Claims

1. A localization and tracking method, comprising:

acquiring one or more images of odd frames that are collected with a preset first exposure duration by one or more tracking cameras arranged on a head-mounted display device as well as one or more images of even frames that are collected with a preset second exposure duration by the one or more tracking cameras arranged on the head-mounted display device, wherein the one or more images of even frames at least contain blobs corresponding to multiple luminous bodies arranged on a gamepad, and the gamepad is in communication connection with the head-mounted display device;

determining Degree of Freedom (DoF) information of the head-mounted display device according to the one or more images of odd frames and attitude information of the head-mounted display device; and

determining DoF information of the gamepad according to the one or more images of even frames, attitude information of the gamepad, and the DoF information of the head-mounted display device.

2. The method according to claim 1, wherein the preset second exposure duration is less than the first exposure duration.

3. The method according to claim 1, wherein the first exposure duration is adaptively adjusted according to an intensity of ambient light.

4. The method according to claim 1, wherein at least two tracking cameras are arranged on the head-mounted display device, exposure durations of respective tracking cameras having a same center point;

acquiring one or more images of odd frames that are collected with a preset first exposure duration by one or more tracking cameras arranged on a head-mounted display device as well as one or more images of even frames that are collected with a preset second exposure duration by the one or more tracking cameras arranged on the head-mounted display device comprises:

collecting, by the at least two tracking cameras, multiple images of odd frames with the preset

first exposure duration, each image of odd frame corresponding to each tracking camera; and collecting, by the at least two tracking cameras, multiple images of even frames with the preset second exposure duration, each image of even frame corresponding to each tracking camera.

5. The method according to claim 1, further comprising:

controlling the luminous bodies on the gamepad to be turned on according to a preset lighting duration when the one or more images of even frames are collected with the preset second exposure duration by the one or more tracking cameras arranged on the head-mounted display device, a middle time of the preset second exposure duration corresponding to a middle time of the preset lighting duration, wherein the preset lighting duration is equal to or greater than the preset second exposure duration.

6. The method according to claim 1, wherein determining DoF information of the head-mounted display device according to the one or more images of odd frames and attitude information of the head-mounted display device comprises:

calculating by a Simultaneous Localization and Mapping (SLAM) algorithm according to spatial structure information in the one or more images of odd frames, to obtain DoF of translational movement of the head-mounted display device;

obtaining DoF of rotational movement of the head-mounted display device according to the attitude information of the head-mounted display device that is collected by a first Inertial Measurement Unit (IMU) integrated in the head-mounted display device; and

determining the DoF information of the head-mounted display device based on the DoF of translational movement and the DoF of rotational movement.

7. The method according to claim 1, wherein after determining DoF information of the head-mounted display device according to the one or more images of odd frames and attitude information of the head-mounted display device, the method further comprises:

performing pre-integration on the DoF information of the head-mounted display device with reference to the attitude information of the head-mounted display device to obtain high-frequency DoF information of the head-mounted display device.

8. The method according to claim 1, wherein determining DoF information of the gamepad according to the one or more images of even frames, attitude information of the gamepad, and the DoF information of the head-mounted display device comprises:

performing blob detection on the one or more images of even frames to determine position information of all the blobs in the one or more images of even frames;

determining three-dimensional coordinates of the luminous bodies on the gamepad corresponding to the blobs in the one or more images of even frames according to the position information of all the blobs in the one or more images of even frames; and

determining the DoF information of the gamepad according to the three-dimensional coordinates of the luminous bodies on the gamepad corresponding to the blobs in the one or more images of even frames, the attitude information of the gamepad, and the DoF information of the head-mounted display device.

9. The method according to claim 8, wherein performing blob detection on the one or more images of even frames to determine position information of all the blobs in the one or more images of even frames comprises:

performing blob detection on each of the one or more images of even frames that are collected by the one or more tracking cameras to determine the position information of all the blobs in each of the one or more images of even frames, wherein the position information of all the blobs at least comprises position information of four blobs;

determining a correspondence between the blobs in the one or more images of even frames and the luminous bodies on the gamepad according to a distribution rule of each blob in the one or more images of even frames; and

determining the three-dimensional coordinates of the corresponding luminous bodies on the gamepad based on two-dimensional coordinates of the corresponding luminous bodies on the gamepad.

10. The method according to claim 8, wherein determining the DoF information of the gamepad according to the three-dimensional coordinates of the luminous bodies on the gamepad corresponding

to the blobs in the one or more images of even frames, the attitude information of the gamepad, and the DoF information of the head-mounted display device comprises:

determining DoF of translational movement according to the three-dimensional coordinates of the luminous bodies on the gamepad corresponding to the blobs in the one or more images of even frames;

determining DoF of rotational movement of the gamepad according to the attitude information of the gamepad;

determining the DoF information of the gamepad according to the DoF of translational movement and the DoF of rotational movement; and

performing coordinate transformation on the DoF information of the gamepad according to the DoF information of the head-mounted display device.

11. The method according to claim 8, wherein after determining the DoF information of the gamepad according to the three-dimensional coordinates of the luminous bodies on the gamepad corresponding to the blobs in the one or more images of even frames, the attitude information of the gamepad, and the DoF information of the head-mounted display device, the method further comprises:

performing pre-integration on the DoF information of the gamepad with reference to the attitude information of the gamepad to obtain high-frequency DoF information of the gamepad.

12. The method according to claim 1, wherein the one or more images of odd frames at least contain hand or hands, and the method further comprises:

determining DoF information of the hand or hands according to the one or more images of odd frames.

13. The method according to claim 12, wherein determining DoF information of the hand or hands according to the one or more images of odd frames comprises:

inputting the one or more images of odd frames to a preset convolutional neural network model to obtain positions of multiple key points of the hand or hands; and

determining the DoF information of the hand or hands according to the positions of the multiple

key points of the hand or hands.

14. The method according to claim 13, further comprising an operation of training the preset convolutional neural network model, which comprises:

acquiring multiple scenario images, the scenario images containing hand or hands;

marking positions of the multiple key points of the hand or hands in the multiple scenario images, and forming an image training sample set with the multiple marked scenario images; and

training the convolutional neural network model according to the image training sample set.

15. A non-transitory computer-readable storage medium, in which a computer program is stored, the computer program being executed by a processor to implement the method according to claim 1.

16. A localization and tracking platform, comprising a memory storing instructions and a processor in communication with the memory, wherein the processor is configured to execute the instructions to:

acquire one or more images of odd frames that are collected with a preset first exposure duration by one or more tracking cameras arranged on a head-mounted display device as well as one or more images of even frames that are collected with a preset second exposure duration by the one or more tracking cameras arranged on the head-mounted display device, wherein the one or more images of even frames at least contain blobs corresponding to multiple luminous bodies arranged on a gamepad, and the gamepad is in communication connection with the head-mounted display device;

determine DoF information of the head-mounted display device according to the one or more images of odd frames and attitude information of the head-mounted display device; and

determine DoF information of the gamepad according to the one or more images of even frames, attitude information of the gamepad, and the DoF information of the head-mounted display device.

17. A head-mounted display system, wherein the head-mounted display system comprises a head-mounted display device, a gamepad connected to the head-mounted display device, and a

localization and tracking platform, at least two tracking cameras are arranged on the head-mounted display device, and the gamepad is provided with multiple luminous bodies, wherein

the localization and tracking platform comprises a memory storing instructions and a processor in communication with the memory, wherein the processor is configured to execute the instructions to:

acquire one or more images of odd frames that are collected with a preset first exposure duration by one or more tracking cameras arranged on a head-mounted display device as well as one or more images of even frames that are collected with a preset second exposure duration by the one or more tracking cameras arranged on the head-mounted display device, wherein the one or more images of even frames at least contain blobs corresponding to multiple luminous bodies arranged on a gamepad, and the gamepad is in communication connection with the head-mounted display device;

determine DoF information of the head-mounted display device according to the one or more images of odd frames and attitude information of the head-mounted display device; and

determine DoF information of the gamepad according to the one or more images of even frames, attitude information of the gamepad, and the DoF information of the head-mounted display device.

18. The system according to claim 17, wherein exposure durations of respective tracking cameras arranged on the head-mounted display device have a same center point.

19. The system according to claim 17, wherein four tracking cameras are arranged on the head-mounted display device, wherein the four tracking cameras are arranged in the head-mounted display device according to a preset first position constraint rule, and the preset first position constraint rule is used for ensuring that each tracking camera meets requirements of the head-mounted display device on Field Angle Of View (FOV).

20. The system according to claim 17, wherein multiple luminous bodies are arranged on the gamepad according to a preset second position constraint rule, wherein being arranged according to the preset second position constraint rule meets the following conditions: coincidence or adhesion of

local pattern information in images collected within a preset distance range from the tracking cameras to the gamepad does not occur, and blobs corresponding to at least four luminous bodies are presented in every image collected by the one or more tracking cameras at any angles.