

**Internal distribution code:**

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

**Datasheet for the decision  
of 5 February 2019**

**Case Number:** T 2123/16 - 3.5.07

**Application Number:** 10763927.0

**Publication Number:** 2483889

**IPC:** G11B27/28, G11B27/32,  
G06F17/30, H04N21/231,  
H04N21/233, H04N21/472,  
H04N21/81, H04N5/445

**Language of the proceedings:** EN

**Title of invention:**

Systems and methods for identifying audio content using an  
interactive media guidance application

**Patent Proprietor:**

Rovi Guides, Inc.

**Headword:**

Audio-content identification/ROVI GUIDES

**Relevant legal provisions:**

EPC Art. 56, 84, 123(2)

**Keyword:**

Amendments - added subject-matter (no)  
Claims - clarity (yes)  
Inventive step - (yes)

**Decisions cited:**

G 0003/14



**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 2123/16 - 3.5.07

**D E C I S I O N**  
**of Technical Board of Appeal 3.5.07**  
**of 5 February 2019**

**Appellant:** Rovi Guides, Inc.  
(Patent Proprietor) 2160 Gold Street  
San Jose, CA 95002 (US)

**Representative:** Pisani, Diana Jean  
Haley Guiliano International LLP  
Central Court  
25 Southampton Buildings  
London WC2A 1AL (GB)

**Decision under appeal:** **Decision of the Opposition Division of the  
European Patent Office posted on 13 July 2016  
revoking European patent No. 2483889 pursuant to  
Article 101(2) and (3)(b) EPC**

**Composition of the Board:**

**Chairman** R. Moufang  
**Members:** R. de Man  
M. Jaedicke

## **Summary of Facts and Submissions**

- I. The patent proprietor (appellant) appealed against the decision of the Opposition Division revoking European patent No. 2 483 889, granted on European patent application No. 10763927.0 which had been published as international application WO 2011/041259.
- II. The opponent (respondent) had opposed the patent as a whole on the basis of Article 100(a) EPC (lack of novelty and lack of inventive step), Article 100(b) EPC (insufficiency of disclosure) and Article 100(c) EPC (added subject-matter).
- III. The decision cited the following documents:
- D1: WO 02/27600 A2, published on 4 April 2002;
- D4: WO 2004/090752 A1, published on 21 October 2004;
- D5: M. Fink et al.: "Mass personalization: social and interactive applications using sound-track identification", *Multimedia Tools and Applications*, January 2008, Vol. 36, No. 1-2, pp. 115-132, published online on 21 December 2006;
- D6: C.-W. Chen et al.: "Content identification in consumer applications", *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo*, 28 June 2009, pp. 1536-1539;
- D7: US 2002/0116195 A1, published on 22 August 2002;
- D8: EP 1 855 216 A2, published on 14 November 2007;
- D9: US 2007/0162436 A1, published on 12 July 2007;
- D10: WO 02/05542 A2, published on 17 January 2002;
- D11: US 2009/0041418 A1, published on 12 February 2009;  
and
- D12: WO 2009/005760 A2, published on 8 January 2009.

The Opposition Division decided that the subject-matter of claim 1 of the patent as granted (main request) lacked inventive step over document D1, that the subject-matter of claim 1 of the first auxiliary request infringed Article 123(2) EPC, that the subject-matter of claim 1 of the second auxiliary request lacked inventive step over document D1, that the third auxiliary request was not to be admitted into the proceedings, and that claim 1 of the fourth auxiliary request did not meet the requirements of Article 84 EPC.

At the oral proceedings, the Opposition Division expressed the view that the grounds of insufficiency of disclosure and added subject-matter (Article 100(b) and (c) EPC) did not prejudice maintenance of the patent as granted.

- IV. In its statement of grounds of appeal, the appellant maintained its main request (patent as granted) and replaced its auxiliary requests with first and second auxiliary requests based on the second and third auxiliary requests considered by the Opposition Division (with minor amendments) and a new third auxiliary request.
- V. In its reply, the respondent stated that it maintained all its arguments as previously submitted with regard to each of paragraphs (a), (b) and (c) of Article 100 EPC.

It argued that the patent as granted included added subject-matter and that the subject-matter of its independent claims lacked inventive step over document D1. With respect to the first auxiliary request, it argued that the combination of claims 1

and 2 (and 7 and 8) lacked any basis in the application as filed and that the subject-matter of the independent claims lacked inventive step over document D1. It also argued that the second and third auxiliary requests should not be admitted into the proceedings, did not have any basis in the application as filed, and complied neither with Article 56 EPC nor with Article 84 EPC.

The respondent also pointed out that it had not been heard at the oral proceedings on any document other than D1. It therefore requested the Board, in case the appeal was not dismissed, to remit the case to the department of first instance for further consideration, in particular with regard to document D8.

- VI. In a letter dated 9 November 2017, the respondent withdrew its opposition.
- VII. In a communication accompanying the summons to oral proceedings, the Board summarised the points to be discussed. In particular, it discussed inventive step with respect to document D1.
- VIII. In a letter dated 4 January 2019, the appellant filed copies of the claims of a main request and first to ninth auxiliary requests. The main request and second to fourth auxiliary requests were said to be identical to the pending main request and first to third auxiliary requests. The fifth to ninth auxiliary requests corresponded to the main request and first to fourth auxiliary requests with an amendment to address an added-matter objection.
- IX. In the course of oral proceedings held on 5 February 2019, the appellant replaced the text of its

third auxiliary request, which inadvertently included dependent claims conflicting with the independent claims, with the text of the second auxiliary request filed with the statement of grounds of appeal and made this request its sole substantive request. At the end of the oral proceedings, the chairman pronounced the Board's decision.

X. The appellant requested that the decision under appeal be set aside and that the patent be maintained in amended form on the basis of the second auxiliary request filed with the statement of grounds of appeal.

XI. Independent claim 1 of the sole substantive request reads as follows:

"A media guidance system comprising:

means (906) for receiving a video program that includes one or more audio assets;

means (924) for receiving user input including a request to identify an audio asset playing within the video program, wherein the audio asset is a song or background music;

means (906) for determining a segment of the audio asset where interference from audio data not related to the audio asset is minimized, wherein the means (906) for determining the segment of the audio asset where interference is minimized includes means for analyzing metadata included with the video program to determine when audio data not related to the audio asset is present, the metadata being closed captioning data;

means (910) for generating an audio signature based on the segment of the audio asset;

means for identifying the audio asset by comparing the generated audio signature to known audio signatures of a plurality of known audio assets within a data

store, wherein each known audio signature is associated with a known audio asset and wherein the data store includes audio asset information associated with a plurality of known audio assets; and

means for providing audio asset information associated with the identified audio asset to a user interface for display."

Claim 2 reads as follows:

"The system of claim 1, wherein the audio asset information includes at least one of audio title, artist, album, album art, genre, type, audio asset location in video program, play time of audio asset in video program, start time of audio asset, end time of audio asset, and audio quality."

Claims 3 to 5 are dependent on claim 1.

Independent claim 6 reads as follows:

"A method for identifying audio using a media guidance application, the method comprising:

receiving a video program that includes one or more audio assets;

receiving user input including a request to identify an audio asset playing within the video program, wherein the audio asset is a song or background music;

determining a segment of the audio asset where interference from audio data not related to the audio asset is minimized, wherein determining the segment of the audio asset where interference is minimized includes analyzing metadata included with the video program to determine when audio data not related to the



audio asset is present, the metadata being closed captioning data;

generating an audio signature based on the segment of the audio asset;

identifying the audio asset by comparing the generated audio signature to known audio signatures of a plurality of known audio assets within a data store, wherein each audio signature is associated with a known audio asset and wherein the data store includes audio asset information associated with a plurality of known audio assets; and

providing audio asset information associated with the identified audio asset to a user interface for display."

Claim 7 reads as follows:

"The method of claim 6, wherein the audio asset information includes at least one of audio title, artist, album, album art, genre, type, audio asset location in video program, play time of audio asset in video program, start time of audio asset, end time of audio asset, and audio quality."

Claims 8 to 10 are dependent on claim 6.

## **Reasons for the Decision**

1. The appeal complies with the provisions referred to in Rule 101 EPC and is therefore admissible.
2. *The invention*

The invention relates to identifying an audio asset included in a video program, in particular a song or

background music, by generating an "audio signature" from a segment of the audio asset and looking up the audio signature in a database of audio signatures of known audio assets.

To improve the identification process, the invention proposes generating the audio signature from a segment of the audio asset "where interference from audio data not related to the audio asset is minimized". This segment is determined by analysing closed-captioning data included in the video program.

As explained in paragraph [0148] of the patent, the closed-captioning data will typically be used to ensure that the segment is chosen to be one where dialogue is not present.

3. *Admission of the sole substantive request*

3.1 The present sole substantive request corresponds to the third auxiliary request filed in the oral proceedings before the Opposition Division with a minor amendment. The latter request was not admitted into the opposition proceedings because it had been filed after the time limit set in accordance with Rule 116(1) and (2) EPC and because the amendments did not *prima facie* overcome the inventive-step objection raised against the then main request.

3.2 Since the respondent has now withdrawn its opposition, and since the Board is able to deal with the request, the Board exercises its own discretion under Article 12(4) RPBA to admit the request into the appeal proceedings.

4. *Added subject-matter*

4.1 Independent system claim 1 is a combination of original claim 30 (corresponding to original independent claim 27 with the additional features of dependent claims 28, 29 and 30), the additional features of dependent claims 35, 36 and 38, and the features "wherein the audio asset is a song or background music" and "the metadata being closed captioning data". The latter two features are based on paragraph [0161] of the international publication.

4.2 Original claim 35 is dependent on original claim 30 only via claim 31. The former respondent had argued that the omission of the additional feature of original claim 31 ("wherein the means for identifying the audio asset includes means for identifying a match between the generated audio asset signature and one of the plurality of known audio signatures") resulted in subject-matter extending beyond the content of the application as filed.

However, the Board takes the view that identifying an audio asset by finding a "match" of audio signatures is equivalent to identifying an audio asset by "comparing" audio signatures. Indeed, the skilled person understands that, in the context of audio signatures, the term "match" cannot be understood as being limited to an identical match; the additional feature of dependent claim 31 does not therefore restrict the additional feature of dependent claim 30.

4.3 In its communication the Board questioned whether paragraph [0161] was a sufficient basis for the addition of the feature "the metadata being closed captioning data", as paragraph [0161] related to step

1306 of Figure 13, whereas the claim feature "determining a segment of the audio asset where interference from audio data not related to the audio asset is minimized" related to paragraph [0162] and step 1308 of Figure 13.

But as the appellant pointed out, the statement in paragraph [0162], that "[a]n optimal audio signal segment may be one in which interference or unwanted audio content (such as background noise) is at a minimum", also applies to paragraph [0161], which discloses that closed-captioning data can be used to determine "an optimal audio signal segment".

- 4.4 Independent method claim 6 finds a basis in the corresponding original claims 56, 61, 62 and 64 and paragraph [0161].
  
- 4.5 Dependent claims 2 and 7 are based on original claims 33 and 59. In its reply, the former respondent had submitted - with respect to the then first auxiliary request - that the combination of claims 1 and 2 infringed Article 123(2) EPC because claim 2 confused "audio asset information" with the "metadata" of claim 1. In particular, whereas start/end/play time is referred to as metadata in paragraph [0161] of the original description, it is referred to as audio asset information in claim 2.

Since claim 2 has a literal basis in original claim 33, and since the application as filed uses the terms "metadata" and "audio asset information" for different purposes and consistent with claims 1 and 2, the Board fails to see why the combination of claims 1 and 2 (or of claims 6 and 7) would infringe Article 123(2) EPC.

4.6 Neither the Opposition Division nor the former respondent argued that the remaining dependent claims added subject-matter, and the Board sees no reason to question that they have a basis in the application as filed.

4.7 The Board is therefore satisfied that the sole substantive request complies with Article 123(2) EPC.

5. *Sufficiency of disclosure*

In its reply, the former respondent had maintained its objection under Article 100(b) EPC, but only by formally maintaining all its arguments submitted in the first-instance proceedings. The Board sees no reason to disagree with the view expressed in the communication annexed to the summons to oral proceedings before the Opposition Division that the patent disclosed the invention sufficiently clear and complete for it to be carried out by the skilled person.

6. *Clarity*

6.1 The former respondent had argued in its reply that the addition to the independent claims of the feature "the metadata being closed captioning data", which had been taken from the description, resulted in a lack of clarity due to missing essential features. In particular, it was not clear from the independent claims how, in general, an audio segment was to be determined by analysing metadata comprising closed-captioning data.

6.2 In the Board's view, the skilled person would have no difficulty in finding ways to use closed-captioning data to determine a segment that can be reasonably

expected to be relatively free from interference. Indeed, closed-captioning data normally directly relates to spoken text in the audio data. The Board therefore does not agree that the amendment is objectionable under Article 84 EPC.

7. *Inventive step*

7.1 In its notice of opposition, the former respondent argued that the subject-matter of claim 1 as granted lacked inventive step starting from any of documents D1, D7 and D8.

7.2 Document D1 discloses a system allowing users in a noisy environment to identify a sound, such as music, from among a number of sound recordings present in a database of recordings (see abstract and pages 1 to 3).

The user may capture a signal sample using a mobile phone from a "media experience (including audio and video)" that he is monitoring on, for example, a television (page 27, lines 8 to 16; Figure 4). The captured sample is relayed by the user to the interactive voice-response (IVR) unit 450 (page 28, lines 1 and 2; Figure 4).

Alternatively, the sample may be captured directly from the media distribution network 420 that transmits the source signal (page 27, lines 18 to 21). In this case, monitoring by the user may not be necessary (page 27, lines 21 to 23).

7.3 The IVR 450 "derives information or characteristics" of the received sample "including the identification of content contained therein (for example, the song ID)" (page 28, lines 13 to 15). This derived

information is returned to the user's mobile phone for display (page 28, lines 15 to 17; page 29, lines 18 and 19; page 35, lines 7 to 14).

- 7.4 The process of identifying the signal is performed in signal identification block 110, which is shown in Figure 2. It involves computing fingerprints from the captured sample at "landmarked" time points (page 30, lines 5 to 13). The computed fingerprints are matched with known song fingerprints stored in a database, which themselves are associated with song landmark and song ID values (page 30, lines 13 to 16). For each song ID, the set of matches is "scanned for linear correspondences in the pairs of landmarks and scored according to best fit". The song ID with the highest score wins.

The Board considers that the set of "landmarked" time points and corresponding fingerprints qualifies as an "audio signature" and that the process of scanning for linear correspondences and scoring according to best fit qualifies as "comparing the generated audio signature to known audio signatures".

- 7.5 The subject-matter of claim 1 hence differs from what is disclosed in document D1 in that, with the help of closed-captioning data included in the video program, a segment of the audio asset, i.e. of a song or background music, is determined (for the purpose of capturing a sample) "where interference from audio data not related to the audio asset is minimized".

By generating the audio signature from a segment of the audio data that is (relatively) free from interference from unrelated audio signals, the chances that the song or background music is correctly identified are

improved. The distinguishing features therefore solve the problem of improving the recognition of songs or background music.

- 7.6 In its decision, the Opposition Division had stated that the analysis of closed-captioning data in general did not address the problem of finding an appropriate segment and that claim 1 therefore defined a desideratum (which was apparently found to be obvious).

The Board notes that claim 1 is limited to determining a segment "where interference from audio data not related to the audio asset is minimized", which means that the claimed analysis of closed-captioning data is to be interpreted as an analysis that at least can be reasonably expected to help in finding such a segment. The distinguishing features therefore do address the problem of finding an appropriate segment. And they do not define a mere desideratum, as they not only claim the problem but also express the specific solution of analysing closed-captioning data.

- 7.7 Since document D1, on page 10, lines 19 to 24, explains that recognition of a song may sometimes fail "due to very short or noisy samples", the Board considers that the skilled person would realise that song recognition in document D1 can be improved by capturing a sample that is relatively free from interference from other audio data. Although document D1 focuses on noisy environments and, on page 24, lines 16 to 19, mentions "background noise (such as that encountered while in a moving car), talking voices, transmission errors and impairments, interference, time warping, compression, quantization, filtering" as sources of signal degradation, the Board also judges that the skilled person was aware that songs or background music in



television programmes or movies often overlaps with actor voices or voice-overs.

7.8 However, the Board is not convinced that the skilled person, on the basis of only his common general knowledge, would consider analysing closed-captioning data included in the video program for the purpose of identifying a segment of audio data free from interference of unrelated audio signals such as actor voices or voice-overs. Closed-captioning data included in video programs was well known at the priority date, but the data served the purpose of informing hearing-impaired or foreign-language viewers of the content of spoken text, not of automatically selecting a segment of audio data free from voices or other types of unrelated audio signal.

7.9 The only documents on file relating to the use of closed-captioning data are documents D4 and D10.

Document D4, on page 18, lines 1 to 11, discloses that the potential boundaries of songs in a multimedia stream can be identified by analysing closed-captioning data. Boundary detection is performed for the purpose of segmenting the stream into separate music videos (page 12, lines 11 to 13). Each segmented music video is identified by detecting and textually analysing the song's chorus (page 12, lines 13 to 17; page 21, line 1, to page 22, lines 26). This process does not involve the generation of audio signatures.

Although in the case of content that includes song announcements at the beginning of a song, knowing when a song starts can be helpful to locate a segment later in the song that is free from announcements, boundary detection in document D4 is not performed for that

purpose. The skilled person, faced with the problem of improving song recognition in document D1, for example by identifying a segment of audio that is free of unrelated audio signals, would not therefore find in document D4 a specific pointer that that problem can be solved by analysing closed-captioning data.

Document D10, on page 10, lines 7 to 13, merely discloses that closed-captioning data can be used to determine the current subject-matter of a program and therefore likewise does not lead the skilled person to the claimed solution.

7.10 Hence, starting from document D1 the skilled person would not arrive at the subject-matter of claim 1 in an obvious manner.

7.11 Document D7, in Figure 3 and paragraphs [0030] to [0036], discloses an audio identification process for automatically identifying audio content on the basis of an audio sample recorded by a user. A unique signature is generated for the audio sample and matched with audio content in an audio-content database (paragraphs [0033] and [0037] to [0070]; Figures 4A to 8). The audio may be obtained from a recorded video program (paragraph [0024]).

Document D8, in paragraphs [0023] to [0037], discloses an audio analysis process that determines when a speaking voice stops and a song starts and that matches the leading portion of a song with leading portions of known songs stored in a song database to extract metadata about the song, including the song title.

Neither of these documents discloses or hints at the use of closed-captioning data for determining a segment

of audio data that is relatively free of interference. These documents are therefore not closer to the invention than document D1.

7.12 The subject-matter of independent claim 1 and corresponding independent claim 6 therefore involves an inventive step (Article 56 EPC).

8. *Remittal*

8.1 Since the amended claims of the sole substantive request comply with the provisions of the EPC, the case is to be remitted to the Opposition Division with the order to maintain the patent in amended form.

8.2 However, the description still needs to be adapted to the amendments made to the independent claims. In this respect, the Board notes that the description as granted appears not to have been fully adapted to the granted claims. This being a question of support of the claims by the description as required by Article 84 EPC, in accordance with decision G 3/14 (OJ EPO 2015, A102) the Opposition Division is to limit its adaptation of the description to the changes necessitated by the amendments made to the claims as granted.

## Order

### For these reasons it is decided that:

1. The decision under appeal is set aside.
2. The case is remitted to the department of first instance with the order to maintain the patent in amended form on the basis of the claims filed as second auxiliary request with the statement of grounds of appeal dated 23 November 2016 and with a description and drawings yet to be adapted.

The Registrar:

The Chairman:



I. Aperribay

R. Moufang

Decision electronically authenticated