Internal distribution code:

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

Datasheet for the decision of 9 July 2024

Case Number: T 0434/21 - 3.4.01

Application Number: 15888197.9

Publication Number: 3282445

IPC: G10L15/22, H04M1/725

Language of the proceedings: EN

Title of invention:
VOICE RECOGNITION METHOD, VOICE WAKE-UP DEVICE, VOICE RECOGNITION DEVICE AND TERMINAL

Applicant:
Honor Device Co., Ltd.

Headword:
Voice recognition method / Honor Device Co.

Relevant legal provisions:
EPC Art. 56
RPBA 2020 Art. 12(4)

Keyword:
Inventive step - Main request, Auxiliary requests 1 and 2 (no)
Amendment to appeal case - Auxiliary request 2 - amendment overcomes issues raised (no)

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 0434/21 - 3.4.01

D E C I S I O N
of Technical Board of Appeal 3.4.01
of 9 July 2024

Appellant: (Applicant)	Honor Device Co., Ltd. Suite 3401, Unit A, Building 6 Shum Yip Sky Park, No. 8089 Hongli West Road Xiangmihu Street Futian District Shenzhen, Guangdong 518040 (CN)
Representative:	Roth, Sebastian Mitscherlich PartmbB Patent- und Rechtsanwälte Karlstraße 7 80333 München (DE)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 25 November 2020 refusing European patent application No. 15888197.9 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chairwoman	A. Medeiros Gaspar
Members:	P. Fontenay
	D. Rogers

Summary of Facts and Submissions

I. This decision relates to the appellant's (applicant's) appeal of the Examining Division's decision to refuse European patent application 15 888 197.

II. In their decision to refuse the application, the Examining Division held that claim 1 of the main request and claim 1 of each of auxiliary requests 1 and 2 did not involve an inventive step in the sense of Article 56 EPC. Reference was made, in this respect, to document

D1: EP-A-2 669 889.

Regarding the main request, it was acknowledged that document D1, relied upon as starting point, did not disclose using a first and second preset duration, the second preset duration being dependent on when the speech recognition function is enabled. However, it was held that no technical effect resulted from this difference. The Examining Division considered that the step of separating the speech into two parts to later join them for recognition of a command word did not have any technical advantage or surprising effect, compared to recording any speech coming in during a first preset during after a wakeup word was recognized and processing voice recognition thereon, as disclosed in D1.

III. Claim 1 of auxiliary request 1 was considered to lack an inventive step, because it was common knowledge for

the skilled person to switch off a device when it had not been used for a certain time. Claim 1 of auxiliary request 2 was considered to lack an inventive step in view of a combination of D1 with document

D2: US-A-2013/289994,

which disclosed that a wakeup recognizer of a system, that was constantly active, should run on a Digital Signal Processor (DSP) and that a more demanding voice command recognizer should run on an Application Processor (AP) or a Central Processor Unit (CPU) for energy consumption reasons (D2, paragraphs 7 and 8).

IV. The Appellant requests that the impugned decision be set aside and a patent granted on the basis of a main request, or of a first, second or third auxiliary request. The main request and auxiliary requests 1 and 2 correspond to the corresponding requests underlying the impugned decision. Auxiliary request 3 is a new request, filed on appeal.

V. In a communication under Article 15 RPBA, that accompanied a summons to oral proceedings, the appellant was informed of the Board's preliminary opinion.

VI. This opinion was negative. In essence, the Board shared the conclusion of the Examining Division regarding the lack of inventive step of claim 1 of the main request and of claim 1 of each of auxiliary requests 1 and 2 vis-à-vis D1. Concerning auxiliary request 3, the Board indicated that it was inclined not to admit it, in

addition to raising doubts of added-matter, the amendments introduced in claim 1 did not appear sufficient to address the issue of lack of inventive step in view of D1.

VII. In reply to the Board's communication, the appellant elaborated on why, in their opinion,all requests were allowable, and auxiliary request 3 should be admitted.

Concerning the main request, the proprietor confirmed the interpretation, that the Board had provided in the provisional opinion, that the first speech information according to the invention comprised both a wakeup information and a first portion of the command word, and that both were buffered. The crucial aspect of the invention was that the buffered information, including first portion of the command word, would then be provided to the speech recognition apparatus, which by itself would listen to the rest of the command (second portion).

Concerning the existence of a technical effect, it was emphasised that the distribution of the recognition processes between a speech wakeup apparatus and a speech recognition apparatus allowed the latter to be disabled while still allowing recognition of the speech command. The problem addressed by the invention consisted in reducing the resources required for reliable speech recognition. The appellant considered that the Board had failed to recognise, in its preliminary assessment, that the claim referred to two distinct recognition apparatuses, namely the speech wakeup apparatus and the speech recognition apparatus.

Regarding auxiliary request 1, the appellant further emphasised that the Board had failed to recognise that only one out of two apparatuses was switched off. D1 simply mentioned that the entire device was switched into a standby state.

Regarding auxiliary request 2, the appellant acknowledged that D2 disclosed a digital signal processor (DSP) and an app processor (AP) or a central processing unit. There was however no teaching or motivation in D2 to combine the distinct recognition processes executed in the separate processing units of D2 in the way proposed by the invention.

As for auxiliary request 3, the appellant reiterated the view that claim 1 was supported by the original disclosure. Arguments as to why auxiliary request 3 should be admitted into the appeal proceedings were also put forward.

VIII. In a later submission, the Appellant informed the Board that they would not participate in the oral proceedings.

IX. The oral proceedings were cancelled.

X. Claim 1 of the main request reads:

A speech control method, comprising:
obtaining, by a speech wakeup apparatus, a first speech information by listening in a surrounding environment, wherein the first speech information

comprises a wakeup information and a first portion of command word;
when determining, by the speech wakeup apparatus, that the wakeup information matches a speech wakeup model, buffering, by the speech wake up apparatus, the first speech information obtained by listening within a first preset duration, and generating, by the speech wakeup apparatus, a trigger signal for a speech recognition apparatus to enable a speech recognition function;
obtaining, by the speech recognition apparatus, a second speech information by listening within a second preset duration after the speech recognition function is enabled, wherein the second speech information comprises a second portion of command word;
obtaining, by the speech recognition apparatus, a speech instruction information according to the first speech information and the second speech information, wherein the speech instruction information matches the command word, the command word comprises the first portion of the command word and the second portion of the command word;
performing, by an execution module, operation according to the speech instruction information.

XI. Claim 1 of the first auxiliary request adds at the end of claim 1 of the main request:

[... speech instruction information;]
when determining that a further trigger signal is not received again within a third preset duration after the previous trigger signal is received, disabling, by the speech recognition apparatus, the speech recognition function automatically.

XII. Claim 1 of auxiliary request 2 adds at the end of claim 1 of auxiliary request 1 the feature

> [...speech recognition function automatically;] *wherein the speech wakeup apparatus is implemented by using a digital signal processor, and the speech recognition apparatus is implemented by using an application processor or a central processing unit.*

XIII. Claim 1 of auxiliary request 3 reads (additions with regard to claim 1 of the main request indicated underlined):

> *A speech control method, comprising:*
> *obtaining, by a speech wakeup apparatus, a first speech information by listening in a surrounding environment, wherein the first speech information comprises a wakeup information and a first portion of command word, wherein a time point for starting to speak the wakeup information is t0, a time point for finishing speaking the wakeup information is t1, and a time point for starting to speak the first portion of command word is t2;*
> *when determining, by the speech wakeup apparatus, that the wakeup information matches a speech wakeup model, buffering, by the speech wakeup apparatus, the first speech information obtained by listening within a first preset duration, and generating, by the speech wakeup apparatus, a trigger signal for a speech recognition apparatus to enable a speech recognition function, wherein a time point for enabling the speech recognition function is t3, and the first present duration is from t0 to t3;*

obtaining, by the speech recognition apparatus, a second speech information by listening within a second preset duration after the speech recognition function is enabled, wherein the second speech information comprises a second portion of command word;
obtaining, by the speech recognition apparatus, a speech instruction information according to the first speech information and the second speech information, wherein the speech instruction information matches the command word, the command word comprises the first portion of the command word and the second portion of the command word;
performing, by an execution module, operation according to the speech instruction information.

Reasons for the Decision

Main request - Inventive step in view of D1

1. The Board concurs with the Examining Division that document D1 (paragraphs [0045], [0046] and [0049]) constitutes a good starting point to decide on the inventive merits of the claimed subject-matter.

2. D1 discloses a speech control method implemented by a system with a speech control system. The system comprises a speech wakeup apparatus configured to obtain a first speech information by listening in a surrounding environment, wherein the first speech information comprises a wakeup information. This speech wakeup apparatus of D1 determines that the wakeup

information matches a speech wakeup model and generates a trigger signal for a speech recognition apparatus, possibly implemented at a server. This trigger signal enables a speech recognition function, to recognise a subsequent second speech information corresponding to a voice command. Following to the recognition of the voice command, an execution module performs then an operation accordingly.

3. In D1, this is the second speech information, i.e. the portion of the voice signal corresponding to the voice command, that is buffered. This corresponds to the portion of the voice signal that is transmitted to the speech recognition apparatus (paragraph [0049], lines 21-23).

4. The appellant argues that D1 does not disclose the step of "buffering, by the speech wakeup apparatus, of the first speech information obtained by listening within a first preset duration".

5. This step was considered by the Examining Division to be implicitly known from D1, since it needs to take place before the voice command is sent to a server for recognition.

6. The Board acknowledge that the disclosed transmission of the portion of the voice signal corresponding to the voice command to a server necessarily implies the buffering of this portion of the signal.

7. However, this corresponds to the second speech information, i.e. the portion of the voice signal corresponding to the voice command, that is buffered before transmission to the speech recognition apparatus at the server, not to the first speech information

comprising the wake up information, as correctly argued by the applicant. D1 is also silent about splitting the command word into a first portion and a second portion, or using a first and second preset duration, the second preset duration being dependent on when the speech recognition function is enabled.

8. The Board fails, however, to see any technical advantage to which the buffering of the wake up information could possibly contribute. Such buffering is even superfluous considering that in D1, as in the subject-matter defined in the claim, it is only after the wakeup model has been recognised in the first place, that the subsequent voice command is processed in order for it to be recognised. This difference cannot thus contribute to an inventive step.

9. As for the separating of the speech in two parts to later join for recognition, one of which is buffered by the speech wakeup apparatus, also this difference does not have any technical advantage or surprising effect compared to simply buffering any voice signal coming in, to later carry out the speech recognition on it, as is the case in D1. Hence, also this difference cannot render the claimed subject-matter inventive.

10. The appellant also underlines that, since according to the claimed invention, the speech wakeup apparatus enables the speech recognition apparatus after the wakeup information is heard, the speech recognition apparatus does not need to stay active all the time. In this respect, the appellant mentions an effect on the energy consumption, that is reduced as a result of the speech recognition apparatus being activated only upon reception of the wakeup signal.

11. The Board recognises this as indeed a technical effect, but notes that also in D1 the voice command is only sent to the server, and hence the voice recognition apparatus of the server "enabled", following the detection of a wakeup signal.

12. The Board further notes that whether the speech recognition is carried out in the wakeup apparatus or in a dedicated speech recognition unit seems to have no bearing on the total power consumption, which depends primarily on the speech recognition method employed by each of these "apparatus".

13. Additionally, no technical effect can be derived from the reference to two units in the claim, which may both correspond to different routines of the speech control algorithm running on a terminal. This is confirmed by the wording of independent claim 10 (erroneously referred to as claim 11 in the communication of the Board) which refers to a terminal with possibly only one processor.

14. The same observation applies to the absence of a need to wait for the speech recognition apparatus to become active, also mentioned by the appellant. Also this does not define any technical contribution over D1, since in D1 the voice command is also buffered, thus allowing the user to speak out the instructions before the voice command is forwarded to the server. The Board further notes that the claimed method requires that the two portions of the command be available before it is recognised, so that no technical contribution in terms of processing speed is identifiable either.

15. In summary, the Board fails to see any of the differences indicated above as contributing to a

technical effect, and, consequently as possibly contributing to an inventive step in view of the disclosure of D1.

16. Hence, the subject-matter of claim 1 is not inventive over D1 (Article 56 EPC).

First auxiliary request

17. Claim 1 of the first auxiliary request further differs from D1 by the step of disabling the speech recognition function automatically, when it is determined that a further trigger signal is not received again within a third pre-set duration after the previous trigger signal is received.

18. The problem solved may be defined as saving resources (energy, memory space...) of the speech recognition unit or the host processor operating said function if no speech signal is received.

19. Concretely, the claimed solution consists in automatically setting the speech recognition apparatus in a standby state after a certain predetermined period.

20. The setting of a device or unit in a stand-by state, in order to save energy in case the functionality of the device is not needed, was part of the common general knowledge.

21. Moreover, as noted by the Examining Division, D1 itself also refers to idle states (paragraphs [0045]), thus suggesting that some function might be switched into a standby state if it is not used for a while.

22. The argument that system of D1 does not comprise two different processing units and thus does not suggest to disable only one of them, cannot be followed because D1 discloses the detection of the wake up command as possibly taking place at one device and the recognition of the voice command as taking place at a server.

23. Hence, the skilled person seeking to solve the objective technical problem would also have recognised that the speech wakeup apparatus needs to stay active and that only the speech recognition apparatus that is not being used for a while is to be switched into standby state.

24. It follows that the subject-matter of claim 1 is not inventive (Article 56 EPC).

Second auxiliary request

25. Claim 1 according to the second auxiliary request further defines that the speech wakeup apparatus is implemented by using a DSP, and the speech recognition apparatus is implemented by using an AP or a CPU.

26. A technical effect of these further distinguishing features is that distinct method steps can be effectively processed by specialised processing units.

27. However, the skilled person seeking to carry out the claimed method and to efficiently recognise wakeup and command words would find in D2, paragraphs [0007] and [0008], the disclosure of what abilities and advantages are provided by the combined use of a DSP and a Central Processor Unit in the context of the invention.

28. Reference is also made to section 18 in the decision under appeal. The Board agrees that the skilled person would have indeed considered document D2 and recognised the advantages resulting from the use of distinct processing units in terms of energy consumption.

29. The subject-matter of claim 1 is not inventive in view of a combination of documents D1 and D2 (Article 56 EPC).

Third auxiliary request

30. Auxiliary request 3 was filed for the first time with the statement of grounds of appeal. It constitutes an amendment to the appellant's case, the admission of which is a matter of discretion. The decision depends, inter alia, on the suitability of the amendments to address the issues which led to the impugned decision (Article 12(4) RPBA 2020).

31. Claim 1 comprises the further indications that

> *a time point for starting to speak the wakeup information is t0, a time point for finishing speaking the wakeup information is t1, and a time point for starting to speak the first portion of command word is t2;* and
> *a time point for enabling the speech recognition function is t3, and the first preset duration is from t0 to t3*.

32. It appears that such a specifications, are not suitable to address the issue of lack of inventive step in view of D1. In fact, D1 also discloses a time sequence of events such as the one that appears intended at being

defined in claim 1 of auxiliary request 3 (paragraph [0049]). In the embodiment disclosed in paragraph [0049], the utterance of the wake-up signal implicitly covers a time period between the start of the wake-up word and its ending. It may be followed by a period of silence that is itself followed by the utterance of the command word, the speaking of the command word being the point in time at which the speech recognition function is enabled.

33. Hence the amendment is not suitable to address the issue of lack of inventive step in view of D1, identified above having regard the main request.

34. The Board decides thus not to admit auxiliary request 3 into the proceedings (Article 12(4) RPBA 2020).

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chair:

D. Meyfarth A. Medeiros Gaspar

Decision electronically authenticated