Internal distribution code:

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

Datasheet for the decision of 14 November 2023

Case Number: T 0415/21 - 3.5.07

Application Number: 14800220.7

Publication Number: 3066584

IPC: G06F17/28

Language of the proceedings: EN

Title of invention:
Presenting translations of text depicted in images

Applicant:
Google LLC

Headword:
Translations of text in images/GOOGLE LLC

Relevant legal provisions:
EPC Art. 56, 113(1)

Keyword:
Inventive step - (no) - mixture of technical and non-technical features
Disputed common general knowledge
Right to be heard - violation (no)

Decisions cited:

T 0939/92, T 0190/03, T 0928/03, T 0154/04, T 1242/04, T 1143/06, T 1741/08, T 2467/09, T 2045/10, T 1526/11, T 2028/11, T 2035/11, T 1273/20

Europäisches Patentamt
European Patent Office
Office européen des brevets

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 0415/21 - 3.5.07

D E C I S I O N
of Technical Board of Appeal 3.5.07
of 14 November 2023

Appellant: (Applicant)	Google LLC 1600 Amphitheatre Parkway Mountain View, CA 94043 (US)
Representative:	Williams, Michael David Marks & Clerk LLP 1 New York Street Manchester M1 4HD (GB)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 6 November 2020 refusing European patent application No. 14800220.7 pursuant to Article 97(2) EPC**

Composition of the Board:

Chair	J. Geschwind
Members:	P. San-Bento Furtado
	R. de Man

Summary of Facts and Submissions

I. The appeal lies from the decision of the examining division to refuse European patent application No. 14800220.7 by means of a "decision according to the state of the file" referring to a communication.

The following prior-art documents were cited in the international search report:
D1: US 2013/0004068 A1, 3 January 2013;
D2: A. Kazmucha: "Word Lens for iPhone review", 5 April 2013.

The examining division decided that the subject-matter of the independent claims of the main request and of claim 1 of the first and second auxiliary requests lacked inventive step over document D1.

II. In its statement of grounds of appeal, the appellant requested that the decision under appeal be set aside and that a patent be granted on the basis of the main request or the first or second auxiliary request, all three requests as considered in the appealed decision.

III. In a communication accompanying a summons to oral proceedings, the board expressed its preliminary opinion that the subject-matter of claim 1 of each of the three requests was not inventive over document D1 in combination with notorious knowledge.

IV. With a letter of reply, the appellant presented further arguments in support of inventive step.

V. Oral proceedings were held as scheduled. During the oral proceedings, the appellant expressed doubt that its right to be heard had been respected but noted that

it was not raising an objection under Rule 106 EPC. At the end of the oral proceedings, the Chair announced the board's decision.

VI. The appellant's final requests were that the decision under appeal be set aside and that a patent be granted on the basis of the claims of the main request or, in the alternative, of one of the first and second auxiliary requests.

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

"A method performed by data processing apparatus, the method comprising:

receiving an image;

identifying text depicted in the image;

selecting, for the image, a presentation context from a plurality of presentation contexts based on an arrangement of the text depicted by the image, wherein each presentation context corresponds to a particular arrangement of text within images and each presentation context has a corresponding user interface for presenting additional information related to the text depicted in the image, wherein the user interface for each presentation context is different from the user interface for other presentation contexts;

identifying the user interface that corresponds to the selected presentation context;

automatically presenting using the identified user interface additional information for a first portion of the text depicted in the image and not automatically presenting additional information for a second portion of the text depicted in the image, the user interface presenting the additional information in an overlay over the image; and

providing a selectable user interface element, selectable by a user to present additional information for the second portion of text depicted in the image."

VIII. Claim 1 of the first auxiliary request differs from claim 1 of the main request in that the following text has been added before "the user interface that corresponds":
", in a user interface index mapping each of the plurality of presentation contexts with a respective user interface,".

IX. Claim 1 of the second auxiliary request differs from claim 1 of the first auxiliary request in that the following text has been added at the end of the claim:
", wherein selecting the presentation context for the image comprises:

determining , based on the arrangement of the text depicted in the image, that a first portion of the text is presented more prominently than at least one other portion of the text and selecting a prominence context from the plurality of presentation contexts in response to the determination; or

determining that the text depicted in the image comprises an address and selecting a map context from the plurality of presentation contexts in response to the de-termination [sic]; or

identifying a plurality of individual text blocks depicted in the image, determining that the plurality of individual text blocks belong to a collection of text based on an arrangement of the individual text blocks and presentation of the individual text blocks, and selecting a collection context from the plurality of presentation context [sic] in response to the determination."

Reasons for the Decision

Application

1. The application concerns a system for identifying text depicted in an image, translating the text, and presenting the translation, the presentation being based on the arrangement and/or other visual characteristics of the text within the image. The purpose is to present the translation in a manner that is useful to a user and avoids cluttering the display (see original description, page 9, lines 11 to 19).

1.1 Based on the arrangement and visual characteristics of the text in the image, the system selects a "presentation context". The system then selects a user interface corresponding to the selected presentation context for presenting additional information about text identified in the image (e.g. a language translation of text identified in the image or a currency translation of a monetary amount identified in the image, see page 11, line 26, to page 12, line 3, and original claim 2).

1.2 In one embodiment, the presentation context is "prominence", and the user interface presents, on top of the original image, an overlay that includes a translation of the most prominent text block and a user-interface element that, when selected, presents a translation of the less prominent text blocks (page 15, lines 26 to 29; page 17, lines 26 to 30).

Main request

2. *Inventive step - claim 1*

2.1 Document D1 discloses techniques for recognising text in an image captured by a mobile device using optical character recognition (OCR), translating the text to a language understandable by the user and replacing the symbols in the image while reducing the artifacts which may result from re-rendering the background image (abstract).

2.2 Therefore, document D1 discloses a method performed by a data processing apparatus comprising the steps of claim 1 of receiving an image and identifying text depicted in the image.

2.3 In the method of document D1, translation of text in images captured by a mobile device may be automatic or performed upon instruction by the user through a function menu or other suitable means (paragraph [0032]). The system detects symbols, such as text or characters, in a captured image and determines the boundaries isolating the symbols in the image from the background. It then generates the pixels for the translated symbols and for the background within the boundaries by interpolating the pixel values between the boundaries (paragraphs [0031], [0033] to [0035], Figures 2 to 4).

2.4 In the method of document D1, the background of the text in the original image (both within and outside the boundary area of the symbols) is maintained as much as possible (see e.g. paragraphs [0033], [0034] and [0042], Figures 2 and 3). The translated text corresponds to "additional information" for "text depicted in the image" within the meaning of claim 1

(as also confirmed by the text of claim 2, "wherein the additional information comprises a language translation of the at least a portion of the identified text."). Since in the system of document D1 the translated text replaces the original text in the image, as shown in Figures 2 and 3, document D1 discloses automatically presenting in an overlay image "additional information for a first portion of the text depicted in the image". The board notes that this presentation arrangement corresponds to a (fixed) "presentation context".

2.5 The board agrees with the appellant that document D1 does not disclose a plurality of (predefined types of) presentation contexts with corresponding user interfaces with a selectable element.

2.6 The subject-matter of claim 1 therefore differs from the method of D1 in that it includes the following features:

(i) the presentation context is selected from a plurality of presentation contexts based on an arrangement of the text depicted by the image, wherein each presentation context corresponds to a particular arrangement of text within images and each presentation context has a corresponding user interface for presenting additional information related to the text depicted in the image, wherein the user interface for each presentation context is different from the user interface for other presentation contexts;

(ii) the user interface that corresponds to the selected presentation context is identified;

(iii) additional information for a second portion of the text is not automatically presented;

(iv) a selectable user-interface element is provided, which is selectable by a user to present

additional information for the second portion of text depicted in the image.

2.7 The distinguishing features concern a graphical user interface (GUI) and presentation of information. The layout of a GUI is usually not considered to be technical (T 1526/11, Reasons 2.5 to 2.9), whereas a user-interface element that the user can activate to trigger an associated action is a technical part of the user interface (T 2028/11, reasons 3.6). Presentation of information is as such not technical (Article 52(2) (d) and (3) EPC). Displaying information for cognitive processing by the user is not a technical use of the information and lowering the cognitive burden of the user is not a technical effect (see e.g. T 1143/06, reasons 3.8 and 5.4; T 1741/08, reasons 2.1.6; T 2045/10, reasons 5.6.2 and 5.6.3; T 1562/11, reasons 2.5 to 2.9 and T 2035/11, reasons 5.1.3).

Features which as such are non-technical can only be taken into account for inventive step if they make a technical contribution by interacting with claimed technical subject-matter for solving a technical problem (T 154/04, reasons 5). For the assessment of inventive step in the case at hand, it has thus to be established whether a technical problem is solved and which distinguishing features (i) to (iv) are technical or make a technical contribution.

2.8 The appellant formulated the objective technical problem as: how to modify the image to include further information while resolving the conflict with the technical constraints of a limited display area and the physical features of the original image.

In support of the technicality of this problem, the appellant cited decision T 928/03, in which the particular manner of conveying to the user the location of the nearest teammate by dynamically displaying a guide mark on the edge of the screen when the teammate was off-screen produced the technical effect of facilitating a continued human-machine interaction by resolving conflicting technical requirements: displaying an enlarged portion of an image and maintaining an overview of a zone of interest larger than the display area. The distinguishing features of claim 1 likewise achieved an enhanced human-machine interaction by resolving conflicting technical requirements.

2.9 For the reasons given in the following, the board does not find the appellant's arguments convincing.

Features (i), (ii) and (iii) concern the layout design of a graphical user interface (GUI) and presentation of information. Feature (iv) specifies a selectable user-interface element which the user may select to view additional information for a second portion of the text which is not automatically displayed (feature (iii)).

It cannot be derived from claim 1 that the choice of presentation of information of features (i) to (iii) is especially determined by constraints of the display area.

The different presentation contexts are used in features (i) and (ii) to distinguish between different types of text layouts (or arrangements) or even content (e.g. whether the text in the image corresponds to a list of items or a map), and to choose a corresponding layout for displaying the information. The second portion of the text is presented only optionally, as

specified in feature (iii), in order to avoid confusing the user with too much information or multiple text blocks displayed simultaneously (see page 9, lines 11 to 19). These are all non-technical aspects. Any improvement caused by distinguishing features (i) to (iii) is merely at the user's cognitive level, which is not a technical effect. Therefore, distinguishing features (i) to (iii) merely reflect non-technical requirements of how to present the translated text.

This case is different from that of T 928/03 because features (i) to (iii) do not solve technical constraints of the display and there is no zone of interest outside of the displayed area.

2.10 The selectable user-interface element of feature (iv) is an interactive element which the user can activate to trigger the display of the additional information. Feature (iv) is thus a technical part of the user interface (see also T 2028/11, reasons 3.6). It solves the problem of modifying the method of D1 to implement the optional display of additional information for a second portion of the text (feature (iii)).

2.11 Selectable user-interface elements such as buttons or links which, when selected, cause further information to be displayed are notoriously known, for example from web applications and from mobile devices. In particular, buttons of a graphical user interface and links (or hyperlinks) to further information in a web application are universally used by members of the general public, e.g. when operating smart phones or laptops.

2.11.1 At the oral proceedings, the appellant expressed doubts that its right to be heard had been respected because the board had alleged that certain features were known

without providing sufficient documentary evidence. The board had apparently based its opinion on evidence which had not been provided to the appellant.

2.11.2 Article 113(1) EPC, which establishes the right to be heard, stipulates that the decisions of the EPO may only be based on grounds or evidence on which the parties concerned have had an opportunity to present their comments. A party's right to be heard is violated if a decision negatively affecting the party relies on (unspecified) facts or evidence not known to the party.

2.11.3 If the common general knowledge relevant for the outcome of a case is disputed by a party, normally it has to be proved like any other fact under contention, for instance by documentary or oral evidence (T 939/92, OJ EPO 1996, 309, Reasons 2.3; Case Law of the Boards of Appeal, 10th edition, 2022, I.C.2.8.5). However, exceptionally it suffices to give cogent reasons based on readily verifiable facts. This applies, for example, to knowledge that is "notorious" or indisputably forms part of the common general knowledge (see decisions T 1242/04, OJ EPO 2007, 421, Reasons 9.2 and T 2467/09, Reasons 4 and 8). Likewise, no specific documentary evidence may be needed to prove knowledge which belongs to the "mental furniture" of the skilled person, such as routine design skills and general principles of system design which are often necessary just to understand the prior art in the relevant field (T 190/03, Reasons 16; T 1273/20, Reasons 5.7).

2.11.4 As explained above, the board considers that buttons of a graphical user interface and links (or hyperlinks) to further information in a web application are universally used by members of the general public. They

are so ubiquitous that documentary evidence of them does not have to be presented.

2.11.5 In the present case, the board has communicated to the appellant the grounds and evidence on which its decision in respect of the main request is based and has given the appellant the opportunity to comment. The appellant's right to be heard has thus been respected. The appellant might not agree with the board's justification for its finding that selectable user-interface elements which, when selected, cause further information to be displayed were known in the art, but that does not mean that this finding was based on specific documents that have been kept hidden from the appellant.

2.11.6 At the oral proceedings, the board has nevertheless additionally called the attention of the appellant to the disclosure of those features in document D2. The system of document D2 recognises and translates text in real-world images captured by a mobile device and displays the translation of the recognised text overlaid on the captured real-world images (page 1). The figure shown on page 1 illustrates on the right-hand side a button "*i*" to further information. This corresponds to distinguishing feature (iv) above. For completeness, the board notes that page 2 describes that the user "can also tap on words to view more information on them" and page 3 mentions in the section titled "The bad" that "Sometimes definitions don't always seem to be right for words that you tap on". This means that an overlaid word of the translated text can be taped on in order to access further information such as definitions. These features correspond to distinguishing feature (iv).

2.12 It would thus have been obvious for the skilled person given the task of implementing the desired non-technical presentation of the optional display of a second portion of the text to use a selectable user-interface element in the system of D1, which runs on a mobile device (see e.g. D1, abstract and paragraph [0002]).

2.13 In view of the above, the subject-matter of claim 1 of the main request is not inventive (Article 56 EPC).

First auxiliary request

3. Compared to the main request, claim 1 of the first auxiliary request further specifies that a "user interface index" maps each of the plurality of presentation contexts to a respective user interface.

4. *Inventive step - claim 1*

4.1 The appellant did not contest that "index" data structures for mapping entities of a first type to entities of a second type were well known in the art.

4.2 When implementing the non-technical user constraint of providing different user interfaces corresponding to different presentation contexts in the method of D1, the skilled person would have immediately recognised that a mapping was necessary between presentation contexts and respective user interfaces. Hence, they would have used a suitable "index" data structure to do so.

4.3 The appellant argued that the "user interface index" feature, by clarifying that the presentation contexts and corresponding user interfaces were predetermined entities, invalidated the argument that each original

image, used as the background for the presentation of the translated text in document D1, corresponded to a "presentation context" within the meaning of claim 1.

However, the board's inventive-step reasoning does not use that argument.

4.4 Therefore, the first auxiliary request does not meet the requirements of Article 56 EPC for lack of inventive step in the subject-matter of claim 1.

Second auxiliary request

5. *Inventive step - claim 1*

5.1 The additional features (see section IX. above) essentially specify that a "prominence context" is selected as the presentation context if a first portion of the text in the image is presented more prominently than one or more other portions, a "map context" is selected if an address is detected, and a "collection context" is selected if a plurality of individual text blocks are identified.

5.2 The appellant argued that, as with the main request, the method of claim 1 contributed to supporting and reducing user input needed to change the way in which received images were processed to extract particular information. The user-interface element enabled the user to extract different information from the received image more easily. This provided a particularly convenient mechanism that allowed for continued and guided human-machine interaction in the technical task of processing or modifying the received images.

5.3 The board is, however, of the opinion that the features added by claim 1 of the second auxiliary request merely reflect non-technical requirements regarding the presentation of the translated text depending on the layout (e.g. list of items) or content (e.g. address) of the text in the captured image. Since the images are displayed only for cognitive processing by the user, the board does not agree that these features, in combination with the other distinguishing features, contribute to assisting the user in performing a technical task. Instead, the features added by claim 1 of the second auxiliary request relate to presentation of information as such, do not contribute to a technical effect and cannot therefore establish an inventive step.

5.4 Hence, the second auxiliary request does not meet the requirements of Article 56 EPC, either.

Conclusion

6. Since none of the requests is allowable, the appeal is to be dismissed.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chair:

S. Lichtenvort J. Geschwind

Decision electronically authenticated