CLAIMS

What is claimed is:

1. A method for training an adaptive rigid prior model, comprising:

initializing model parameters of the adaptive rigid prior model;

acquiring a plurality of frames of facial data of a same face by face tracking on training video data using the adaptive rigid prior model, wherein the facial data comprises facial sizes of faces and displacements of a plurality of facial regions of the faces;

updating the model parameters based on the plurality of frames of facial data;

determining whether a condition of stopping updating the model parameters is satisfied;

stopping, in response to the condition of stopping updating the model parameters being satisfied, updating the model parameters of the adaptive rigid prior model, and acquiring a final adaptive rigid prior model, wherein model parameters of the final adaptive rigid prior model are functions with the facial sizes as independent variables, an input of the final adaptive rigid prior model is facial sizes of faces in video data for face tracking, and an output of the final adaptive rigid prior model is weights of the plurality of facial regions of faces at different facial sizes; and

returning, in response to the condition of stopping updating the model parameters being not satisfied, to the step of acquiring the plurality of frames of facial data of the same face by face tracking on the training video data using the adaptive rigid prior model.

2. The method according to claim 1, wherein acquiring the plurality of frames of facial data of the same face by face tracking on the training video data using the adaptive rigid prior model comprises:

extracting facial key points and facial expression parameters of faces in a plurality of frames of the training video data by face tracking on the training video data;

determining the facial sizes of the faces based on the facial key points; and

calculating displacements of the plurality of facial regions of the faces relative to a neutral face of the faces using the adaptive rigid prior model, the facial expression parameters, and a predetermined blend shape deformer.

3. The method according to claim 1 or claim 2, wherein

updating the model parameters based on the plurality of frames of facial data comprises:

organizing the faces to a plurality of face containers based on the facial sizes;

calculating average values of the facial sizes of the faces in the plurality of face containers;

acquiring, in a case that the facial sizes are the average values, data points of the model parameters of the adaptive rigid prior model by analyzing the displacements of the plurality of facial regions of the faces in the plurality of face containers; and

acquiring the model parameters by interpolating the data points of the model parameters corresponding to different average values of the facial sizes in the plurality of face containers, and determining the model parameters as new model parameters of the adaptive rigid prior model, wherein the model parameters are functions of the facial sizes.

4. The method according to claim 3, wherein organizing the faces corresponding to facial image data to the plurality of face containers based on the facial sizes comprises:

determining a maximum facial size and a minimum facial size based on the facial image data;

evenly dividing a size range from the maximum facial size to the minimum facial size into a plurality of size ranges, and determining the plurality of size ranges as facial size ranges of the plurality of face containers;

determining target facial size ranges of the facial sizes of the faces; and

organizing the faces to face containers within the target facial size ranges.

5. The method according to claim 3, wherein acquiring, in the case that the facial sizes are the average values, the data points of the model parameters of the adaptive rigid prior model by analyzing the displacements of the plurality of facial regions of the faces in the plurality of face containers comprises:

determining, based on facial image data of the faces, maximum displacements of the plurality of facial regions for the plurality of facial regions of the faces in the plurality of face containers;

organizing the displacements of the plurality of facial regions to a plurality of displacement containers;

determining a displacement container containing a greatest number of displacements in the plurality of displacement containers as a maximum displacement container;

determining a middle displacement of the plurality of displacements in the maximum displacement container as a data point of a first model parameter; and

calculating a data point of a second model parameter and a data point of a third model parameter based on the maximum displacements of the plurality of facial regions, the middle displacement, a predetermined maximum weight, a predetermined minimum weight, and the data point of the first model parameter;

wherein in the case that the facial sizes are the average values, the data points of the model parameters of the adaptive rigid prior model comprise: the data point of the first model parameter, the data point of the second model parameter, and the data point of the third model parameter.

6. The method according to claim 5, wherein acquiring the model parameters by interpolating the data points of the model parameters corresponding to the different average values of the facial sizes in the plurality of face containers, and determining the model parameters as the new model parameters of the adaptive rigid prior model comprise:

acquiring the first model parameter, the second model parameter, and the third model parameter by spline interpolating the data point of the first model parameter, the data point of the second model parameter, and the data point of the third model parameter in a case that the facial sizes in the plurality of face containers are the different average values, and determining the first model parameter, the second model parameter, and the third model parameter as the new model parameters of the adaptive rigid prior model.

7. The method according to claim 1, wherein determining whether the condition of stopping updating the model parameters is satisfied comprises:

calculating difference values between model parameters acquired by two updates;

determining whether the difference values are less than a first predetermined threshold; and

determining, in response to the difference values being less than the first predetermined threshold, that the condition of stopping updating the model parameters is satisfied.

8. The method according to claim 1, wherein determining whether the condition of stopping updating the model parameters is satisfied comprises:

counting a number of updates of the model parameters;

determining whether the number of updates is greater than a second predetermined threshold; and

determining, in response to the number of updates being greater than the second predetermined threshold, that the condition of stopping updating the model parameters is satisfied.

9. The method according to any one of claims 1, 2, 7, and 8, wherein the adaptive rigid prior model is:

$$w_s^k(d^k) = \begin{cases} \alpha^k(s) exp\left(-\frac{\gamma^k(s)}{\beta^k(s)}\right) & d^k < \gamma^k(s) \\ \alpha^k(s) exp\left(-\frac{d^k}{\beta^k(s)}\right) & d^k \geq \gamma^k(s) \end{cases};$$

wherein s represents the facial size of the face, $w_s^k(\cdot)$ represents a weight of a kth facial region in the face with the facial size being s, $\alpha^k(s)$, $\beta^k(s)$, and $\gamma^k(s)$ represent the model parameters, $\alpha^k(s)$, $\beta^k(s)$, and $\gamma^k(s)$ are functions of two-dimensional facial sizes s, and d^k represent a displacement of the kth facial region.

10. A method for tracking faces, comprising:

extracting a plurality of facial key points of a same face, initial facial expression parameters, and initial rigid transformation parameters in a plurality of frames of video data by face tracking on the video data;

determining facial sizes of faces in the plurality of frames of video data based on the plurality of facial key points;

acquiring weights of a plurality of facial regions of the faces by inputting the facial sizes to a pre-trained adaptive rigid prior model; and

acquiring optimal facial expression parameters based on the weights of the plurality of facial regions, the plurality of facial key points, the initial rigid transformation parameters, and the initial facial expression parameters, and determining the optimal facial expression parameters as results of face tracking on the plurality of frames of video data;

wherein the adaptive rigid prior model is trained by the method for training the adaptive rigid prior model as defined in any one of claims 1 to 9.

11. The method according to claim 10, wherein acquiring the optimal facial expression parameters based on the weights of the plurality of facial regions, the plurality of facial key points, the initial rigid transformation parameters, and the initial facial expression parameters, and determining the optimal facial expression parameters as the results of face tracking on the plurality of frames of video data comprise:

acquiring neutral face data of the faces;

generating three-dimension face models of the faces in the plurality of frames video data based on the neutral face data, the initial facial expression parameters, and a predetermined blend shape deformer, and rigidly transforming the three-dimension face models;

acquiring two-dimensional facial data by performing two-dimensional projection on the three-dimension face models based on the initial rigid transformation parameters, wherein the two-

dimensional facial data comprises a plurality of projected two-dimensional key points in one-to-one correspondence to the plurality of facial key points;

calculating distances between the plurality of projected two-dimensional key points and the corresponding plurality of facial key points;

constructing a rigid transformation optimization function and a facial expression optimization function based on the weights of the plurality of facial regions of the plurality of projected two-dimensional key points and the distances;

acquiring optimal rigid transformation parameters by optimizing the rigid transformation optimization function; and

acquiring optimal facial expression parameters by optimizing the facial expression optimization function based on the optimal rigid transformation parameters, and determining the optimal facial expression parameters as the results of face tracking on the plurality of frames of video data.

12. An apparatus for training an adaptive rigid prior model, comprising:

a model parameter initializing module, configured to initialize model parameters of the adaptive rigid prior model;

a facial data acquiring module, configured to acquire a plurality of frames of facial data of a same face by face tracking on training video data using the adaptive rigid prior model, wherein the facial data comprises facial sizes of faces and displacements of a plurality of facial regions of the faces;

a model parameter updating module, configured to update the model parameters based on the plurality of frames of facial data;

a stopping training condition determining module, configured to determine whether a condition of stopping updating the model parameters is satisfied;

a model parameter stopping updating module, configured to stop updating the model parameters of the adaptive rigid prior model, and acquire a final adaptive rigid prior model, wherein model parameters of the final adaptive rigid prior model are functions with the facial sizes as independent variables, an input of the final adaptive rigid prior model is facial sizes of faces in video data for face tracking, and an output of the final adaptive rigid prior model is weights of the plurality of facial regions of faces at different facial sizes.

13. An apparatus for tracking faces, comprising:

a face tracking module, configured to extract a plurality of facial key points of a same face, initial facial expression parameters, and initial rigid transformation parameters in a plurality of frames of video data by face tracking on the video data;

a facial size determining module, configured to determine facial sizes of faces in the plurality of frames of video data based on the plurality of facial key points;

a weight predicting module, configured to acquire weights of a plurality of facial regions of the faces by inputting the facial sizes to a pre-trained adaptive rigid prior model; and

an optimizing module, configured to acquire optimal facial expression parameters based on the weights of the plurality of facial regions, the plurality of facial key points, the initial rigid transformation parameters, and the initial facial expression parameters, and determine the optimal facial expression parameters as results of face tracking on the plurality of frames of video data;

wherein the adaptive rigid prior model is trained by the method for training the adaptive rigid prior model as defined in any one of claims 1 to 9.

14. An electronic device, comprising:

one or more processors;

a memory configured to store one or more programs;

wherein the one or more processors, when loading and executing the one or more programs, are caused to perform at least one of: the method for training the adaptive rigid prior model as defined in any one of claims 1 to 9, and the method for tracking the faces as defined in any one of claims 10 to 11.

15. A computer-readable storage medium, storing one or more computer programs, wherein the one or more programs, when loaded and run by a processor, causes the processor to perform at least one of: the method for training the adaptive rigid prior model as defined in any one of claims 1 to 9, and the method for tracking the faces as defined in any one of claims 10 to 11.