

**Internal distribution code:**

- (A) [ - ] Publication in OJ
- (B) [ - ] To Chairmen and Members
- (C) [ - ] To Chairmen
- (D) [ X ] No distribution

**Datasheet for the decision  
of 23 October 2019**

**Case Number:** T 1170/15 - 3.5.05

**Application Number:** 08252314.3

**Publication Number:** 2040156

**IPC:** G06F3/01, G06F3/03

**Language of the proceedings:** EN

**Title of invention:**

Image processing for issuing commands

**Applicant:**

Sony Corporation

**Headword:**

Image processing of hand gestures for issuing commands / Sony

**Relevant legal provisions:**

EPC Art. 56, 123(2)

**Keyword:**

Inventive step - auxiliary request (yes) - non-obvious  
modification

**Decisions cited:**



**Beschwerdekammern**  
**Boards of Appeal**  
**Chambres de recours**

Boards of Appeal of the  
European Patent Office  
Richard-Reitzner-Allee 8  
85540 Haar  
GERMANY  
Tel. +49 (0)89 2399-0  
Fax +49 (0)89 2399-4465

Case Number: T 1170/15 - 3.5.05

**D E C I S I O N**  
**of Technical Board of Appeal 3.5.05**  
**of 23 October 2019**

**Appellant:** Sony Corporation  
(Applicant) 1-7-1 Konan, Minato-ku  
Tokyo (JP)

**Representative:** J A Kemp LLP  
14 South Square  
Gray's Inn  
London WC1R 5JJ (GB)

**Decision under appeal:** **Decision of the Examining Division of the  
European Patent Office posted on 5 January 2015  
refusing European patent application No.  
08252314.3 pursuant to Article 97(2) EPC.**

**Composition of the Board:**

**Chair** A. Ritzka  
**Members:** N. H. Uhlmann  
D. Prietzel-Funk

## **Summary of Facts and Submissions**

- I. The appeal lies from the decision of the examining division to refuse European patent application No. 08252314.3 for lack of clarity of claim 1 of the main and auxiliary requests and for lack of inventive step in the subject-matter of claim 1 of the auxiliary request.
- II. The reasons for the decision refer to the following prior-art documents:  
  
D1: EP 0 823 683;  
D5: US 2006/0209021.
- III. In the statement setting out the grounds of appeal, the appellant submitted amended sets of claims according to a main request and an auxiliary request 1 and requested that the decision be set aside and that a patent be granted on the basis of the main request or auxiliary request 1.
- IV. The board arranged for oral proceedings to be held.
- V. In the summons, the board set out its provisional view of the case. It considered that neither of the pending requests met the requirements of Articles 84, 123(2) and 56 EPC.
- VI. In response, by letter dated 20 August 2019, the appellant re-filed the main request and first auxiliary request and submitted further second to fifth auxiliary requests.
- VII. Oral proceedings were held on 23 October 2019 and attended by the appellant's representatives. In the course thereof, a sixth auxiliary request was submitted.

VIII. The appellant requested that the decision under appeal be set aside and that a patent be granted based on the claims of the main request or on the first auxiliary request, both submitted with the statement setting out the grounds of appeal, or on the second to fifth auxiliary requests submitted with the letter dated 20 August 2019, or on the sixth auxiliary request submitted during the oral proceedings.

IX. Claim 1 of the main request reads as follows:

"An imaging processing apparatus (111) comprising:

an image obtaining unit (31) that supplies captured images frame-by-frame to a frame buffer (32);

a trigger detection unit (121) reads captured images from the frame buffer and detects a gesture of a user's hand as a trigger, and supplies to a feature point extraction unit (122) first area information of said images corresponding to a first trigger, and supplies second area information corresponding to a second trigger;

wherein the feature point extraction unit (122) extracts a feature point corresponding to a hand of a user in said images defined by said first area information, and extracts a feature point of a hand of a user in said images defined by said second area information,

display control means (127) for performing control, based on the current positions of the feature points, to cause a display means (13) to display feature-point pointers (141, 142) indicating the current positions of the feature points in the captured images;

means for recognizing a user performing a grabbing gesture with their hand and locating a respective

feature point in an image of the hand grabbing the feature-point pointer;  
a calculation unit (123) arranged to read consecutive frames of the captured image from the frame buffer (32) and calculate the optical flow of each feature point to recognize the current position of each feature point;  
issuing means (125, 126) for issuing for each of the feature points, based on the recognized position of the feature point in the captured image, a command corresponding to the position of the feature point, or for issuing a command for each of the feature points based on the motion of the feature point from the optical flow of the feature point corresponding to a predefined motion."

- X. Claim 1 of each of the first to fifth auxiliary requests comprises the same wording with regard to the trigger detection unit:

"a trigger detection unit (121) reads captured images from the frame buffer and detects a gesture of a user's hand as a trigger, and supplies to a feature point extraction unit (122) first area information of said images corresponding to a first trigger, and supplies second area information corresponding to a second trigger".

- XI. Claim 1 of the sixth auxiliary request reads as follows:

"An imaging processing method for an image processing apparatus, the image processing method comprising the steps of:

supplying captured images frame-by-frame to a frame buffer (32);

reading captured images from the frame buffer and detecting a first gesture of a user's hand as a trigger and detecting a second gesture of a user's hand as a second trigger within a predetermined period of time after the detection of the first trigger, and supplying to a feature point extraction unit (122) first area information of said images corresponding to the first trigger, and supplies second area information corresponding to the second trigger;

extracting, by the feature point extraction unit (122), a feature point corresponding to a hand of a user in said images defined by said first area information, and extracting a feature point of a hand of a user in said images defined by said second area information,

performing control, based on the current positions of the feature points, to cause a display means (13) to display feature-point pointers (141, 142) indicating the current positions of the feature points in the captured images and predetermined user interfaces (151, 152);

recognizing feature points that are indicated by the feature-point pointers grabbed with a user's hand;

reading, following the step of recognizing, consecutive frames of the captured image from the frame buffer (32) and calculating the optical flow of each feature point to recognize the current position of each feature point;

issuing for each of the feature points, based on the recognized position of the feature point in the captured image, a command corresponding to the position of the feature point, or issuing a command for each of the feature points based on the motion of the feature

point from the optical flow of the feature point corresponding to a predefined motion,

said command corresponds to the user interface (151, 152) that is being displayed at the position of the feature point, and

wherein the predetermined user interfaces are shared by the plurality of feature points."

### **Reasons for the Decision**

1. The invention as described uses augmented-reality techniques for supporting user interaction with a device. It concerns an "image processing apparatus" with a camera which captures images of one or more users and which can be controlled by the user by gestures made with the user's hand "in the real world". Those gestures are captured by the camera and translated into commands to control the apparatus.

2. Prior art

Document D1 discloses a user interface for appliances having a display, e.g. a computer, a word processor or a television, which recognizes and simultaneously displays the shape and move of the hand of an operator, and in which the user gestures control the appliance (abstract).

As the user faces the recognizing means of the interface apparatus and shows a hand, "a special shape corresponding to the shape of the hand" and its motion are displayed on the screen. A virtual switch or the like displayed on the display screen can be "selected" by a hand gesture. The display object displayed on the screen "can be grabbed or carried depending on the

purpose". In this way, "a very simple manipulation of appliance is realized" without requiring a mouse or another input device (column 2, lines 29 to 44).

**Main request**

3. Amendments

3.1 It is common ground that the independent claims, in particular the aspects pertaining to two trigger gestures, are based on pages 35 to 48 of the description.

3.2 The second paragraph on page 36 teaches that the second trigger takes place, and is detected, within a predetermined period of time after the detection of the first trigger. Figure 24, steps S42 and S43 and the corresponding passages of the description (second and third full paragraphs on page 45) confirm that the detection of a second trigger is carried out after the detection of a first trigger.

The current wording of claims 1 and 4 (lines 6 to 11 of claim 1) does not specify this sequence of detecting actions. Hence, the subject-matter claimed encompasses simultaneous detection, without any basis in the original application documents.

3.3 The appellant argued that the sequence of detecting actions is nowhere specified as an essential feature of the invention.

In this regard, the board holds that whether a feature is specified as being essential is not the proper criteria to be applied in the present situation. Instead, it has to be established whether the sequence of detecting actions is inextricably linked to the presence of a first and a second trigger. The board judges that indeed it is, because the passages referred



to in the previous section make it unmistakably clear that they are two triggers which are detected one after another (see also the Case Law of the Boards of Appeal, 9th edition 2019, section II E 1.9).

- 3.4 Consequently, claims 1 and 4 of the main request relate to subject-matter extending beyond the content of the application as originally filed (Article 123(2) EPC), and the main request is not allowable for this reason.

Hence, the issue of compliance with the patentability requirements of the EPC does not need to be addressed.

#### **First to fifth auxiliary requests**

4. Claim 1 of each of these requests comprises the wording objected to in section 3.2 above. Consequently, the first to fifth auxiliary requests are not allowable because they do not meet the requirements of Article 123(2) EPC.

#### **Sixth auxiliary request**

5. Amendments

The objection set out in section 3.2 has been overcome by a corresponding amendment.

The "recognizing" step is based on the first paragraph on page 21.

Finally, the addition that the reading takes place "following the step of recognizing" is based on the second paragraph on page 21.

Therefore, the amended claims comply with the requirements of Article 123(2) EPC.

6. Patentability

6.1 Document D1 forms a suitable starting point for inventive-step analysis. It is common ground that this document does not disclose first and second gestures of a user's hand, first and second triggers and two feature points according to a first and second area information.

6.2 Additionally, document D1 does not disclose that "the calculating the optical flow of each feature point to recognize the current position of each feature point" takes place "following the step of recognizing" (claim 1, last paragraph on page 66).

In this regard, the board holds that pointing out one finger of a user's hand (D1, Figures 8 and 10(A), column 9, first eight lines) corresponds to the first gesture and first trigger as claimed, while gripping the hand to form a fist (Figure 10(B), column 9, lines 9 to 11) corresponds to the grabbing with a user's hand. From these passages, in particular from column 9, lines 6 to 11, it is apparent that immediately after the detection of a hand pointing out one finger, an arrow cursor is moved on the display. Furthermore, after the detection of a hand forming a fist, an instruction is given to the computer, but the cursor is not moved afterwards.

Differently, according to claim 1, the recognition of the current position of each feature point and the displaying of feature-point pointers is performed following the recognition of a grabbing gesture.

6.3 Document D1, in the fourth embodiment, discloses grabbing and subsequent movement of a virtual object (column 22, lines 8 to 27). However, in this embodiment the cursor is also moved on the screen before grabbing

is detected, in order to approach the virtual object (column 14, lines 4 to 12). Hence, this part of document D1 does not anticipate the claimed features referred to in section 6.2.

- 6.4 Document D5 pertains to a virtual mouse driving apparatus which processes gestures based on two-handed gesture information obtained by a video camera. D5 discloses the determination of a selection gesture and the moving of an object so selected (page 3, paragraph 29). Before the selection gesture is recognized, a cursor is moved according to the movement of the right hand and following a pointing gesture (*ibidem*). Therefore, document D5 does not disclose the distinguishing features set out in section 6.2.
- 6.5 That "the calculating the optical flow of each feature point to recognize the current position of each feature point" takes place "following the step of recognizing" leads to the technical effect of reduced computational effort because the movement of the hand is only detected following, and not before, the recognition of the grabbing gesture.
- 6.6 Therefore, the objective technical problem to be solved is how to modify the method known from document D1 to reduce the computational effort needed for detecting a hand's movement.
- 6.7 Facing this problem and considering the teaching of documents D1 and D5, the person skilled in the art would not arrive at these distinguishing features.

In both documents the displaying of a cursor which follows the movement of a hand is needed to inform the user about the position on which a fist gesture (D1) or selection gesture (D5) will take effect. Hence, the detection of the hand's movement before these gestures

are recognized is indispensable for the user interaction disclosed in those documents. Consequently, the skilled person would not limit the detection of the hand's movement to the phase following the recognition of the grabbing gesture.

Moreover, the board is not aware of any common general knowledge which would lead the person skilled in the art to modify the method of document D1 accordingly.

- 6.8 For these reasons, the subject-matter of claim 1 involves an inventive step within the meaning of Article 56 EPC.
- 6.9 Independent claim 2 comprises all the features of claim 1. Therefore, the subject-matter of claim 2 involves an inventive step for the same reasons.

## Order

### For these reasons it is decided that:

1. The decision under appeal is set aside.
2. The case is remitted to the examining division with the order to grant a patent based on claims 1 and 2 of the sixth auxiliary request submitted during the oral proceedings and the description and drawings to be adapted.

The Registrar:

The Chair:



K. Götz-Wein

A. Ritzka

Decision electronically authenticated