

**(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)**

**(19) World Intellectual Property Organization**  
International Bureau



WIPO | PCT

**(43) International Publication Date  
3 July 2014 (03.07.2014)**

**(10) International Publication Number**



6) International Publication Number:

WO 2014/103167 A1

(51) **International Patent Classification:**  
*G06F 3/01* (2006.01)      *G06F 3/0481* (2013.01)

(21) **International Application Number:**  
PCT/JP2013/006979

(22) **International Filing Date:**  
27 November 2013 (27.11.2013)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**  
2012-285025 27 December 2012 (27.12.2012)      JP

(71) **Applicant:** SONY CORPORATION [JP/JP]; 1-7-1 Konan, Minato-ku, Tokyo, 1080075 (JP).

(72) **Inventors:** OHMURA, Junki; c/o SONY CORPORATION, 1-7-1 Konan, Minato-ku, Tokyo, 1080075 (JP). KOHNO, Michinari; c/o SONY CORPORATION, 1-7-1 Konan, Minato-ku, Tokyo, 1080075 (JP). IKEDA, Takuo; c/o SONY CORPORATION, 1-7-1 Konan, Minato-ku, Tokyo, 1080075 (JP). OKADA, Kenichi; c/o SONY CORPORATION, 1-7-1 Konan, Minato-ku, Tokyo, 1080075 (JP).

(74) **Agents:** KAMEYA, Yoshiaki et al.; HAZUKI INTERNATIONAL YOTSUYA, Daiichi Tomizawa Building, 3-1-3, Yotsuya, Shinjuku-ku, Tokyo, 1600004 (JP).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

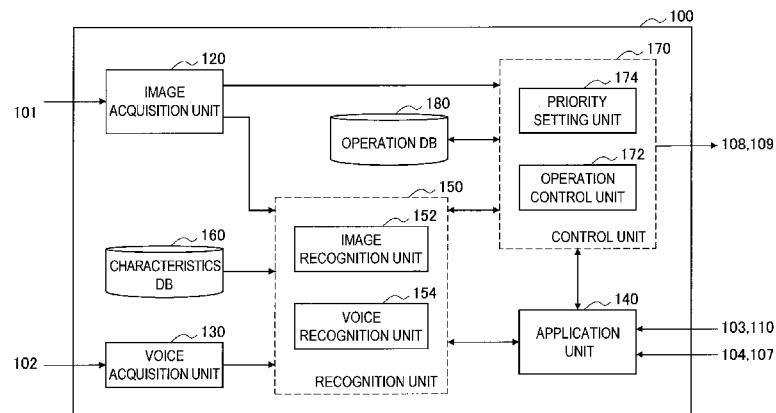
(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**  
— with international search report (Art. 21(3))

**Published:**

— with international search report (Art. 21(3))

**(54) Title: INFORMATION PROCESSING APPARATUS, INFORMATION PROCESSING METHOD, AND PROGRAM**



**(57) Abstract:** An information processing system includes processing circuitry configured to control a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, and the predetermined displayed feature is an image derived from a camera captured image.

## Description

### Title of Invention: INFORMATION PROCESSING APPARATUS, INFORMATION PROCESSING METHOD, AND PROGRAM

#### Technical Field

[0001] The present disclosure relates to an information processing apparatus, an information processing method, and a program.

The present disclosure contains subject matter related to that disclosed in Japanese Priority Patent Application JP 2012-285025 filed in the Japan Patent Office on December 27, 2012, the entire content of which is hereby incorporated by reference.

#### Background Art

[0002] In recent years, a user interface (UI) that receives an user input via an image from a video camera has been proposed.

[0003] As one example, JP 2002-196855A discloses a method of superimposing object images for UI purposes on an image in which a mirror image of a user appears and carries out an application process associated with an object image selected by movement of the user's hand. JP 2005-216061A meanwhile discloses a method which eliminates the trouble of making initial settings, such as setting a camera angle, in a UI which uses camera images by determining the position of the user's head and hands in an input image and automatically displaying an object image in the vicinity of the determined position.

#### Citation List

#### Patent Literature

[0004] PTL 1: JP 2002-196855A  
PTL 2: JP 2005-216061A

#### Summary

#### Technical Problem

[0005] However, the screen region in the vicinity of the head or hands of the user is limited. This means that with the method disclosed in JP 2005-216061A, when a large number of selectable objects are provided in a UI, the screen becomes crowded with such objects, which can conversely cause a drop in usability.

[0006] Accordingly, it is desirable to realize an improved UI capable of avoiding a drop in usability due to crowding of the screen, even when a large number of selectable objects are provided.

#### Solution to Problem

[0007] According to one aspect, an information processing system is described that includes

processing circuitry configured to control a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, wherein the predetermined displayed feature is an image derived from a camera-captured image.

[0008] According to another aspect, an information processing method is described that includes

controlling with processing circuitry a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, and the predetermined displayed feature is an image derived from a camera-captured image.

[0009] According to another aspect, a non-transitory computer readable medium is described that includes computer readable instructions that when executed by a processing circuitry perform a method, the method including

controlling with the processing circuitry a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, and the predetermined displayed feature is an image derived from a camera-captured image.

### **Advantageous Effects of Invention**

[0010] According to the above embodiments of the present disclosure, a UI capable of avoiding a drop in usability due to crowding of the screen, even when a large number of selectable objects are provided, is realized.

### **Brief Description of Drawings**

[0011] [fig.1]Fig. 1 is a diagram useful in explaining an overview of an information processing apparatus according to a first embodiment of the present disclosure. [fig.2]Fig. 2 is a diagram useful in explaining an overview of an information processing apparatus according to a second embodiment of the present disclosure. [fig.3]Fig. 3 is a block diagram showing an example hardware configuration of the information processing apparatus according to the first embodiment. [fig.4]Fig. 4 is a block diagram showing an example configuration of logical functions of the information processing apparatus according to the first embodiment. [fig.5]Fig. 5 is a diagram useful in explaining one example of the result of image

recognition.

[fig.6]Fig. 6 is a diagram useful in explaining a first example of another gesture that may be recognized.

[fig.7]Fig. 7 is a diagram useful in explaining a second example of another gesture that may be recognized.

[fig.8]Fig. 8 is a diagram useful in explaining a third example of another gesture that may be recognized.

[fig.9]Fig. 9 is a diagram useful in explaining a fourth example of another gesture that may be recognized.

[fig.10]Fig. 10 is a diagram useful in explaining a fifth example of another gesture that may be recognized.

[fig.11]Fig. 11 is a diagram useful in explaining a sixth example of another gesture that may be recognized.

[fig.12]Fig. 12 is a diagram useful in explaining a seventh example of another gesture that may be recognized.

[fig.13]Fig. 13 is a diagram useful in explaining a first example of UI objects.

[fig.14]Fig. 14 is a diagram useful in explaining a second example of UI objects.

[fig.15]Fig. 15 is a diagram useful in explaining a first example of a mode of approach toward the user of UI objects.

[fig.16]Fig. 16 is a diagram useful in explaining a second example of a mode of approach toward the user of UI objects.

[fig.17]Fig. 17 is a diagram useful in explaining a third example of a mode of approach toward the user of UI objects.

[fig.18]Fig. 18 is a diagram useful in explaining a fourth example of a mode of approach toward the user of UI objects.

[fig.19]Fig. 19 is a diagram useful in explaining an example of priority data corresponding to the example shown in Fig. 17.

[fig.20]Fig. 20 is a diagram useful in explaining an example of priority data corresponding to the example shown in Fig. 18.

[fig.21]Fig. 21 is a diagram useful in explaining a fifth example of a mode of approach toward the user of UI objects.

[fig.22]Fig. 22 is a diagram useful in explaining a first example of an operation event.

[fig.23]Fig. 23 is a diagram useful in explaining a second example of an operation event.

[fig.24]Fig. 24 is a diagram useful in explaining a third and fourth examples of an operation event.

[fig.25]Fig. 25 is a diagram useful in explaining a fifth example of an operation event.

[fig.26]Fig. 26 is a diagram useful in explaining a sixth example of an operation event.

[fig.27]Fig. 27 is a diagram useful in explaining a seventh and eighth examples of an operation event.

[fig.28]Fig. 28 is a diagram useful in explaining a ninth example of an operation event.

[fig.29]Fig. 29 is a diagram useful in explaining a first example of an operation scenario involving a plurality of operation objects.

[fig.30]Fig. 30 is a diagram useful in explaining a second example of an operation scenario involving a plurality of operation objects.

[fig.31]Fig. 31 is a diagram useful in explaining a third example of an operation scenario involving a plurality of operation objects.

[fig.32]Fig. 32 is a diagram useful in explaining a first example of the window composition of an output image.

[fig.33]Fig. 33 is a diagram useful in explaining a second example of the window composition of an output image.

[fig.34]Fig. 34 is the former half of a flowchart showing an example of the flow of processing according to the first embodiment.

[fig.35]Fig. 35 is the latter half of the flowchart showing an example of the flow of processing according to the first embodiment.

[fig.36]Fig. 36 is a block diagram showing an example hardware configuration of the information processing apparatus according to the second embodiment.

[fig.37]Fig. 37 is a diagram useful in explaining an example of a operation scenario in the second embodiment.

## Description of Embodiments

[0012] Hereinafter, preferred embodiments of the present disclosure will be described in detail with reference to the appended drawings. Note that, in this specification and the appended drawings, structural elements that have substantially the same function and structure are denoted with the same reference numerals, and repeated explanation of these structural elements is omitted.

[0013] The following description is given in the order indicated below.

1. Overview
2. First Embodiment
  - 2-1. Example Hardware Configuration
  - 2-2. Example Functional Configuration
  - 2-3. Mode Of Approach of UI Objects
  - 2-4. Various Examples of Operation Events
  - 2-5. Incorporation of a Plurality of Operation Objects
  - 2-6. Example Window Compositions
  - 2-7. Example Processing Flow

## 3. Second Embodiment

## 4. Conclusion

[0014] &lt;1. Overview&gt;

First, an overview of an information processing apparatus to which the technology according to an embodiment of the present disclosure can be applied will be described with reference to Figs. 1 and 2. The technology according to an embodiment of the present disclosure can be applied to a variety of apparatuses and systems that use an image in which a user appears as part of a user interface. As examples, the technology according to an embodiment of the present disclosure can be applied to a digital home appliance such as a television apparatus. The technology according to an embodiment of the present disclosure can also be applied to a terminal apparatus such as a PC (Personal Computer), a smartphone, a PDA (Personal Digital Assistant), or a game console. The technology according to an embodiment of the present disclosure can also be applied to a special-purpose apparatus such as an amusement apparatus.

[0015] Fig. 1 is a diagram useful in explaining an overview of an information processing apparatus 100 according to a first embodiment of the present disclosure. As shown in Fig. 1, the information processing apparatus 100 is a television apparatus. The information processing apparatus 100 includes a camera 101, a microphone 102, and a display 108. The camera 101 picks up images of users who are looking at the display 108 of the information processing apparatus 100. The microphone 102 picks up voice samples produced by such users. The display 108 displays images generated by the information processing apparatus 100. The images displayed by the display 108 may include user interface (UI) images in addition to content images. In the example in Fig. 1, users Ua and Ub are looking at the display 108. On the display 108, a UI image W01 is displayed. The UI image W01 is generated using a picked-up image picked up by the camera 101 and realizes a so-called "mirror image" display. A plurality of UI objects Obj are superimposed on the UI image W01. The users Ua and Ub may interact with the information processing apparatus 100 by operating the UI objects Obj via various gestures using their bodies. A voice command inputted via the microphone 102 into the information processing apparatus 100 may also be used to supplement interaction with the information processing apparatus 100.

[0016] Fig. 2 is a diagram useful in explaining an overview of an information processing apparatus 200 according to a second embodiment of the present disclosure. As shown in Fig. 2, the information processing apparatus 200 is a tablet PC. The information processing apparatus 200 includes a camera 201, a microphone 202, and a display 208. The camera 201 picks up images of users who are looking at the display 208 of the information processing apparatus 200. The microphone 202 picks up voice samples produced by such users. The display 208 displays images generated by the information

processing apparatus 200. The images displayed by the display 208 may include user interface (UI) images in addition to content images. In the example in Fig. 2, a user Uc is looking at the display 208. On the display 208, a UI image W02 is displayed. The UI image W02 is generated using a picked-up image picked up by the camera 201 and realizes a so-called "mirror image" display. A plurality of UI objects Obj are also superimposed on the UI image W02. The user Uc may interact with the information processing apparatus 200 by operating the UI objects Obj via various gestures that use the body.

[0017] According to existing methods, the UI objects operated by the user may be automatically laid out in the vicinity of the head or hand of the user in the image. However, since the screen region in the vicinity of the head or hand of the user is limited, when a plurality of UI objects are provided, there is the risk of such UI objects being congested in the vicinity of the user. If the UI objects are congested in a limited screen region, it becomes difficult to select the individual UI objects, which can conversely cause a drop in usability. For this reason, the information processing apparatuses 100 and 200 avoid such drop in usability in accordance with the framework described in detail in the following sections.

[0018] <2. First Embodiment>

<2-1. Example Hardware Configuration>

Fig. 3 is a block diagram showing an example hardware configuration of the information processing apparatus 100. As shown in Fig. 3, the information processing apparatus 100 includes the camera 101, the microphone 102, an input device 103, a communication interface (I/F) 104, a memory 105, a tuner 106, a decoder 107, a display 108, a speaker 109, a remote control I/F 110, a bus 111, and a processor 112.

[0019] (1) Camera

The camera 101 includes an image pickup element such as a CCD (Charge Coupled Device) or a CMOS (Complementary Metal Oxide Semiconductor) and picks up images. The images picked up by the camera 101 (frames that construct video) are treated as input images for processing by the information processing apparatus 100.

[0020] (2) Microphone

The microphone 102 picks up a voice sample produced by a user and generates a voice signal. The voice signal generated by the microphone 102 is able to be treated as an input voice intended for voice recognition by the information processing apparatus 100. The microphone 102 may be an omnidirectional microphone or a microphone with fixed or variable directionality.

[0021] (3) Input Device

The input device 103 is a device used by the user to directly operate the information processing apparatus 100. As examples, the input device 103 may include buttons,

switches, dials, and the like disposed on the housing of the information processing apparatus 100. On detecting a user input, the input device 103 generates an input signal corresponding to the detected user input.

[0022] (4) Communication Interface

The communication I/F 104 acts as an intermediary for communication between the information processing apparatus 100 and another apparatus. The communication I/F 104 supports an arbitrary wireless communication protocol or wired communication protocol and establishes a communication connection with the other apparatus.

[0023] (5) Memory

The memory 105 is constructed of a storage medium such as a semiconductor memory or a hard disk drive and stores programs and data for processing by the information processing apparatus 100, as well as content data. As one example, the data stored by the memory 105 may include characteristic data used for image recognition and voice recognition, described later. Note that some or all of the programs and data described in the present specification may not be stored by the memory 105 and instead may be acquired from an external data source (as examples, a data server, network storage, or an externally-attached memory).

[0024] (6) Tuner

The tuner 106 extracts and demodulates a content signal on a desired channel from a broadcast signal received via an antenna (not shown). The tuner 106 then outputs the demodulated content signal to the decoder 107.

[0025] (7) Decoder

The decoder 107 decodes content data from the content signal inputted from the tuner 106. The decoder 107 may decode content data from a content signal received via the communication I/F 104. Content images may be generated based on the content data decoded by the decoder 107.

[0026] (8) Display

The display 108 has a screen constructed of an LCD (Liquid Crystal Display), an OLED (Organic Light-Emitting Diode), a CRT (Cathode Ray Tube), or the like and displays images generated by the information processing apparatus 100. As examples, content images and UI images that were described with reference to Figs. 1 and 2 may be displayed on the screen of the display 108.

[0027] (9) Speaker

The speaker 109 has a diaphragm and circuit elements such as an amplifier and outputs audio based on an output voice signal generated by the information processing apparatus 100. The volume of the speaker 109 is variable.

[0028] (10) Remote Control Interface

The remote control I/F 110 is an interface that receives a remote control signal (an

infrared signal or other wireless signal) transmitted from a remote controller used by the user. On detecting a remote control signal, the remote control I/F 110 generates an input signal corresponding to the detected remote control signal.

[0029] (11) Bus

The bus 111 connects the camera 101, the microphone 102, the input device 103, the communication I/F 104, the memory 105, the tuner 106, the decoder 107, the display 108, the speaker 109, the remote control I/F 110, and the processor 112 to each other.

[0030] (12) Processor

As examples, the processor 112 may be a CPU (Central Processing Unit) or a DSP (Digital Signal Processor). By executing a program stored in the memory 105 or on another storage medium, the processor 112 causes the information processing apparatus 100 to function in various ways as described later.

[0031] <2-2. Example Functional Configuration>

Fig. 4 is a block diagram showing an example configuration of logical functions realized by the memory 105 and the processor 112 of the information processing apparatus 100 shown in Fig. 3. As shown in Fig. 4, the information processing apparatus 100 includes an image acquisition unit 120, a voice acquisition unit 130, an application unit 140, a recognition unit 150, a characteristics database (DB) 160, a control unit 170, and an operation DB 180. The recognition unit 150 includes an image recognition unit 152 and a voice recognition unit 154. The control unit 170 includes an operation control unit 172 and a priority setting unit 174.

[0032] (1) Image Acquisition Unit

The image acquisition unit 120 acquires an image picked up by the camera 101 as an input image. The input image is typically an individual frame in a series of frames that construct video in which users appear. The image acquisition unit 120 then outputs the acquired input image to the recognition unit 150 and the control unit 170.

[0033] (2) Voice Acquisition Unit

The voice acquisition unit 130 acquires the voice signal generated by the microphone 102 as an input voice. The voice acquisition unit 130 then outputs the acquired input voice to the recognition unit 150. Note that processing of an input voice may be omitted from the present embodiment.

[0034] (3) Application Unit

The application unit 140 carries out various application functions of the information processing apparatus 100. As examples, a television program reproduction function, an electronic program guide display function, a recording setting function, a content reproduction function, a content searching function, and an Internet browsing function may be carried out by the application unit 140. The application unit 140 outputs application images (which may include content images) and audio which have been

generated via the application function to the control unit 170.

[0035] In the present embodiment, at least some of the processes carried out by the application unit 140 are associated with UI objects laid out on a UI image. Such processes may be carried out in response to operation events that involve the associated UI objects. The processes that may be carried out via UI objects may include arbitrary processes, such as setting a channel and volume for a television program reproduction function, setting a channel and time period for an electronic program guide display function, selecting content for a content reproduction function, and designating a search keyword and carrying out a search for a content search function.

[0036] (4) Image Recognition Unit

The image recognition unit 152 recognizes an operation object used by the user in an input image inputted from the image acquisition unit 120. In the present embodiment, the operation object is the user's hand. A user's hand that makes a specified shape (such as a shape where the hand is open, a gripping shape, or a shape of pointing with a finger) may be used as the operation object. In other embodiments, instead of the user's hand, the user's foot or a known actual object held by the user may be used as the operation object. As one example, the image recognition unit 152 may recognize the hand region in the input image by matching image characteristic values extracted from the input image and image characteristic values of an operation object stored in advance by the characteristics DB 160. In the same way, the image recognition unit 152 may recognize a face region in the input image.

[0037] Fig. 5 is a diagram useful in explaining one example of the result of image recognition by the image recognition unit 152. As shown in Fig. 5, the user Ua appears in the input image W03. The user Ua is facing the camera 101 and raising his left hand. By matching image characteristic values or using another known method, the image recognition unit 152 is capable of recognizing a hand region A01 and a face region A02 in the input image W03. The image recognition unit 152 then outputs position data showing the positions in the image of such recognized regions to the control unit 170.

[0038] As one example, the image recognition unit 152 may identify the user by matching an image part (facial image) of the face region recognized in an input image against facial image data of known users stored in advance by the characteristics DB 160. As examples, the user identification result produced by the image recognition unit 152 can be used to personalize menus displayed in a UI image, or by the application unit 140 to recommend content, and to make adjustments to the voice recognition.

[0039] In the present embodiment, the image recognition unit 152 also recognizes gestures of the users appearing in an input image. In the example in Fig. 5, the image recognition unit 152 recognizes a gesture G0 by monitoring movement of a hand

region A01 of the user. The gesture G0 is a gesture ("Hand Up") of raising the hand. Note that in the present specification, the expression "gestures" is assumed to also include static "poses" (forms) that do not involve dynamic movement of the user's body. On recognizing a user gesture, the image recognition unit 152 outputs gesture data showing the recognized type of gesture to the control unit 170.

- [0040] Example of gestures that can be recognized by the image recognition unit 152 will now be described with reference to Figs. 6 to 12.
- [0041] In Fig. 6, a hand region A11 is shown. The hand region A11 is being waved to the left and right at short intervals. From such movement of the user's hand, the image recognition unit 152 can recognize a gesture G1. The gesture G1 is a gesture ("Waving") of waving the hand.
- [0042] In Fig. 7, a hand region A12 is shown. The hand region A12 is substantially stationary in the input image for a specific length of time. From such stationary state of the user's hand, the image recognition unit 152 can recognize a gesture G2. The gesture G2 ("Keep Still") is a gesture of keeping the hand still.
- [0043] In Fig. 8, a hand region A13 is shown. The user's hand appearing in the hand region A13 is rotating in the counterclockwise direction around a center point in a vicinity of the wrist. From such movement of the user's hand, the image recognition unit 152 can recognize a gesture G3a. The gesture G3a ("Rotation") is a gesture of rotating the hand.
- [0044] In Fig. 9, a hand region A14 is shown. The user's hand appearing in the hand region A14 is in a shape where fingers aside from the index finger are bent over and the hand is rotating in the counterclockwise direction around a center point in a vicinity of the wrist. From such movement of the user's hand, the image recognition unit 152 recognizes a gesture G3b. The gesture G3b ("Rotation") is also a gesture of rotating the hand.
- [0045] In Fig. 10, a hand region A15 is shown. The hand of the user appearing in the hand region A15 is moving so as to bend the wrist to the right. From such movement of the user's hand, the image recognition unit 152 can recognize a gesture G4a. The gesture G4a ("Touch") is a gesture of touching an object.
- [0046] In Fig. 11, a hand region A16 is shown. The hand of the user appearing in the hand region A16 is moving so as to bend the wrist forward with the hand in a shape where all of the fingers aside from the index finger are bent over. From such movement of the user's hand, the image recognition unit 152 can recognize a gesture G4b. The gesture G4b ("Touch") is also a gesture of touching an object.
- [0047] In Fig. 12, a hand region A17 is shown. The hand of the user appearing in the hand region A17 is changing from a shape where the palm of the hand is open to a shape where the palm is closed. From such movement of the user's hand, the image

recognition unit 152 can recognize a gesture G5. The gesture G5 ("Grasp") is also a gesture of grasping an object.

[0048] Note that the gestures described here are mere examples. It is not necessary for the image recognition unit 152 to recognize a number of such gestures and/or the image recognition unit 152 may additionally recognize other types of gestures.

[0049] (5) Voice Recognition Unit

The voice recognition unit 154 carries out voice recognition on the voice of the user based on an input voice inputted from the voice acquisition unit 130. If, for example, an application being carried out or a UI receives the inputting of a voice command, the voice recognition unit 154 recognizes a voice command from the user's voice and outputs an identifier of the recognized voice command to the application unit 140 or the control unit 170.

[0050] (6) Characteristics DB

The characteristics DB 160 stores in advance image characteristics data which is to be used in image recognition by the image recognition unit 152. As one example, the image characteristics data may include known image characteristic values for an operation object (such as the hand) used by the user and the face of the user. The image characteristics data may also include facial image data for each user. The image characteristics data may also include gesture definition data defining gestures to be recognized by the image recognition unit 152. The characteristics DB 160 may also store in advance voice characteristics data to be used for voice recognition by the voice recognition unit 154.

[0051] (7) Operation Control Unit

The operation control unit 172 generates a UI image by superimposing at least one UI object on the input image, and displays a generated UI image (an output image corresponding to the input image) on the screen of the display 108. The input image for generating the UI image may differ to the input image to be used by the image recognition unit 152 for recognizing the operation object (as one example, an image with reduced resolution may be used to recognize the operation object). The operation control unit 172 then controls the displaying and operation of at least one UI object based on the recognition result of the operation object inputted from the image recognition unit 152.

[0052] Fig. 13 is a diagram useful in explaining a first example of UI objects. In the first example, the UI objects are menu items of an application menu. As shown in Fig. 13, a UI image W04 includes UI objects B11 to B16. The UI object B11 is a menu item (Volume) for setting the volume when reproducing a television program. The UI object B12 is a menu item (Channel) for setting the channel to be reproduced. The UI object B13 is a menu item (TV Guide) for launching an electronic program guide display

function. The UI object B14 is a menu item (Apps) for launching other application functions. The UI object B15 is a menu item (Internet) for launching an Internet browsing function. The UI object B16 is a menu item (Settings) for launching an appliance setting function of the information processing apparatus 100. Such menu items may be defined in a hierarchy, and as one example subitems "Increase volume" and "Decrease volume" may be present below a UI object B11 that is the menu item for setting volume. If individual users are identified using input images, a set of UI objects that are personalized for individual users may be displayed.

[0053] Fig. 14 is a diagram useful in explaining a second example of UI objects. In this second example, the UI objects are content items. As shown in Fig. 14, a UI image W05 includes UI objects B21 to B26. The UI objects B21 to B26 respectively express thumbnails of photographic content. Note that instead of photographic content, the UI objects may be other types of content items, such as video content, music content, or text content.

[0054] As shown in Figs. 13 and 14, in a state where an operation object has not been recognized, the operation control unit 172 lays out the UI objects at default display positions. The default display positions may be positions provided so as to be fixed or may be positions that move (as one example, that move so as to float) in accordance with some type of algorithm. The operation control unit 172 then causes the display positions of at least one UI object displayed before recognition of the operation object to approach toward the user after recognition of the operation object. The operation control unit 172 may cause the display position of a UI object to approach toward the recognized operation object or to approach toward part of the body of a different user to the recognized operation object. Thus, a position of the UI-object is spatially related to the recognized operation object since it is moved toward the recognized operation object, such as the user's hand.

[0055] In a certain scenario, the mode of approach of UI objects toward the user is uniform. That is, the operation control unit 172 sets the approach speeds of the UI objects that are to make the approach at the same value so that such UI objects all approach toward the user at the same approach speed.

[0056] In another scenario, the mode of approach of UI objects toward the user is non-uniform. That is, the operation control unit 172 sets the approach speeds of the UI objects that are to make the approach at different values so that the UI objects approach toward the user at different approach speeds. In addition to (or instead of) the approach speed, other attributes of the UI objects may be set non-uniformly. As examples, such other attributes may include at least one of approach start timing, post-approach display positions (hereinafter referred to as the "target positions"), display size, transparency, and depth.

[0057] As one example, when causing the display positions of the UI objects to approach toward the user, the operation control unit 172 may vary the mode of approach of the respective objects in accordance with priorities set for the respective objects. Such priorities are set in advance by the priority setting unit 174 in accordance with a specific priority setting standard and stored by an operation DB 180. A first example of the priority setting standard is a standard relating to an operation history of the UI objects. As one example, the priorities may be set higher for UI objects with a higher operation frequency (the number of past operations per specific period) and the priorities may be set lower for UI objects with a lower operation frequency. It is also possible to set the priority higher for UI objects that were operated at more recent timing in the past. A second example of a priority setting standard is a standard relating to user attributes. As one example, out of a plurality of content items, the priorities of UI objects corresponding to content items with a high recommendation score calculated according to a known recommendation technology based on the user attributes may be set at higher values. The operation control unit 172 may provide the user with a UI for switching the priority setting standard at desired timing between a plurality of candidates. Such UI may be realized by any method such as user gestures or voice commands.

[0058] Typically, the operation control unit 172 sets the approach speed and other attributes of the UI objects so as to make objects that have higher priorities easier to operate for the user. More specifically, as one example, the operation control unit 172 may set the approach speed toward the user higher for objects with higher priorities. The operation control unit 172 may also set the approach start timing of objects earlier for objects with higher priorities. Also, for an object with higher priority, the operation control unit 172 may set the target position closer to the user, the display size larger, the transparency lower, or the depth shallower.

[0059] Regardless of whether the mode of approach toward the user of the UI objects is uniform or non-uniform, the operation control unit 172 controls operations of UI objects by the user in response to a number of operation events defined in advance. The operation events typically include recognition of a user gesture, and recognition of voice commands may be used to complement such recognition. At least one operation event is recognition of a new operation object. Recognition of a new operation object may trigger UI objects approaching toward the user. Another operation event may trigger execution (launching) of a process associated with a UI object. A number of specific examples of operation events that can be used in the present embodiment are described later in this specification.

[0060] The operation control unit 172 also controls the displaying of a UI image via the display 108. As one example, the operation control unit 172 may display only an UI

image on which UI objects are superimposed on the screen of the display 108. Alternatively, the operation control unit 172 may display a single output image generated by combining a UI image and an application image generated by the application unit 140 on the screen. A number of examples of window compositions of output images that can be used in the present embodiment are described later.

[0061] (8) Priority Setting Unit

The priority setting unit 174 sets the priority of each UI object in accordance with the priority setting standard described earlier. As one example, in accordance with a priority setting standard relating to the operation histories of UI objects, the priority setting unit 174 may set the priorities higher for UI objects with a higher operation frequency. Also, in accordance with a priority setting standard relating to user attributes, the priority setting unit 174 may set the priorities higher for UI objects corresponding to content items with a higher recommendation score. The priority setting unit 174 may also set the priorities of UI objects randomly to add an element of surprise to the UI. The priority setting unit 174 may update the priority data for example when a UI object has been operated or when the user attributes have changed.

[0062] (9) Operation DB

The operation DB 180 stores data used by the operation control unit 172 to control displaying and operations of UI objects. The data stored by the operation DB 180 includes object data showing a default display position and default values of other attributes for each UI object. The data stored by the operation DB 180 may also include priority data showing priorities set by the priority setting unit 174, operation history data showing an operation history for each user, and user data showing attributes (such as age, sex, occupation, and tastes) for each user.

[0063] <2-3. Mode Of Approach of UI Objects>

This section will describe a number of examples of modes of approach toward the user of UI objects with reference to Figs. 15 to 21.

[0064] (1) First Example

Fig. 15 is a diagram useful in explaining a first example of a mode of approach toward the user of UI objects. In this first example, the mode of approach of the UI objects is uniform. After sufficiently approaching the user, the respective UI objects move away from the vicinity of the user without remaining at such target positions.

[0065] In Fig. 15, four UI images ST11 to ST14 are shown along a time axis.

[0066] A user Ud appears in the UI image ST11 and a mirror image display is realized. UI objects B11 to B16 are laid out at default display positions. It is assumed here that as a result of the user Ud raising his hand, the image recognition unit 152 recognizes the gesture G0.

[0067] In the next UI image ST12, the user Ud is raising his hand. In response to the

recognition of the gesture G0, the operation control unit 172 causes the UI objects B11 to B16 to start approaching toward the user Ud (as one example, the hand that is the operation object of the user Ud). In the UI image ST12, the UI object B13 is positioned closest to the user Ud.

- [0068] In the next UI image ST13, the UI object B13 returns to the default display position and in place of the UI object B13, the UI objects B12 and B14 are positioned in the vicinity of the user Ud's hand. In the next UI image ST14, the UI objects B12 and B14 return to the default display positions and in place of the UI objects B12 and B14, the UI objects B11, B15 and B16 are positioned in the vicinity of the user Ud's hand.
- [0069] According to this mode of approach, it is possible for the user to touch the UI object B13 at the time of the UI image ST12, to touch any of the UI objects B12 and B14 at the time of the UI image ST13, and to touch any of the UI objects B11, B15, and B16 at the time of the UI image ST14 with a simple operation of merely moving his hand or arm. At such time, since the screen region in the vicinity of the user is not crowded by a large number of UI objects, the risk of an erroneous operation being carried out, such as touching the wrong UI object, is reduced.
- [0070] (2) Second Example
  - Fig. 16 is a diagram useful in explaining a second example of a mode of approach toward the user of UI objects. In this second example, the mode of approach of the UI objects is non-uniform. After sufficiently approaching the user, the respective UI objects remain at their target positions.
- [0071] In Fig. 16, four UI images ST21 to ST24 are shown along a time axis.
- [0072] The user Ud appears in the UI image ST21 and a mirror image display is realized. UI objects B11 to B16 are laid out at default display positions. It is assumed here that as a result of the user Ud raising his hand, the image recognition unit 152 recognizes the gesture G0.
- [0073] In the next UI image ST22, the user Ud is raising his hand. In response to the recognition of the gesture G0, the operation control unit 172 causes the UI objects B11 to B16 to start approaching toward the user Ud. In the UI image ST22, the UI object B13 is positioned closest to the user Ud.
- [0074] In the next UI image ST23, the UI object B13 remains at its target position and the UI objects B12 and B14 also reach the vicinity of the user Ud's hand. In the next UI image ST24, the UI objects B12, B13, and B14 remain at their target positions and the UI objects B11, B15 and B16 also reach the vicinity of the user Ud's hand.
- [0075] According to this mode of approach, it is possible for the user to touch the UI object B13 at the time of the UI image ST22, to touch any of the UI objects B12, B13, and B14 at the time of the UI image ST23, and to touch any of the UI objects B11 to B16 at the time of the UI image ST24 with a simple operation of merely moving his hand or

arm. In particular, at the time of the UI images ST22 and ST23, crowding of the screen region in the vicinity of the user by a large number of UI objects is avoided. Also, as described later, if an operation event of moving the display position of a designated UI object away from the user is used, it is possible to ease the crowding of the screen region in the vicinity of the user from a state where the UI objects B11 to B16 have reached the vicinity of the user.

[0076] (3) Third Example

Fig. 17 is a diagram useful in explaining a third example of a mode of approach toward the user of UI objects. In this third example, the mode of approach of the UI objects is non-uniform and the operation control unit 172 sets the approach speeds of the UI objects that are to make an approach at different values.

[0077] In Fig. 17, three UI images ST31 to ST33 are shown along a time axis.

[0078] The user Ud appears in the UI image ST31 and a mirror image display is realized. The UI objects B11 to B16 are laid out at default display positions. It is assumed here that as a result of the user Ud raising his hand, the image recognition unit 152 recognizes the gesture G0.

[0079] In the next UI image ST32, the user Ud is raising his hand. In response to the recognition of the gesture G0, the operation control unit 172 sets the respective approach speeds  $V_{11}$  to  $V_{16}$  of the UI objects B11 to B16 in accordance with the priorities set by the priority setting unit 174. In the example in Fig. 17, it is assumed that priorities that were set by the priority setting unit 174 in accordance with a priority setting standard relating to the operation histories of the respective UI objects are used by the operation control unit 172.

[0080] Fig. 19 is a diagram useful in explaining an example of priority data corresponding to the third example in Fig. 17. As shown in Fig. 19, priority data 182a includes three data items, namely "object ID", "operation frequency", and "priority". The "object ID" is an identifier for uniquely identifying each UI object. The "operation frequency" shows the number of past operations per specific period (for example, one week, one day, or a few hours) of each UI object. The "priority" shows a priority set for each UI object. Here, as one example, the priority is set at one of three levels, namely "High", "Middle" and "Low". In the example in Fig. 19, the operation frequencies of the UI objects B11 and B15 are the first and second highest, and accordingly the priorities of the UI objects B11 and B15 are set at High. The operation frequency of the UI object B12 is the third highest, and accordingly the priority of the UI object B12 is set at Middle. The priorities of the remaining UI objects B13, B14, and B16 are set at Low.

[0081] The operation control unit 172 refers to such priority data 182a and sets the approach speeds  $V_{11}$  and  $V_{15}$  of the UI objects B11 and B15 at the fastest speed, the approach speed  $V_{12}$  of the UI object B12 at next fastest speed, and the approach speeds  $V_{13}$ ,  $V_{14}$ ,

and  $V_{16}$  of the UI objects B13, B14, and B16 at the slowest speed. As shown in Fig. 17, in the UI image ST33, the UI objects B11 and B15 reach the vicinity of the user Ud's hand the earliest.

[0082] According to this mode of approach, it is possible for the user to rapidly operate the UI objects that are operated most frequently. Also, since the screen region in the vicinity of the user does not become crowded with a large number of UI objects, the risk of an erroneous operation being carried out, such as touching the wrong UI object, is reduced.

[0083] (4) Fourth Example

Fig. 18 is a diagram useful in explaining a fourth example of a mode of approach toward the user of UI objects. In this fourth example, the mode of approach of the UI objects is non-uniform and the operation control unit 172 sets the approach speeds of the UI objects that are to make an approach at different values.

[0084] In Fig. 18, three UI images ST41 to ST43 are shown along a time axis.

[0085] The user Ud appears in the UI image ST41 and a mirror image display is realized. The UI objects B21 to B26 are laid out at default display positions. It is assumed here that as a result of the user Ud raising his hand, the image recognition unit 152 recognizes the gesture G0.

[0086] In the next UI image ST42, the user Ud is raising his hand. In response to the recognition of the gesture G0, the operation control unit 172 sets the respective approach speeds  $V_{21}$  to  $V_{26}$  of the UI objects B21 to B26 in accordance with the priorities set by the priority setting unit 174. In the example in Fig. 18, it is assumed that priorities that were set by the priority setting unit 174 in accordance with a priority setting standard relating to user attributes are used by the operation control unit 172.

[0087] Fig. 20 is a diagram useful in explaining an example of priority data corresponding to the fourth example in Fig. 18. As shown in Fig. 20, priority data 182b includes three data items, namely "object ID", "recommendation score", and "priority". The "object ID" is an identifier for uniquely identifying each UI object. The "recommendation score" shows a recommendation score calculated (by an arbitrary known recommendation algorithm) based on attributes of the user and information relating to content. The "priority" shows a priority set for each UI object. Here, as one example, it is assumed that the priority is set at one of three levels, namely "High", "Middle" and "Low". In the example in Fig. 20, the recommendation score of the UI object B21 is the highest, and accordingly the priority of the UI object B21 is set at High. The recommendation scores of the UI objects B22 and B25 are the second and third highest, and accordingly the priorities of the UI objects B22 and B25 are set at Middle. The priorities of the remaining UI objects B23, B24, and B26 are set at Low.

[0088] The operation control unit 172 refers to such priority data 182b and sets the approach

speed  $V_{21}$  of the UI object B21 at the fastest speed, the approach speeds  $V_{22}$  and  $V_{25}$  of the UI objects B22 and B25 at the next fastest speed, and the approach speeds  $V_{23}$ ,  $V_{24}$ , and  $V_{26}$  of the UI objects B23, B24 and B26 at the slowest speed. Returning to Fig. 18, in the UI image ST42, the UI object B21 has reached the vicinity of the user Ud's hand the earliest. However, if the UI object B21 (and the following UI objects B22, B25) is not operated by the user Ud, in the next UI image ST43, the UI objects B23, B24, and B26 that have a slow approach speed may reach the vicinity of the user Ud.

[0089] According to this mode of approach, it is possible for the user to rapidly operate a UI object determined by a recommendation algorithm to be more suited to such user. Also, since the screen region in the vicinity of the user is not crowded by a large number of UI objects, the risk of an erroneous operation being carried out, such as touching the wrong UI object, is reduced.

[0090] (5) Fifth Example

Fig. 21 is a diagram useful in explaining a fifth example of a mode of approach toward the user of UI objects. In this fifth example, the mode of approach of the UI objects is non-uniform and the operation control unit 172 sets the approach speeds of the UI objects that are to make an approach at different values. The operation control unit 172 also sets the display sizes of the UI objects to make an approach at different values.

[0091] In Fig. 21, three UI images ST51 to ST53 are shown along a time axis.

[0092] The user Ud appears in the UI image ST51 and a mirror image display is realized. The UI objects B11 to B16 are laid out at default display positions. It is assumed here that as a result of the user Ud raising his hand, the image recognition unit 152 recognizes the gesture G0.

[0093] In the next UI image ST52, the user Ud is raising his hand. In response to the recognition of the gesture G0, the operation control unit 172 sets the respective approach speeds and display sizes of the UI objects B21 to B26 in accordance with the priorities set by the priority setting unit 174. In the example in Fig. 21, it is assumed that the priority data 182a illustrated in Fig. 19 is used.

[0094] The operation control unit 172 refers to such priority data 182a and sets the approach speeds  $V_{11}$  and  $V_{15}$  of the UI objects B11 and B15 at the fastest speed, the approach speed  $V_{12}$  of the UI object B12 at the next fastest speed, and the approach speeds  $V_{13}$ ,  $V_{14}$ , and  $V_{16}$  of the UI objects B13, B14 and B16 at the slowest speed. The operation control unit 172 also sets the display size of the UI objects B11 and B15 the largest, the display size of the UI object B12 the next largest, and the display sizes of the UI objects B13, B14 and B16 the smallest.

[0095] In the next UI image ST53, most of the UI objects have already reached the vicinity of the user Ud's hand and out of such UI objects, the UI objects B11 and B15 have

larger display sizes than the other UI objects.

[0096] According to this mode of approach, it is possible for the user to operate a UI object that has a higher priority more rapidly and more accurately.

[0097] <2-4. Various Examples of Operation Events>

In this section, a number of examples of operation events relating to control by the operation control unit 172 will be described with reference to Figs. 22 to 28.

[0098] (1) First Example

Fig. 22 is a diagram useful in explaining a first example of an operation event. In Fig. 22, the user's hand that is the operation object and a UI object B13 are shown. From a movement of the user's hand that bends the wrist, the image recognition unit 152 can recognize the gesture G4b of touching an UI object. By comparing a position in the UI image of the hand region at such time and the display position of the respective UI objects, the operation control unit 172 can determine that the UI object B13 has been designated (that is, touched). If the position of the hand region coincides with the display position of the same UI object for a specified length of time, the operation control unit 172 may also determine that such UI object is designated. In response to such an operation event, the operation control unit 172 causes the application unit 140 to carry out a process associated with the designated UI object B13. In the example in Fig. 22, since the UI object B13 is a menu item for launching an electronic program guide display function, an electronic program guide can be then displayed on the screen by the application unit 140.

[0099] As a result of using such an operation event, the user is capable of remotely controlling the information processing apparatus 100 even when a remote controller is not at hand. At this time, since the only movement necessary by the user is a simple gesture, the user is capable of having the information processing apparatus 100 carry out a desired process (for example, a menu process or an application process) without feeling stress.

[0100] Note that the process associated with a UI object may be a process for UI control. For example, opening a submenu item from the designated menu item, calling a setting screen corresponding to the designated menu item, and the like may be carried out in response to recognition of a gesture of touching a UI object.

[0101] (2) Second Example

Fig. 23 is a diagram useful in explaining a second example of an operation event. As shown in Fig. 23, in a UI image ST15 that follows the UI image ST14 shown in Fig. 15, the user Ud's left hand that is the operation object coincides with the UI object B16. When this happens, the operation control unit 172 determines that the UI object B16 has been designated by the user Ud. In addition, it is assumed that a specific voice command VR1 issued by the user has been recognized by the voice recognition unit

154. In the example in Fig. 23, the voice command VR1 corresponds to a voice input of "Go!". The operation control unit 172 causes the application unit 140 to carry out a process associated with the designated UI object B16 in response to an operation event corresponding to recognition of the voice command VR1 by the voice recognition unit 154.

- [0102] In the UI image ST15, the operation control unit 172 sets display attributes (for example, at least one of texture, color, transparency, display size, and depth) of the designated UI object B16 at different attribute values to other UI objects. By doing so, it is possible for the user to grasp that the UI object B16 was appropriately designated.
- [0103] Even when such an operation event is used, it is possible for the user to remotely control the information processing apparatus 100 even when a remote controller is not at hand. At this time, since the only movement necessary for the user is a simple gesture, it is possible to reduce the user's stress. Note that for the other examples of operation events described in this section, recognition of the user's gesture may be substituted with recognition of a voice command.

[0104] (3) Third and Fourth Examples

Fig. 24 is a diagram useful in explaining a third example and a fourth example of operation events. In Fig. 24, the UI image ST24 shown in Fig. 16 and following UI images S25 and ST26 are also shown. In the UI image ST24, the UI objects B11 to B16 have display positions laid out in a ring in the vicinity of the user Ud's hand that is the operation object. After this, the image recognition unit 152 may recognize the gesture G2 of keeping the hand still from movement of the user's hand where the hand remains substantially still in the input image for a specific length of time. The operation control unit 172 stops the movement on the screen of the UI objects B11 to B16 in response to an operation event corresponding to recognition of the gesture G2. In the next UI image ST25 also, the display positions of the UI objects B11 to B16 are not updated.

- [0105] In the next UI image ST26, the image recognition unit 152 may recognize the gesture G3a of rotating the hand from a movement where the user Ud's hand, which is operation object, rotates. In response to such operation event, the operation control unit 172 rotates (in the direction D1 in the image) the display positions of the UI objects B11 to B16 around a reference point in the image. As examples, the reference point referred to here may be a center of gravity of the hand region, a center of the UI objects B11 to B16, or any other arbitrary point.
- [0106] As a result of using an operation event of stopping the movement of the UI objects described earlier, it is possible for the user to stop further movement of the objects when a desired UI object has reached a display position that is suited to operation and to then accurately operate a desired UI object.

[0107] Also, as a result of using an operation event of rotating the display positions of the UI objects described earlier, it is possible for the user to move the display positions of UI objects that have approached the vicinity of the user to positions that is easier to handle. Instead of rotating the display positions of the UI objects in response to recognition of the gesture of rotating the hand, the display positions of the UI objects may move in parallel in response to movement of the user's hand. Note that instead of all of the displayed UI objects rotating or moving as shown in the example in Fig. 24, only some of the UI objects may be rotated or moved. As one example, at least one UI object to be rotated or moved may be designated by a gesture of the user tracing the UI objects with his/her hand.

[0108] (4) Fifth Example

Fig. 25 is a diagram useful in explaining a fifth example of an operation event. In the upper half of Fig. 25, the user's hand that is the operation object is shown together with the UI objects B13 and B14. The image recognition unit 152 may recognize the gesture G1 of waving the hand from a movement where the user's hand is waved to the left and right. By comparing the position in the UI image of the hand region at such time and the display positions of the respective UI objects, the operation control unit 172 can determine that the UI object B13 has been designated. In response to such operation event, the operation control unit 172 moves the display position of the designated UI object B13 away from the user. In the lower half of Fig. 25, the display position of the UI object B13 is moved away from the user and in place of the UI object B13 the UI object B14 approaches the user's hand.

[0109] As a result of using such an operation event, even if a UI object that is not necessary has approached the user earlier than a desired UI object, it is still possible for the user to remove the UI object that is not necessary from the screen region in the vicinity of the user and thereby prevent crowding of such screen region.

[0110] (5) Sixth Example

Fig. 26 is a diagram useful in explaining a sixth example of an operation event. In Fig. 26, the UI images ST51 and ST52 shown in Fig. 21 and a following UI image ST54 are shown along a time axis.

[0111] The user Ud appears in the UI image ST51 and a mirror image display is realized. The UI objects B11 to B16 are also laid out at default display positions. The UI objects B11 to B16 are objects belonging to a first category out of a plurality of categories defined in advance. As one example, the first category is a category relating to a television program reproduction function. Note that in an initial state, UI objects are not necessarily visible to the user. For example, in an initial state, UI objects may be positioned outside the screen, or may be transparent or translucent. The UI objects may change from a non-active state (undisplayed or translucent) to an active state

(displayed or non-transparent) at timing where the user raises his/her hand.

- [0112] In a UI image ST52, as a result of the user Ud raising his hand, the UI objects B11 to B16 start to approach the user Ud. In the UI image ST52, the image recognition unit 152 may recognize the gesture G1 of waving the hand from a movement where the user's hand is waved to the left and right. In response to an operation event that corresponds to recognition of the gesture G1, the operation control unit 172 replaces the objects B11 to B16 laid out in the UI image with the UI objects belonging to a second category. The second category may be an arbitrary category (such as a category relating to a content reproduction function) that differs to the first category.
- [0113] In the UI image ST54, the objects B11 to B16 are removed from the screen and new UI objects B31 to B37 are laid out on the screen.
- [0114] As a result of using such operation event, the information processing apparatus 100 is capable of displaying only some of the UI objects on the screen without displaying all of the UI object candidates that can be displayed on the screen. Accordingly, crowding of the screen region is significantly eased. It is also possible for the user to have a desired UI object, which is not presently displayed at such time, displayed on the screen via a simple gesture and to appropriately operate such UI object.
- [0115] Note that selection of the category of UI objects to be displayed in an UI image may depend on the shape of the user's hand. For example, the UI objects that have been displayed so far may be replaced with UI objects that belong to any of the first to fifth categories in response to recognition of five types of hand shape that respectively express the numbers one to five.
- [0116] Also, the gesture G1 may be defined as not as a gesture for switching the category of UI objects to be displayed but as a gesture for switching the priority setting standard used to set the approach speeds. In this case, in response to the operation event corresponding to the recognition of the gesture G1, the operation control unit 172 resets the priorities of at least one of the UI objects being displayed in accordance with the new priority setting standard.
- [0117] (6) Seventh and Eighth Examples  
Fig. 27 is a diagram useful in explaining a seventh example and an eighth example of operation events. In Fig. 27, the UI image ST42 shown in Fig. 18 and following UI images ST44, ST45, and ST46 are shown along a time axis.
- [0118] The user Ud appears in the UI image ST42 and is raising his hand. The UI objects B21 to B26 are approaching toward the user Ud at approach speeds respectively set by the operation control unit 172.
- [0119] In the next UI image ST44, the image recognition unit 152 may recognize the gesture G5 of grasping an object from a movement of the user's hand that changes from a shape where the palm of the hand is open to a shape where the hand is closed. By

comparing the position in the UI image of the hand region at such time with the display positions of the respective UI objects, the operation control unit 172 may determine that the UI object B25 is designated (that is, grasped). In response to an operation event corresponding to the recognition of the gesture G5, the operation control unit 172 thereafter has the display position of the designated UI object B25 track the position of the hand region (that is, has the UI object B25 move together with the operation object).

- [0120] In the next UI image ST45, UI objects aside from the designated UI object B25 are removed. Also, two screen regions R11 and R12 are set in the image. As one example, in a UI image, the operation control unit 172 may set an equal number of screen regions to the number of processes associated with the designated UI object B25. As one example, if the UI object B25 is a content item for photographic content, the screen region R11 may be associated with launching an image viewer and the screen region R12 may be associated with transmitting a message to which photographic content is appended.
- [0121] In the next UI image ST46, as a result of the user's hand region moving to a position that coincides with the screen region R12, the display position of the UI object B25 also moves to a position that coincides with the screen region R12. In response to the operation event corresponding to the movement of the UI object B25 to such specific screen region, the operation control unit 172 causes the application unit 140 to carry out a process associated with the UI object B25. Here, as one example, a message transmission function is launched by the application unit 140 and photographic content may be appended to a new message.
- [0122] As a result of using such operation event, it is possible for the user to launch a desired process for a UI object with an easy and intuitive operation, even when a variety of processes are related to a single UI object.
- [0123] (7) Ninth Example  
Fig. 28 is a diagram useful in explaining a ninth example of an operation event. In Fig. 28, the four UI images ST61 to ST64 are shown along a time axis.
- [0124] The user Ud appears in the UI image ST61 and a mirror image display is realized. The UI objects B11 to B16 are also laid out at default display positions. The UI objects B11 to B16 are objects belonging to a first category out of a plurality of categories defined in advance. As one example, the first category is a category relating to a television program reproduction function. In addition, four screen regions R21 to R24 are set in the image. The screen regions R21 to R24 may be associated with respectively different categories.
- [0125] In the UI image ST62, as a result of the user Ud raising her hand, the UI objects B11 to B16 start to approach toward the user Ud. The position of the user Ud's hand

coincides with the screen region R23. It is assumed that the screen region R23 is associated with the first category.

- [0126] In the UI image ST63, as a result of the position of the hand region being lowered in a downward direction, the position of the user Ud's hand region coincides with the screen region R24. It is assumed that the screen region R24 is associated with a different category to the first category. In response to an operation event corresponding to movement of the hand that is the operation object between screen regions, the objects B11 to B16 laid out in the UI image are replaced with objects belonging to another category.
- [0127] In the UI image ST63, the objects B11 to B16 are removed from the screen and new UI objects B41 to B45 are laid out in the image. In the next UI image ST64, the UI objects B41 to B45 start to approach toward the user Ud.
- [0128] In the same way as the sixth example described earlier, as a result of such operation event being used, the information processing apparatus 100 is capable of displaying only some of the UI objects on the screen instead of all of the UI object candidates that can be displayed. Accordingly, crowding of the screen region is significantly eased. It is also possible for the user to have desired UI objects that are not on display at the present time displayed on the screen via a simple operation of simply moving the hand and to appropriately operate such UI objects.
- [0129] <2-5 Incorporation of a Plurality of Operation Objects>  
Examples where a single operation object is recognized in the input image have mainly been described so far. However, with the technology according to the present disclosure, a plurality of operation objects may be recognized in the input image. In this section, a number of example operation scenarios involving a plurality of operation objects will be described with reference to Figs. 29 to 31.
- [0130] (1) First Example  
Fig. 29 is a diagram useful in explaining a first example of an operation scenario involving a plurality of operation objects. In this first example, the left hand and the right hand of a single user are recognized as separate operation objects.
- [0131] In Fig. 29, four UI images ST71 to ST74 are shown along a time axis.
- [0132] A user Ud appears in the UI image ST71 and a mirror image display is realized. UI objects B51 to B58 are laid out at default display positions. It is assumed here that the UI objects B51 to B58 are grouped into a plurality of groups in accordance with a grouping standard. As examples, the UI objects may be grouped according to a standard relating to the priorities described earlier, the types of corresponding menu items or content items, or display positions, or may be randomly grouped.
- [0133] In the next UI image ST72, the user Ud raises his left hand and the hand region A21 is recognized. In response to recognition of the gesture G0 for the hand region A21, the

operation control unit 172 causes the UI objects B53 to B56 included in the first group to start approaching toward the user Ud.

- [0134] In the next UI image ST73, the user Ud further raises his right hand and the hand region A22 is recognized. In response to recognition of the gesture G0 for the hand region A22, the operation control unit 172 causes the UI objects B51, B52, B57, and B58 included in the second group to start approaching toward the user Ud.
- [0135] In the next UI image ST74, the UI objects B53 to B56 are laid out in a ring in the vicinity of the user Ud's left hand and the UI objects B51, B52, B57, and B58 are laid out in a ring in the vicinity of the user Ud's right hand. As one example, in response to recognition of a gesture of the user Ud bringing his hands together, the operation control unit 172 may form a single ring by merging the two rings of such UI objects. If the hand regions are positioned at edge portions of the screen, the operation control unit 172 may distort the shapes of the rings.
- [0136] According to this operation scenario, it is possible for the user to have a large number of UI objects approach the vicinity of the user using both hands. A range produced by combining the regions that can be reached by two hands is wider than a region that can be reached by one hand. This means that the user can rapidly designate a desired UI object out of a larger number of UI objects in the vicinity and operate the desired UI object.
- [0137] (2) Second Example
  - Fig. 30 is a diagram useful in explaining a second example of an operation scenario involving a plurality of operation objects. In this second example, the hands of two users are recognized as separate operation objects.
- [0138] In Fig. 30, two UI images ST81 to ST82 are shown along a time axis.
- [0139] The user Ud and a user Ue appear in the UI image ST81 and a mirror image display is realized. UI objects B61 to B68 are also displayed. The user Ud raises his left hand and a hand region A21 is recognized. In response to recognition of the gesture G0 for the hand region A21, the UI objects B61 to B68 are approaching toward the user Ud.
- [0140] In the next UI image ST82, the user Ue is raising her right hand and a hand region A31 is recognized. In response to recognition of the gesture G0 for the hand region A31, the operation control unit 172 has the UI objects B61, B64, B65 and B68 start to approach toward the user Ue. As one example, the UI objects may be grouped into a plurality of groups in accordance with a grouping standard relating to an operation history for each user or user attributes. In the example in Fig. 30, although the UI objects B62, B63, B66 and B67 are included in a first group intended for the user Ud, the UI objects B61, B64, B65 and B68 are included in a second group intended for the user Ue.
- [0141] In the UI image ST82, by setting display attributes (for example, color) of the UI

objects B61, B64, B65 and B68 at different attribute values to the other UI objects, the operation control unit 172 expresses that the UI objects B61, B64, B65 and B68 are included in the second group intended for the user Ue. If the target positions of the two groups interfere with one another, the operation control unit 172 may shift the target positions to eliminate such interference.

[0142] According to this operation scenario, it is possible to share UI objects among a plurality of users. When doing so, it is possible for each user to have suitable UI objects for such user approach the user and to rapidly operate a desired UI object. The fifth example of an operation event described with reference to Fig. 25 may be defined as an operation event of passing a UI object to another user, with it being possible to use such operation in this operation scenario.

[0143] (3) Third Example

Fig. 31 is a diagram useful in explaining a third example of an operation scenario involving a plurality of operation objects. In this third example also, the hands of two users are recognized as separate operation objects.

[0144] In Fig. 31, the UI image ST81 shown in Fig. 30 and a following UI image ST83 are shown along a time axis.

[0145] The user Ud and the user Ue appear in the UI image ST81 and a mirror image display is realized. The UI objects B61 to B68 are also displayed. The user Ud raises his left hand and a hand region A21 is recognized. In response to recognition of the gesture G0 for the hand region A21, the UI objects B61 to B68 continue to approach toward the user Ud.

[0146] In the next UI image ST83, the user Ue uses her right hand to touch the UI object B65. In response to recognition of the gesture G4a for the hand region A32 of the user Ue's right hand, the operation control unit 172 may determine that the UI object B65 is designated. In response to such operation event, the operation control unit 172 has the application unit 140 carry out a process associated with the designated UI object B65.

[0147] According to this operation scenario, it is possible for a different user to easily operate a UI object that has approached toward a given user.

[0148] Note that examples where UI objects are two-dimensionally laid out in a UI image have mainly been described so far. However, the respective UI objects are not limited to having two-dimensional display positions and may have an attribute corresponding to depth. If the information processing apparatus 100 is capable of recognizing the distance between a camera and an operation object using a known method such as parallax, the operation control unit 172 may determine which UI object has been designated by the user also based on such recognized distance.

[0149] <2-6. Example Window Compositions>

Figs. 32 and 33 show examples of the window composition of output images that

may be used by the present embodiment. In such drawings, a UI window  $W_{UI}$  and an application window  $W_{APP}$  are displayed by the display 108. The UI window  $W_{UI}$  displays a UI image generated by the operation control unit 172. The application window  $W_{APP}$  displays an application image (for example, a content image) inputted from the application unit 140. In the first example in Fig. 32, the application window  $W_{APP}$  is combined at the bottom right corner of the UI window  $W_{UI}$ . In the second example in Fig. 33, the UI window  $W_{UI}$  is blended with one part of the application window  $W_{APP}$ . By using such window composition, as one example it is possible, even when the user does not have a remote controller to hand, for the user to control functions of the information processing apparatus 100 via UI objects while viewing content images.

[0150] <2-7. Example Processing Flow>

The flowcharts in Figs. 34 and 35 show an example of the flow of processing that may be carried out by the information processing apparatus 100 according to the present embodiment. The processing described here is repeated for each frame in a series of frames that construct video picked up by the camera 101.

[0151] As shown in Fig. 34, first the image acquisition unit 120 acquires an image picked up by the camera 101 as an input image (step S100). The image acquisition unit 120 then outputs the acquired input image to the recognition unit 150 and the control unit 170.

[0152] Next, the image recognition unit 152 recognizes the operation object appearing in the input image inputted from the image acquisition unit 120 (step S105). It is assumed here that the operation object is the user's hand. For example, the image recognition unit 152 recognizes a hand region in the input image and outputs position data showing the position of such recognized hand region to the control unit 170. The image recognition unit 152 also recognizes a user gesture based on movement of the hand region. In addition, a voice command may also be recognized by the voice recognition unit 154 based on an input voice.

[0153] Next, the operation control unit 172 determines an operation event based on an image recognition result inputted from the image recognition unit 152 and a voice recognition result that may be inputted as necessary from the voice recognition unit 154 (step S110). The subsequent processing branches in accordance with the operation event determined here.

[0154] In step S115, the operation control unit 172 determines whether a new set of UI objects is to be displayed (step S115). As examples, if a UI image is to be newly displayed or if the operation event described with reference to Fig. 26 or Fig. 28 has been detected, the operation control unit 172 determines that a new set of UI objects is to be displayed. If it is determined here that a new set of UI objects is not to be displayed, the UI objects that were displayed in the previous frame are maintained and

the processing proceeds to step S120. Meanwhile, if it is determined here that a new set of UI objects is to be displayed, the processing proceeds to step S135.

- [0155] In step S120, the operation control unit 172 determines whether any of the UI objects has been selected (step S120). As one example, if the operation event described with reference to Fig. 22, Fig. 23, Fig. 27, or Fig. 31 has been detected, the operation control unit 172 determines that a UI object has been selected. If it is determined here that a UI object has been selected, the processing proceeds to step S125. If not, the processing proceeds to step S145.
- [0156] In step S125, in response to an operation event that selects a UI object, the operation control unit 172 causes the application unit 140 to carry out a process associated with the selected UI object (step S125). By increasing the operation frequency of the selected UI object for example, the priority setting unit 174 then updates the priority data (step S130). After this, the processing returns to step S100.
- [0157] In step S135, the operation control unit 172 sets up the new set of UI objects (step S135). As one example, the operation control unit 172 specifies a set of UI objects belonging to a different category to the set of UI objects that were displayed in the previous frame. The operation control unit 172 then lays out the UI objects included in the new set at the default display positions (step S140). After this, the processing proceeds to step S145.
- [0158] In step S145, the operation control unit 172 determines whether an operation object has been newly recognized (step S145). As examples, if the gesture G0 described with reference to Figs. 15 to 18 has been detected, the operation control unit 172 determines that an operation object has been newly recognized. Here, if it is determined that an operation object has been newly recognized, the processing proceeds to step S150. If not, the processing in step S150 is skipped.
- [0159] In step S150, the operation control unit 172 sets the approach speeds and other attributes of the UI objects (step S150). As one example, the operation control unit 172 may set the approach speed toward the user of an object with a higher priority at a higher speed. The operation control unit 172 may also set the display size of an object with a higher priority at a larger size.
- [0160] As shown over in Fig. 35, the operation control unit 172 then determines whether the display positions of the UI objects should be updated (step S155). As one example, if the gesture G2 described with reference to Fig. 24 has been detected, the operation control unit 172 determines that updating of the display positions is not necessary. Here, if it is determined that the display positions of the UI objects should be updated, the processing proceeds to step S160. If not, the processing in steps S160 and S165 is skipped.
- [0161] In step S160, the operation control unit 172 updates the display positions of UI

objects related to a special event (step S160). As one example, if the operation control unit 172 has detected the operation event described with reference to Fig. 25, the display position of the designated UI object is moved away from the user. Also, if the gesture G3a described with reference to Fig. 24 has been detected, the operation control unit 172 rotates the display positions of the UI objects.

- [0162] The operation control unit 172 then updates the display positions of other UI objects based on their approach speeds (step S165). As one example, the display positions of UI objects that have faster approach speeds may be moved much closer toward the user.
- [0163] After this, the operation control unit 172 generates a UI image by superimposing at least one UI object on the input image in accordance with the display positions and attributes decided via the processing so far (step S170). The operation control unit 172 displays an output image including a generated UI image on the screen of the display 108 (step S175). After this, the processing returns to step S100.
- [0164] <3. Second Embodiment>  
As described earlier, the technology according to an embodiment of the present disclosure is not limited to a television apparatus and can be applied to various types of apparatus. For this reason, an example where the technology according to an embodiment of the present disclosure has been applied to the information processing apparatus 200 that includes the internet will now be described as a second embodiment. As was described with reference to Fig. 2, the information processing apparatus 200 is a tablet PC.
- [0165] (1) Example Hardware Configuration  
Fig. 36 is a block diagram showing an example hardware configuration of the information processing apparatus 200. As shown in Fig. 36, the information processing apparatus 200 includes the camera 201, the microphone 202, an input device 203, a communication I/F 204, a memory 205, the display 208, a speaker 209, a bus 211, and a processor 212.
- [0166] The camera 201 includes an image pickup element such as a CCD or a CMOS and picks up images. The images picked up by the camera 201 (frames that construct video) are treated as input images for processing by the information processing apparatus 200.
- [0167] The sensor 202 may include various sensors such as a measurement sensor, an acceleration sensor, and a gyro sensor. The sensor data generated by the sensor 202 may be used by an application function of an information processing apparatus 200.
- [0168] The input device 203 is a device used by the user to directly operate the information processing apparatus 200 or to input information into the information processing apparatus 200. As one example, the input device 103 may include a touch panel,

buttons, switches, and the like. On detecting a user input, the input device 203 generates an input signal corresponding to the detected user input.

- [0169] The communication I/F 204 acts as an intermediary for communication between the information processing apparatus 200 and another apparatus. The communication I/F 204 supports an arbitrary wireless communication protocol or wired communication protocol and establishes a communication connection with the other apparatus.
- [0170] The memory 205 is constructed of a storage medium such as a semiconductor memory or a hard disk drive and stores programs and data for processing by the information processing apparatus 200, as well as content data. Note that some or all of the programs and data may not be stored by the memory 205 and instead may be acquired from an external data source (as examples, a data server, network storage, or an externally attached memory).
- [0171] The display 208 has a screen constructed of an LCD, an OLED, or the like and displays images generated by the information processing apparatus 200. As one example, the same UI images as those described in the first embodiment may be displayed on the screen of the display 208.
- [0172] The speaker 209 has a diaphragm and circuit elements such as an amplifier and outputs audio based on an output audio signal generated by the information processing apparatus 200. The volume of the speaker 209 is variable.
- [0173] The bus 211 connects the camera 201, the microphone 202, the input device 203, the communication I/F 204, the memory 205, the display 208, the speaker 209, and the processor 212 to each other.
- [0174] As examples, the processor 112 may be a CPU or a DSP. By executing a program stored in the memory 205 or on another storage medium, in the same way as the processor 112 of the information processing apparatus 100 according to the first embodiment, the processor 212 causes the information processing apparatus 200 to function in various ways. Aside from differences in the application function, the configuration of the logical functions realized by the memory 205 and the processor 212 of the information processing apparatus 200 may be the same as the configuration of the information processing apparatus 100 illustrated in Fig. 4.
- [0175] (2) Example Operation Scenario  
Fig. 37 is a diagram useful in explaining an example operation scenario for the second embodiment. In Fig. 37, four output images ST91 to ST94 are shown along a time axis. In this scenario, the respective output images are composed of an application image  $W_{APP}$  of an Internet browser in the left half and an UI image  $W_{UI}$  in the right half.
- [0176] In the output image ST91, the application image  $W_{APP}$  includes text written in a Web page. In Fig. 37, three keywords "XXX Computer Entertainment Inc.", "GameStation",

and "Christmas" extracted from the text of a Web page are surrounded by rectangular frames. The user Ud appears in the UI image WUI and a mirror image display is realized.

- [0177] The next output image ST92 may be displayed after the hand of the user Ud that is the operation object is recognized, for example. In the output image ST92, UI objects B71 to B73 are superimposed on the UI image. The UI object B71 is associated with the keyword "XXX Computer Entertainment Inc.". The UI object B72 is associated with the keyword "GameStation". The UI object B73 is associated with the keyword "Christmas".
- [0178] In the next output image ST93, the user Ud's hand coincides with the UI object B72. Three screen regions R41, R42, and R43 are set in the UI image. The screen region R41 is associated with a Web search (text search) process. The screen region R42 is associated with an image search process. The screen region R43 is associated with a movie search process.
- [0179] In the next output image ST94, the UI object B72 has moved so as to track movement of the user Ud's hand and has moved to a position that coincides with the screen region R41. In response to such operation event, the operation control unit 172 of the information processing apparatus 200 causes the application unit 140 to carry out a Web search function that uses the keyword "Gamestation" shown by the UI object B72.
- [0180] <4. Conclusion>

Embodiments of the present disclosure have been described in detail so far with reference to Figs. 1 to 37. According to the embodiments described above, a plurality of UI objects are displayed in a UI image that displays a mirror image of the user and, in UI in which operation of a UI object is controlled based on an image recognition result, the display positions of UI objects that are displayed before an operation object such as the user's hand is recognized approach toward the user after such operation object has been recognized. Accordingly, since the limited screen region in the vicinity of the user is not filled by the low number of UI objects, it is possible to avoid a drop in usability due to the screen being crowded.
- [0181] Also, according to the embodiments described above, the mode of approach toward the user of the UI objects may vary according to the priorities set for the respective UI objects. Accordingly, the user is capable of rapidly operating a UI object that has a higher priority (as examples, a UI object operated with higher frequency or a UI object determined to be suited to the user).
- [0182] According to the embodiments described above, various operation events triggered by user gestures may be realized. Accordingly, the user is capable of flexibly operating an information appliance using UI objects that have approached the vicinity of the

user, even when the user does not have a remote controller or other physical operation device.

[0183] Note that the series of processes carried out by the various apparatuses described as embodiments of the present disclosure are typically realized using software. As one example, programs composed of software that realizes such series of processes are stored in advance on a storage medium (non-transitory medium) provided internally in or externally to such apparatuses. As one example, during execution, such programs are then written into RAM (Random Access Memory) and executed by a processor such as a CPU.

[0184] Although preferred embodiments of the present disclosure are described in detail above with reference to the appended drawings, the technical scope of the disclosure is not limited thereto. It should be understood by those skilled in the art that various modifications, combinations, sub-combinations and alterations may occur depending on design requirements and other factors insofar as they are within the scope of the appended claims or the equivalents thereof.

[0185] Additionally, the present technology may also be configured as below.

(1) An information processing system comprising:  
processing circuitry configured to  
control a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, and wherein the predetermined displayed feature is an image derived from a camera-captured image.

(2) The information processing system of (1), wherein  
the processing circuitry is configured to vary a mode of approach of the UI object in accordance with a parameter related to the UI object.

(3) The information processing system of (2), wherein  
the mode of approach is non-uniform for the displayed object and other displayed objects such that respective speeds of approach are different for different displayed objects.

(4) The information processing system of (2), wherein  
the parameter is a priority.

(5) The information processing system of (4), wherein the priority is based on an operation frequency or a recommendation

(6)

The information processing system of (2), wherein the mode of approach is non-uniform for the displayed object and other displayed objects such that respective post-recognition displayed positions are different for different displayed objects.

(7)

The information processing system of (2), wherein a trigger for the movement for the displayed object and another displayed object are different.

(8)

The information processing system of (7), wherein a first detected gesture triggers a movement of the displayed object and a second detected gesture triggers a movement of the other displayed object.

(9)

The information processing system of (7), wherein the displayed object and the other displayed object are displayed in a ring around the operation object.

(10)

The information processing system of (1), wherein the processing circuitry is configured to control a movement of a plurality of UI objects.

(11)

The information processing system of (1), wherein the post-recognition position for the UI object is different for the UI object than for a different UI object.

(12)

The information processing system of (11), wherein the post-recognition position is closer to the operation object when the UI object is identified as a higher priority than the different UI object, and further from the operation object when the UI object is identified as a lower priority than the different UI object.

(13)

The information processing system of (1), wherein the predetermined displayed feature is a body part of the user.

(14)

The information processing system of (1), wherein the predetermined displayed feature is the operation object.

(15)

The information processing system of (1), wherein

the predetermined displayed feature is a feature of a user image.

(16)

The information processing system of (1), wherein  
the predetermined displayed feature is a feature of an action of a user.

(17)

The information processing system of (16), wherein  
the processing circuitry is also configured to implement an image recognition unit that  
recognizes a feature of the user as the operation object.

(18)

The information processing system of (1), wherein  
the post-recognition position of the displayed object to the operation object is a closer  
distance than the pre-recognition position such that the displayed object moves toward  
the operation object.

(19)

An information processing method comprising: controlling with processing circuitry a  
movement of a UI object on a display screen from a pre-recognition position toward a  
post-recognition position in response to recognition of an operation object initiated by  
a user, wherein the post-recognition position is spatially related to a displayed position  
of a predetermined displayed feature, and the predetermined displayed feature is an  
image derived from a camera-captured image.

(20)

A non-transitory computer readable medium having computer readable instructions  
that when executed by a processing circuitry perform a method, the method  
comprising:

controlling with the processing circuitry a movement of a UI object on a display screen  
from a pre-recognition position toward a post-recognition position in response to  
recognition of an operation object initiated by a user, wherein the post-recognition  
position is spatially related to a displayed position of a predetermined displayed  
feature, and the predetermined displayed feature is an image derived from a camera-  
captured image.

[0186] Additionally, the present technology may also be configured as below.

(1)

An information processing apparatus including:  
an image acquisition unit acquiring an input image;  
a recognition unit recognizing, in the input image, an operation object used by a user;  
and  
a control unit displaying, on a screen, an output image which corresponds to the input  
image and on which a plurality of objects to be operated by the user are superimposed

and controlling displaying of at least one of the objects based on a recognition result of the operation object,

wherein the control unit causes display positions of the plurality of objects being displayed on the screen before recognition of the operation object to respectively approach toward the user after the recognition of the operation object.

(2)

The information processing apparatus according to (1),

wherein the control unit varies a mode of approach of the respective objects toward the user in accordance with priorities set for the respective objects.

(3)

The information processing apparatus according to (2),

wherein the control unit sets at least one out of an approach speed, an approach start timing, a display position after approach, a display size, a transparency, and a depth of the plurality of objects so as to make operation of an object that has a higher priority easier for the user.

(4)

The information processing apparatus according to (3),

wherein the control unit sets the approach speed of the object that has the higher priority toward the user to a higher speed.

(5)

The information processing apparatus according to any one of (2) to (4), further including:

a priority setting unit setting the priority for each of the plurality of objects in accordance with a setting standard relating to any of an operation history for each of the objects and an attribute of the user.

(6)

The information processing apparatus according to any one of (1) to (5), further including:

an application unit carrying out processes associated with the respective objects, wherein the control unit is operable in response to a first event, to cause the application unit to carry out a process associated with an object designated by the operation object.

(7)

The information processing apparatus according to any one of (1) to (6),

wherein the control unit is operable in response to a second event, to stop movement of the plurality of objects on the screen.

(8)

The information processing apparatus according to (7),

wherein the control unit is operable in response to a third event after the movement of

the plurality of objects is stopped, to rotate the display positions of the plurality of objects around a reference point in the image.

(9)

The information processing apparatus according to any one of (1) to (8), wherein the control unit is operable in response to a fourth event, to move the display position of at least one object near the operation object away from the user.

(10)

The information processing apparatus according to any one of (1) to (9), wherein the plurality of objects belong to a first category out of a plurality of categories defined in advance, and wherein the control unit is operable in response to a fifth event, to replace the plurality of objects superimposed on the input image with objects belonging to a second category that is different from the first category.

(11)

The information processing apparatus according to any one of (6) to (10), wherein the operation object is a hand of the user, and wherein the first event is recognition by the recognition unit of a specific gesture of the user.

(12)

The information processing apparatus according to any one of (6) to (10), wherein the first event is recognition of a specific voice command issued by the user.

(13)

The information processing apparatus according to (6), wherein the control unit moves the display position of the object designated by the operation object together with the operation object, and wherein the first event is movement of the object to a specific screen region.

(14)

The information processing apparatus according to (13), wherein the control unit sets a number of screen regions on the screen equal to a number of processes associated with the object designated by the operation object.

(15)

The information processing apparatus according to (10), wherein the first category and the second category are associated with different screen regions, and wherein the fifth event is movement of the operation object from a first screen region associated with the first category to a second screen region associated with the second category.

(16)

The information processing apparatus according to any one of (1) to (15), wherein the operation object is a right hand and a left hand of the user, and wherein the control unit is operable, after recognition of one of the right hand and the left hand, to cause a first group out of the plurality of objects to approach toward the recognized one of the right hand and the left hand and is operable, after recognition of another of the right hand and the left hand, to cause a second group out of the plurality of objects to approach toward the recognized other of the right hand and the left hand.

(17)

The information processing apparatus according to any one of (1) to (15), wherein the operation object is a hand of a first user and a hand of a second user, and wherein the control unit is operable, after recognition of the hand of the first user, to cause a first group out of the plurality of objects to approach toward the first user and is operable, after recognition of the hand of the second user, to cause a second group out of the plurality of objects to approach toward the second user.

(18)

The information processing apparatus according to (6), wherein the control unit is operable in response to a sixth event designating at least one object, to cause the application unit to carry out a process associated with the designated object, and wherein the sixth event is recognition of a specific gesture of another user for the designated object.

(19)

An information processing method carried out by an information processing apparatus, the information processing method including:  
acquiring an input image;  
recognizing, in the input image, an operation object used by a user;  
displaying, on a screen, an output image which corresponds to the input image and on which a plurality of objects to be operated by the user are superimposed; and  
causing display positions of the plurality of objects being displayed on the screen before recognition of the operation object to respectively approach toward the user after the recognition of the operation object.

(20)

A program for causing a computer controlling an information processing apparatus to function as:  
an image acquisition unit acquiring an input image;  
a recognition unit recognizing, in the input image, an operation object used by a user; and  
a control unit displaying, on a screen, an output image which corresponds to the input

image and on which a plurality of objects to be operated by the user are superimposed and controlling displaying of at least one of the objects based on a recognition result of the operation object,

wherein the control unit causes display positions of the plurality of objects being displayed on the screen before recognition of the operation object to respectively approach toward the user after the recognition of the operation object.

### **Reference Signs List**

- [0187] 100, 200 Information processing apparatus
- 120 image acquisition unit
- 140 application unit
- 152 image recognition unit
- 154 voice recognition unit
- 172 operation control unit
- 174 priority setting unit

## Claims

[Claim 1] An information processing system comprising:  
processing circuitry configured to  
control a movement of a UI object on a display screen from a pre-  
recognition position toward a post-recognition position in response to  
recognition of an operation object initiated by a user, wherein the post-  
recognition position is spatially related to a displayed position of a pre-  
determined displayed feature, and the predetermined displayed feature  
is an image derived from a camera-captured image.

[Claim 2] The information processing system of claim 1, wherein  
the processing circuitry is configured to vary a mode of approach of the  
UI object in accordance with a parameter related to the UI object.

[Claim 3] The information processing system of claim 2, wherein  
the mode of approach is non-uniform for the displayed object and other  
displayed objects such that respective speeds of approach are different  
for different displayed objects.

[Claim 4] The information processing system of claim 2, wherein  
the parameter is a priority.

[Claim 5] The information processing system of claim 4, wherein the priority is  
based on an operation frequency or a recommendation.

[Claim 6] The information processing system of claim 2, wherein  
the mode of approach is non-uniform for the displayed object and other  
displayed objects such that respective post-recognition displayed  
positions are different for different displayed objects.

[Claim 7] The information processing system of claim 2, wherein  
a trigger for the movement for the displayed object and another  
displayed object are different.

[Claim 8] The information processing system of claim 7, wherein  
a first detected gesture triggers a movement of the displayed object and  
a second detected gesture triggers a movement of the other displayed  
object.

[Claim 9] The information processing system of claim 7, wherein  
the displayed object and the other displayed object are displayed in a  
ring around the operation object.

[Claim 10] The information processing system of claim 1, wherein the processing  
circuitry is configured to control a movement of a plurality of UI  
objects.

[Claim 11] The information processing system of claim 1, wherein the post-recognition position for the UI object is different for the UI object than for a different UI object.

[Claim 12] The information processing system of claim 11, wherein the post-recognition position is closer to the operation object when the UI object is identified as a higher priority than the different UI object, and further from the operation object when the UI object is identified as a lower priority than the different UI object.

[Claim 13] The information processing system of claim 1, wherein the predetermined displayed feature is a body part of the user.

[Claim 14] The information processing system of claim 1, wherein the predetermined displayed feature is the operation object.

[Claim 15] The information processing system of claim 1, wherein the predetermined displayed feature is a feature of a user image.

[Claim 16] The information processing system of claim 1, wherein the predetermined displayed feature is a feature of an action of a user.

[Claim 17] The information processing system of claim 16, wherein the processing circuitry is also configured to implement an image recognition unit that recognizes a feature of the user as the operation object.

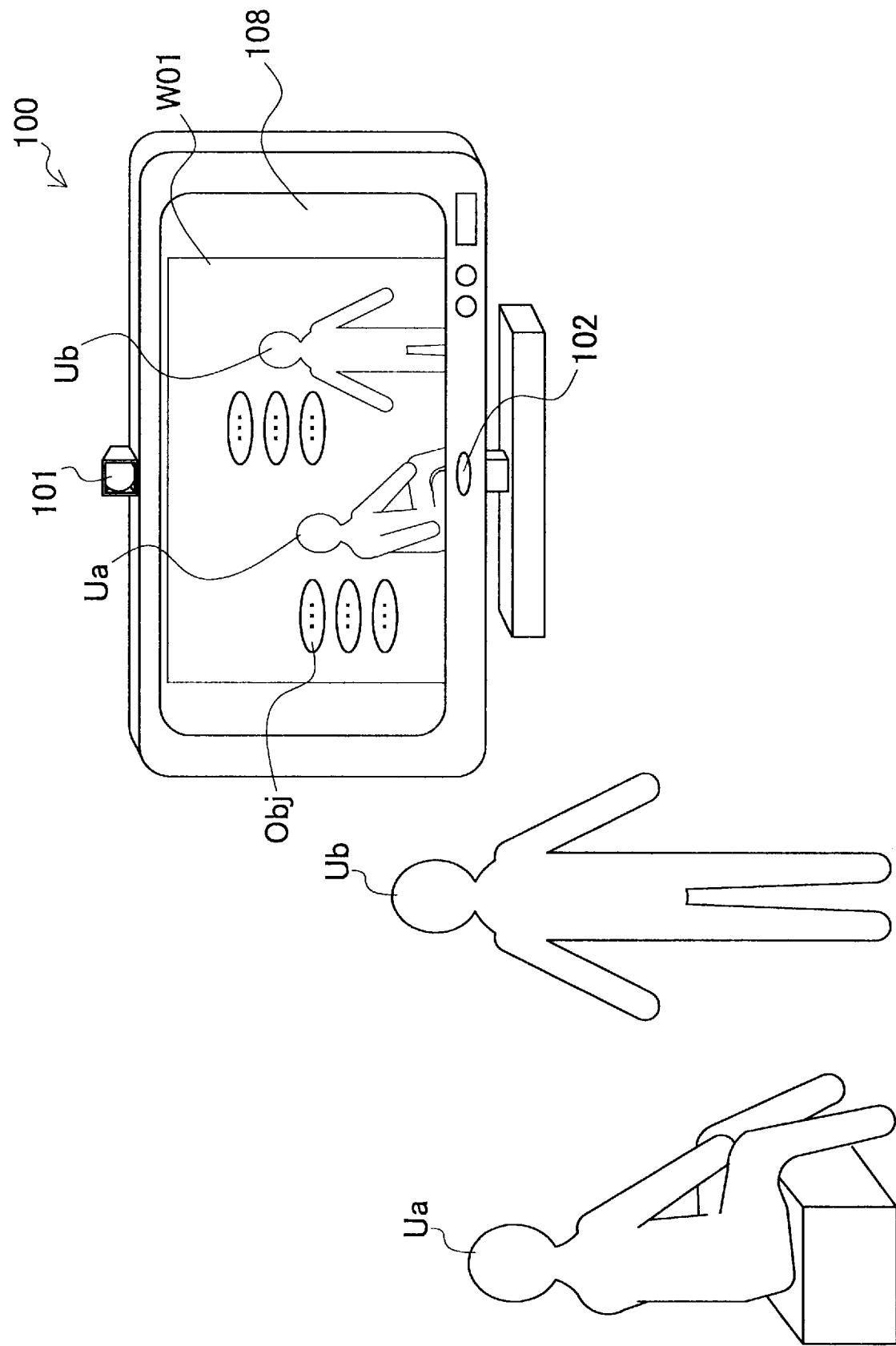
[Claim 18] The information processing system of claim 1, wherein the post-recognition position of the displayed object to the operation object is a closer distance than the pre-recognition position such that the displayed object moves toward the operation object.

[Claim 19] An information processing method comprising:  
controlling with processing circuitry a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, and the predetermined displayed feature is an image derived from a camera-captured image.

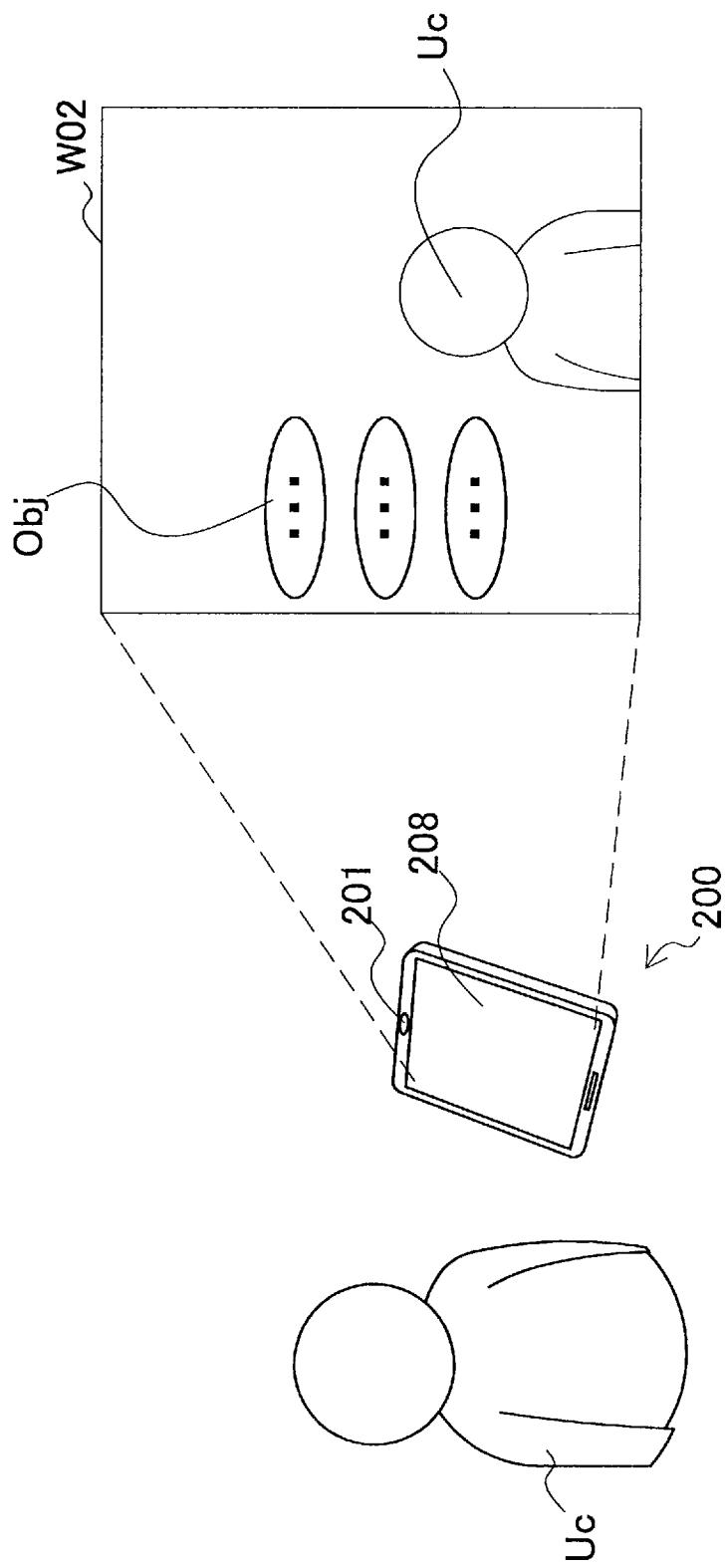
[Claim 20] A non-transitory computer readable medium having computer readable instructions that when executed by a processing circuitry perform a method, the method comprising:  
controlling with the processing circuitry a movement of a UI object on a display screen from a pre-recognition position toward a post-recognition position in response to recognition of an operation object

initiated by a user, wherein the post-recognition position is spatially related to a displayed position of a predetermined displayed feature, the predetermined displayed feature is an image derived from a camera-captured image.

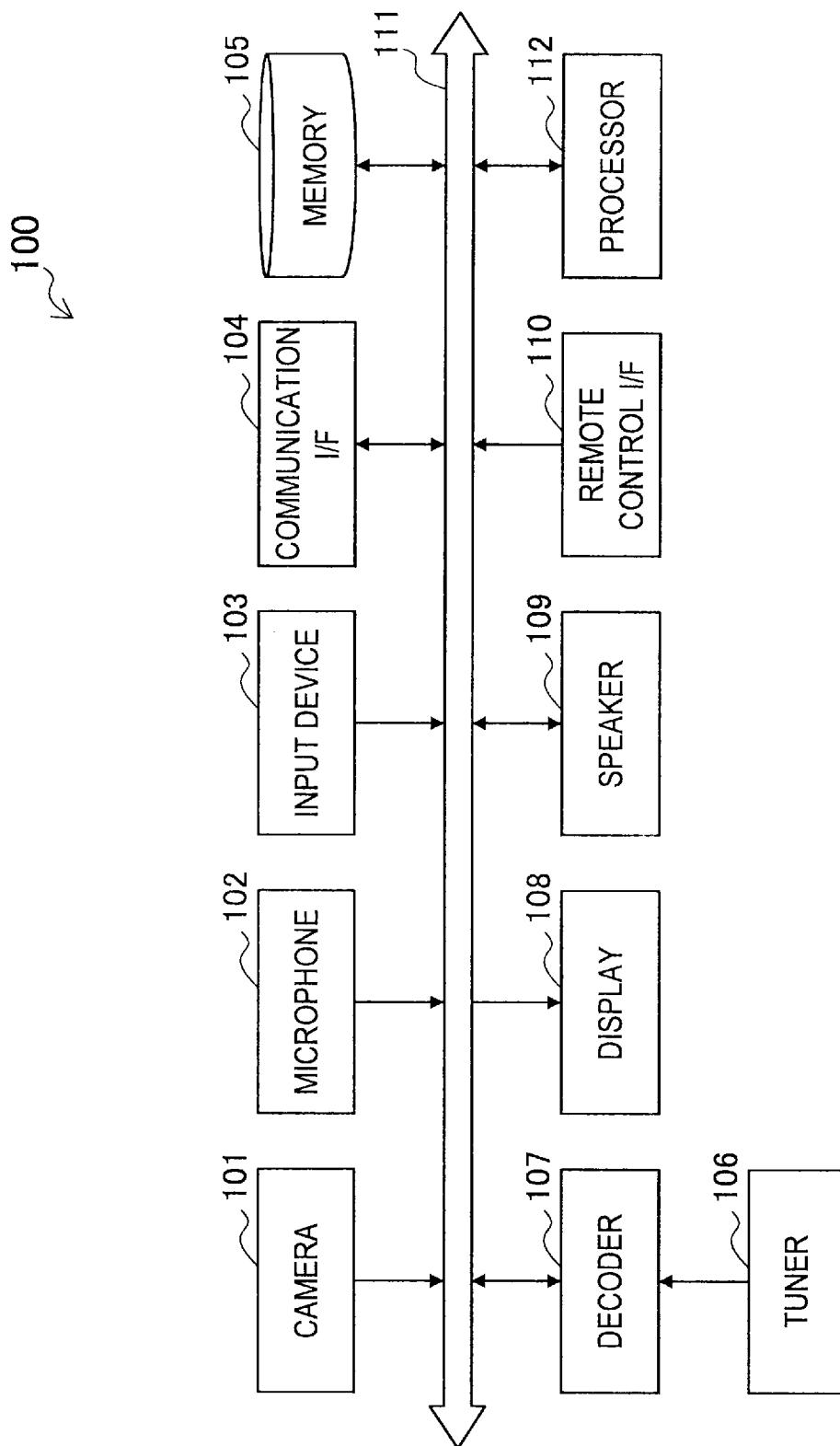
[Fig. 1]



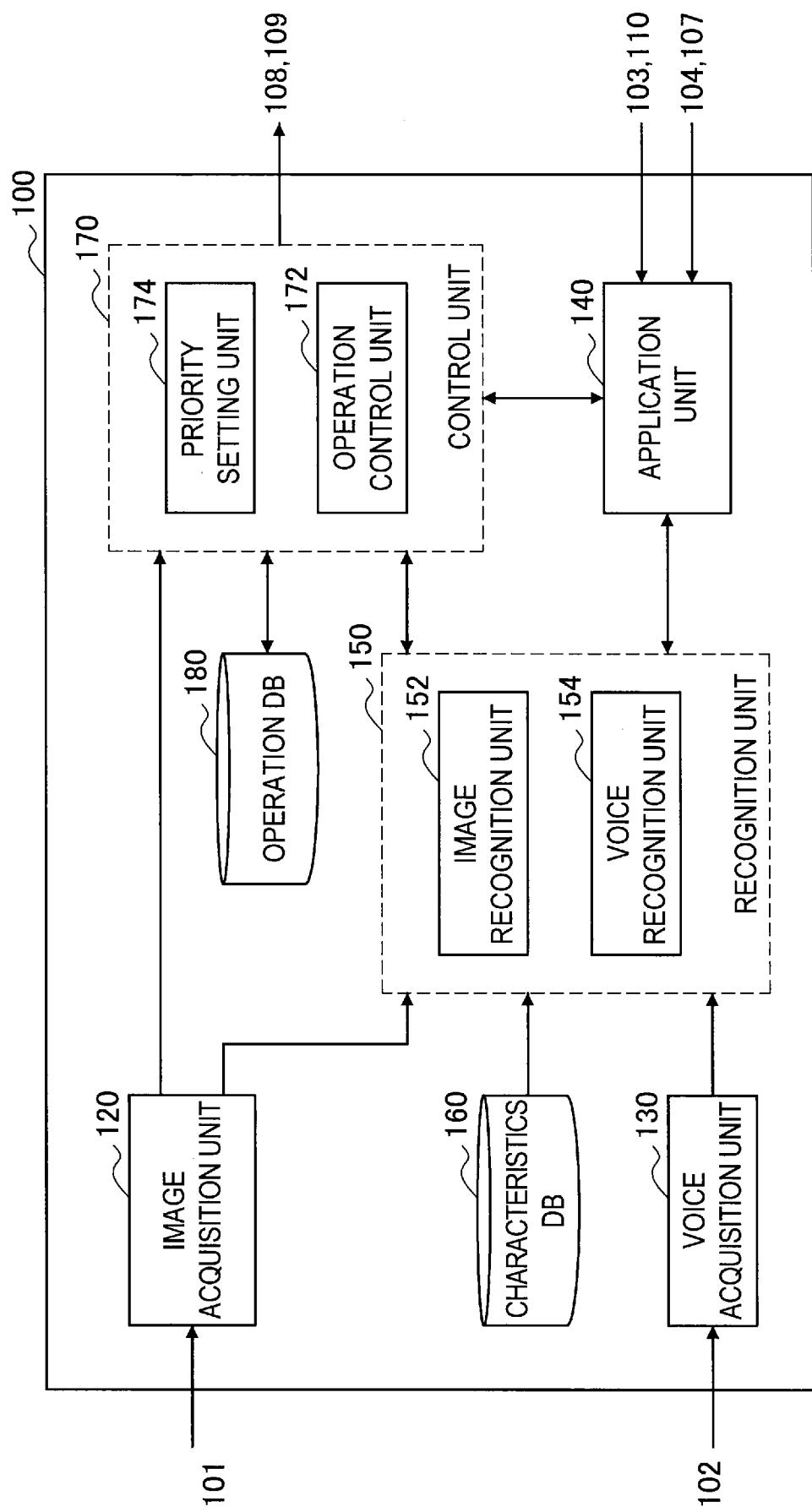
[Fig. 2]



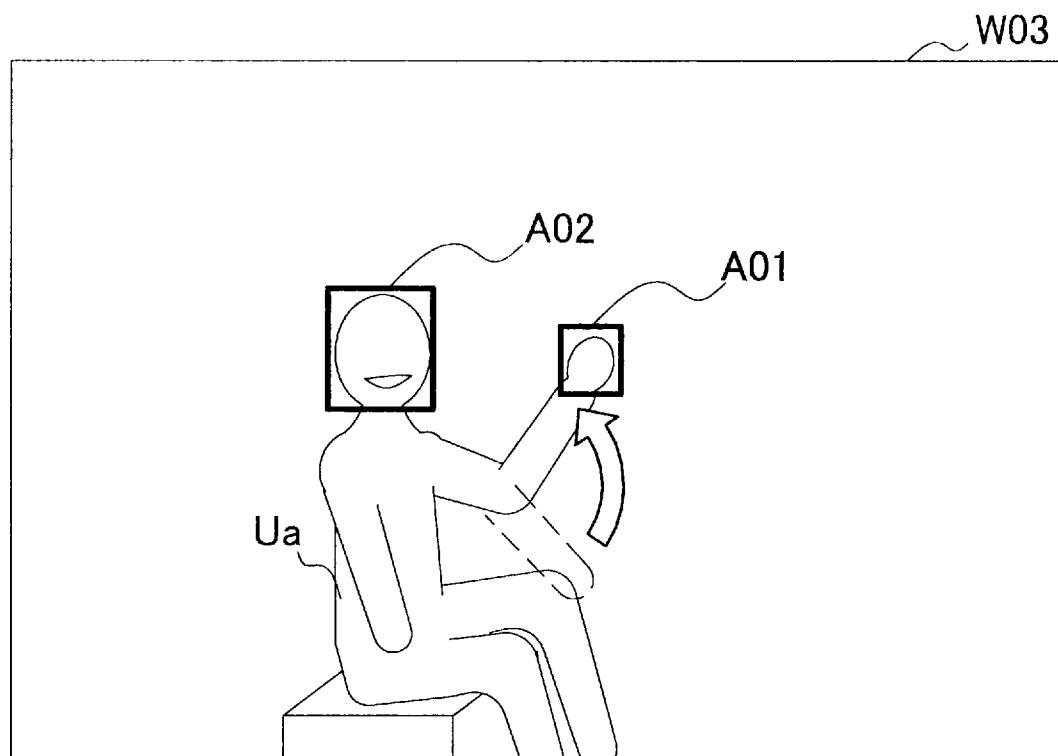
[Fig. 3]



[Fig. 4]

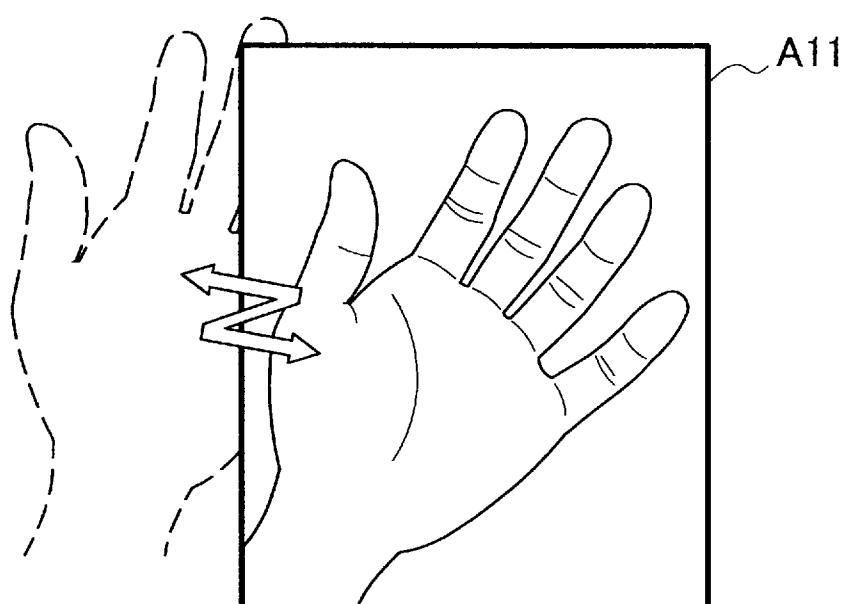


[Fig. 5]



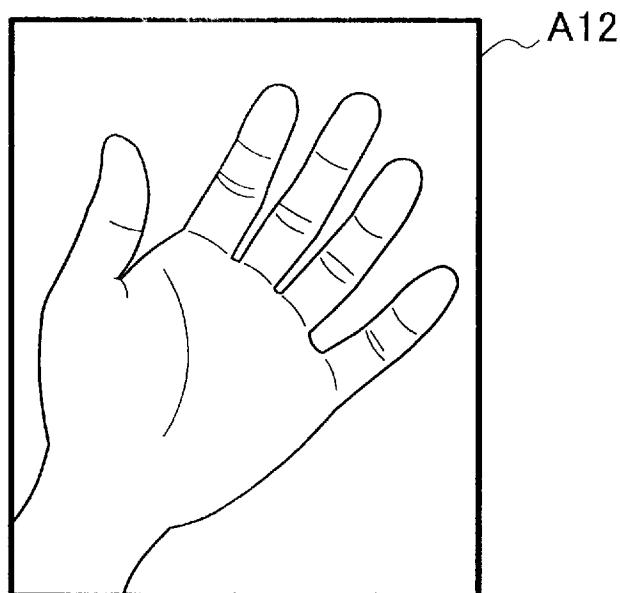
G0 : Hand Up

[Fig. 6]



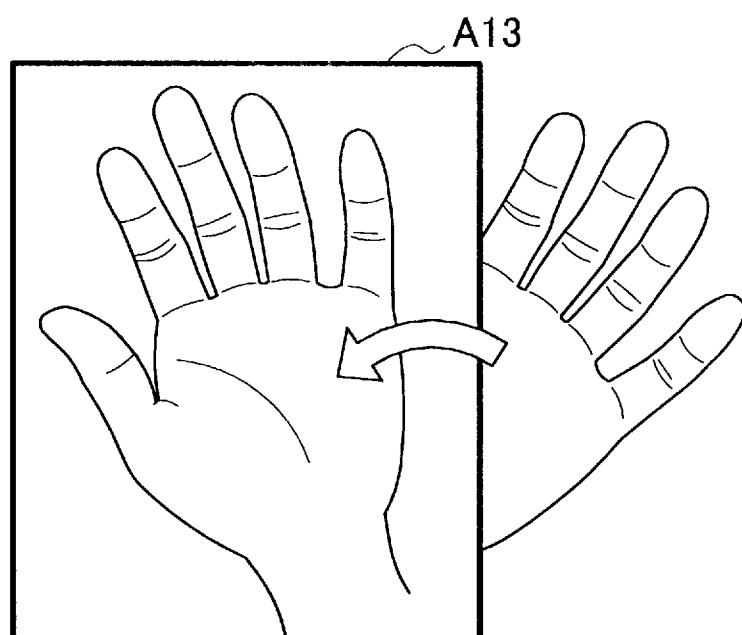
G1 : Waving

[Fig. 7]



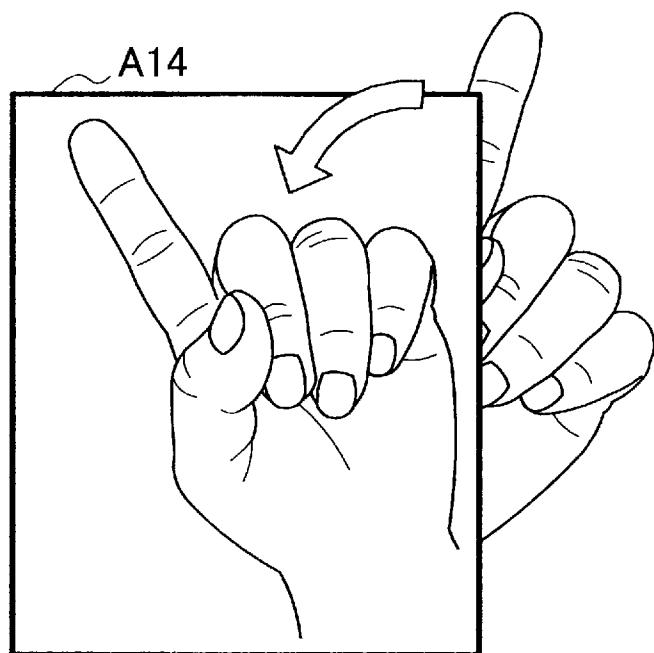
G2: Keep Still

[Fig. 8]



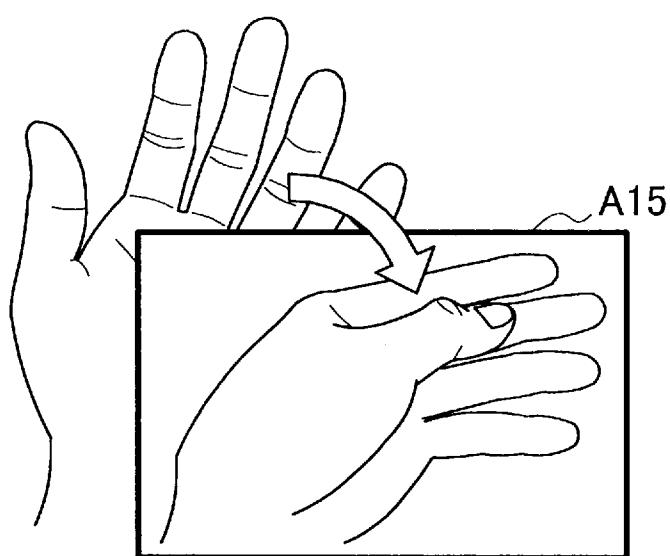
G3a: Rotation

[Fig. 9]



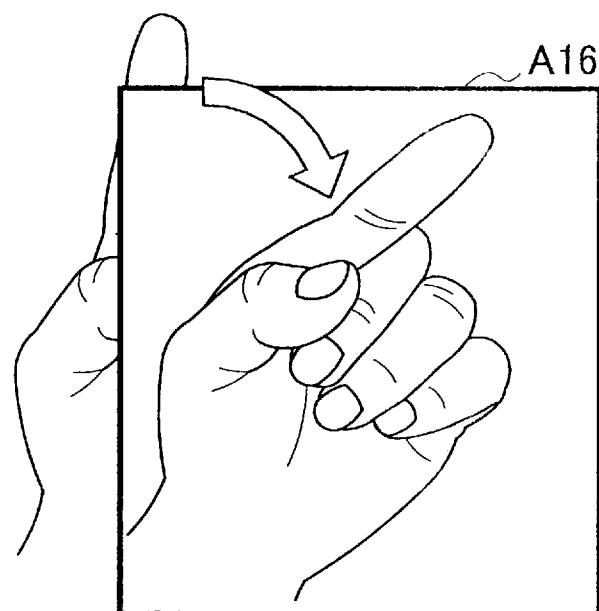
G3b : Rotation

[Fig. 10]



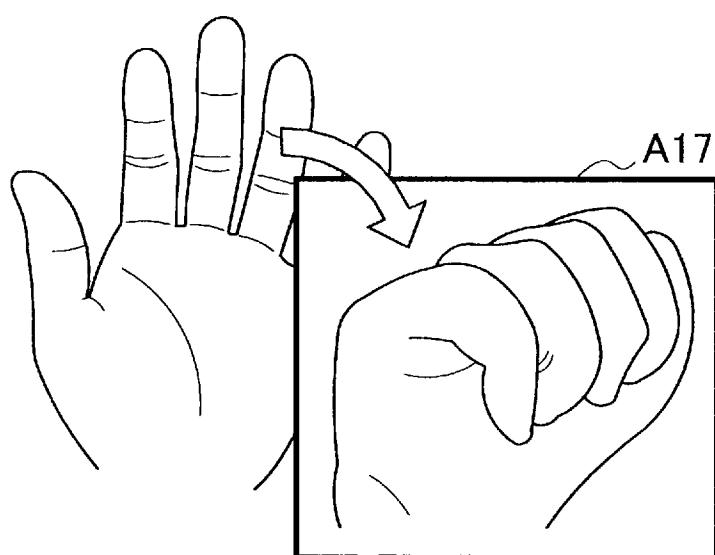
G4a : Touch

[Fig. 11]



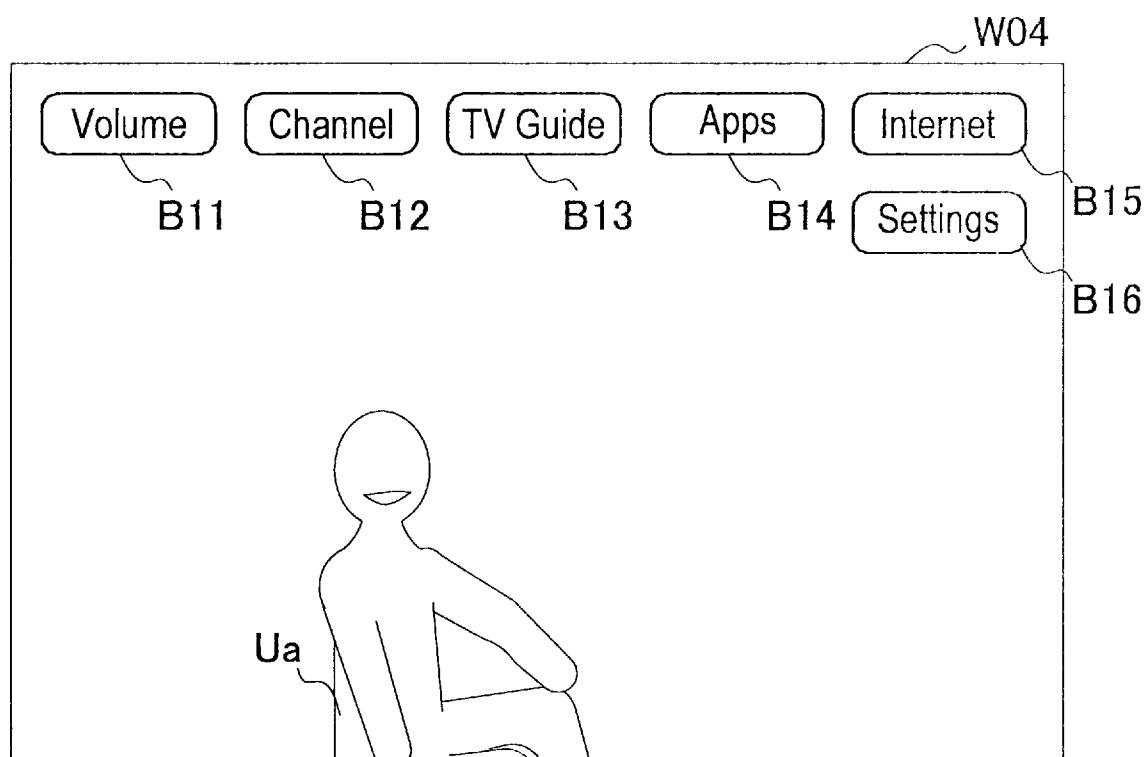
G4b : Touch

[Fig. 12]

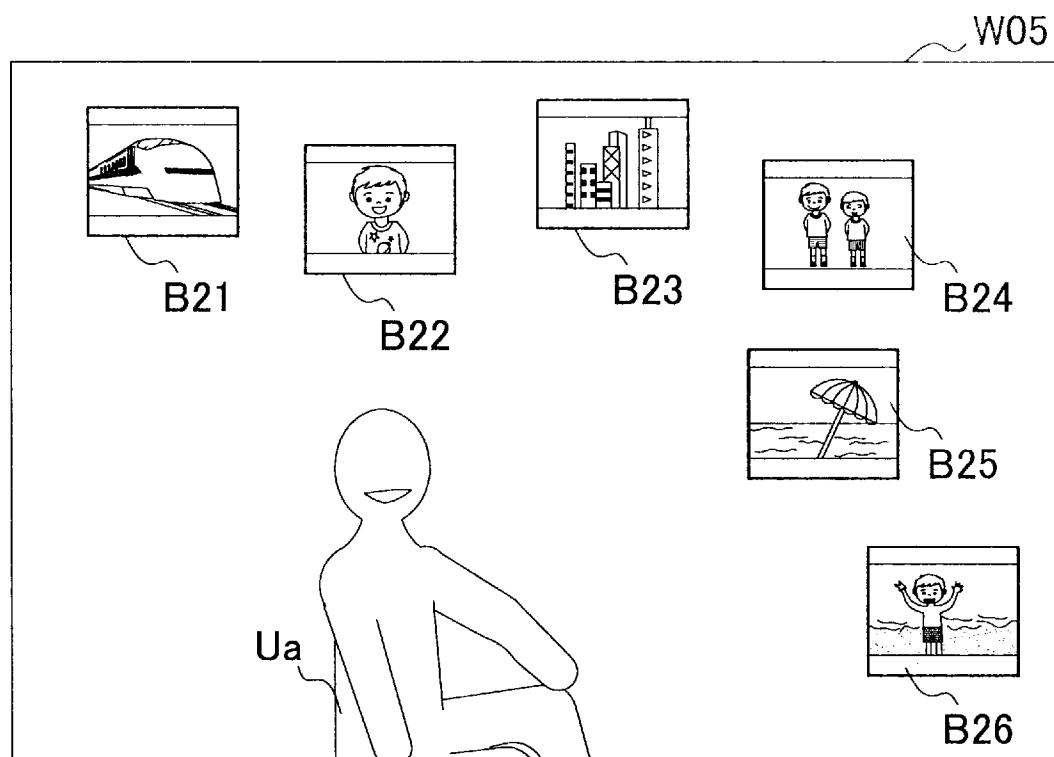


G5 : Grasp

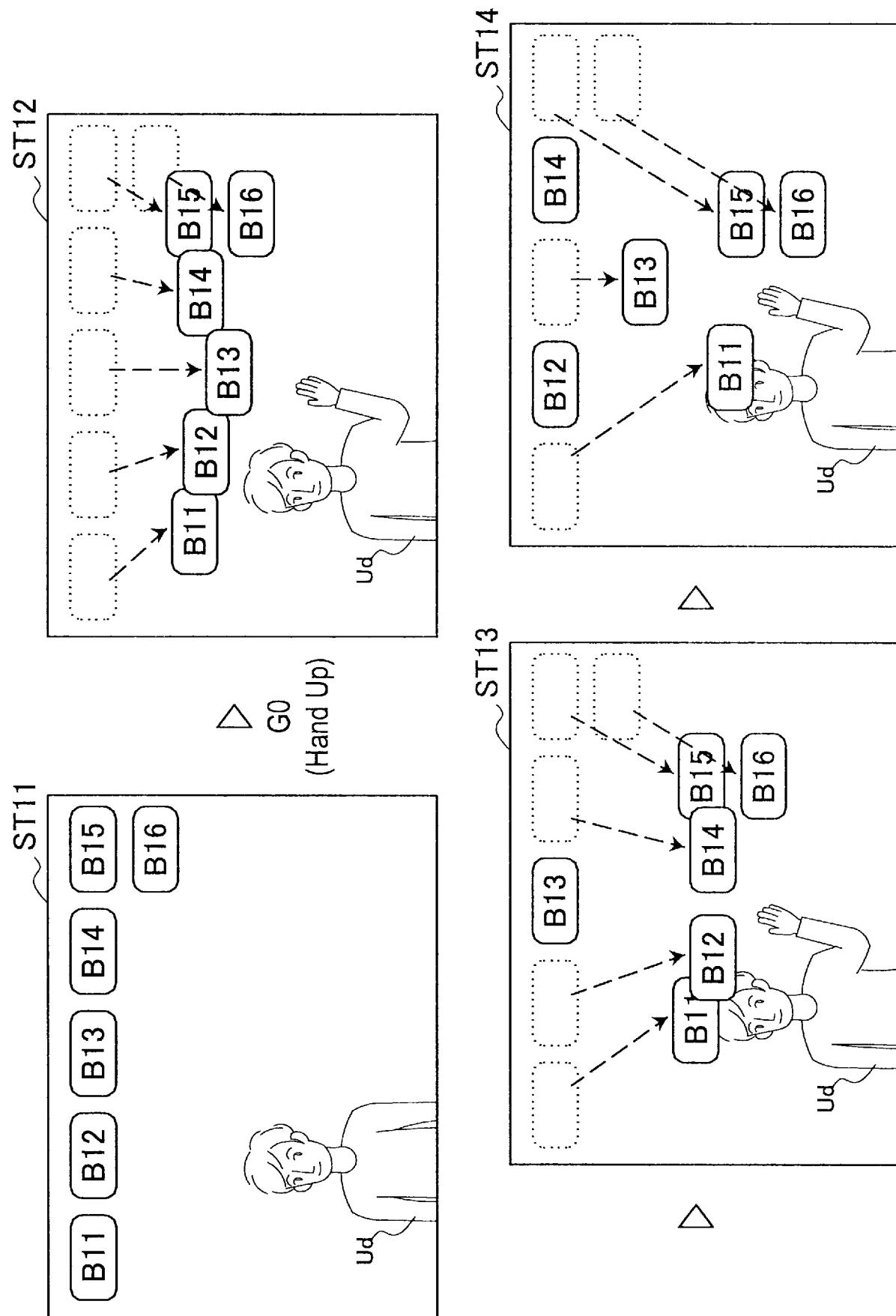
[Fig. 13]



[Fig. 14]

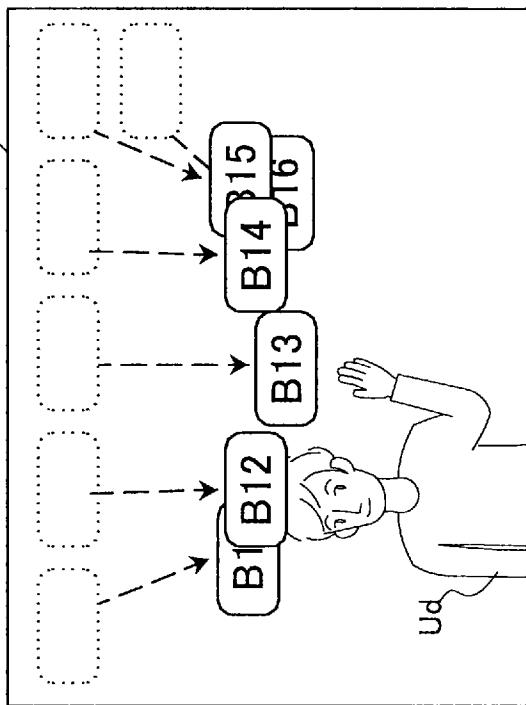


[Fig. 15]



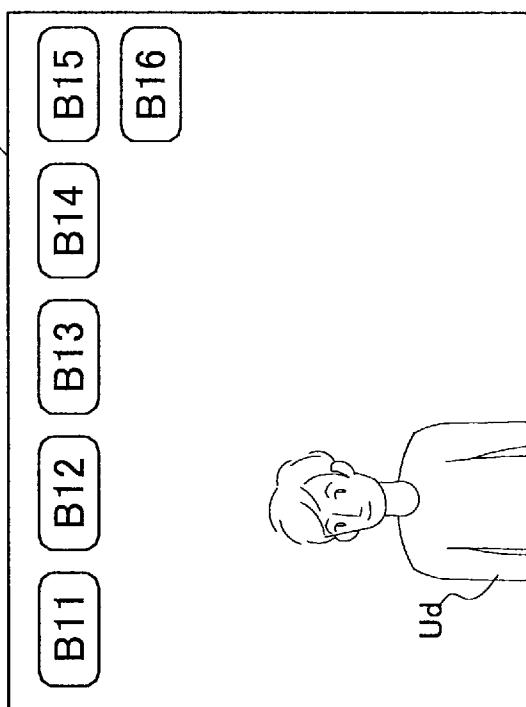
[Fig. 16]

ST22

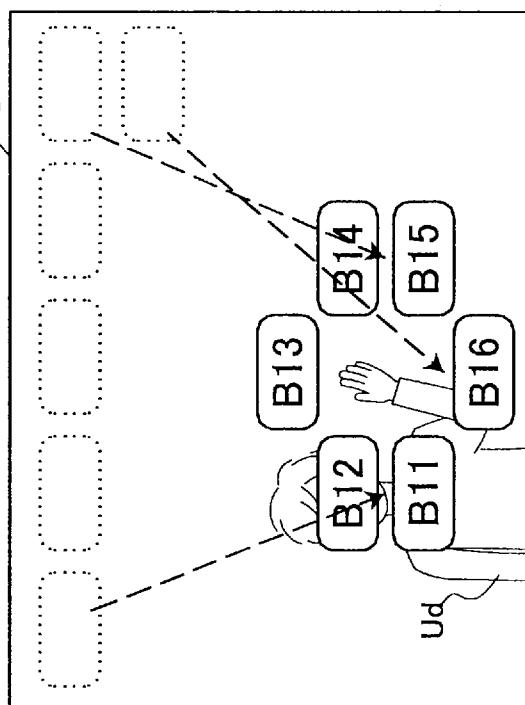


▷  
G0  
(Hand Up)

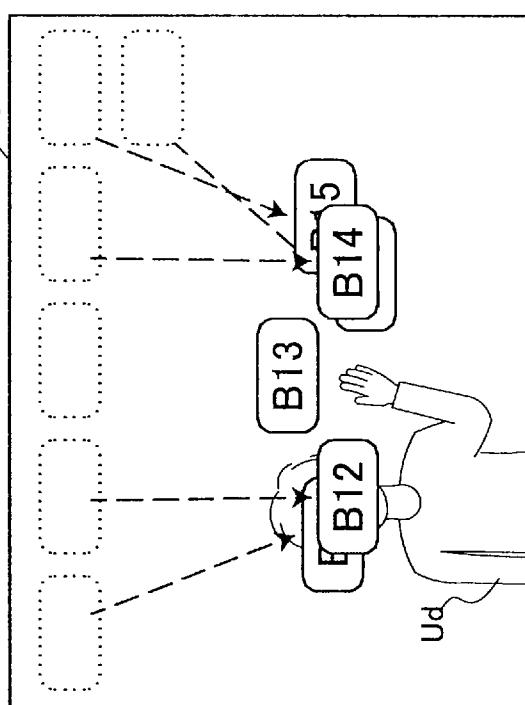
ST21



ST24

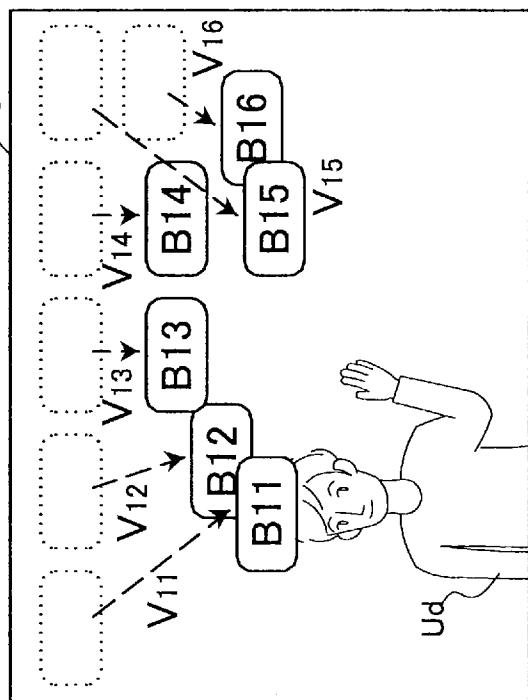


ST23

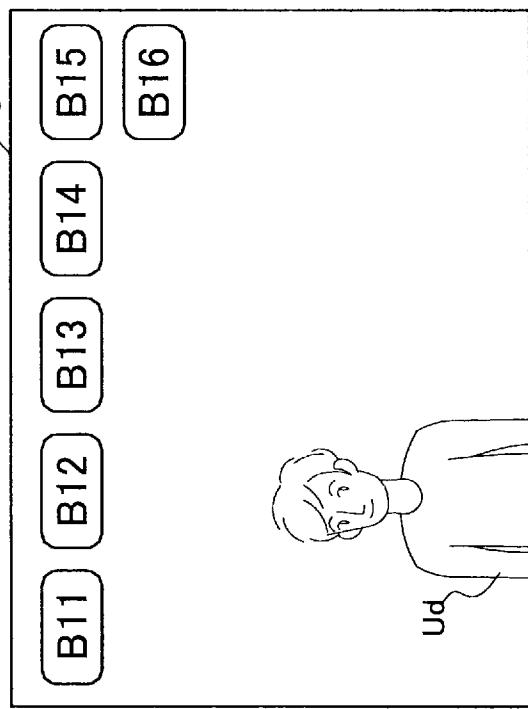


[Fig. 17]

ST32



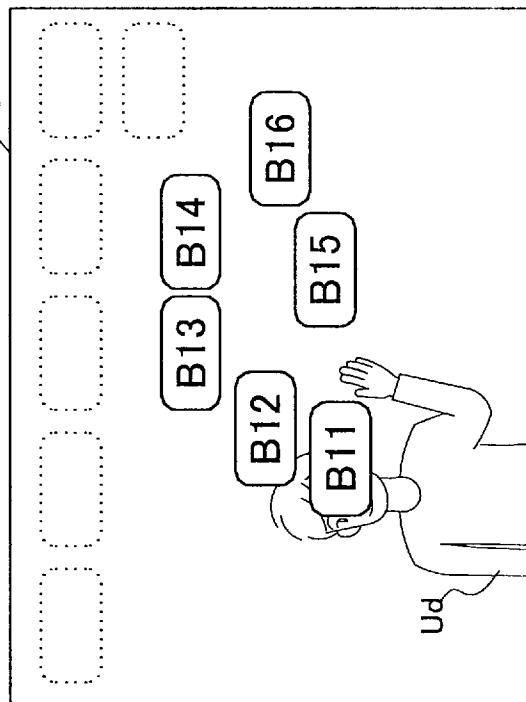
ST31



APPROACH SPEED

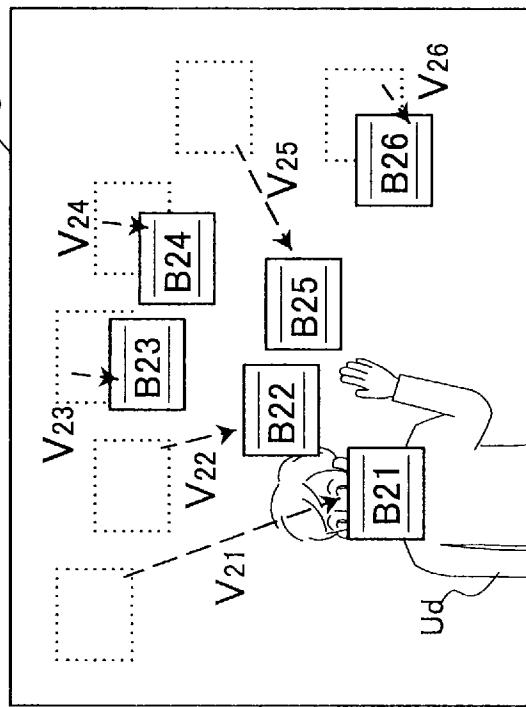
$$V_{11} = V_{15} > V_{12} > V_{13} = V_{14} = V_{16}$$

ST33



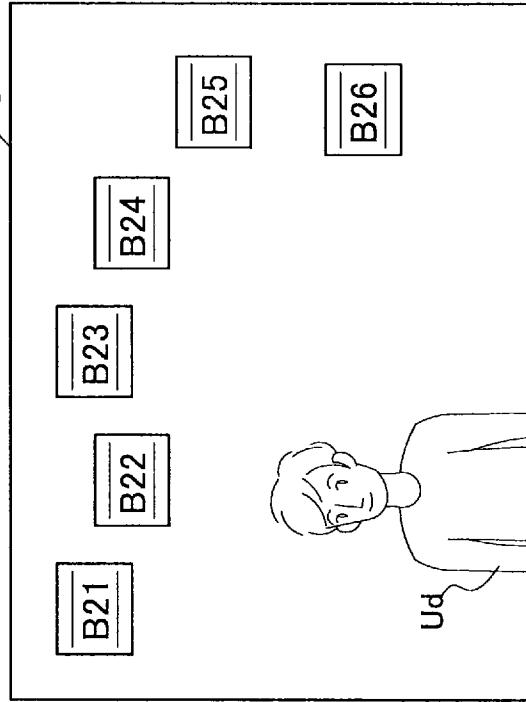
[Fig. 18]

ST42



▷  
GO  
(Hand Up)

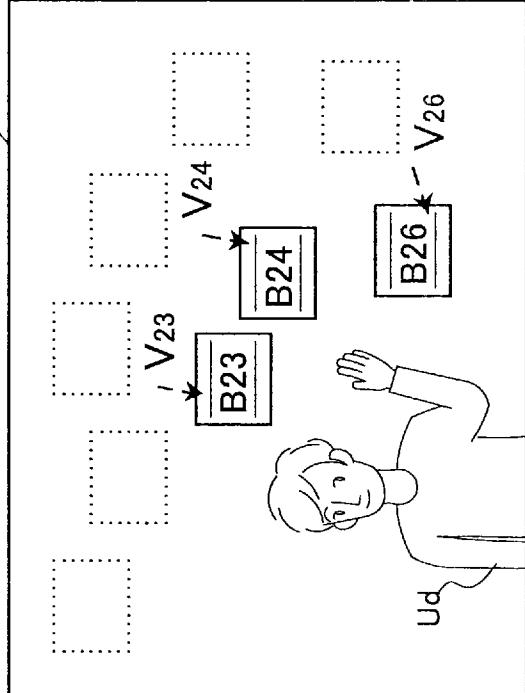
ST41



APPROACH SPEED

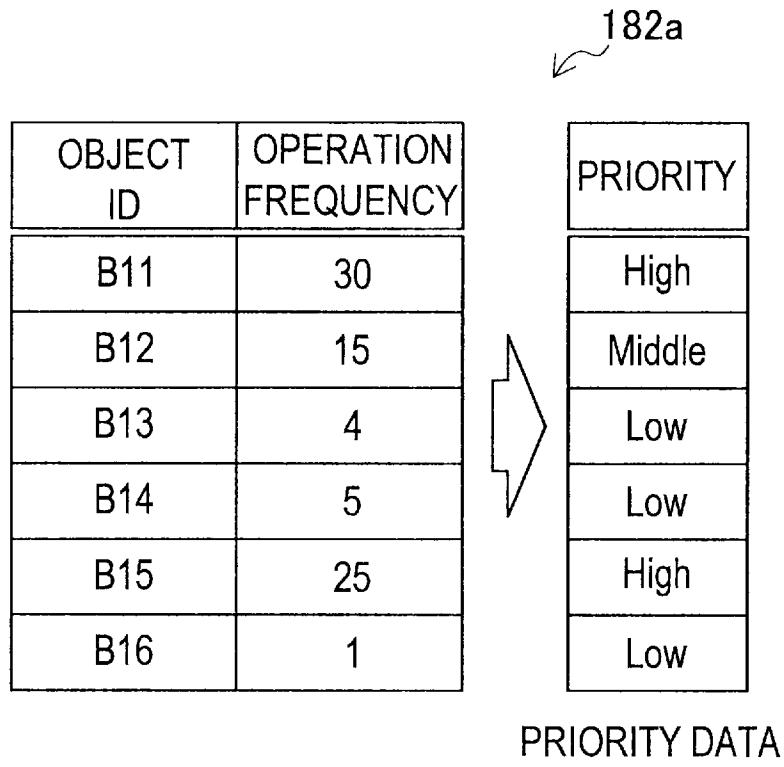
$V_{21} > V_{22} = V_{25} > V_{23} = V_{24} = V_{26}$

ST43

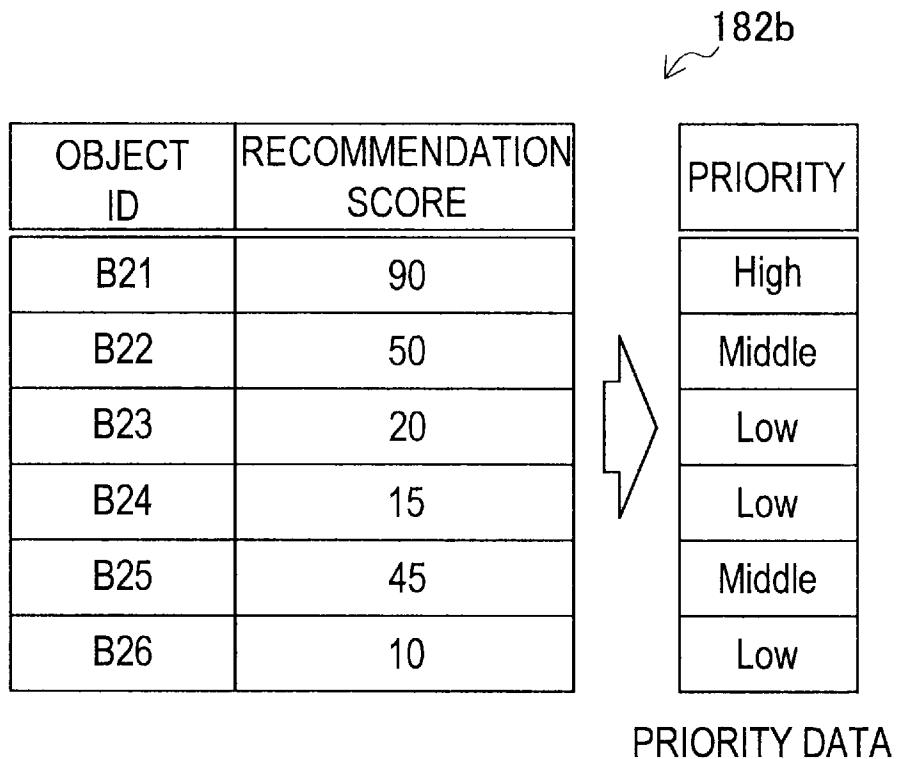


▷

[Fig. 19]

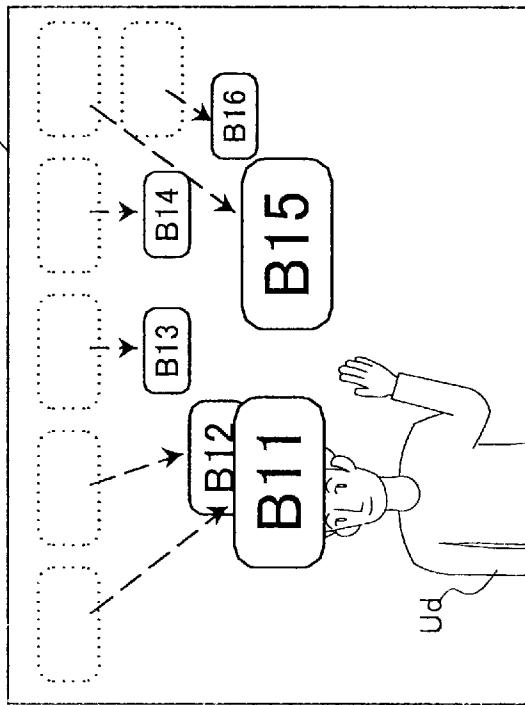


[Fig. 20]

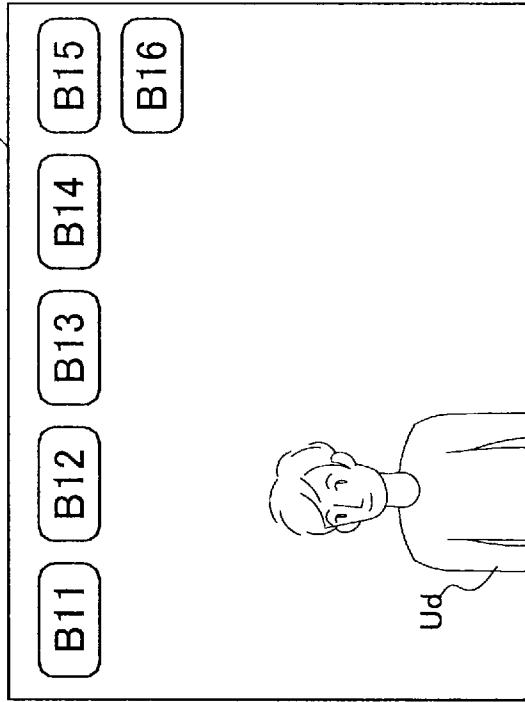


[Fig. 21]

ST52

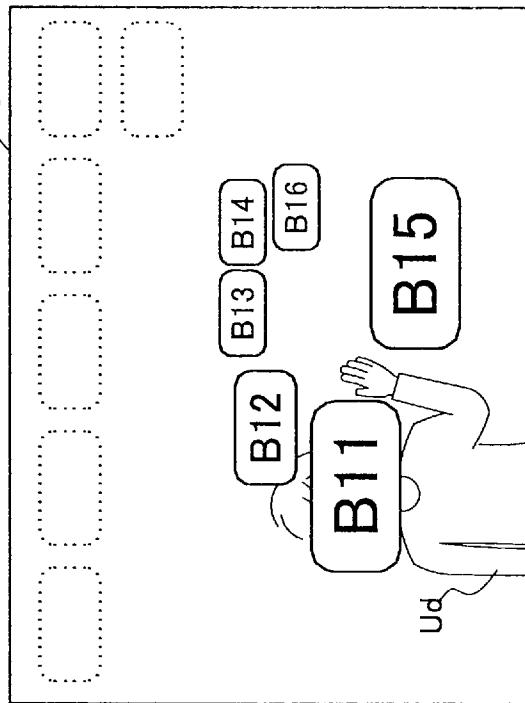


ST51

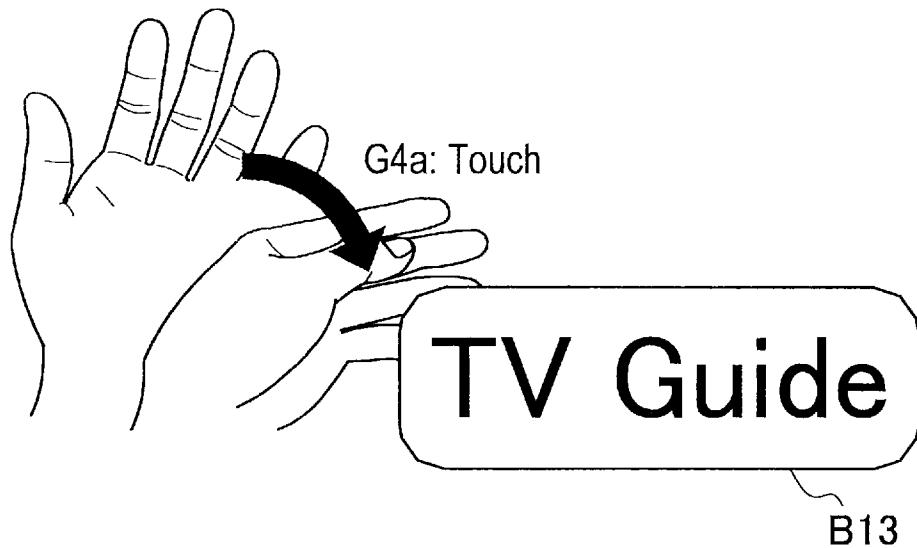


DISPLAY SIZE  
 $B_{11} = B_{15} > B_{12} > B_{13} = B_{14} = B_{16}$

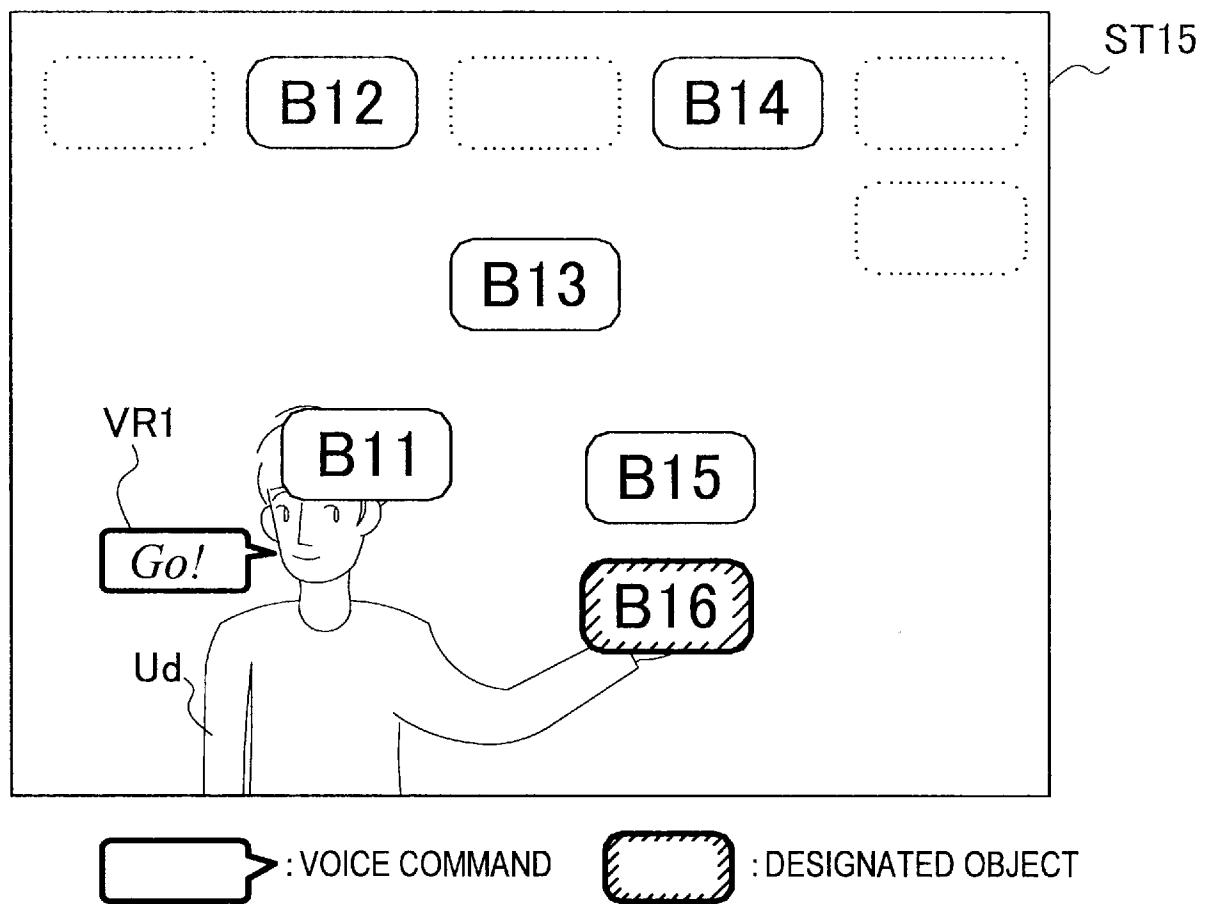
ST53



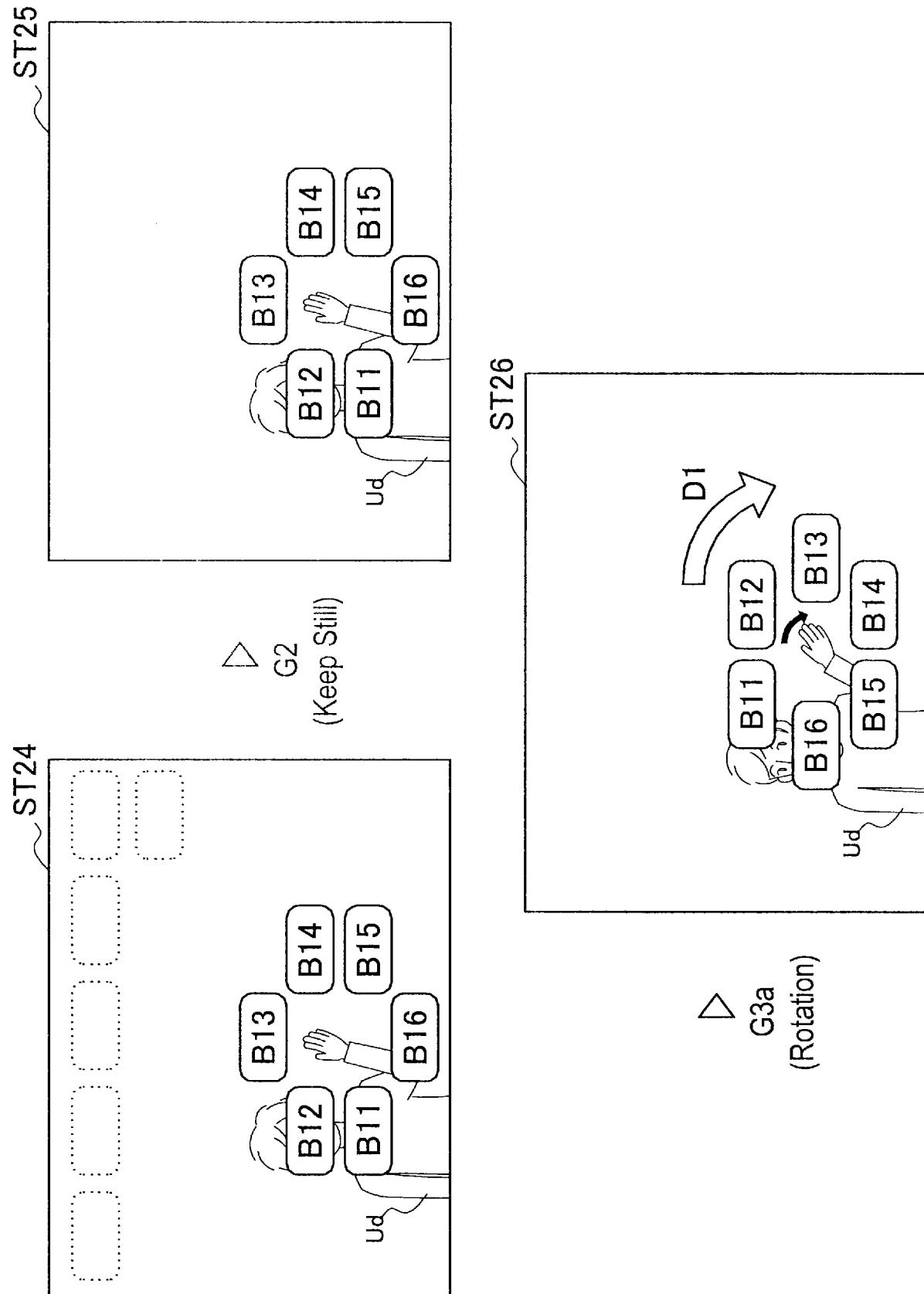
[Fig. 22]



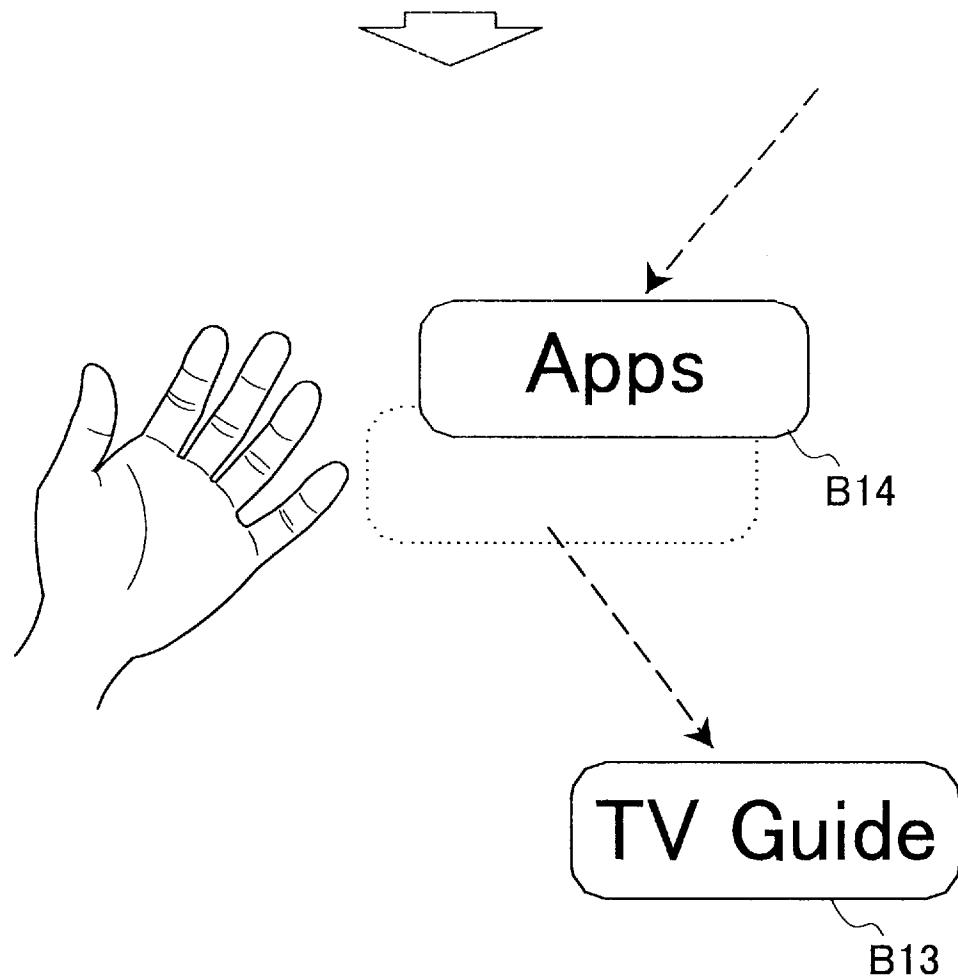
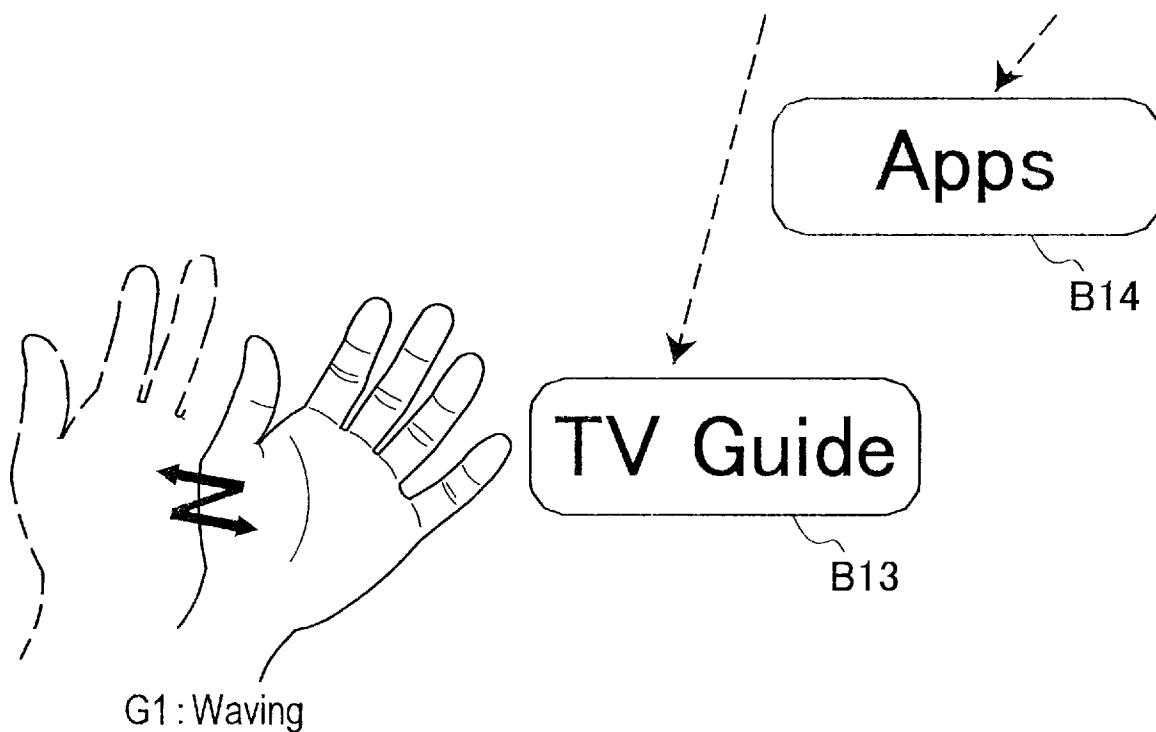
[Fig. 23]



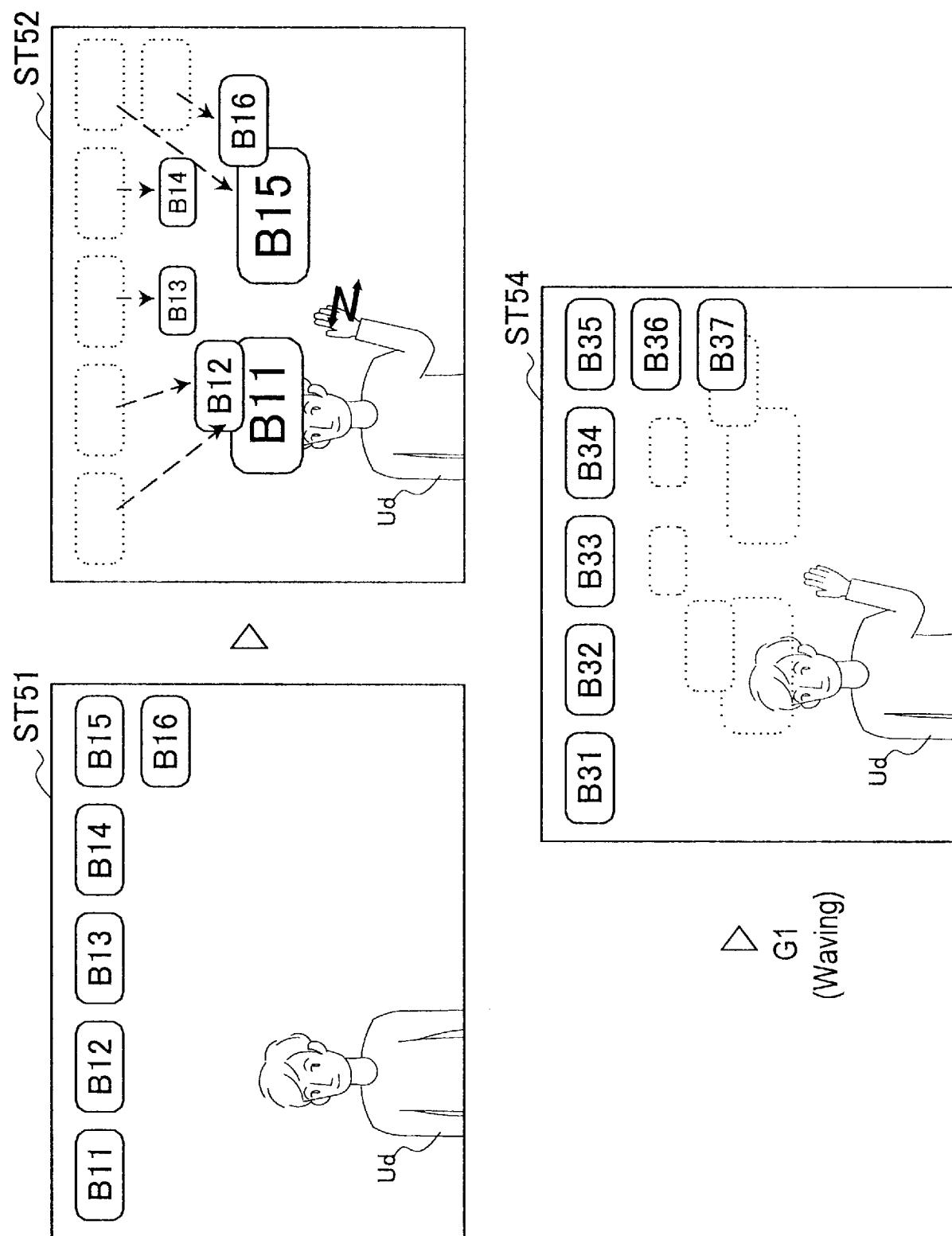
[Fig. 24]



[Fig. 25]

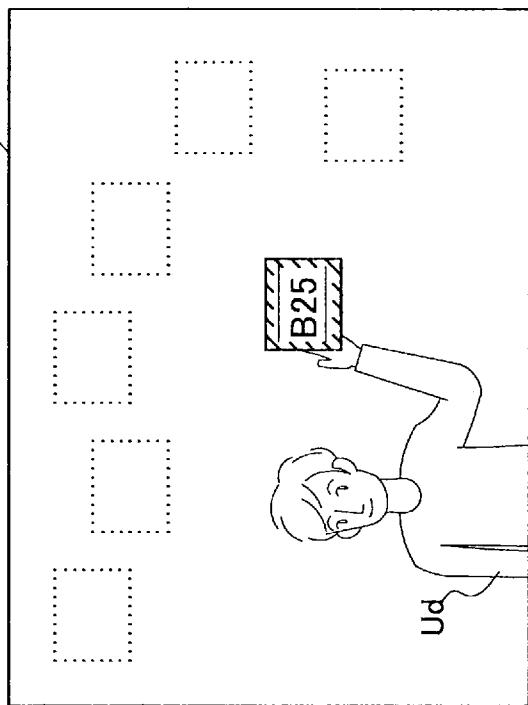


[Fig. 26]

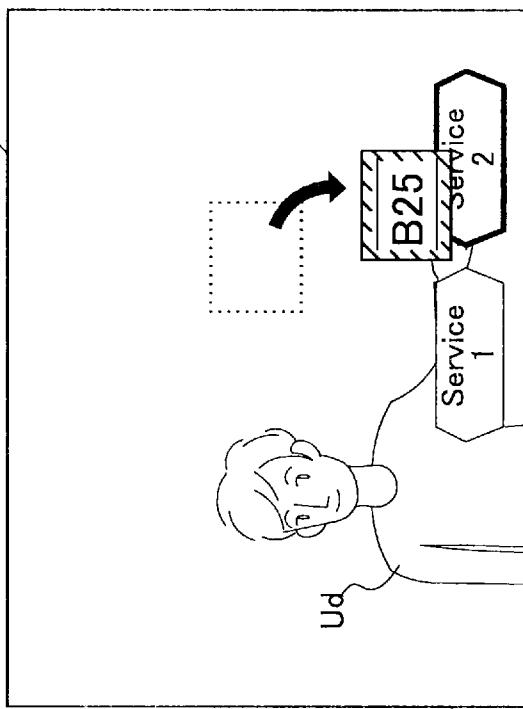


[Fig. 27]

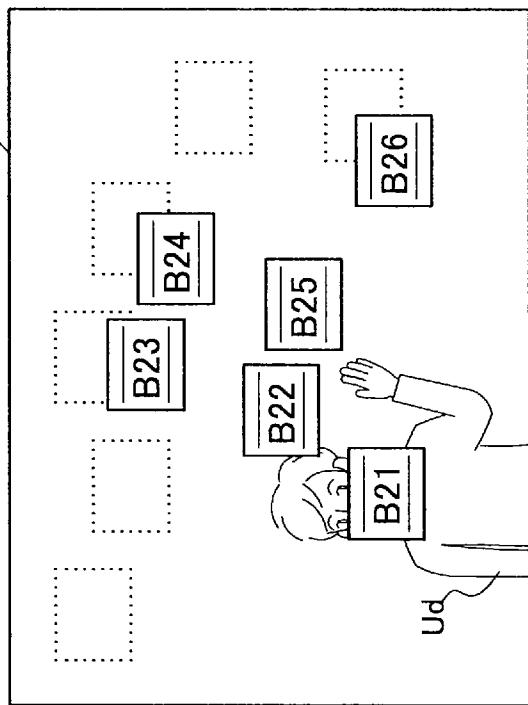
ST44



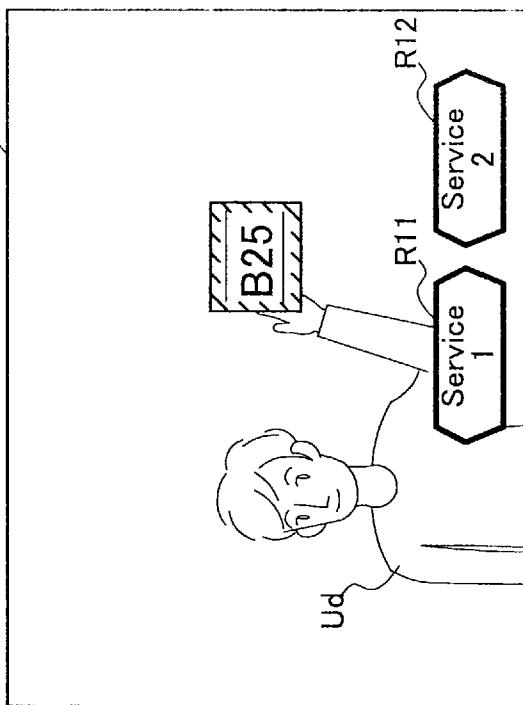
ST46



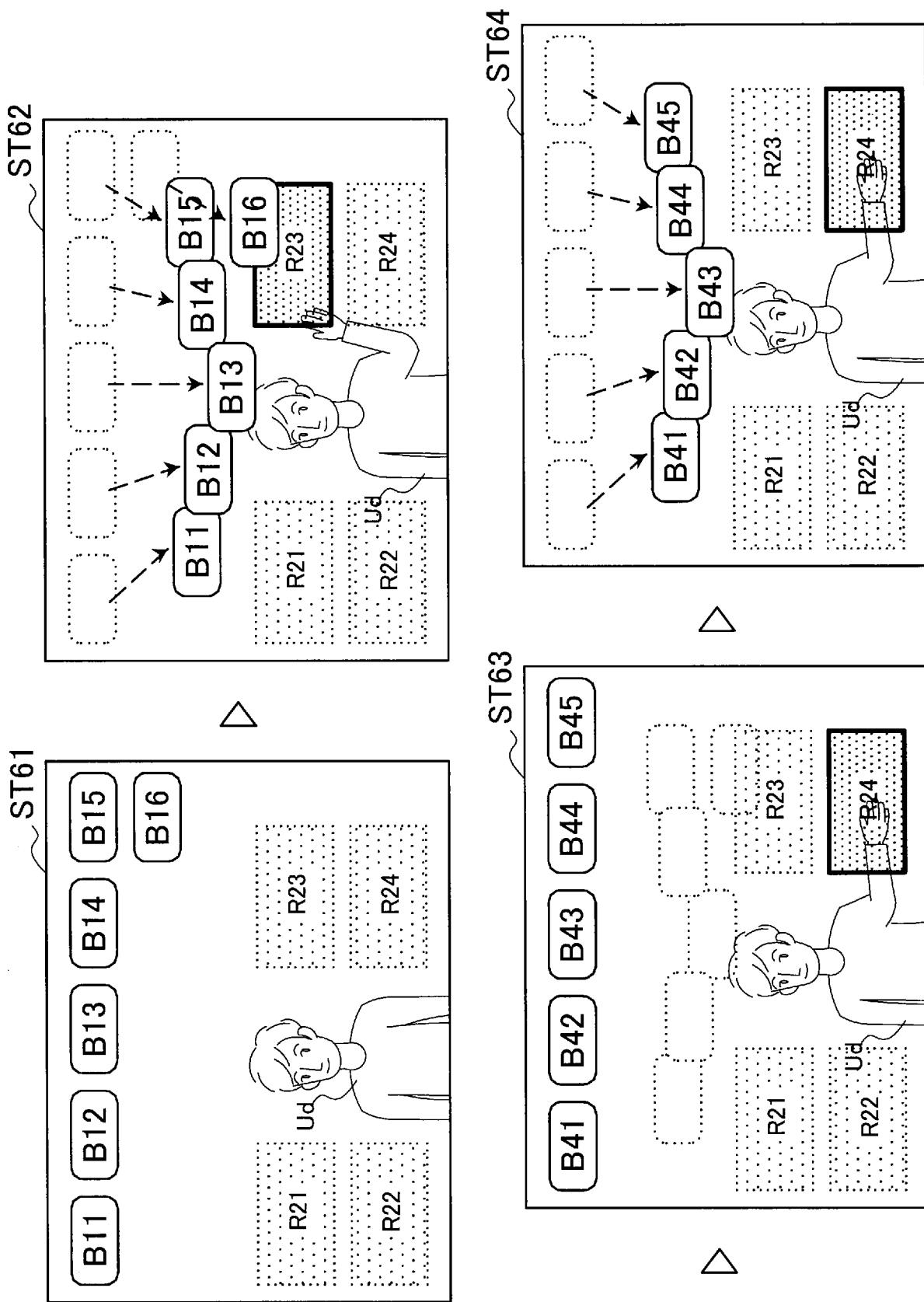
ST42



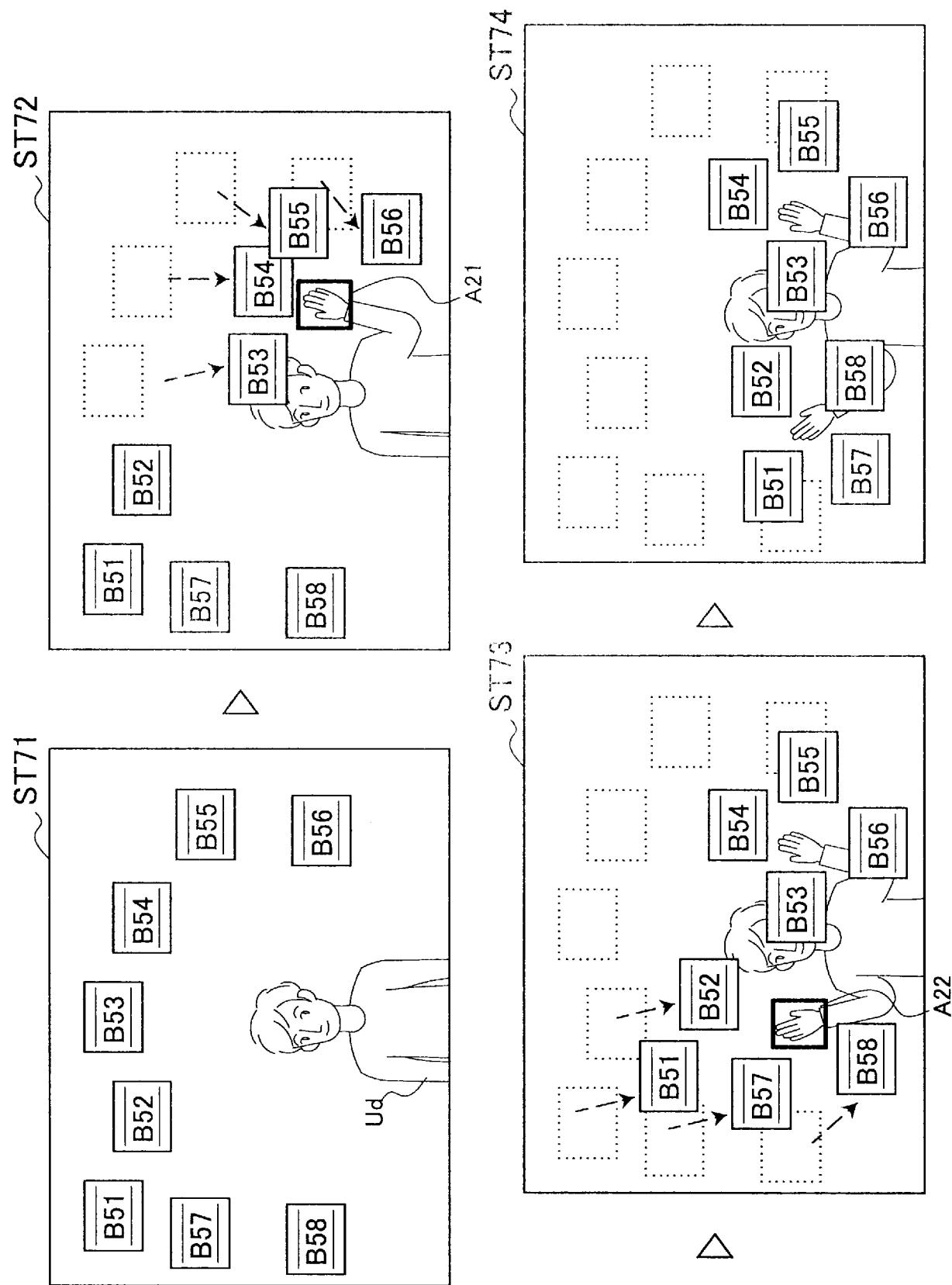
ST45



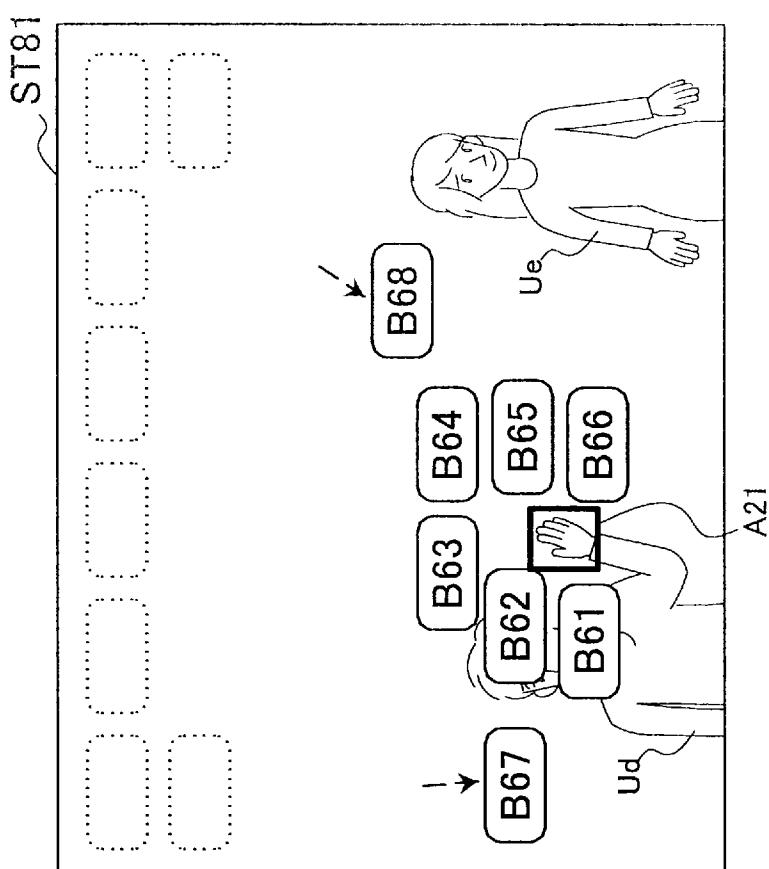
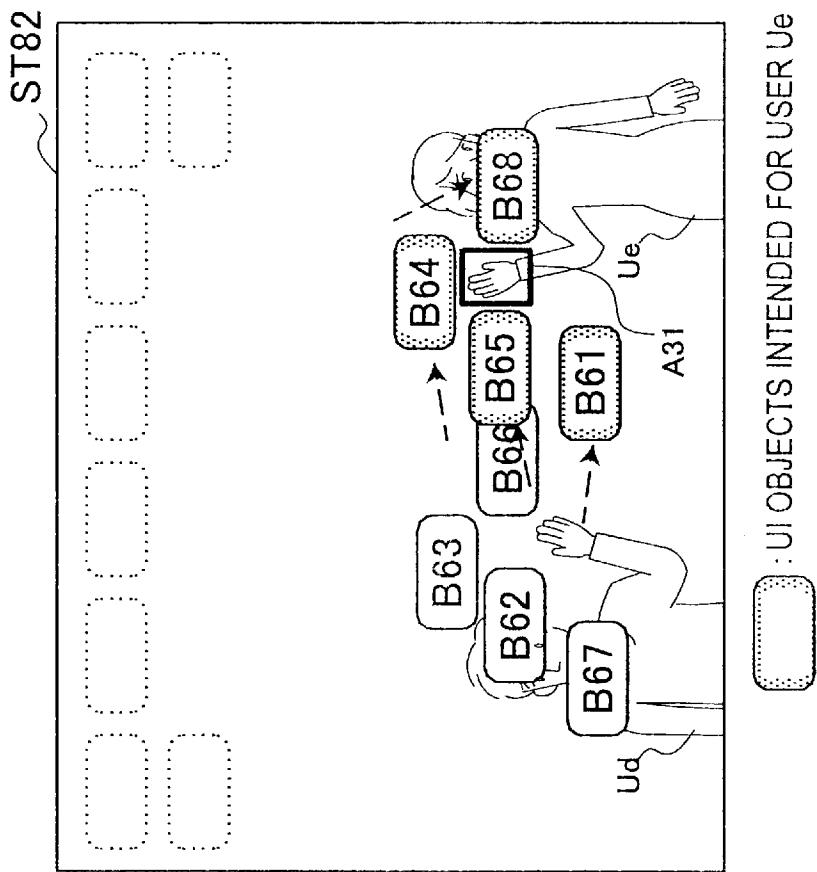
[Fig. 28]



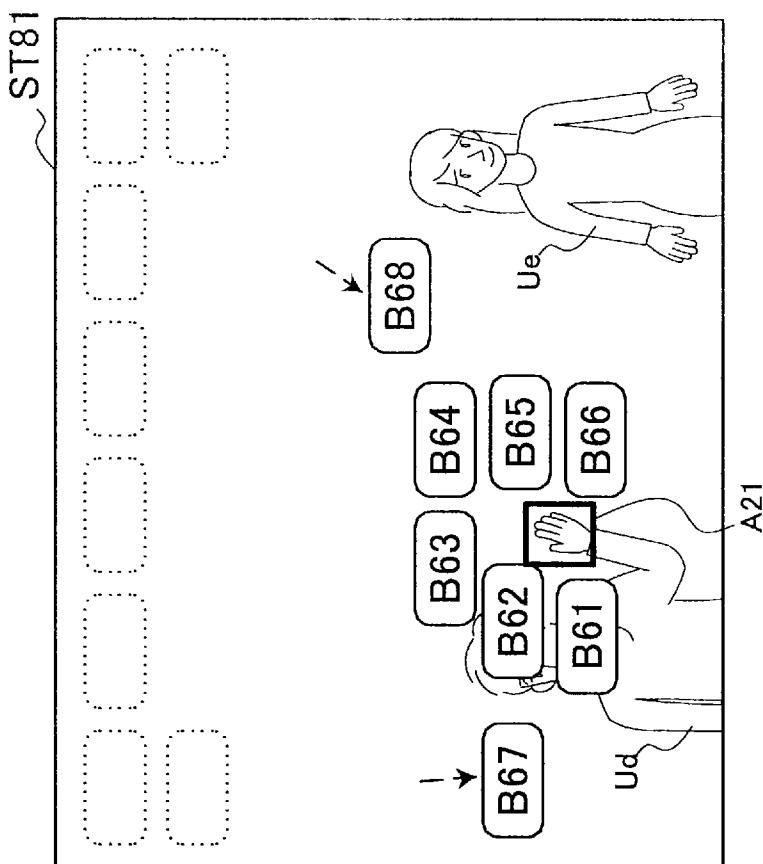
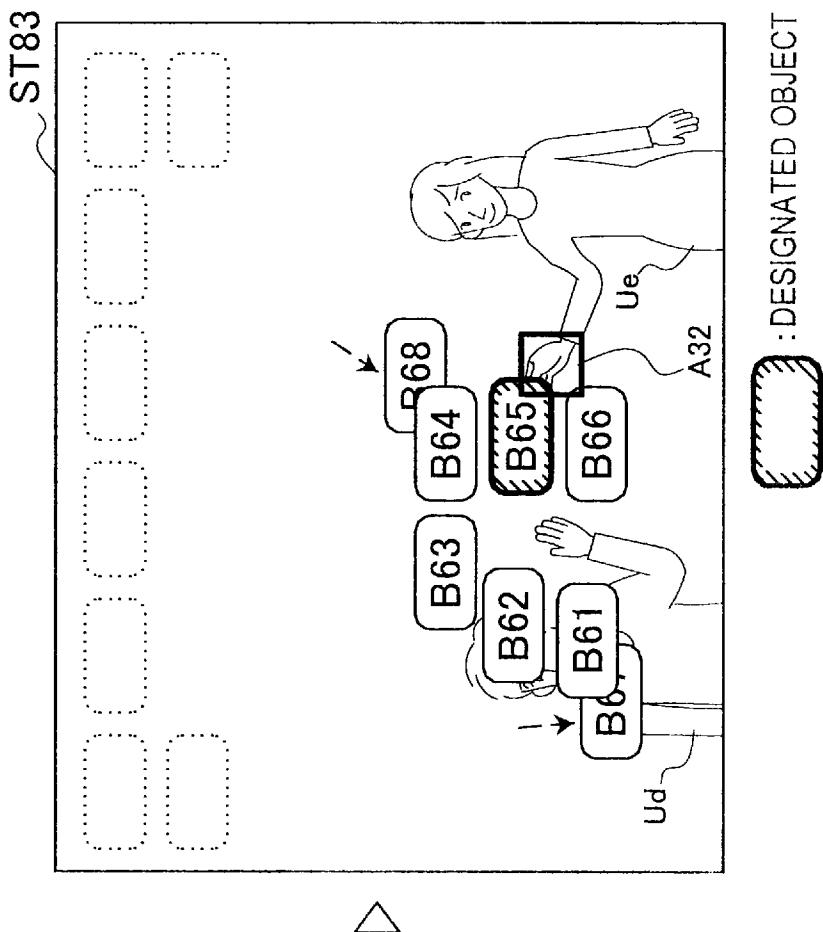
[Fig. 29]



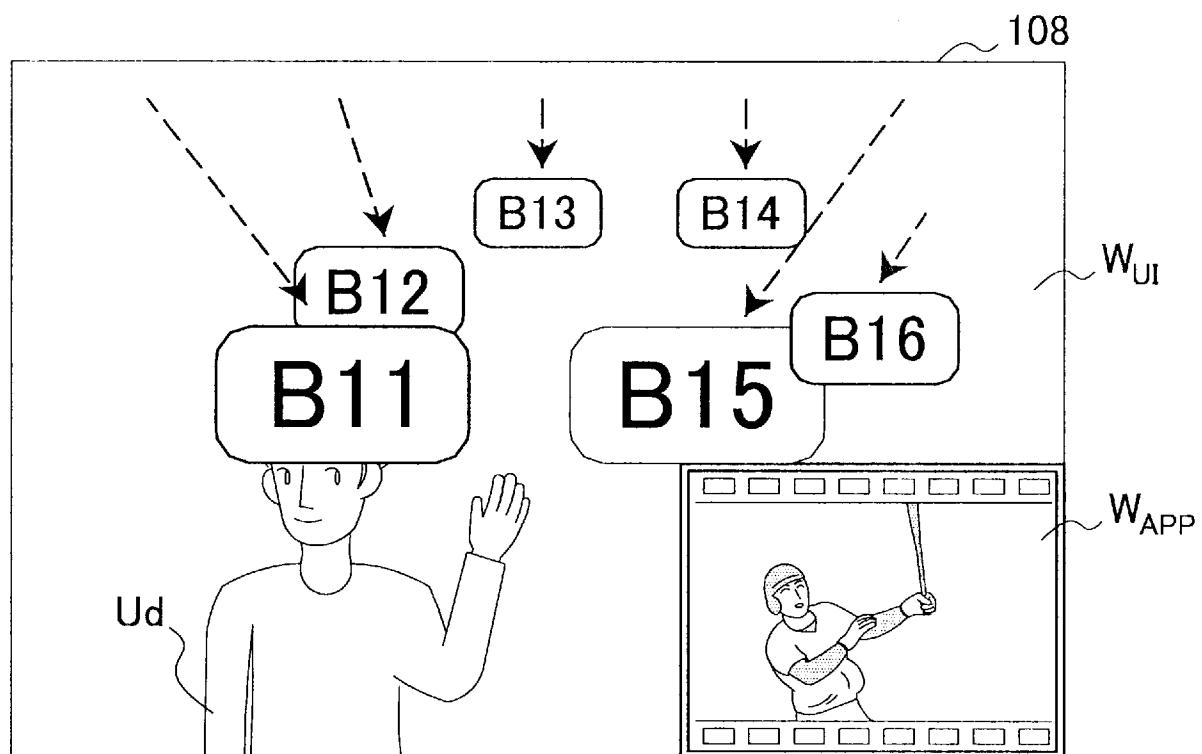
[Fig. 30]



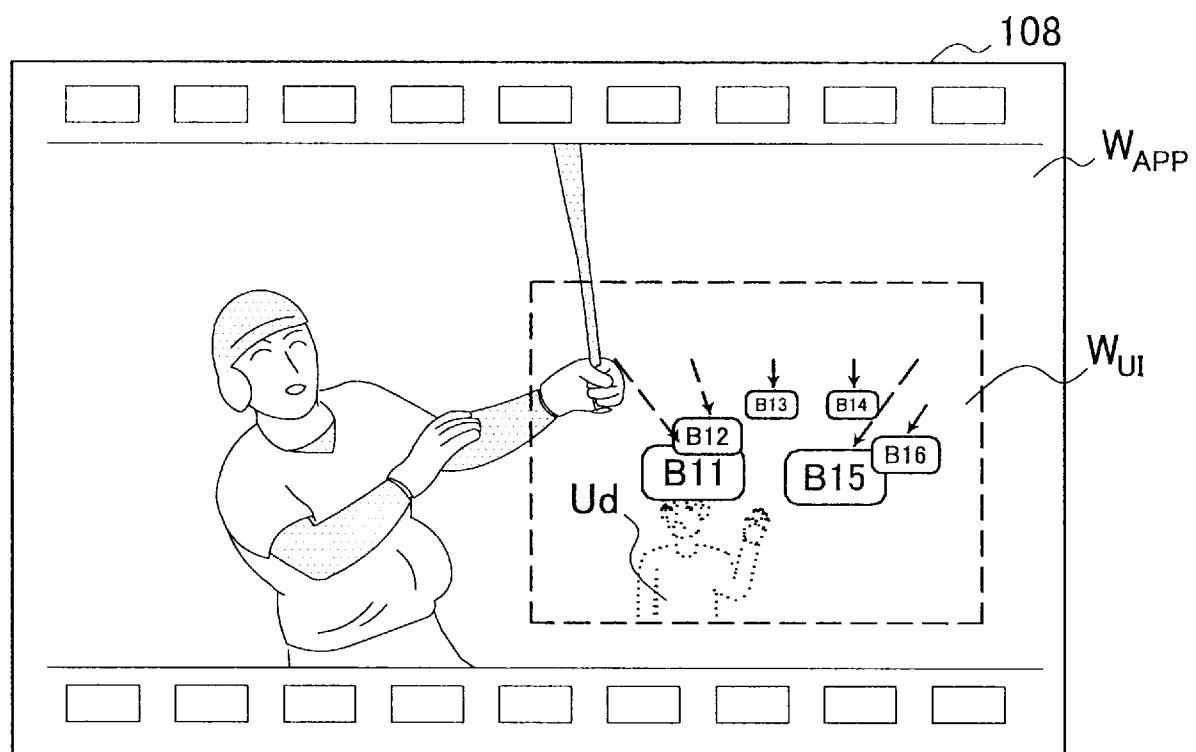
[Fig. 31]



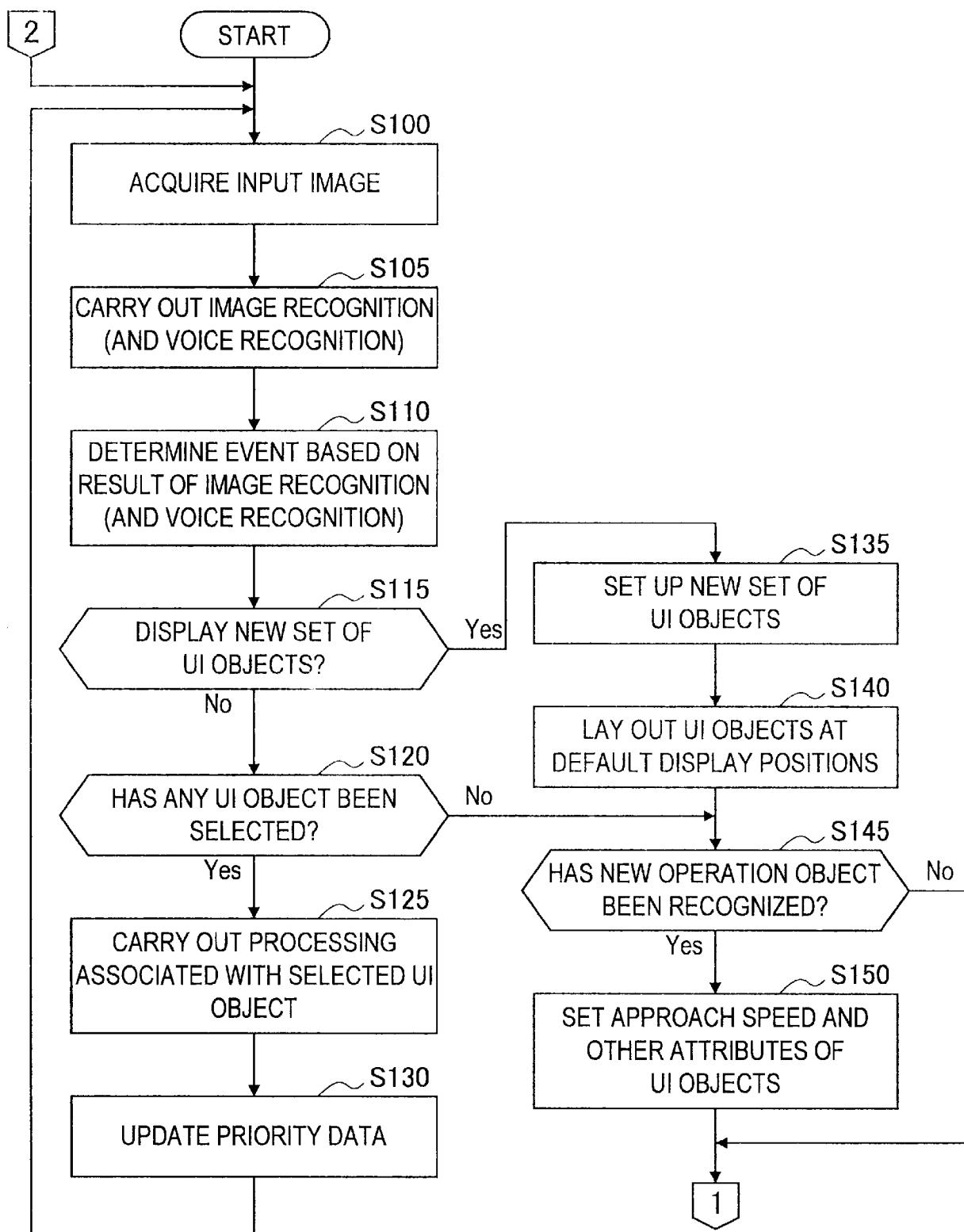
[Fig. 32]



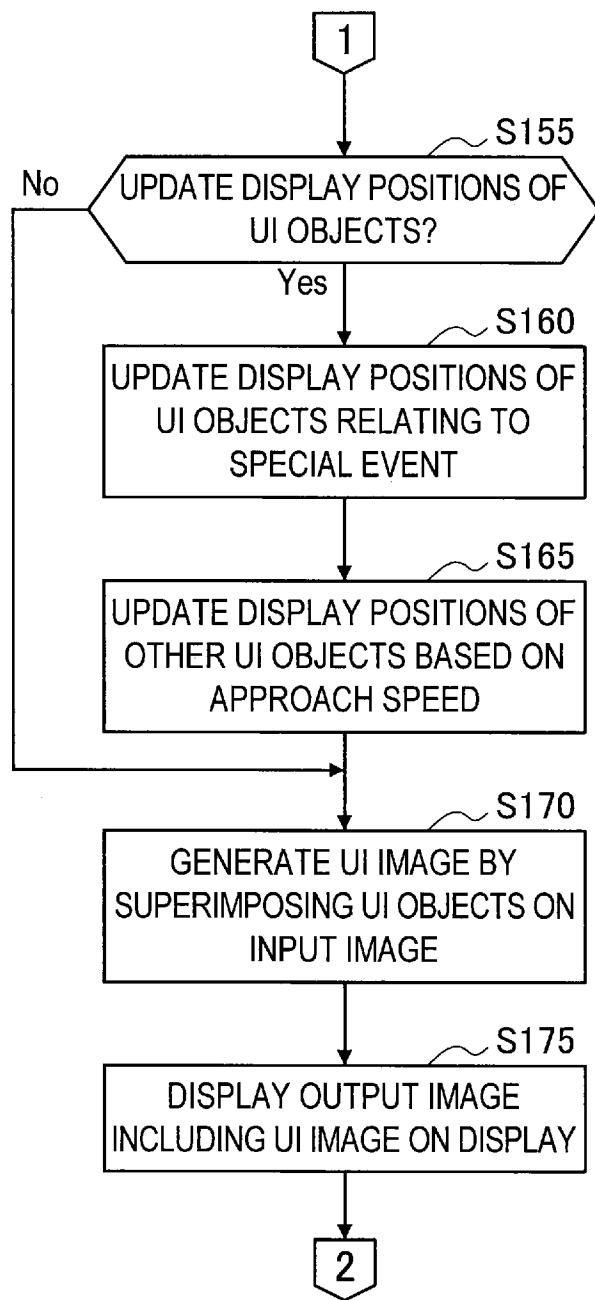
[Fig. 33]



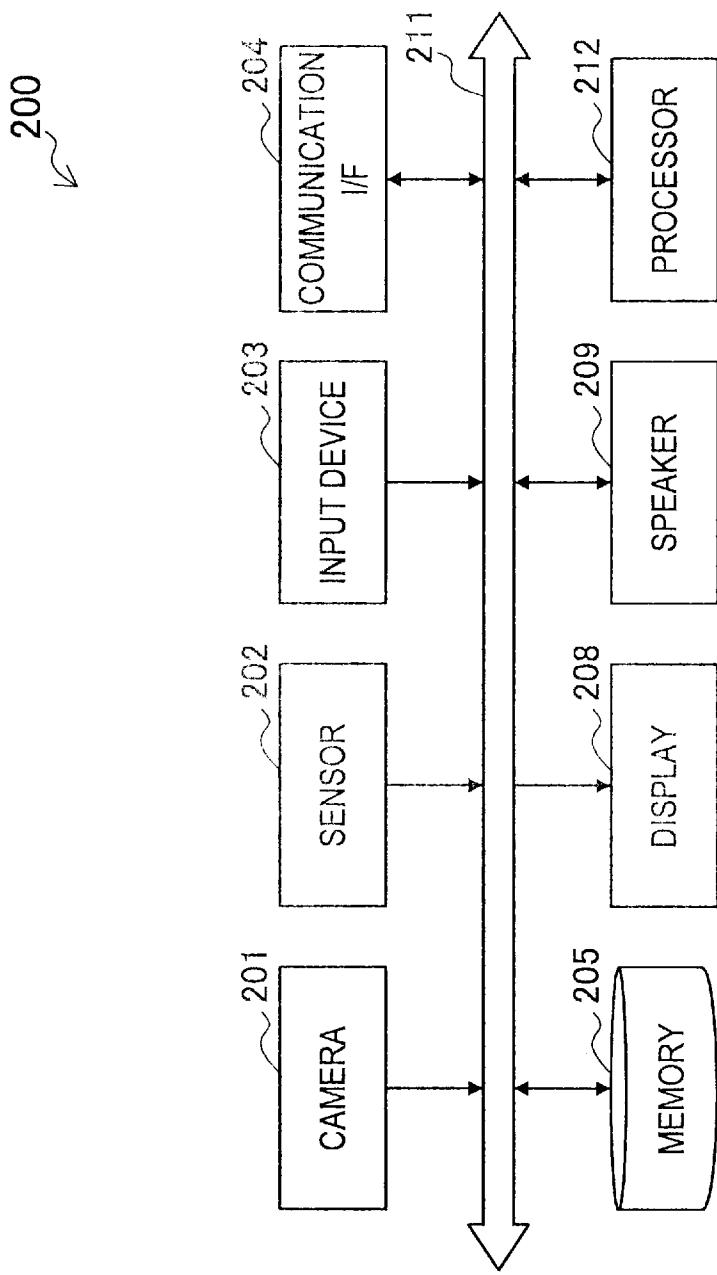
[Fig. 34]



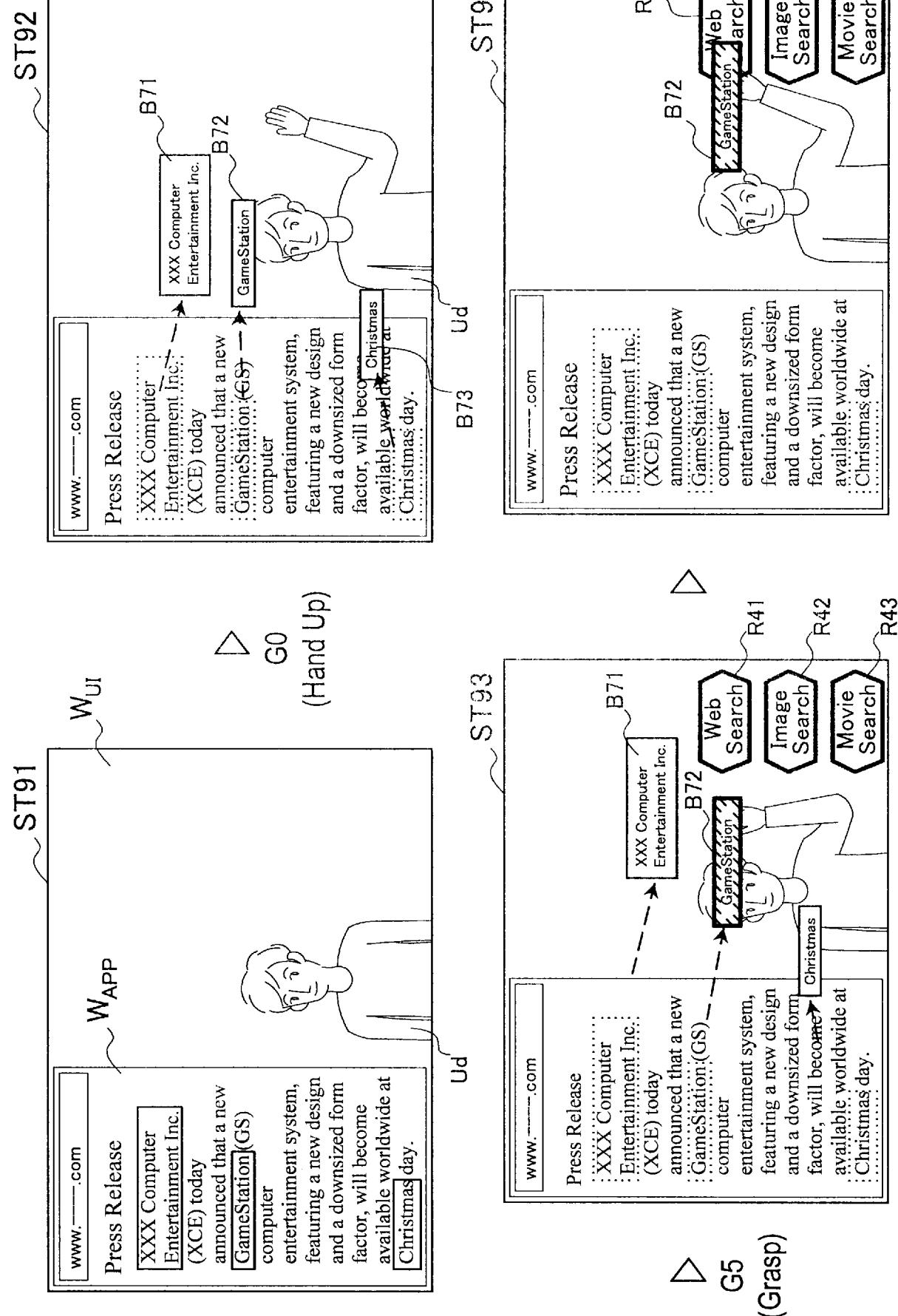
[Fig. 35]



[Fig. 36]



[Fig. 37]



# INTERNATIONAL SEARCH REPORT

International application No  
PCT/JP2013/006979

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. G06F3/01 G06F3/0481  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012/302289 A1 (KANG HEEJOON [KR]) 29 November 2012 (2012-11-29) paragraph [0057] paragraph [0077] paragraph [0080] paragraph [0082] - paragraph [0083] paragraph [0100] paragraph [0106] paragraph [0111] figures 6-7 figure 10 figure 15 figure 19 ----- -/- -----	1-20

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier application or patent but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 January 2014	05/02/2014
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  <b>Thibault, Guillaume</b>

## INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2013/006979

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2010/053151 A1 (MARTI STEFAN [US] ET AL) 4 March 2010 (2010-03-04) paragraph [0025] - paragraph [0028] paragraph [0032] figure 1 figure 3 paragraph [0046] -----	1,13-17, 19,20
A	US 2007/057966 A1 (OHNO TOMOYUKI [JP] ET AL) 15 March 2007 (2007-03-15) paragraph [0190] - paragraph [0203] paragraph [0137] - paragraph [0138] figure 7A figure 9 -----	1-4,6-9, 19,20
A	US 2002/080180 A1 (MANDER RICHARD [US] ET AL) 27 June 2002 (2002-06-27) paragraph [0082] paragraph [0086] paragraph [0094] figure 13 figure 15 -----	1,5,19, 20

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/JP2013/006979

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
US 2012302289	A1	29-11-2012	KR	20120132096	A	05-12-2012
			US	2012302289	A1	29-11-2012
-----						
US 2010053151	A1	04-03-2010	KR	20100027976	A	11-03-2010
			KR	20110082636	A	20-07-2011
			US	2010053151	A1	04-03-2010
			US	2010053164	A1	04-03-2010
			US	2010053322	A1	04-03-2010
			US	2010053324	A1	04-03-2010
			US	2010134618	A1	03-06-2010
			WO	2010027193	A2	11-03-2010
-----						
US 2007057966	A1	15-03-2007	CN	1928888	A	14-03-2007
			EP	1762950	A2	14-03-2007
			JP	2007079641	A	29-03-2007
			KR	20070029574	A	14-03-2007
			US	2007057966	A1	15-03-2007
-----						
US 2002080180	A1	27-06-2002	AU	3969093	A	29-11-1993
			EP	0592638	A1	20-04-1994
			EP	0955591	A2	10-11-1999
			JP	3566720	B2	15-09-2004
			JP	4128130	B2	30-07-2008
			JP	H06511580	A	22-12-1994
			JP	2004348704	A	09-12-2004
			US	6243724	B1	05-06-2001
			US	2002080180	A1	27-06-2002
			US	2004128277	A1	01-07-2004
			US	2007033537	A1	08-02-2007
			WO	9322738	A1	11-11-1993
-----						