

ALGORITHM OF FACIAL METRIC FORMATION IN THE APPLICATION OF DETERMINING THE EMOTIONAL STATE OF A PERSON BY VIDEO SEQUENCE

Siarhei Sadau

Department of Radio Physics and Computer Technologies, Belarusian State University, Belarus
seregasadov@gmail.com

Most of the known computer systems which determine the emotional state of a person are based on obtaining data from an image. This approach assumes the availability of a sufficient amount of information about the object being analyzed. A more promising direction is based on the analysis of video sequences. The proposed algorithm allows to obtain all the necessary information about the object in real time.

The task of recognizing facial expressions in a video sequence includes the following main subtasks: search and tracing of individuals; selection and processing of informative facial characteristics; video segmentation; classification of facial expressions.

We give a brief description of the components of the facial expression recognition system for the video stream:

1. The module of sustainable search and face tracking. Such a module is the primary and most important part of the system. Its main tasks are deciding whether a person is in a video stream and narrowing the processing area in frames of a video sequence from the full space to the space containing the face.

2. Video segmentation module. Video segmentation methods allow to detect and calculate the movement characteristics of human facial muscles, and also to increase the efficiency of recognition of changes in facial expressions.

3. Classifier of the emotional state of a person. The input of the algorithm receives a set of sequences of facial expressions (of various durations) and a set of parameters characterizing emotional states. The task of the algorithm, in accordance with the proposed information model for describing the emotional state of a person, is to determine the form of the function F , which relates each sequence to one of the human states.

Face Search and Tracking Module

The classical problem of finding a face is usually solved using such well-known methods as the Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), artificial neural networks, Haar cascades, etc. These approaches of solving the problem have a number of disadvantages: low recognition accuracy with changes in face position in front of the camera. For this emotion recognition system module, it was decided to apply a hybrid algorithm that combines several approaches, such as the Haar cascades and the AdaBoost algorithm, as well as the use of neural networks. The hybrid scheme is constructed as follows: the components of the adaptive gain algorithm decide whether there is a face in the frame of the video sequence and rejects frames without faces. This makes it possible to use it in real time.

Video segmentation

The proposed video segmentation unit is a temporal segmentation of the entire video sequence into sections that contain frames of a single facial expression. This video segmentation algorithm contains the following steps: calculation of movements of pixels of a frame; building a time series; selection and presentation of the analyzed facial characteristics; application of the classification algorithm.

Classification of the emotional state of a person

The state array of facial objects obtained at the previous step of the algorithm is the basis for searching in it the compositions of objects describing a certain emotional state. Such search can be carried out using the methods of fuzzy logic and using neural networks. This a task is not unambiguous, since a person does not experience only one specific emotional state in a period of time. The development and description of methods for analyzing the complex emotional state of a person is the task of further research.

[1] Sochman J. AdaBoost with totally corrective updates for fast face detection, URL: <http://cmp.felk.cvut.cz/sochmj1>.

[2] Viola P., Jones J. Robust real-time object detection // Proc. of IEEE Workshop on Statistical and Computational Theories of Vision, 2001. P. 324 - 332.