

WHAT IS CLAIMED IS:

1. An electronic device, comprising: a memory, a processor and a photographic device, an action unit (AU) feature recognition program being stored in the memory and the AU feature recognition program being executed by the processor to implement the following steps of:

a real-time image capturing step: acquiring a real-time image shot by the photographic device and extracting a real-time facial image from the real-time image by use of a face recognition algorithm;

a facial feature point recognition step: inputting the real-time facial image into a pretrained facial mean shape and recognizing t facial feature points from the real-time facial image by use of the facial mean shape, wherein the training steps for the facial mean shape comprising: establishing a first sample library with n facial images and marking t facial feature points at positions of eyes, eyebrows, noses, mouths and facial outer contours in each of the facial images; and training a facial feature recognition model by use of the t facial feature points to obtain the facial mean shape, where t represents a concatenation sequence number;

a local feature extraction step: determining feature regions matched with each AU in the real-time facial image according to positions of the t facial feature points, extracting local features from the feature regions and generating multiple feature vectors; and

an AU feature prediction step: inputting the multiple feature vectors into pretrained AU classifiers matched with the feature regions respectively to obtain a prediction result of recognition of the corresponding AU features from the feature regions.

2. The electronic device of claim 1, wherein the AU feature recognition program is executed by the processor to further implement the following step:

a judgment step: judging whether probabilities of each AU feature in the prediction result are higher than a preset threshold value or not.

3. The electronic device of claim 2, wherein the judgment step further comprises:

a prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.

4. The electronic device of claim 1, wherein training steps for the AU classifier comprises:

a sample preparation step: collecting a facial sample image, cropping an image region matched with each AU from the facial sample image as a positive sample image of the AU and preparing a negative sample image for each AU;

a local feature extraction step: extracting local features of the positive sample image and negative sample image of each AU and generating corresponding feature vectors; and

a model training step: performing learning training on a Support Vector Machine (SVM) by use of the local feature of the positive/negative sample image of each AU to obtain the corresponding AU classifier.

5. The electronic device of claim 4, wherein the AU feature recognition program is executed by the processor to further implement the following step:

the judgment step: judging whether the probabilities of each AU feature in the prediction result are higher than the preset threshold value or not.

6. The electronic device of claim 5, wherein the judgment step further comprises:

the prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.

7. An action unit (AU) feature recognition method, comprising:

a real-time image capturing step: acquiring a real-time image shot by the photographic device and extracting a real-time facial image from the real-time image by use of a face recognition algorithm;

a facial feature point recognition step: inputting the real-time facial image into a pretrained facial mean shape and recognizing t facial feature points from the real-time facial image by use of the facial mean shape, wherein the training steps for the facial mean shape comprising: establishing a first sample library with n facial images and marking t facial feature points at positions of eyes, eyebrows, noses, mouths and facial outer contours in each of the facial images; and training a facial feature recognition model by use of the t facial feature points to obtain the facial mean shape, where t represents a concatenation sequence number;

a local feature extraction step: determining feature regions matched with each AU in the real-time facial image according to positions of the t facial feature points, extracting local features from the feature regions and generating multiple feature vectors; and

an AU feature prediction step: inputting the multiple feature vectors into pretrained AU classifiers matched with the feature regions respectively to obtain a prediction result of recognition of the corresponding AU features from the feature regions.

8. The AU feature recognition method of claim 7, wherein the facial mean shape is obtained by training a facial feature recognition model, and the facial feature recognition model is an ensemble of regression trees (ERT) algorithm and is represented by a formula as follows:

$$\hat{S}^{t+1} = \hat{S}^t + \tau_t(I, \hat{S}^t),$$

where $\tau_t(\cdot, \cdot)$ represents a regressor of a present stage and $S(t)$ is a shape estimate of a present model; each regressor $\tau_t(\cdot, \cdot)$ predicts an increment $\tau_t(I, \hat{S}^t)$ according to an input present image I and $S(t)$ and adds the increment to the present shape estimate to improve the present model; and in a model training process, part of feature points of all sample pictures are extracted to train a first regression tree, a second tree is trained by use of residuals between predicted values of the first regression tree and true values of the part of feature points, and so on, when residuals of predicted values of an Nth trained tree and the true values of the part of feature points are close to zero, all regression trees of the ERT algorithm are obtained, and the facial mean shape is obtained according to these regression trees.

9. The AU feature recognition method of claim 8, further comprising:

a judgment step: judging whether probabilities of each AU feature in the prediction result are higher than the preset threshold value or not.

10. The AU feature recognition method of claim 9, wherein the judgment step further comprises:

a prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.

11. The AU feature recognition method of claim 7, wherein training steps for the AU classifier comprise:

a sample preparation step: collecting a facial sample image, cropping an image region matched with each AU from the facial sample image as a positive sample image of the AU and preparing a negative sample image for each AU;

a local feature extraction step: extracting local features of the positive sample image and negative sample image of each AU and generating corresponding feature vectors; and

a model training step: performing learning training on a Support Vector Machine (SVM) by use of the local feature of the positive/negative sample image of each AU to obtain the corresponding AU classifier.

12. The AU feature recognition method of claim 11, further comprising:

the judgment step: judging whether the probabilities of each AU feature in the prediction result are higher than the preset threshold value or not.

13. The AU feature recognition method of claim 12, wherein the judgment step further comprises:

the prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.

14. A computer-readable storage medium, in which an action unit (AU) feature recognition program is stored, the AU feature recognition program being executed by a processor to implement the following steps of:

a real-time image capturing step: acquiring a real-time image shot by the photographic device and extracting a real-time facial image from the real-time image by use of a face recognition algorithm;

a facial feature point recognition step: inputting the real-time facial image into a pretrained facial mean shape and recognizing t facial feature points from the real-time facial image by use of the facial mean shape, wherein the training steps for the facial mean shape comprising: establishing a first sample library with n facial images and marking t facial feature points at positions of eyes, eyebrows, noses, mouths and facial outer contours in each of the facial images; and training a facial feature recognition model by use of the t facial feature points to obtain the facial mean shape, where t represents a concatenation sequence number;

a local feature extraction step: determining feature regions matched with each AU in the real-time facial image according to positions of the t facial feature points, extracting local features from the feature regions and generating multiple feature vectors; and

an AU feature prediction step: inputting the multiple feature vectors into pretrained AU classifiers matched with the feature regions respectively to obtain a prediction result of recognition of the corresponding AU features from the feature regions.

15. The computer-readable storage medium of claim 14, wherein the facial mean shape is obtained by training a facial feature recognition model, and the facial feature recognition model is an ensemble of regression trees (ERT) algorithm and is represented by a formula as follows:

$$\hat{S}^{t+1} = \hat{S}^t + \tau_t(I, \hat{S}^t),$$

where $\tau_t(\cdot, \cdot)$ represents a regressor of a present stage and $S(t)$ is a shape estimate of a present model; each regressor $\tau_t(\cdot, \cdot)$ predicts an increment $\tau_t(I, \hat{S}^t)$ according to an input present image I and $S(t)$ and adds the increment to the present shape estimate to improve the present model; and in a model training process, part of feature points of all sample pictures are extracted to train a first regression tree, a second tree is trained by use of residuals between predicted values of the first regression tree and true values of the part of feature points, and so on, when residuals of predicted values of an N th trained tree and the true values of the part of feature points are close to zero, all regression trees of the ERT algorithm are obtained, and the facial mean shape is obtained according to these regression trees.

16. The computer-readable storage medium of claim 15, wherein the AU feature recognition program may be executed by the processor to further implement the following step:

a judgment step: judging whether probabilities of each AU feature in the prediction result are higher than the preset threshold value or not; and

a prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.

17. The computer-readable storage medium of claim 14, wherein training steps for the AU classifier comprise:

a sample preparation step: collecting a facial sample image, cropping an image region matched with each AU from the facial sample image as a positive sample image of the AU and preparing a negative sample image for each AU;

a local feature extraction step: extracting local features of the positive sample image and negative sample image of each AU and generating corresponding feature vectors; and

a model training step: performing learning training on a Support Vector Machine (SVM) by use of the local feature of the positive/negative sample image of each AU to obtain the corresponding AU classifier.

18. The computer-readable storage medium of claim 17, wherein the AU feature recognition program may be executed by the processor to further implement the following step:

the judgment step: judging whether the probabilities of each AU feature in the prediction result are higher than the preset threshold value or not; and

the prompting step: when it is judged that an AU feature of which the probability is higher than the preset threshold value exists in the prediction result, making such a prompt that the AU feature is recognized from the real-time facial image.