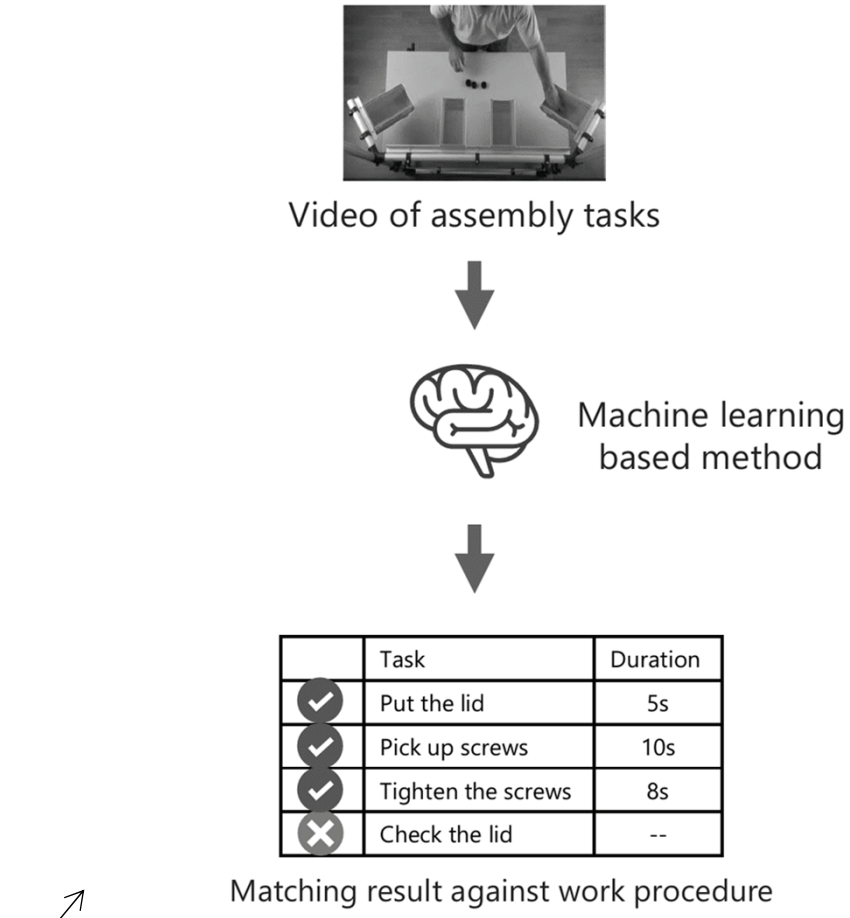

Video of assembly tasks
Machine learning based method
Task
Duration
Put the lid
5s
Pick up screws
10s
Tighten the screws
8s
Check the lid
--
Matching result against work procedure
100

Figure 1

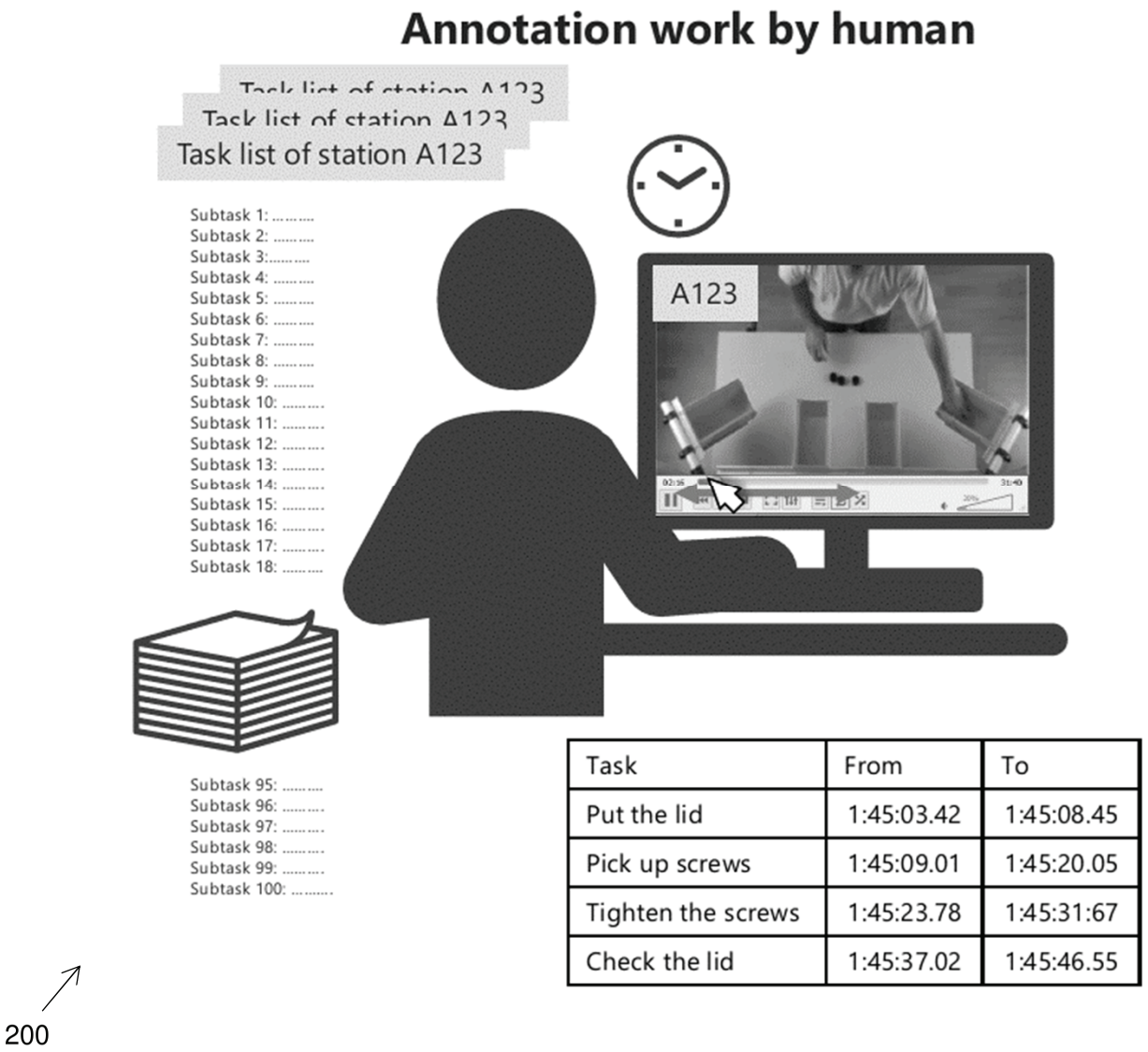

Annotation work by human
Task list of station A123
Task list of station A123
Task list of station A123
Subtask 1: .....
Subtask 2: .....
Subtask 3: .....
Subtask 4: .....
Subtask 5: .....
Subtask 6: .....
Subtask 7: .....
Subtask 8: .....
Subtask 9: .....
Subtask 10: .....
Subtask 11: .....
Subtask 12: .....
Subtask 13: .....
Subtask 14: .....
Subtask 15: .....
Subtask 16: .....
Subtask 17: .....
Subtask 18: .....
Subtask 95: .....
Subtask 96: .....
Subtask 97: .....
Subtask 98: .....
Subtask 99: .....
Subtask 100: .....
A123
Task
From
To
Put the lid
1:45:03.42
1:45:08.45
Pick up screws
1:45:09.01
1:45:20.05
Tighten the screws
1:45:23.78
1:45:31:67
Check the lid
1:45:37.02
1:45:46.55
200

Figure 2

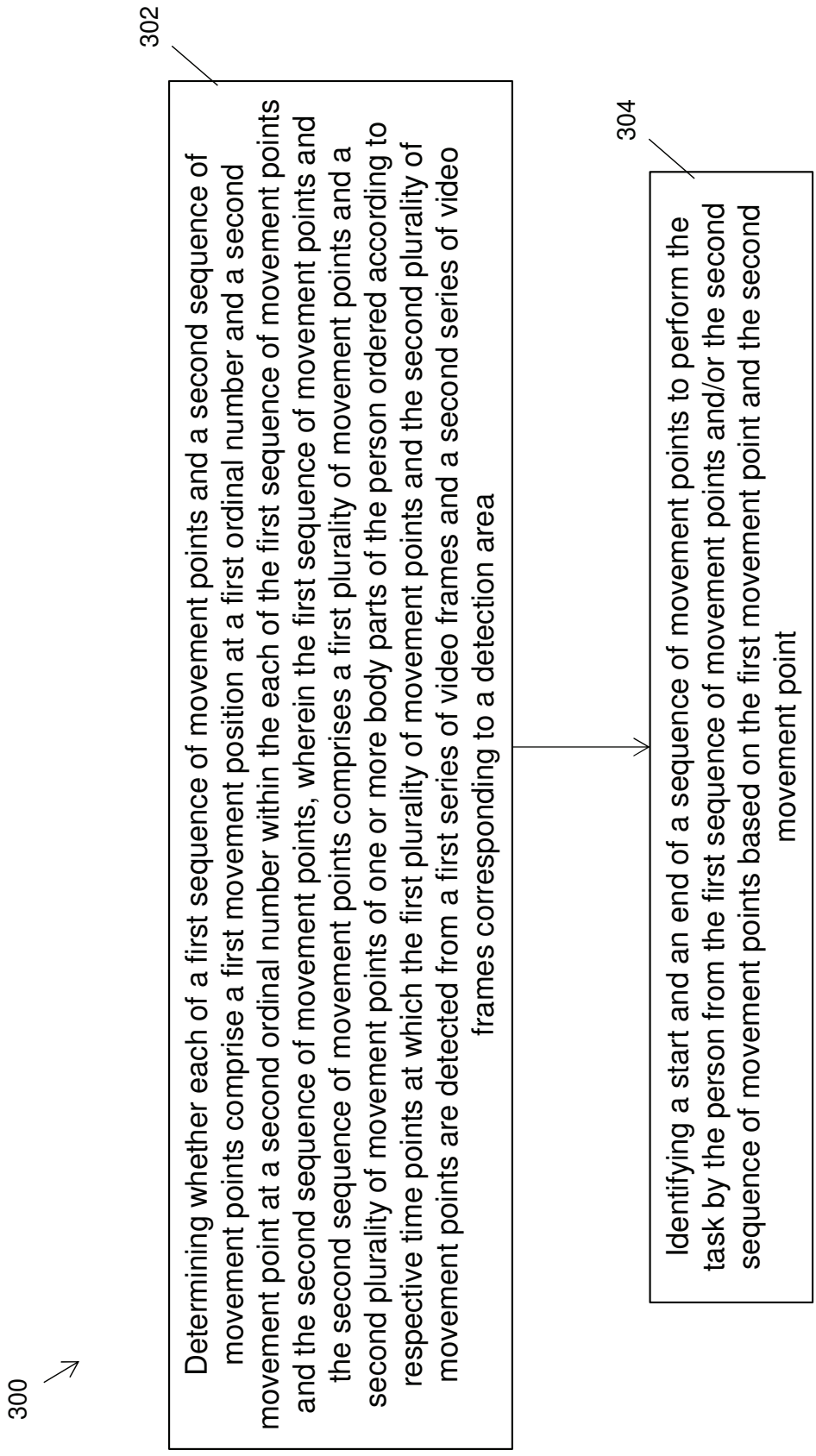

300
302
Determining whether each of a first sequence of movement points and a second sequence of movement points comprise a first movement position at a first ordinal number and a second movement point at a second ordinal number within the each of the first sequence of movement points and the second sequence of movement points, wherein the first sequence of movement points and the second sequence of movement points comprises a first plurality of movement points and a second plurality of movement points of one or more body parts of the person ordered according to respective time points at which the first plurality of movement points and the second plurality of movement points are detected from a first series of video frames and a second series of video frames corresponding to a detection area
304
Identifying a start and an end of a sequence of movement points to perform the task by the person from the first sequence of movement points and/or the second sequence of movement points based on the first movement point and the second movement point

Figure 3

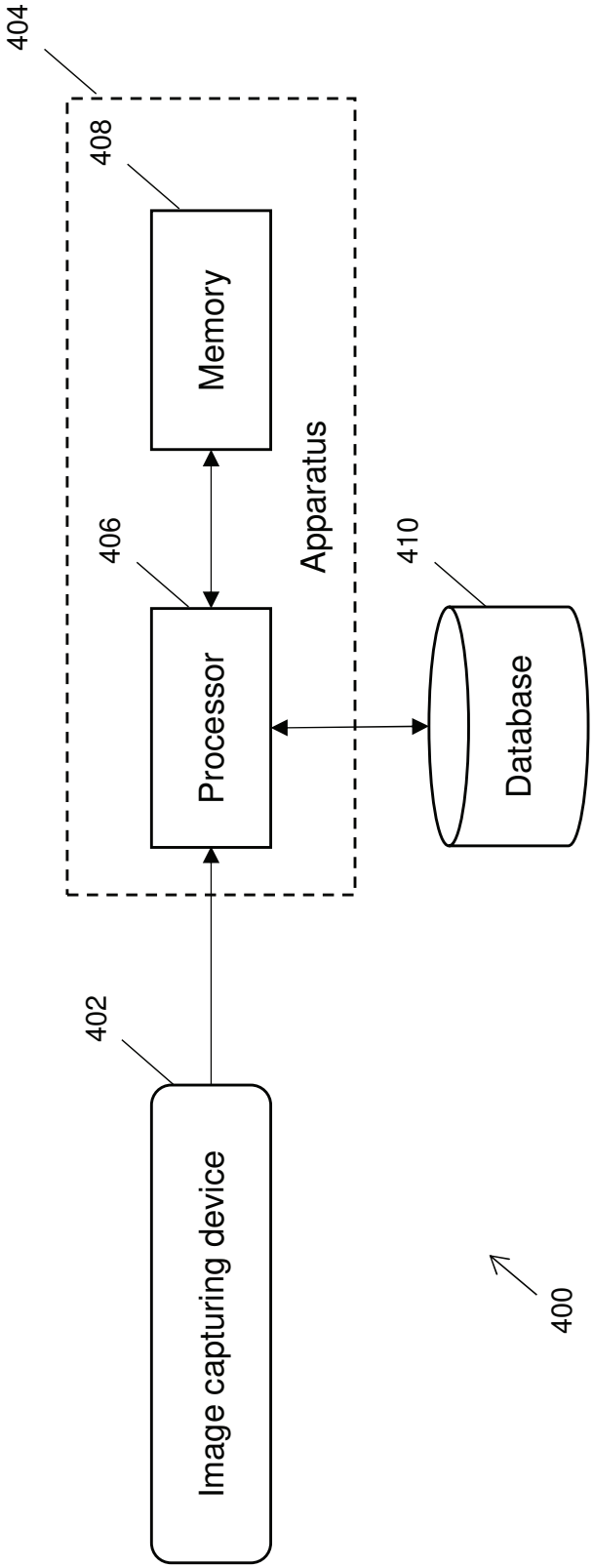

404
408
Memory
Apparatus
406
410
Processor
Database
402
Image capturing device
400

Figure 4

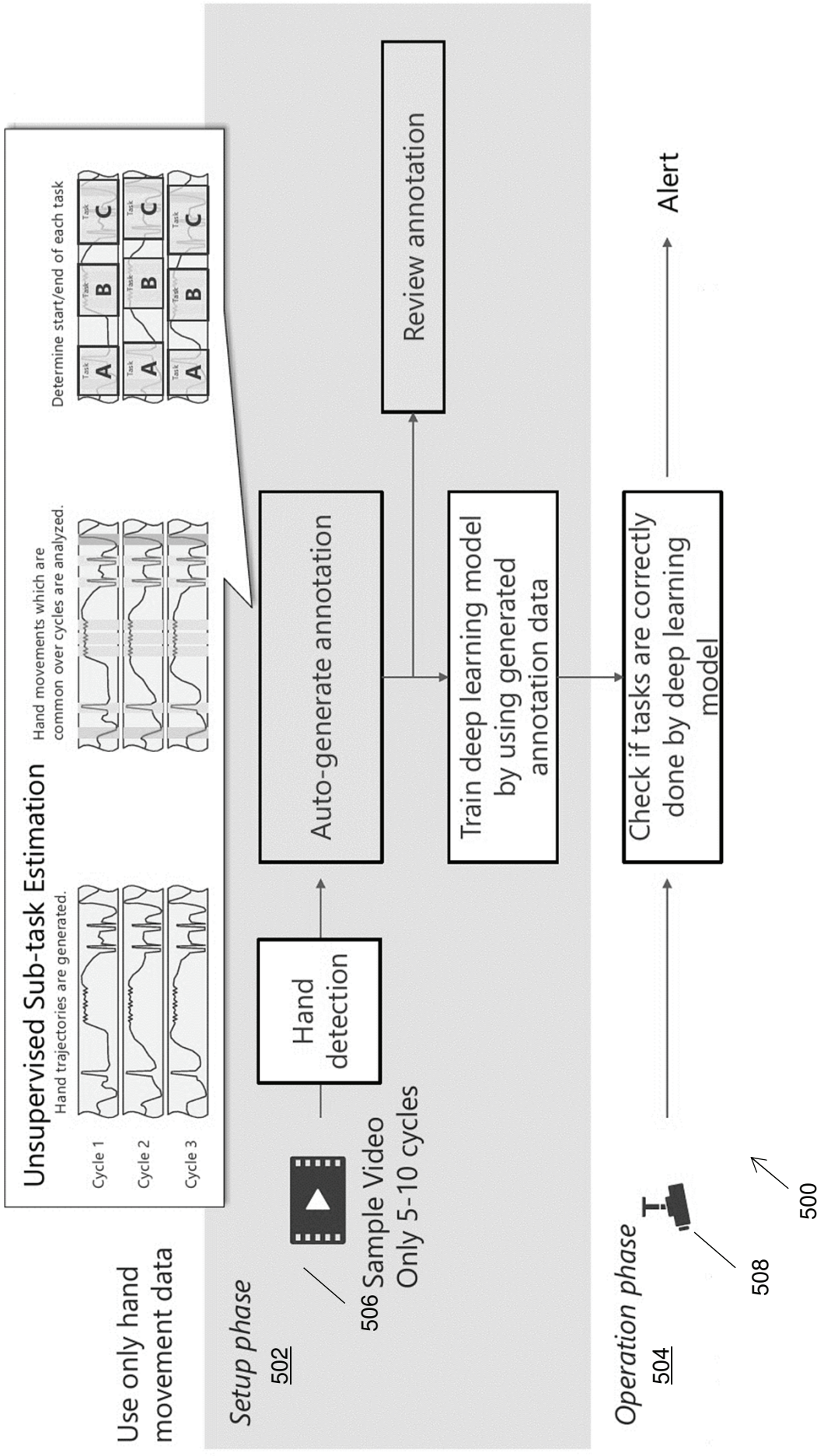
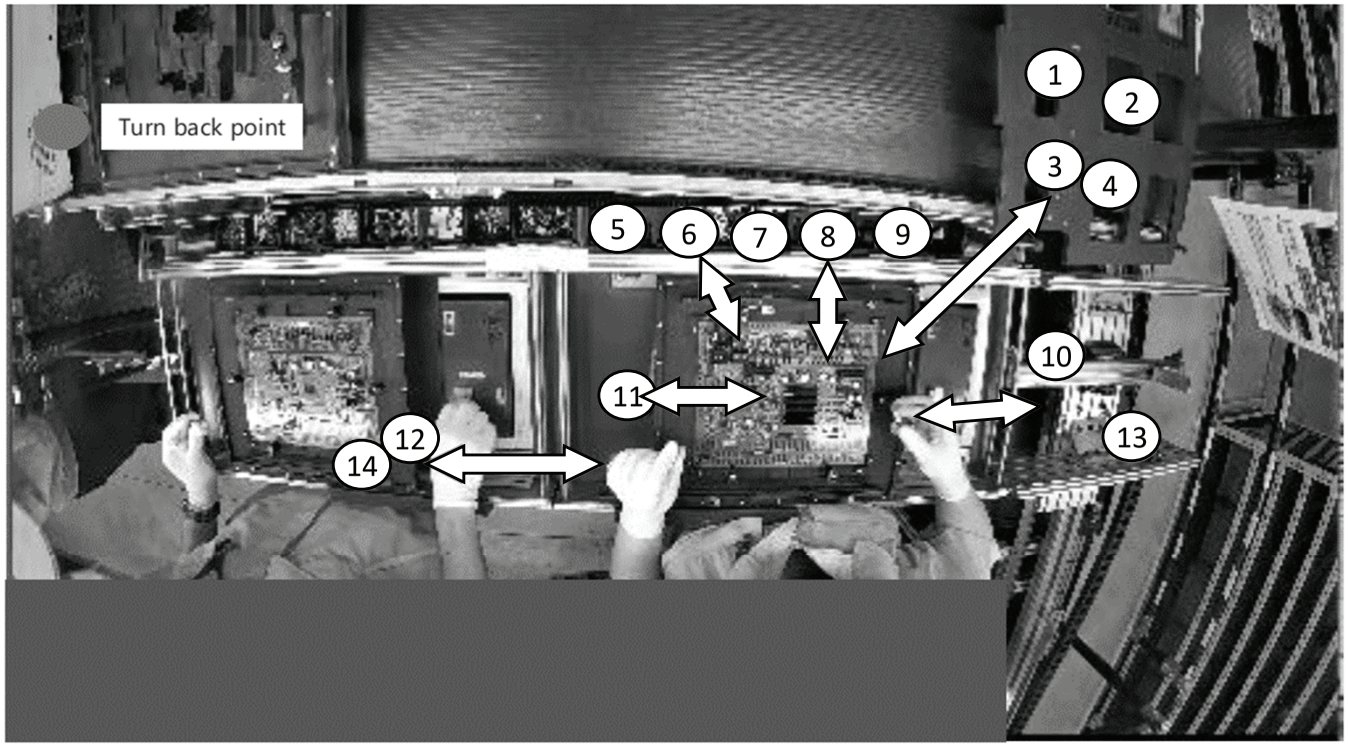
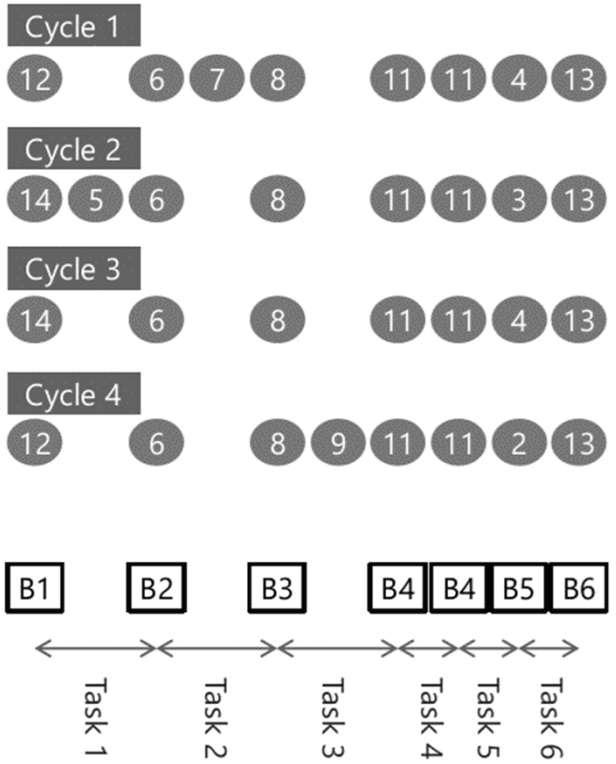

Use only hand movement data
Unsupervised Sub-task Estimation
Hand trajectories are generated.
Hand movements which are common over cycles are analyzed.
Determine start/end of each task
Cycle 1
Cycle 2
Cycle 3
A
B
C
Setup phase
502
Hand detection
Auto-generate annotation
Review annotation
Train deep learning model by using generated annotation data
506 Sample Video
Only 5-10 cycles
Operation phase
504
Check if tasks are correctly done by deep learning model
Alert
508
500

Figure 5


Turn back point
1
2
3
4
5
6
7
8
9
10
11
12
13
14


Cycle 1
12 6 7 8 11 11 4 13
Cycle 2
14 5 6 8 11 11 3 13
Cycle 3
14 6 8 11 11 4 13
Cycle 4
12 6 8 9 11 11 2 13
B1 B2 B3 B4 B4 B5 B6
Task 1
Task 2
Task 3
Task 4
Task 5
Task 6

600

Figure 6

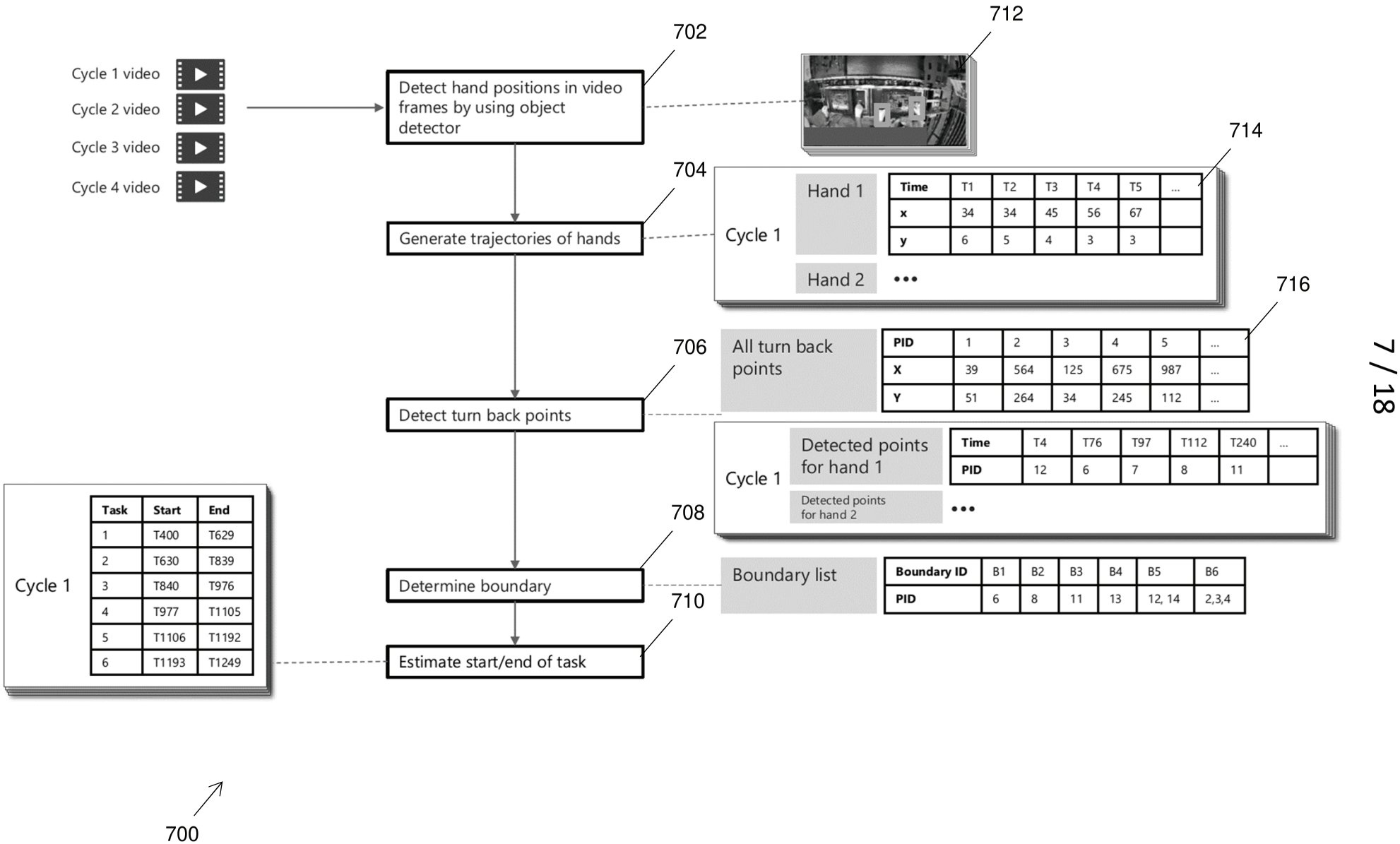

Cycle 1 video
Cycle 2 video
Cycle 3 video
Cycle 4 video
Detect hand positions in video frames by using object detector
702
712
Generate trajectories of hands
704
Cycle 1
Hand 1
Hand 2
714
Time T1 T2 T3 T4 T5 ...
x 34 34 45 56 67
y 6 5 4 3 3
Detect turn back points
706
All turn back points
716
PID 1 2 3 4 5 ...
X 39 564 125 675 987 ...
Y 51 264 34 245 112 ...
Cycle 1
Detected points for hand 1
Detected points for hand 2
Time T4 T76 T97 T112 T240 ...
PID 12 6 7 8 11
Determine boundary
708
Boundary list
Boundary ID B1 B2 B3 B4 B5 B6
PID 6 8 11 13 12, 14 2,3,4
Estimate start/end of task
710
Cycle 1
Task Start End
1 T400 T629
2 T630 T839
3 T840 T976
4 T977 T1105
5 T1106 T1192
6 T1193 T1249
700

Figure 7

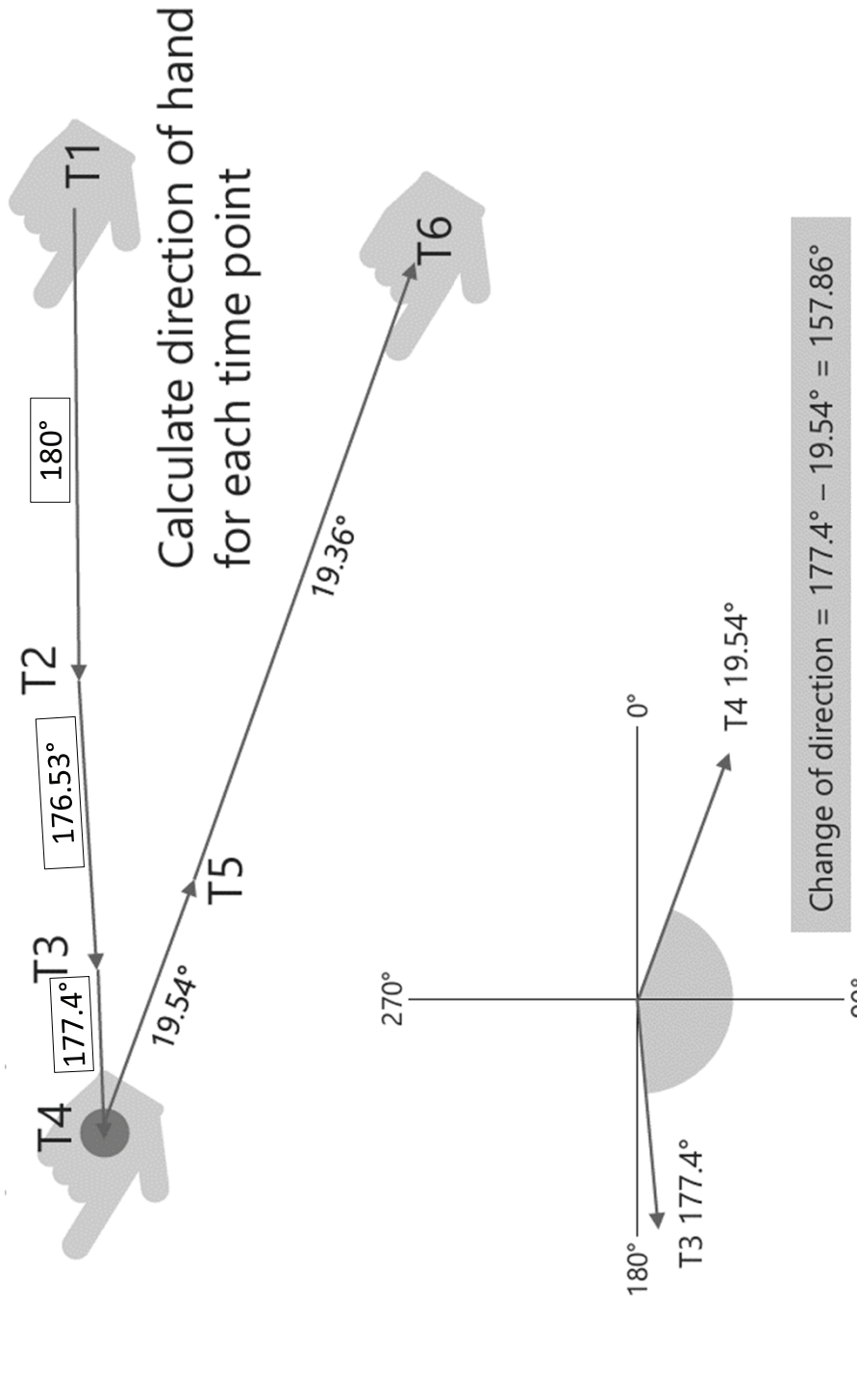

Calculate direction of hand for each time point
T1
T2
T3
T4
T5
T6
180°
176.53°
177.4°
19.54°
19.36°
0°
90°
180°
270°
T4 19.54°
T3 177.4°
Change of direction = 177.4° – 19.54° = 157.86°
800

Figure 8

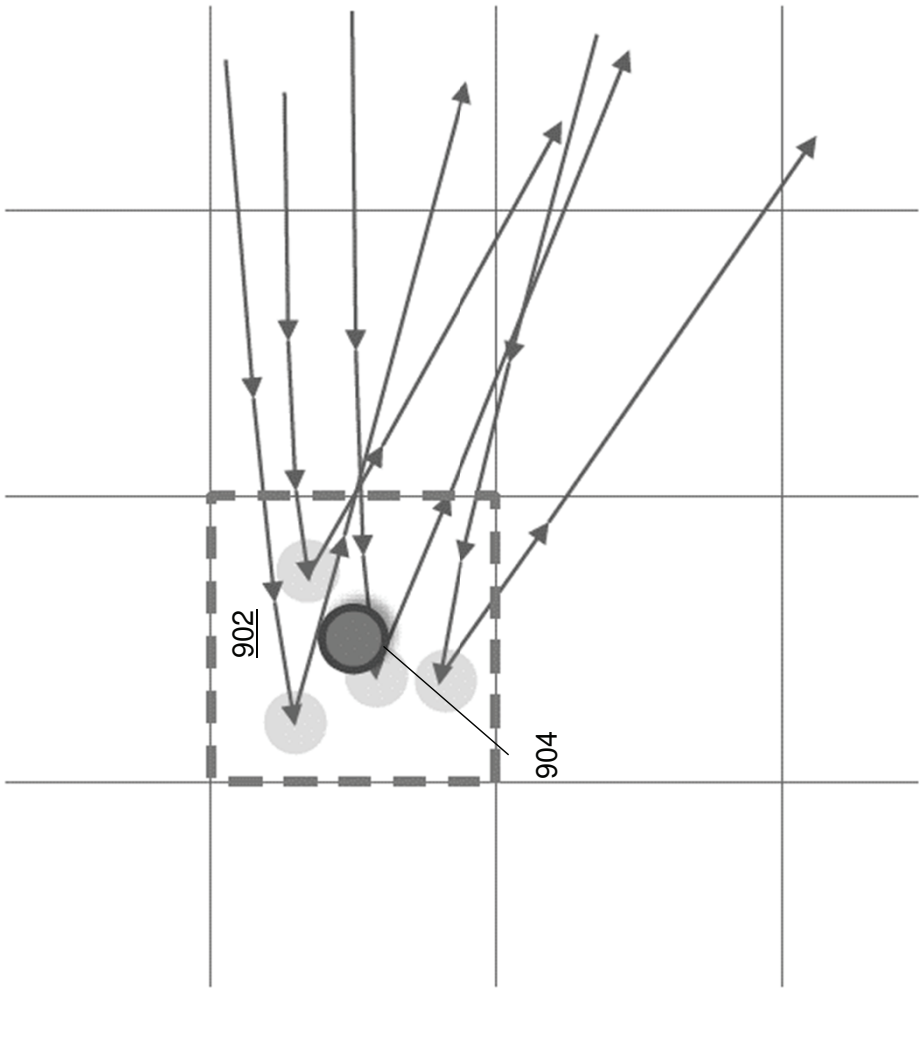


Figure 9

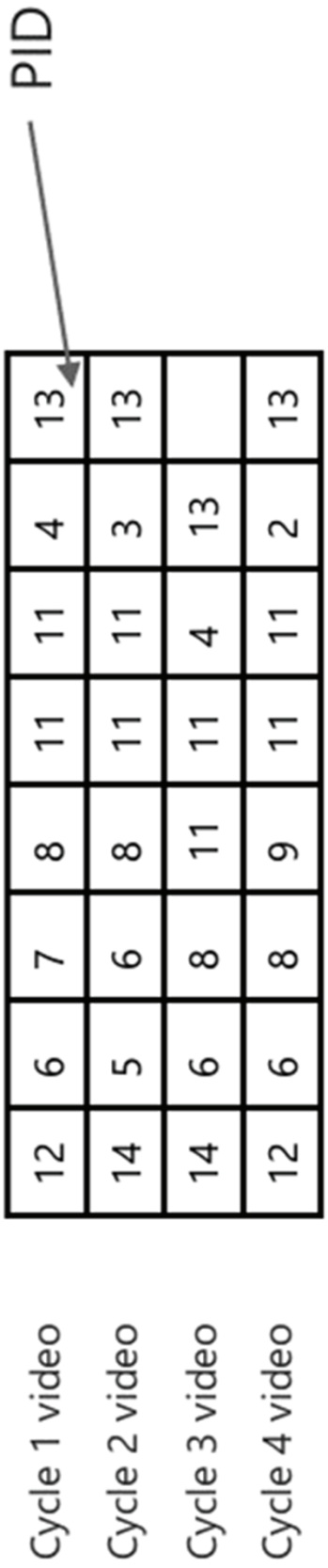
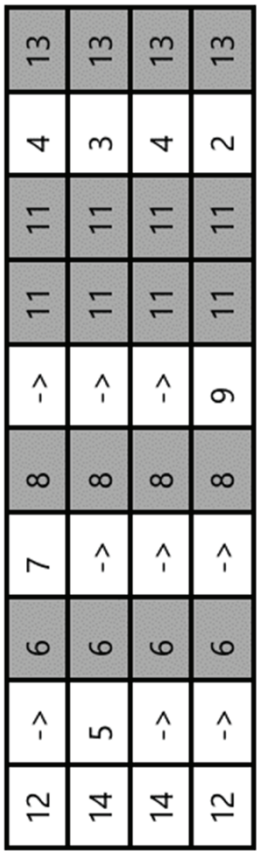
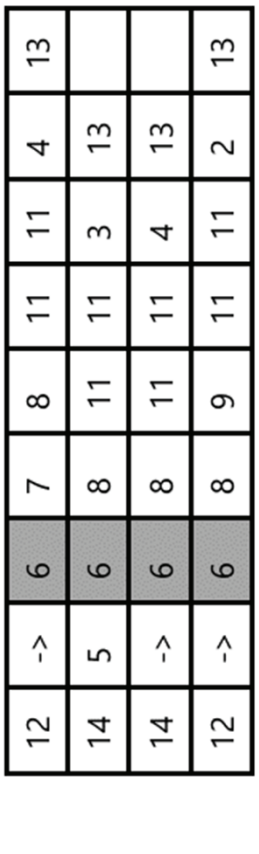

PID
Cycle 1 video
Cycle 2 video
Cycle 3 video
Cycle 4 video
12 6 7 8 11 11 4 13
14 5 6 8 11 3 13
14 6 8 11 11 4 13
12 6 8 9 11 11 2 13
1000

Figure 10

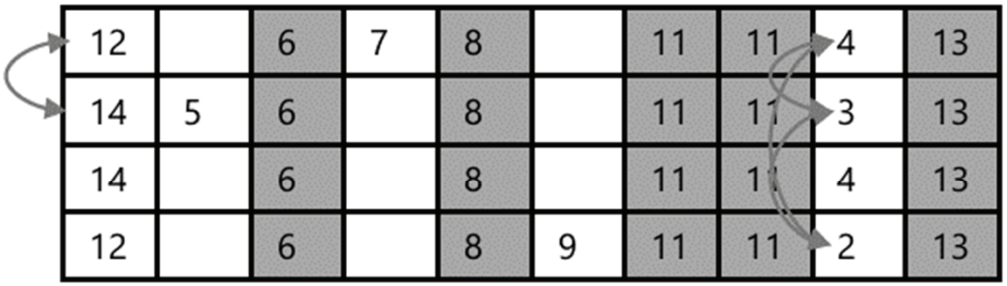


1104

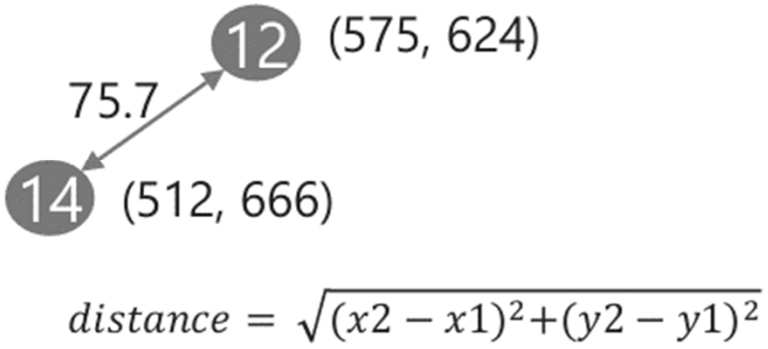


1102

Figure 11



1200



1202

Figure 12

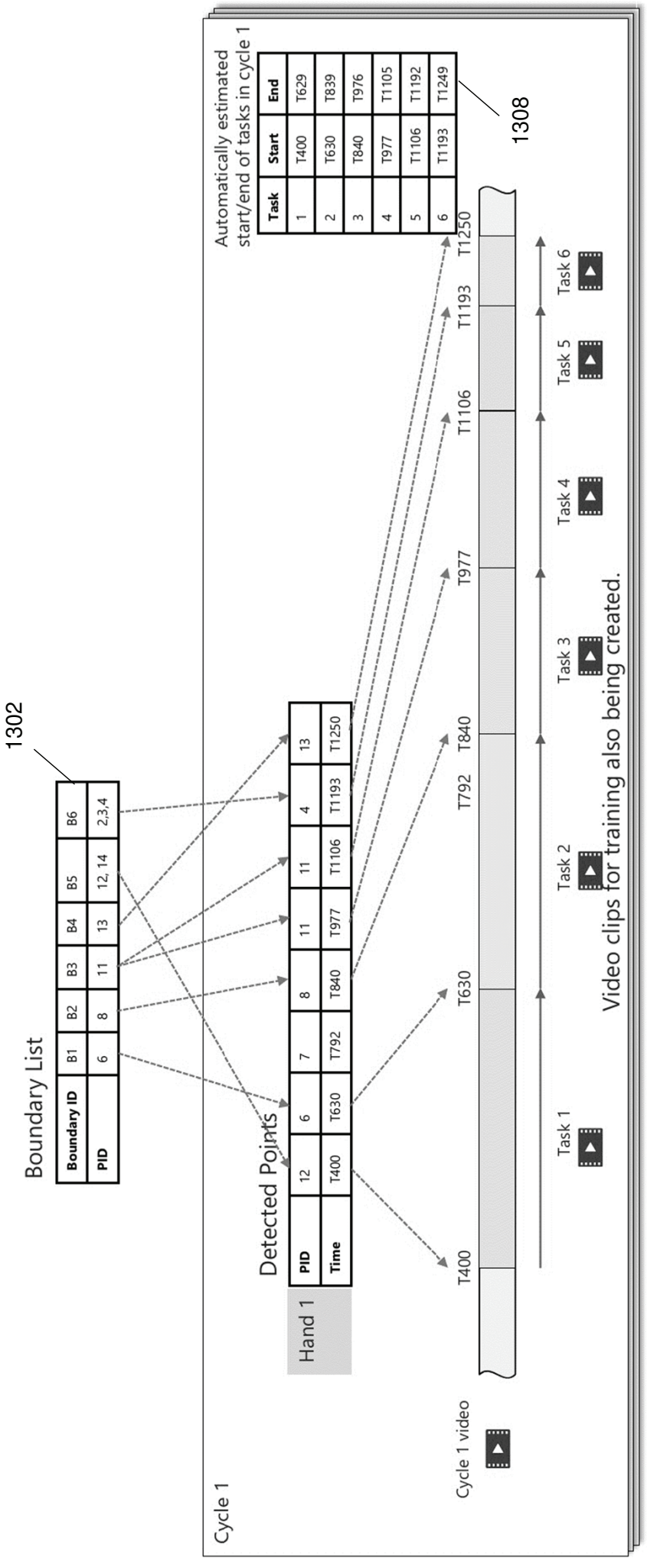

1300
Cycle 1
Cycle 1 video
Boundary List
1302
Boundary ID
PID
B1
B2
B3
B4
B5
B6
6
8
11
13
12, 14
2,3,4
Detected Points
Hand 1
PID
Time
12
6
7
8
11
11
4
13
T400
T630
T792
T840
T977
T1106
T1193
T1250
Automatically estimated start/end of tasks in cycle 1
1308
Task
Start
End
1
T400
T629
2
T630
T839
3
T840
T976
4
T977
T1105
5
T1106
T1192
6
T1193
T1249
Task 1
Task 2
Task 3
Task 4
Task 5
Task 6
Video clips for training also being created.

Figure 13

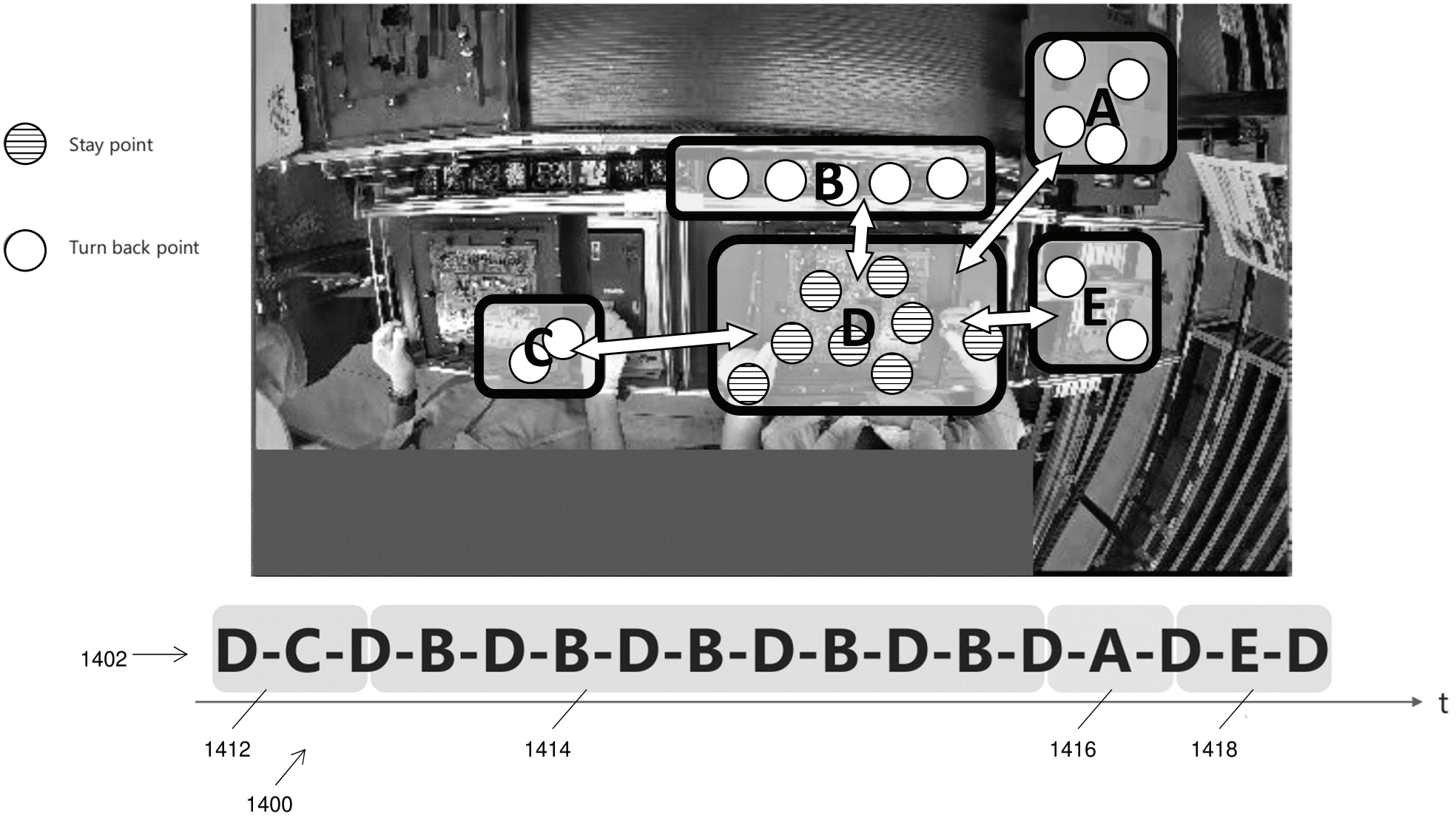

Stay point
Turn back point
A
B
C
D
E
1402
D-C-D-B-D-B-D-B-D-B-D-B-D-A-D-E-D
t
1412
1414
1416
1418
1400

Figure 14

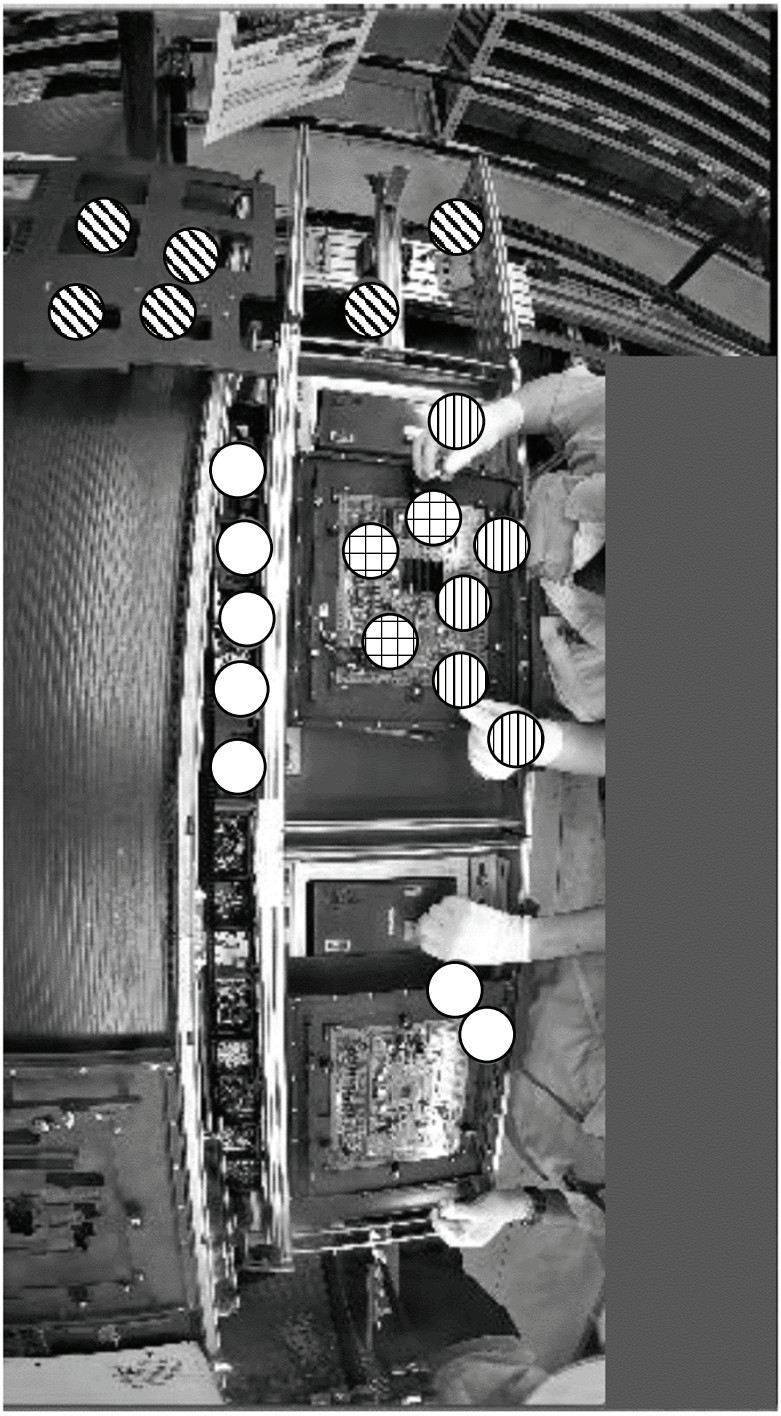

Stay point 1 (without movement)
Turn back point 1 (with short stay period)
Stay point 2 (with movement)
Turn back point 2 (immediate)
1500

Figure 15

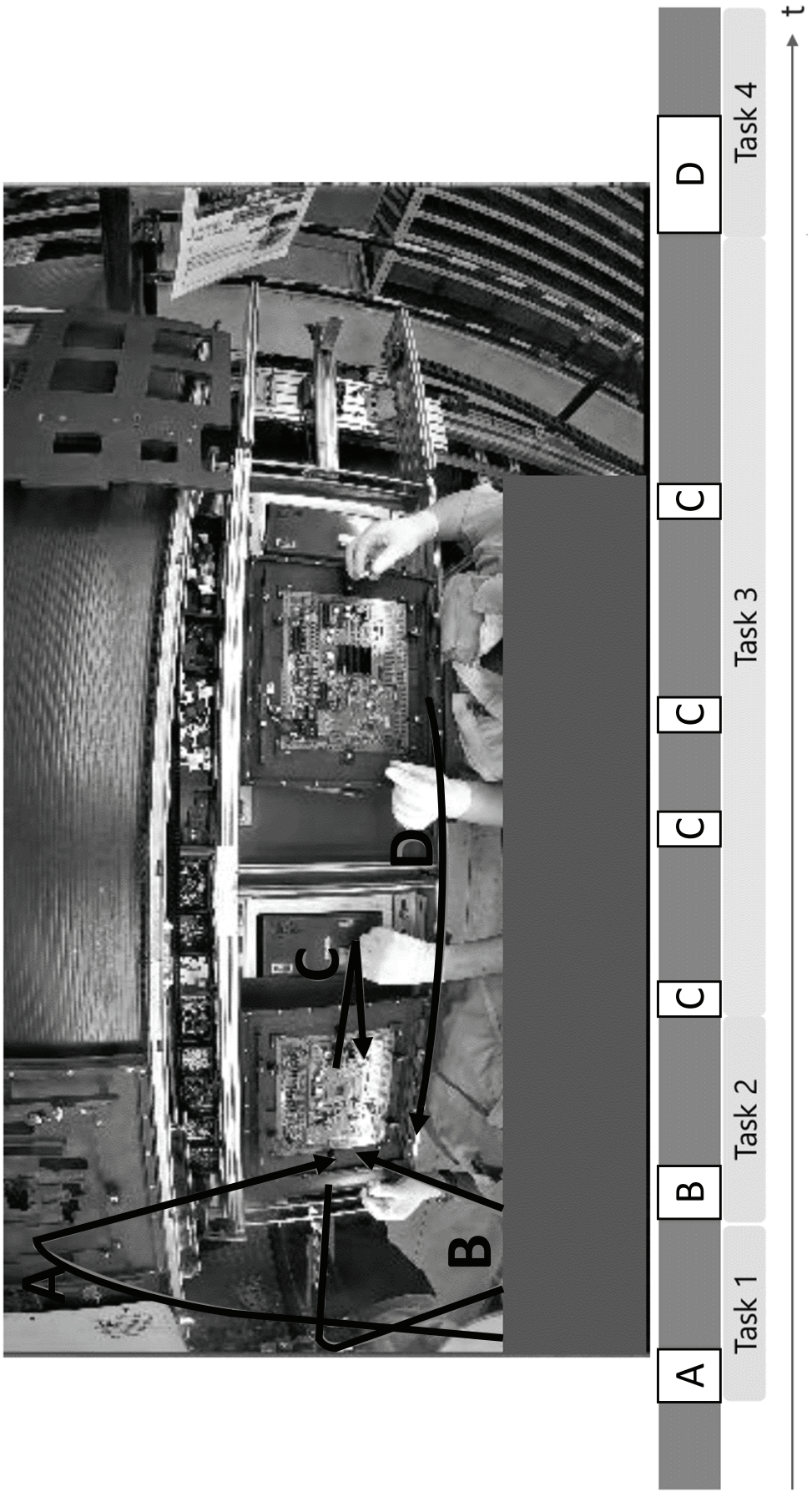


Figure 16

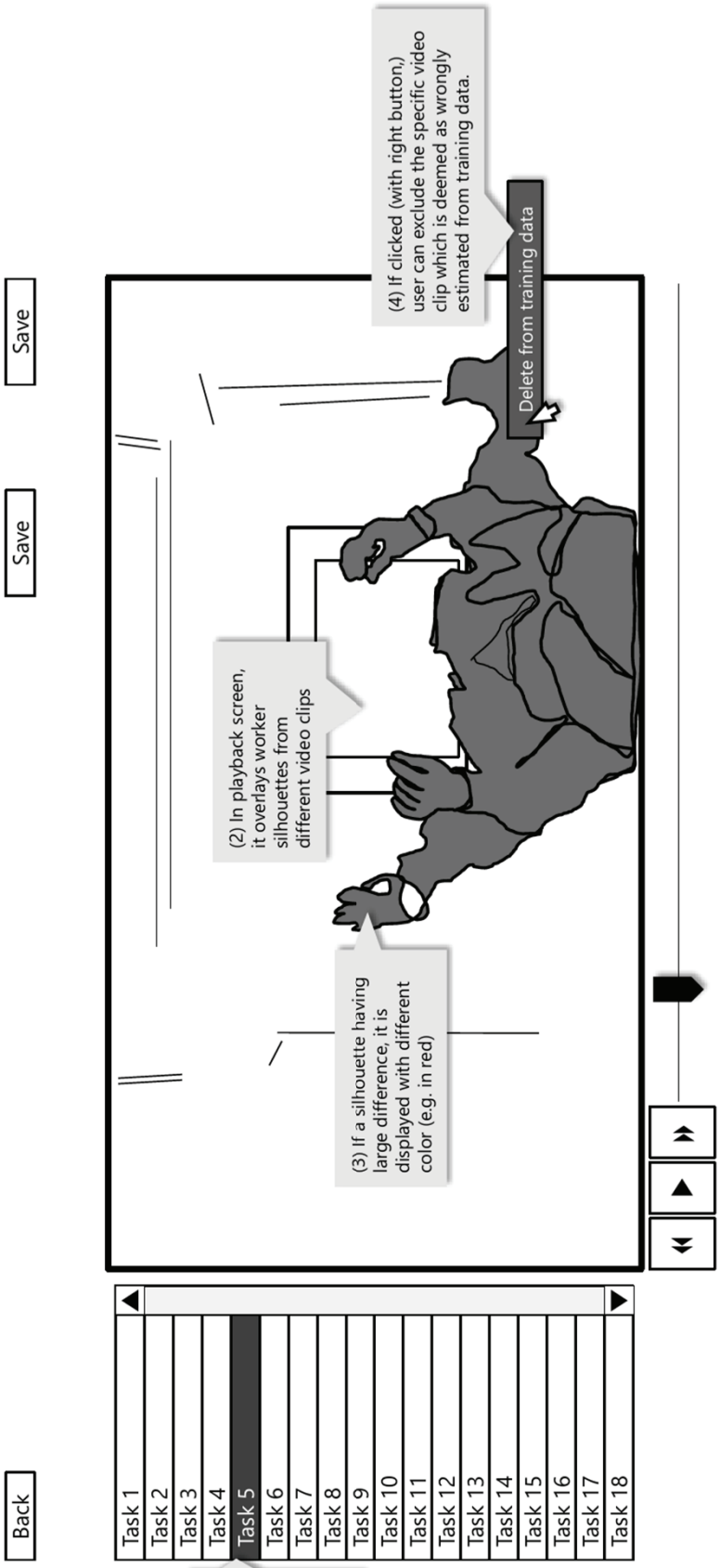

Back
Save
Save
Task 1
Task 2
Task 3
Task 4
Task 5
Task 6
Task 7
Task 8
Task 9
Task 10
Task 11
Task 12
Task 13
Task 14
Task 15
Task 16
Task 17
Task 18
(1) When task is selected, video clip which are specific for the task are shown in the playback screen.
(2) In playback screen, it overlays worker silhouettes from different video clips
(3) If a silhouette having large difference, it is displayed with different color (e.g. in red)
(4) If clicked (with right button,) user can exclude the specific video clip which is deemed as wrongly estimated from training data.
Delete from training data
1700

Figure 17

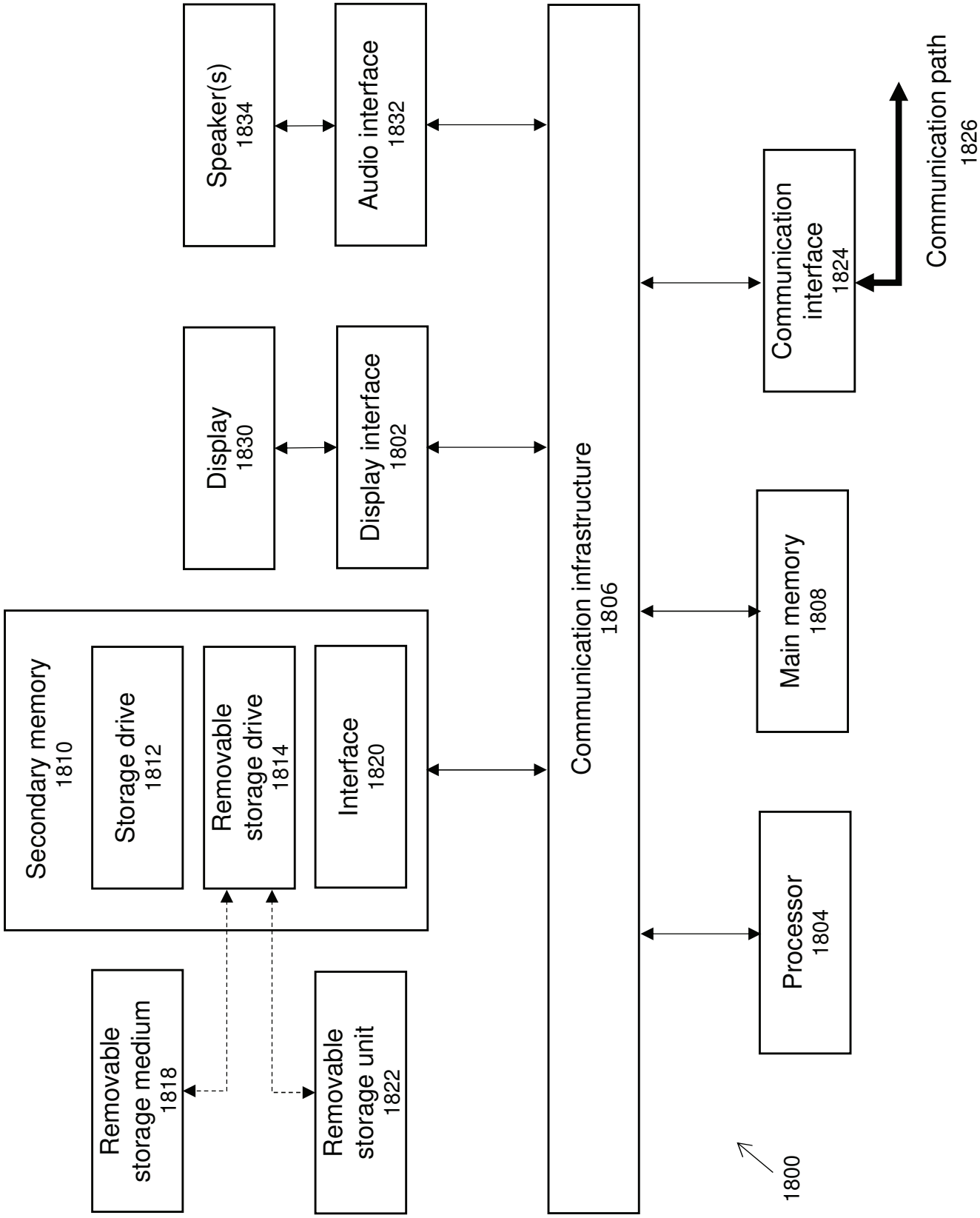

Speaker(s)
1834
Audio interface
1832
Display
1830
Display interface
1802
Secondary memory
1810
Storage drive
1812
Removable storage drive
1814
Interface
1820
Communication infrastructure
1806
Main memory
1808
Communication interface
1824
Communication path
1826
Removable storage medium
1818
Removable storage unit
1822
Processor
1804
1800

Figure 18