WHAT IS CLAIMED IS:

1. A method for identifying an operation event, comprising:

performing object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames, wherein the object is an operable object; and

determining an occurred object-operation-event based on the object-change-information,

wherein performing the object detection and tracking on the at least two image frames of the video to obtain the object-change-information of the object involved in the at least two image frames comprises:

detecting a respective object position of an object in each of the at least two image frames of the video;

identifying the object detected in each of the at least two image frames to obtain respective object identification results;

based on the respective object positions and the respective object identification results of objects detected in different image frames, comparing the objects detected in the different image frames to establish a correspondence between the objects in the at least two image frames; and

obtaining the object-change-information of the object involved in the at least two image frames by comparing the objects involved in the at least two image frames,

wherein comparing the objects detected in the different image frames to establish the correspondence between the objects in the at least two image frames comprises:

establishing a position similarity matrix between the objects in the different

image frames based on the object positions;

establishing an identification result similarity matrix between the objects in the different image frames based on the object identification results;

obtaining an object similarity matrix based on the position similarity matrix and the identification result similarity matrix; and

determining the correspondence between the objects in the at least two image frames based on the object similarity matrix.

2. The method of claim 1, wherein determining an occurred object-operation-event based on the object-change-information comprises:

in response to that the object-change-information satisfies a preset event change condition, and an object operator is detected in at least a part of the at least two image frames, and a distance between a position of the object operator and a position of the object is within a preset distance threshold, determining that an object-operation-event corresponding to the event change condition is occurred by the object operator operating on the object.

3. The method of claim 2, wherein the object operator comprises a hand or an object holding tool.

4. The method of claim 1, wherein

performing object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames comprises:

detecting a first object newly appeared in the at least two image frames, and

determining an object position where the first object appeared in the at least two image frames as a first target area; and

determining an occurred object-operation-event based on the object-change-information comprises:

determining that the occurred object-operation-event is moving the first object into the first target area.

5. The method of claim 1, wherein

performing object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames comprises:

detecting a second object that has disappeared from the at least two image frames, and

determining an object position where the second object appeared in the at least two image frames before the second object disappears, as a second target area; and

determining an occurred object-operation-event based on the object-change-information comprises:

determining that the occurred object-operation-event is removing the second object out of the second target area.

6. The method of claim 1, wherein

performing object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames comprises:

detecting a change in an object identification result with respect to a third

object involved in the at least two image frames; and

determining an occurred object-operation-event based on the object-change-information comprises:

determining that an object-operation-event corresponding to the change in the object identification result has occurred.

7. The method of claim 6, wherein

the third object comprises a plurality of stackable object components, and each of the object components has a respective component attribute;

the object identification result comprises at least one of: a number of object components, and component attribute of respective object components; and

detecting a change in an object identification result with respect to a third object involved in the at least two image frames comprises:

detecting a change in a number of object components contained in the third object, and

detecting whether the third object has an object component of which the component attribute is same before and after the change;

determining that an object-operation-event corresponding to the change in the object identification result has occurred comprises:

in response to detecting that a change has occurred in the number of object components contained in the third object and the third object has an object component of which the component attribute is same before and after the change, determining that the occurred object-operation-event is increasing or decreasing the number of object components contained in the third object.

8. The method of claim 1, wherein

the object has at least two object states, and an object involved in each of the at least two image frames is in one of the object states;

the object-change-information comprises object-change-information on object state of an object; and

determining an occurred object-operation-event based on the object-change-information comprises:

determining, according to the object-change-information on object state, that the occurred object-operation-event is an operation event of controlling change of object states.

9. The method of claim 8, wherein

the object comprises stackable object components, and

the object-change-information comprises stacking state information of the object components.

10. An apparatus for identifying an operation event, comprising:

a detection processing module configured to perform object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames, wherein the object is an operable object; and

an event determining module configured to determine an occurred object-operation-event based on the object-change-information,

wherein the detection processing module is configured to:

detect a respective object position of an object in each of the at least two image frames of the video;

identify the object detected in each of the at least two image frames to obtain respective object identification results;

based on the respective object positions and the respective object identification results of objects detected in different image frames, compare the objects detected in the different image frames to establish a correspondence between the objects in the at least two image frames; and

obtain the object-change-information of the object involved in the at least two image frames by comparing the objects involved in the at least two image frames,

wherein comparing the objects detected in the different image frames to establish the correspondence between the objects in the at least two image frames comprises:

establishing a position similarity matrix between the objects in the different image frames based on the object positions;

establishing an identification result similarity matrix between the objects in the different image frames based on the object identification results;

obtaining an object similarity matrix based on the position similarity matrix and the identification result similarity matrix; and

determining the correspondence between the objects in the at least two image frames based on the object similarity matrix.

11. The apparatus of claim 10, wherein

when the event determining module is configured to determine an occurred object-operation-event based on the object-change-information, in response to that the object-change-information satisfies a preset event change condition, and an object operator is detected in at least a part of the at least two image frames, and a distance between a

position of the object operator and a position of the object is within a preset distance threshold, determines that an object-operation-event corresponding to the event change condition is occurred by the object operator operating on the object.

12. The apparatus of claim 10, wherein

when the detection processing module is configured to perform object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames, detects a first object newly appeared in the at least two image frames, and determines an object position where the first object appeared in the at least two image frames as a first target area; and

the event determining module is configured to determine that the occurred object-operation-event is moving the first object into the first target area.

13. The apparatus of claim 10, wherein

when the detection processing module is configured to perform object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames, detects a second object that has disappeared from the at least two image frames, and determines an object position where the second object appeared in the at least two image frames before the second object disappears, as a second target area; and

the event determining module is configured to determine that the occurred object-operation-event is removing the second object out of the second target area.

14. The apparatus of claim 10, wherein

when the detection processing module is configured to perform object detection and tracking on at least two image frames of a video to obtain object-change-information of an object involved in the at least two image frames, detects a change in an object identification result with respect to a third object involved in the at least two image frames; and

the event determining module is configured to determine that an object-operation-event corresponding to the change in the object identification result has occurred.

15. The apparatus of claim 14, wherein

when the detection processing module is configured to detect a change in an object identification result with respect to a third object involved in the at least two image frames, detects a change in a number of object components contained in the third object, and detects whether the third object has an object component of which the component attribute is same before and after the change, wherein the third object comprises a plurality of stackable object components, and each of the object components has a respective component attribute; the object identification result comprises at least one of: a number of object components, and component attribute of respective object components;

when the event determining module is configured to determine that an object-operation-event corresponding to the change in the object identification result has occurred, in response to detecting that a change has occurred in the number of object components contained in the third object and the third object has an object component of which the component attribute is same before and after the change, determines that the occurred object-operation-event is increasing or decreasing the number of object components contained in the third object.

16. The apparatus of claim 10, wherein

when the event determining module is configured to determine an occurred object-operation-event based on the object-change-information, determines, according to object-change-information on object state, that the occurred object-operation-event is an operation event of controlling change of object states, wherein the object has at least two object states, and an object involved in each of the at least two image frames is in one of the object states, the object-change-information comprises object-change-information on object state of an object.

17. An electronic device, comprising: a memory and a processor, wherein the memory is configured to store computer readable instructions, and the processor is configured to invoke the computer instructions to implement the method of any one of claims 1 to 9.

18. A computer-readable storage medium having a computer program stored thereon, wherein the program is executed by a processor to implement the method of any one of claims 1 to 9.