring the integrated safety of technological facilities. It is important to monitor the state of the operator of automated process control systems in soft real-time mode in order to reduce the risk of attention and concentration losses. Paper discusses the issues of increasing the efficiency of monitoring system of the operator's state by using algorithms for assessing the psycho-emotional state. These algorithms apply methods of intelligent analysis of video sequence data without the use of additional contact sensors, which reduces probability of making a wrong decision due to the timely detection of unstable psycho-emotional states. The accuracy of detecting unstable psycho-emotional states on a test data set is 79%.

1. Introduction

One of the key components of integrated security systems for automated process control systems (APCS) is the operator and it is necessary to consider the influence of the human factor on the stable system functioning. The price of human factor (mistake or delay in deciding) can be very high. Like any other person, operator could be influenced by emotions that can have a destructive impact on his cognitive abilities: narrowing the focus of attention, demobilization of physical strength, lowering working capacity and, as a result, decreasing the potential for making timely decisions on managing a technical object.

Therefore, it is urgent to develop intelligent systems for monitoring the operator's psycho-emotional state based on the analysis of video data, which could inform in advance about an unsafe decrease in the potential for making a timely decision. Video data analysis is a non-contact way of obtaining the state parameters of a human operator and does not require special additional equipment.

The purpose of the work is to increase the efficiency of the monitoring system of the operator's state through the use of algorithms for assessing the psycho-emotional state using the methods of intellectual analysis of video sequence data.

To achieve this goal the following tasks were formulated:

1) analysis of existing methods and solutions in the task of monitoring the psycho-emotional state;

2) development of a structural diagram of an operator's emotional state analysis system;

3) development of algorithms for analyzing the psycho-emotional state of the operator on the basis of the methods of intelligent analysis of video sequence data and evaluating their effectiveness on the field data.

2. Problem statement of assessing the operator's psycho-emotional state

Modern psychology distinguishes 6 basic emotions: "happiness", "sadness", "wonderment", "fear", "disgust" and "anger", the external manifestation of which is expressive facial movements and voice changes [1].

The term “affective computing” proposed in [2] as “calculations that relate to emotions, are their result and/or affect them”. Today, “emotion recognition” term is more widely used.

Emotion recognition systems are aimed to recognize the psychic component – the psycho-emotional state – through the external and/or physiological manifestation of emotions. That is the foundation and root cause of the emergence of emotions and their manifestations.

In this paper, the concept of emotions is used in general meaning, which includes all of the above components, as well as in the meaning of external manifestations (facial movements, gestures, etc.). In the meaning of internal sensations, experiences and feelings, the term “psycho-emotional state” is used.

Foreign studies [3, 4] prove that risk of making a mistake due to the unstable (above the threshold values “joy” and “sadness”) psycho-emotional state increases by at least 15%. Thus, we can conclude that it is necessary to monitor the emotional state of the APCS operator in order to reduce the probability of making a wrong decision.

The following patterns of unstable psycho-emotional state are suggested for analysis:

- sudden manifestation of anger;
- prolonged alternation of anger to sadness and back;
- prolonged alternation of joy to sadness and back;
- prolonged suppressed state (“sadness”);
- sudden manifestation of fear (“panic”).

The analysis of video data containing the image of the subject’s face is a non-contact way of obtaining parameters of the operator’s state and does not require special additional equipment – detectors and sensors. The analysis of the image of a person’s face and its comparison with a particular psycho-emotional pattern is based on the work of facial action coding system (FACS) [5]. FACS consists of 44 motor units, 30 of which are anatomically associated with the contraction of certain groups of facial muscles.

2.1. Analysis of existing methods and algorithms for recognizing emotions and the psycho-emotional state of the operator

The general approach to recognizing emotions from a person’s face image consists of three stages: determining the area of the image containing the face, highlighting key facial features and classifying the emotional state (Figure 1).

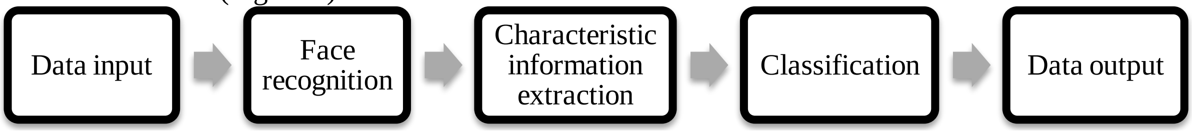


Figure 1. Emotion recognition process diagram.

A comparative analysis of characteristic feature extraction algorithms is shown in Table 1.

Table 1. Advantages and disadvantages of methods for extracting characteristic features.

No.	Method	Advantages	Disadvantages
1	Viola-Jones object detection [6]	1) Quick calculation of characteristic values. 2) An effective set of features. 3) Features are scaled instead of scaling the image.	1) The inability to recognize the desired object when it is rotated by 30 or more degrees. 2) Sensitivity to the angle of light incidence. 3) Long training time for classifiers.
2	Hidden Markov Model (HMM) [7]	1) Simple mathematical structure. 2) The ability to simulate a complex observation chain. 3) The ability to select model parameters to describe an existing dataset for training [8].	1) It is necessary to select model parameters for each database. 2) The learning algorithm only maximizes the response of each image to its model, but does not minimize the response to other models [8].
3	Principal component	1) Stability to the position of the face.	1) Unstable to changes in lighting.

	analysis (PCA)	2) The ability to add a new object without retraining. 3) Computational efficiency [7].	
5	Local Binary Patterns (LBP) [9]	1) Ease of calculation. 2) Resistance to noise. 3) Resistance to texture variations. 4) A significant reduction in the dimension of the problem. [10].	1) Instability to turns. 2) The importance of pixels is neglected, only the difference in values is considered [11].
4	Neural Network Methods Testing Convolutional Neural Network on ORL database [12] showed recognition accuracy of 96% [7].	1) Resistance to scale changes. 2) Resistance to displacement, rotation and change of angle [7].	1) Adding a new reference person to the database requires a complete retraining of the network on the entire existing set. 2) Learning problems: hitting the local optimum, choosing the optimal learning step, retraining, etc. 3) It is difficult to formalize the stage of choosing a network architecture (the number of neurons, layers, and the nature of connections) [7].
6	Histogram of Oriented Gradients (HOG) [13]	1) Resistance to geometric and photometric transformations, except for the orientation of the object [14].	1) Significant computational complexity. 2) The need to normalize the selected feature vectors.
7	Scale-invariant feature transform (SIFT) [15]	1) SIFT features show the highest accuracy of correspondence for an affine transformation of 50 degrees. 2) SIFT-based descriptors are superior to other modern local descriptors for both textured and structural scenes, with higher efficiency for texture scenes [15].	1) Fuzzy selection of the object relative to the background and a low percentage of correct recognition of image elements of objects without a pronounced texture [16].

Table 3 shows the comparative characteristics of the algorithms and databases on which the tests were conducted. The abbreviations are listed in Table 2.

Table 2. Methods of classification and extraction of characteristic features.

Method of extracting characteristic features		Classification method	
Action based	Action based method	ID3 decision tree	ID3 Algorithm (Decision Tree)
GF	Gabor Filter	LVQ	Learning vectors quantization
GASM	Active shape model, based on graphics processing	SVM	Support vector machine
LBP	Local binary patterns	LBP-TOP	Local binary patterns on three orthogonal planes
Patch based	Patch-based method	GL Wavelet	Gabor wavelets (Gabor builds)
CNN	Convolutional neural network	Bayesian NN	Bayesian Neural Network
KNN	k-nearest neighbors algorithm	HOG	Histogram of Oriented Gradients
MFFNN	Multilayer neural network of direct propagation	OSLEM	EM-algorithm with optimal stride length
Bayesian NN	Bayesian Neural Network	Steerable pyramid	Steerable (turning) pyramid

2.2. Analysis of modern solutions in the field of analysis of the emotional state of the operator and their comparative characteristics

Table 4 shows the parameters of some of the most common solutions.

Table 3. Comparative characteristics of emotion recognition algorithms [17, 18].

Method	Database	Recognition accuracy (%)	Number of recognized emotions	Advantage
Action based, ID3 decision tree	JAFFE	75	6	Cost effective in terms of speed and accuracy
GF, LVQ	JAFFE	88.86	Unknown	More accurate in determining fear
GASM, SVM	CK	93.85	6	Flexible selection of facial features
LBP-TOP, SVM	JAFFE, CK, Realtime	86.85	7	More resistant to light changes
SVM	JAFFE	87.5	Unknown	More effective emotion recognition
Patch based, SVM	JAFFE, CK	82.5	6	Effective recognition
GL Wavelet, KNN	JAFFE, CK, MMI	91.9	6	Great features for texture analysis
LBP, SVM	CK, MMI	95.84	6	Effective recognition of emotions by image
GF, MFFNN	JAFFE, Yale	94.16	7	Least Computing Cost
LCT, OSLEM	JAFFE, CK	94.41	7	Robust recognition algorithm
GF, SVM	JAFFE, CK	82.5	7	High stability and fast processing
HOG, SVM	JAFFE	85	7	Resistant to rotation, interference and noise
Streerable pyramid, Bayesian NN	JAFFE, CK	95.73	7	Effective recognition

Table 4. Comparative characteristics of solutions.

No.	Name	The ability to analyze video stream	The ability to recognize faces when turning the head	Photosensitivity	Price	Real time operation
1	FaceReader	+	–	–	600$ and by agreement	+
2	EmoDetect	+	–	–	–	+
3	Affectiva Products	–	+	+	by 2000$	+
4	Microsoft Oxford Project Emotion Recognition	–	–	–	by 1000$	–
5	VibraImage	+	+	+	5000$	+

Based on the results of the study, a data set was selected for further experiments by Cohn-Kanade (CK) and Cohn-Kanade Extended (CK+) [19], as well as FER 2013 [20].

2.2.1. Description of the databases CK and CK+

Facial expressions were recorded from 210 people aged 18 to 50 years:

- 69% of women, 31% of men;
- 81% Europeans, 13% African Americans and 6% others.

One camera was located directly in front of the subject, and the other was located 30 degrees to the right of the subject. Each of them was instructed to depict 23 facial expressions, which included both single motor units and their combinations [21].

For database testing “Base System CK” was developed. This CK system uses active appearance models (AAM) to detect faces and extract their features. Next, the support vector method is used to classify facial expressions and emotions.

With the help of the CK+ database and the existing base system, two tests were conducted: for recognition of motor units and for recognition of emotions.

2.2.2. Description of FER 2013 databases and comparison with CK

The FER 2013 database consists of 3589 images 48x48 pixels in size. Images are preprocessed: translated into shades of gray and cut out areas of the face [20].

There are 123 subjects in the CK + database, and more than a thousand in the FER 2013 database. As a result, the test of the first of them shows a much higher percentage of accuracy in recognizing the emotional state [21]. To obtain a balanced database, both of above were combined, and 40% of the images from the CK+ database were deleted, since they were not related to the task and did not contain the correct layout of individual images.

3. Development of algorithms for the analysis of the psycho-emotional state of the operator on the basis of data mining methods

3.1. Development of a structural-functional diagram

Figure 2 shows the structural-functional diagram of the recognition system of operator’s psycho-emotional state.

Video stream capture block receives data from the webcam (1), where the video is divided into frames, and the resulting frames sequence with time reference goes to the image preprocessing block (2). Preprocessing module includes:

- image normalization (auto white level correction);
- noise filtering (median filter);
- cast to color scheme (grayscale or RGB).

Next, the prepared image (3) is analyzed in order to face searching. The output is the ROI (region of interest) and the timestamp of the frame (4). Processing is performed by one of two algorithms:

- Viola-Jones object detection;
- convolutional neural network (CNN).

The next block is feature extraction, the output of which is a multicomponent feature vector. Extraction of signs is possible at two levels: I and II. Signs of the first level are local structural features of the image of a person’s face and key facial features. For each ROI and timestamp, a multicomponent output feature vector of the first level has the form (5):

$$\left\{V_1^{HOG}, V_2^{SIFT}, V_3^{LBP}, V_4^{Eig}\right\}, \tag{1}$$

and for each vector V_i there is a weight of the feature vector, where $i = 1 \cdots n$; $n = 4$.

Feature extraction is carried out by one of the following algorithms: HOG, SIFT, LBP, PCA (Eigenfaces).

Signs of the second level are the key points of the face (anthropometric points – the centers of the eye-apple, corners of the eyes, etc.). The output of this block is a vector of the form (unit 6):

$$\left\{V_1^{AAM}, V_2^{ASM}, V_3^{Haar}\right\}. \tag{2}$$

Feature extraction is carried out by one of the following algorithms:

- AAM – active appearance models;
- ASM – active shape model;
- PCA (Haar-like features) – PCA-based Haar features selection.

The multicomponent vector of signs of I or II level is the initial data set (7) for the block of selection of significant signs, which is implemented using PCA or ICA (Independent component analysis). Thus, the source data for the next block is a vector of significant features for each ROI and timestamp (8).

The next step is classification. It can be implemented by the following methods:

- XGBoost;
- SVM;
- MLP;
- LR (Logistic regression).

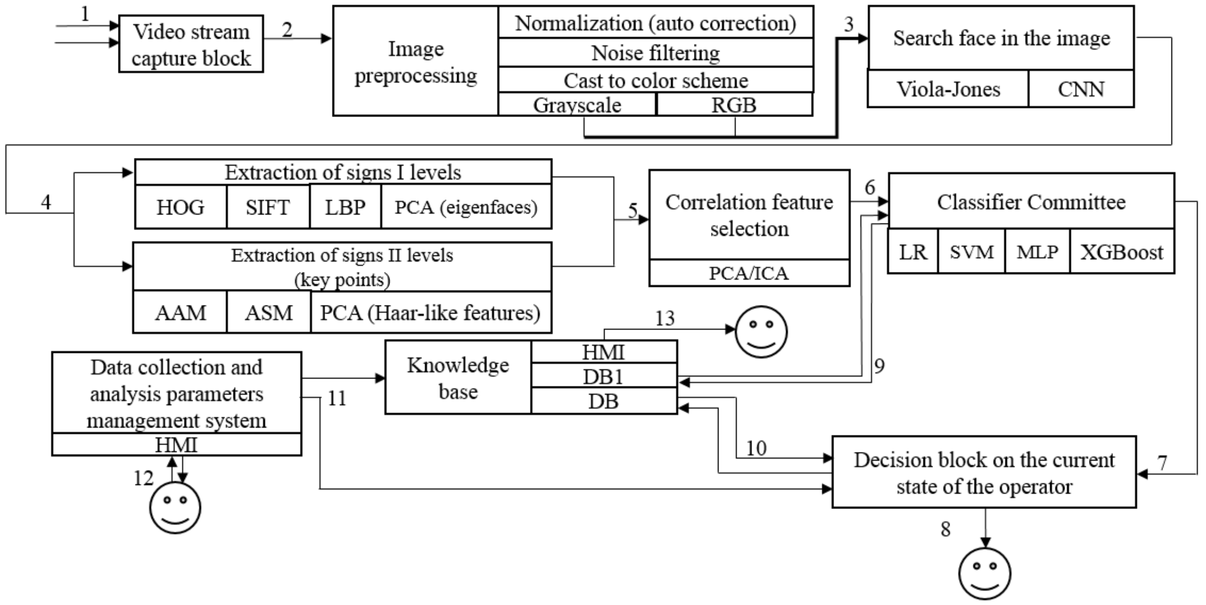

1
Video stream capture block
2
Image preprocessing
Normalization (auto correction)
Noise filtering
Cast to color scheme
Grayscale
RGB
3
Search face in the image
Viola-Jones
CNN
4
Extraction of signs I levels
HOG
SIFT
LBP
PCA (eigenfaces)
Extraction of signs II levels (key points)
AAM
ASM
PCA (Haar-like features)
5
Correlation feature selection
PCA/ICA
6
Classifier Committee
LR
SVM
MLP
XGBoost
13
Data collection and analysis parameters management system
HMI
11
Knowledge base
HMI
DB1
DB
9
12
10
Decision block on the current state of the operator
7
8

Figure 2. Structural-functional diagram of the recognition system of the operator’s psycho-emotional state.

The result stage is the decision-making block of the current psycho-emotional state of the operator. Initial data for this block is the vector of probabilistic assessments of *k*-th psycho-emotional state in the allocated ROI, taking into account timestamps for this operator (9). At this stage, the presence of specific patterns for assessing the psycho-emotional state of the operator in the context of changing estimates of the probabilities of the *k*-th emotional state over time occurs. Output of this block is the vector of assessments of the psycho-emotional state (10) which is controlled by the supervisor (11).

There are also two blocks that implement continuous monitoring and collection of information:

- control system for the parameters of data collection and analysis (14);
- knowledge base.

The knowledge base, in turn, contains two databases. Database No. 1 – a set of training parameters for classifiers (12); database No. 2 – parameters of typical patterns of psycho-emotional state (13).

The administrator of the control system (15) sets the parameters for data analysis of the knowledge base blocks and decision making. Data mining specialist (16) implements the management of identified patterns.

3.2. Search algorithm for image area containing face

The process of finding face in an image using the HOG method in the Dlib library.

Step 1. Video sequence preprocessing. The algorithm is shown in Figure 3.

Step 2. Calculation of gradients. This step involves the calculation of horizontal and vertical pixel gradients, after which the histograms of the gradients are calculated.

In each pixel, the gradient has magnitude and direction. The gradient value in a pixel is the maximum gradient value of the three channels, and the angle is the angle corresponding to the maximum gradient.

Step 3. Calculation of histograms of cell gradients 8x8 pixels. At this stage, the image is divided into cells (8x8) and histograms of gradients for each of their cells are calculated.

Step 4. Normalization of blocks. Groups of 4 cells (8x8 pixels) are combined into blocks.

Step 5. Calculation of the HOG feature vector. Classification using SVM.

3.3. Algorithm for constructing essential features for facial recognition and emotional state

There are many methods for determining the reference points of a face. All of them are essentially trying to locate the following areas of the face: mouth, right eyebrow, left eyebrow, right eye, left eye, nose, jaw.

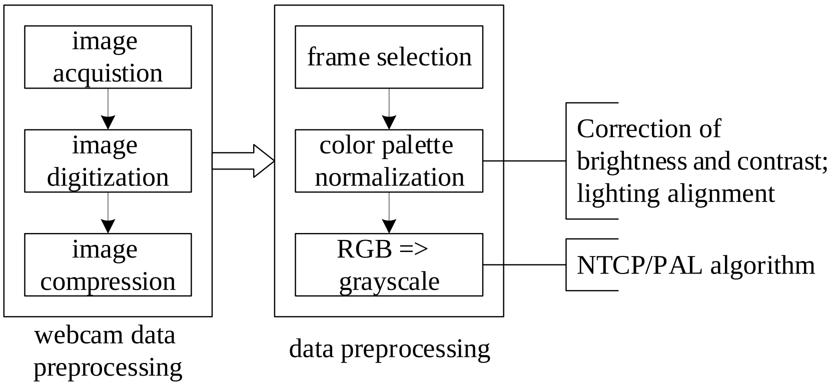


Figure 3. Video sequence preprocessing.

The method for determining essential features included in the Dlib library is an implementation of the method proposed by Kazemi and Sullivan in 2014 [22]. The end result is a face reference point detector that can be used to determine face points in real time with high-quality predictors [23]. This pre-trained detector is used to approximate the location of the 68 x and y coordinates [24], which form a map of parts of the face (Figure 4).

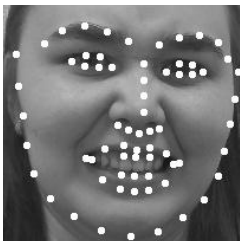


Figure 4. Block diagram of face search in the image.

Next, Euclidean distances between points are used as a function to predict the target variable – the probability of one of four emotions: “happiness”, “sadness”, “anger”, “wonderment”.

Then, the obtained set of coordinates of 68 reference points is converted into a matrix of pairwise distances:

$k = 68$ – number of control points; $P = \{(x_1, x_1), (x_2, x_2), \cdots, (x_1, x_1)\}$ – set of reference points; D – matrix of pairwise distances between reference points; $dist$ – Euclidean distance:

$$d_{i,j} = dist\left(p_i, p_j\right).$$

The resulting matrix D is normalized:

$$d_{\max} = \max_{i,j=1,k}\left\{d_{i,j}\right\} \to d_{i,j}^{n} = \frac{d_{i,j}}{d_{\max}}$$

and is converted line by line into a feature vector $l = k \cdot k$.

3.4. Face recognition and psycho-emotional state algorithm

Classifier training took place on a sample of 13,136 objects, testing - on 3,284 objects. Figure 5 below shows the classifier training process.

Emotions that were marked out in the training set: “happiness”, “joy”, “sadness”, “anger”, “wonderment”.

For encoding the output vector, the “one-hot” scheme was used – a binary code of fixed length containing only one 1 – direct unitary code or only one 0 – inverse (inverse) unitary code. The code length is determined by the number of encoded objects, that is, each object corresponds to a separate code bit, and the code value is position 1 or 0 in the code word.

The structure of the classifier is shown in Table 5, the training parameters in Table 6.

Table 5. Classifier structure.

Layer	**Activation function**
1-4	ReLU
Output layer	SoftMax

Table 6. Training options.

Parameter	Value
Packet size (batch)	64
Number of learning eras	45
Learning Speed Coefficient	0.0001

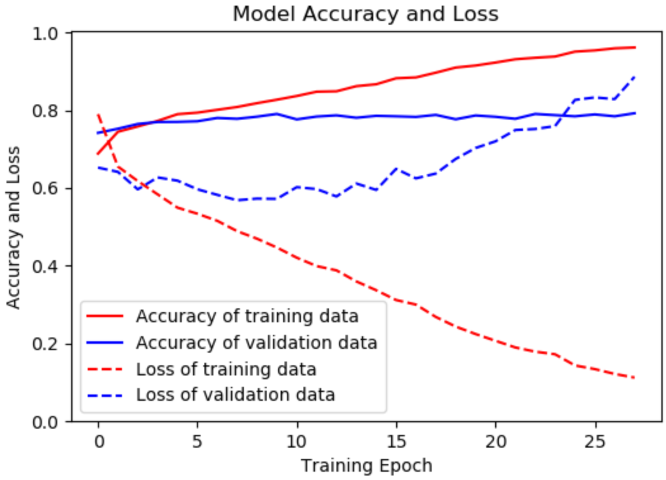


Figure 5. Changing the accuracy of the classifier in the training and test samples depending on the number of training eras.

On the training sample, the accuracy of the classifier was 97.62%, and on the test sample –79.48%. The results of the cross-validation are shown in Table 7.

Table 7. Recognition results using cross validation.

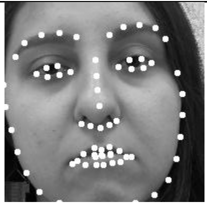
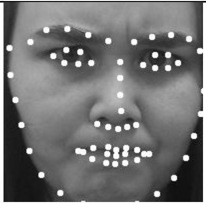
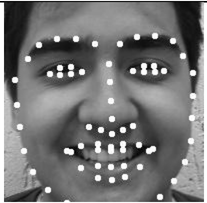
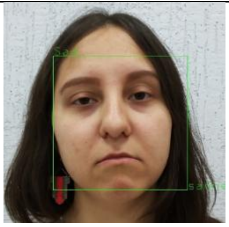
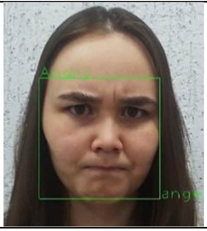
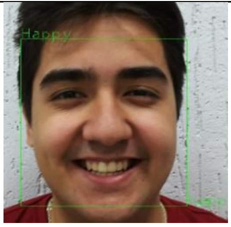
Parameter	Training sample				Test sample			
Accuracy	82.4 %				82.9 %			
Precision	82.1 %				82.7 %			
F_1	82.2 %				82.8 %			
Confusion matrix	75.67	3.93	15.94	4.87	75.78	3.62	15.58	4.26
	4.20	88.42	7.18	5.03	5.13	88.36	6.68	7.10
	16.17	5.06	73.98	5.15	14.67	5.25	74.77	2.84
	3.95	2.59	2.91	84.96	4.42	2.77	2.97	85.80

4. Conclusion

Human factor plays a significant role in ensuring the integrated safety of technological facilities. It is important to monitor the state of APCS operator in soft real time mode in order to reduce the risk of loss of attention and concentration. The risk of making a mistake due to the unstable (above the threshold values “joy” and “sadness”) of the psycho-emotional state increases by at least 15%.

Paper discusses the issues of increasing the efficiency of the monitoring system of the operator’s state by using algorithms for assessing the psycho-emotional state using the methods of intelligent analysis of video sequence data. The development and integration of algorithms for assessing the psycho-emotional state of the operator based on the data of the video sequence without the use of additional contact sensors can reduce the probability of making a wrong decision due to the timely detection of unstable psycho-emotional states.

Table 8. Recognition examples.

	Example 1	Example 2	Example 3
Highlighted features			
Recognized emotion			
True mark	Sadness	Anger	Joy

The following tasks have been solved:

1) Formalized patterns of the operator's psycho-emotional state have been developed for analyzing the state according to the video sequence.

2) The structural diagram of an operator's emotional state analysis system is developed.

3) Algorithm has been developed for analyzing the psycho-emotional state of the operator based on the methods of intellectual analysis of video sequence data.

The effectiveness of the proposed solution lies in the implementation of a functional in the operator's state monitoring system, which makes it possible to identify unstable psycho-emotional states without using additional contact sensors in soft real-time mode based on the technology of intelligent analysis of video sequence data. The accuracy of detecting unstable psycho-emotional states is 79%. The probability of timely detection of an unstable state without the use of video sequence analysis algorithms is significantly lower.

5. Acknowledgments

The reported study was funded by RFBR, project number 20-08-00668.

6. References

[1] Gorbunova, M.Yu. Emotions as a result and as a managerial resource of socialization // Vestnik SPbGU. – 2010. – Vol. 12(3). – P. 298-303.

[2] Picard, R.W. Affective computin – MIT Press, 2000. – 306 p.

[3] Jung, N. How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety / N. Jung, Ch. Wranke, K. Hamburger, M. Knauff // Frontiers in psychology. – 2014. – Vol. 5. – P. 1-12.

[4] Emotional states [Electronic resource]. – Access mode: https://syntone.ru/psy_lib/emotsionalnye-sostoyaniya/ (21.12.2019) (in Russian).

[5] Facial Action Coding System – Paul Ekman Group [Electronic resource]. – Access mode: https://www.paulekman.com/facial-action-coding-system/ (21.12.2019).

[6] Agrawal, S. Facial expression detection techniques: based on Viola and Jones algorithm and principal component analysis / S. Agrawal, P. Khatri // Fifth International Conference on Advanced Computing & Communication Technologies. – IEEE, 2015. – P. 108-112.

[7] Zhigalov, K.Yu. A review of existing approaches to face recognition and other related parameters based on neurointelligence / K.Yu. Zhigalov, A.V. Bogdanov // Estestvennye i tehnicheskie nauki. – 2017. – Vol. 12(114). – P. 278-287. (in Russian).

[8] Avsentyev, A.O. Application of Hidden Markov Models for speaker speech recognition / A.O. Avsentyev, A.S. Lukyanov // Problemy obespechenija bezopasnosti pri likvidacii posledstvij chrezvychajnyh situacij. – 2015. – Vol. 2. (in Russian).

[9] Petruk, V. Application of local binary patterns to the solution of the face recognition problem / V. Petruk, A.V. Samorodov, I.N. Spiridonov // Vestnik MGTU. Ser. “Priborostroenie”. – 2011. – P. 58-63.

[10] Li, W. Local binary patterns and extreme learning machine for hyperspectral imagery classification / W. Li, C. Chen, H. Su, Q. Du // IEEE Transactions on Geoscience and Remote Sensing. – 2015. – Vol. 53(7). – P. 3681-3693.

[11] Local Binary Patterns (LBP) [Electronic resource]. – Access mode: http://biomisa.org/uploads/2016/10/Lect-15.pdf (21.12.2019).

[12] The Database of Faces [Electronic resource]. – Access mode: https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html (21.12.2019).

[13] Zaboleeva-Zotova, A.V. Development of a system for automated determination of emotions and possible areas of application // Otkrytoe obrazovanie. – 2011. – Vol. 2-2. – P. 59-62.

[14] Deniz, O. Face recognition using histograms of oriented gradients / O. Deniz, G. Bueno, J. Salido, F. De la Torre // Pattern Recognition Letters. – 2011. – Vol. 32(12). – P. 1598-1603.

[15] Lindeberg, T. Scale invariant feature transform, 2012.

[16] Finogeev, A.G. The recognition method of images based on random trees in augmented reality computer-aided design systems / A.G. Finogeev, M.V. Chetvergova // Sovremennye problemy nauki i obrazovanija. – 2012. – Vol. 5. [Electronic resource]. – Access mode: http://www.science-education.ru/ru/article/view?id=7110 (21.12.2019).

[17] Ko, B. A brief review of facial emotion recognition based on visual information // Sensors (Basel). – 2018. – Vol. 18(2). – P. 401.

[18] Revina, I.M. A survey on human face expression recognition techniques / I.M. Revina, W.R.S. Emmanuel // Journal of King Saud University-Computer and Information Sciences. – 2018. – P. 1–10.

[19] EmoDetect [Electronic resource]. – Access mode: http://emodetect.ru/ (21.12.2019).

[20] Cohn-Kanade (CK and CK+) database Download Site [Electronic resource]. – Access mode: http://www.consortium.ri.cmu.edu/ckagree/ (21.12.2019).

[21] Lucey, P. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression // IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops. – 2010. – P. 94-101.

[22] Kazemi, V. One millisecond face alignment with an ensemble of regression trees / V. Kazemi, J. Sullivan // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2014. – P. 1867-1874.

[23] Facial landmarks with dlib, OpenCV, and Python – PyImageSearch [Electronic resource]. – Access mode: https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/ (21.12.2019).

[24] Sagonas, C. 300 faces in-the-wild challenge: Database and results / C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, M. Pantic // Image and vision computing. – 2016. – Vol. 47. – P. 3-18.