

CLAIMS

What is claimed is:

1. A method for detecting line-of-sight, performed by an apparatus for detecting line-of-sight, comprising:

determining, based on a key feature point in a face image, a face posture and an eye pupil rotational displacement corresponding to the face image, wherein the eye pupil rotational displacement is a displacement of a pupil center relative to an eyeball center in the face image; and

acquiring a line-of-sight direction of an actual face by back-projecting, based on a preset projection function and the face posture, the eye pupil rotational displacement to a three-dimensional space where the actual face is located;

wherein acquiring the line-of-sight direction of the actual face by back-projecting, based on the preset projection function and the face posture, the eye pupil rotational displacement to the three-dimensional space where the actual face is located comprises:

constructing a corresponding line-of-sight optimization function based on the preset projection function, the face posture, and the eye pupil rotational displacement; and

acquiring a line-of-sight direction that reaches a preset optimization target in the line-of-sight optimization function in the three-dimensional space where the actual face is located, and taking an acquired line-of-sight direction as the line-of-sight direction of the actual face.

2. The method according to claim 1, wherein prior to determining, based on the key feature point in the face image, the face posture and the eye pupil rotational displacement corresponding to the face image, the method further comprises:

acquiring face data corresponding to the face image by scanning the face image;

acquiring a reconstructed face grid model by reconstructing a preset three-dimensional face grid template with the face data; and

extracting the key feature point in the reconstructed face grid model, and taking the key feature point in the reconstructed face grid model as the key feature point in the face image.

3. The method according to claim 2, wherein determining, based on the key feature point in the face image, the face posture corresponding to the face image comprises:

determining a posture position of the key feature point in the face image; and

determining the face posture in the face image based on the posture position.

4. The method according to claim 2, wherein determining, based on the key feature point in the face image, the eye pupil rotational displacement corresponding to the face image comprises:

determining, based on the key feature point in the reconstructed face grid model, the eyeball center corresponding to the face image;

acquiring the pupil center corresponding to the eye image by recognizing an eye image in the face image; and

determining, based on a position of the eyeball center and a position of the pupil center, the eye pupil rotational displacement corresponding to the face image.

5. The method according to claim 4, wherein acquiring the pupil center corresponding to the eye image by recognizing the eye image in the face image comprises:

capturing the eye image in the face image; and

acquiring the pupil center corresponding to the eye image by inputting the eye image into a pre-constructed deep network model.

6. The method according to any of claims 2 to 5, wherein upon acquiring the line-of-sight direction of the actual face by back-projecting, based on the preset projection function and the face posture, the eye pupil rotational displacement to the three-dimensional space where the actual face is located, the method further comprises:

constructing a corresponding three-dimensional eye model in the reconstructed face grid model based on the line-of-sight direction of the actual face.

7. The method according to claim 1, wherein the face posture comprises a rotation parameter and a translation parameter of the face in the face image.

8. The method according to claim 1, wherein prior to acquiring the line-of-sight direction that reaches the preset optimization target in the line-of-sight optimization function in the three-dimensional space where the actual face is located, and taking the acquired line-of-sight direction as the line-of-sight direction of the actual face, the method further comprises:

acquiring an associated image of the face image, wherein the associated image carries the line-of-sight direction corresponding to the associated image; and

updating the line-of-sight optimization function based on the line-of-sight direction of the associated image, a preset associated smoothing parameter, and a preset anti-shake parameter.

9. The method according to claim 8, wherein the updated line-of-sight optimization function is: $\min \left\{ \left\| \Pi(Rx + t) - d \right\|^2 + \alpha \|x - x_0\|^2 + \beta \|x\|^2 \right\};$

wherein $\Pi(Rx + t)$ represents the preset projection function, R represents the rotation parameter in the face posture, t represents the translation parameter in the face posture, x represents the line-of-sight direction of the face image, d represents the eye pupil rotational displacement, x_0 represents the line-of-sight direction of the associated image, α represents the preset associated smoothing parameter, and β represents the preset anti-shake parameter.

10. The method according to any of claims 1 to 5, wherein prior to acquiring the line-of-sight direction of the actual face by back-projecting, based on the preset projection function and the face posture, the eye pupil rotational displacement to the three-dimensional space where the actual face is located, the method further comprises:

determining the preset projection function based on a parameter of a capture device corresponding to the face image.

11. The method according to claim 10, wherein the parameter of the capture device is an imaging focal length of the capture device.

12. A method for processing videos, comprising:

acquiring a video frame in a to-be-processed video; and

acquiring the line-of-sight direction of the actual face corresponding to the video frame by performing the method for detecting line-of-sight as defined in any one of claims 1 to 11.

13. The method according to claim 12, wherein upon acquiring the line-of-sight direction corresponding to the video frame, the method further comprises:

determining, based on the line-of-sight direction corresponding to an adjacent video frame in the to-be-processed video, a line-of-sight offset corresponding to the adjacent video frame; and

performing, based on the line-of-sight offset, a video edit operation corresponding to the line-of-sight offset.

14. An apparatus for detecting line-of-sight, comprising:

a parameter determining module, configured to determine, based on a key feature point in a face image, a face posture and an eye pupil rotational displacement corresponding to the face image, wherein the eye pupil rotational displacement is a displacement of a pupil center relative to an eyeball center in the face image; and

a line-of-sight detecting module, configured to acquire a line-of-sight direction of an actual face by back-projecting, based on a preset projection function and the face posture, the eye pupil rotational displacement to a three-dimensional space where the actual face is located by: constructing a corresponding line-of-sight optimization function based on the preset projection function, the face posture, and the eye pupil rotational displacement; and acquiring a line-of-sight direction that reaches a preset optimization target in the line-of-sight optimization function in the three-dimensional space where the actual face is located, and taking an acquired line-of-sight direction as the line-of-sight direction of the actual face.

15. An apparatus for processing videos, comprising:

a video frame acquiring module, configured to acquire a video frame in a to-be-processed video; and

a line-of-sight detecting module, configured to acquire the line-of-sight direction of the actual face corresponding to the video frame by performing the method for detecting line-of-sight as defined in any one of claims 1 to 11.

16. A system for processing line-of-sight, comprising: a capture device and a processing device that are connected in communication, wherein the capture device is disposed on the processing device; and

the capture device captures a to-be-detected face image and a to-be-processed video, and sends the to-be-detected face image and the to-be-processed video to the processing device;

and the processing device is provided with the apparatus for detecting line-of-sight as defined in claim 14 and the apparatus for processing videos as defined in claim 15.

17. A computer device, comprising:
at least one processor; and
a memory configured to store at least one program; wherein
the at least one program, when run by the at least one processor, causes the at least one processor to perform the method for detecting line-of-sight as defined in any of claims 1 to 11, or perform the method for processing videos as defined in claim 12 or 13.

18. A non-volatile computer-readable storage medium storing a computer program, wherein the computer program, when run by a processor, causes the processor to perform the method for detecting line-of-sight as defined in any one of claims 1 to 11, or perform the method for processing videos as defined in claim 12 or 13.