

**Internal distribution code:**

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

**Datasheet for the decision  
of 15 December 2023**

**Case Number:** T 0748/19 - 3.5.06

**Application Number:** 16205865.5

**Publication Number:** 3340103

**IPC:** G06K9/00, G06K9/46, G06K9/62

**Language of the proceedings:** EN

**Title of invention:**  
METHOD FOR IDENTIFYING EVENTS IN A MOTION VIDEO

**Applicant:**  
Axis AB

**Headword:**  
Camera processing data/Axis

**Relevant legal provisions:**  
EPC Art. 83  
RPBA 2020 Art. 13

**Keyword:**  
Sufficiency of disclosure - effect not established over the  
full claimed scope - see reasons 10 to 13

**Decisions cited:**  
T 2773/18, T 1983/19, T 0814/20, T 0149/21

**Catchword:**



**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 0748/19 - 3.5.06

**D E C I S I O N**  
**of Technical Board of Appeal 3.5.06**  
**of 15 December 2023**

**Appellant:** Axis AB  
(Applicant) Gränden 1  
223 69 Lund (SE)

**Representative:** AWA Sweden AB  
Box 5117  
200 71 Malmö (SE)

**Decision under appeal:** **Decision of the Examining Division of the  
European Patent Office posted on 28 September  
2018 refusing European patent application No.  
16205865.5 pursuant to Article 97(2) EPC.**

**Composition of the Board:**

**Chairman** M. Müller  
**Members:** T. Alecu  
B. Müller

## **Summary of Facts and Submissions**

- I. The appeal lies from the decision of the Examining Division to refuse the application.
- II. With the statement of grounds of appeal the Appellant requested that the decision be set aside and that a patent be granted on the basis of one of the two auxiliary requests underlying the decision under appeal. These requests were refused for a lack of inventive step starting from D1: US 2016/0171852 A1.
- III. In a communication accompanying a summons to oral proceedings the Board informed the Appellant of its preliminary opinion that, for both requests,
- (a) the subject matter claimed was not obvious starting from D1; but that
  - (b) the claimed invention was not sufficiently disclosed in the sense of Article 83 EPC.
- IV. In a reply to this communication the Appellant filed one main request and three auxiliary requests. The main request corresponded to the first auxiliary request underlying the decision, and the first auxiliary request to the second auxiliary request underlying the decision. The second and the third auxiliary request were new. During the oral proceedings the Appellant filed a fourth and a fifth auxiliary request.
- V. Claim 1 of the main request defines:
- Method for identifying events in a scene captured by a monitoring motion video camera, the method comprising two identification processes, a temporary*

*identification process and a long-term identification process;*

*the temporary identification process includes:*

*identifying events in the captured scene by analysing pixel data from captured image frames using an event identifying operation module;*

*storing camera processing data relating to time periods of the capturing of the image frames subjected to the pixel data based identification of events, wherein the camera processing data comprises one or more of: an amount of data generated by a temporally compressing video encoder, a value indicating an auto focus distance, a value indicating the setting of an auto white balance function, values relating to auto exposure settings, such as: aperture, shutter time, or gain, electronic image stabilisation data, a value of signal to noise ratio, a value indicating the contrast in the captured frame, a value indicating the data rate of encoded data sent to a communication network, a value indicating CPU usage, and a value indicating memory usage and position data from a PTZ head connected to the camera; and*

*training a neural network based event identifying operation using the stored camera processing data relating to the time periods of the capturing of the image frames subjected to the identification of events in the captured scene by analysing pixel data from captured image frames as input and the identified events as the correct classification of an event resulting from the neural network based event identifying operation; and*

*the long-term identification process includes:*

*storing camera processing data relating to time periods of capturing of image frames captured subsequent to the image frames used for the analysis of pixel data; and*

*identifying events in the captured scene by inputting the stored camera processing data relating to the time periods of capturing of image frames captured subsequent to the image frames used for the analysis of pixel data to the trained neural network based event identifying operation;*  
*wherein the temporary identification process is executed during a predetermined time period and wherein the long-term identification process is executed after the predetermined time period has expired.*

- VI. In claim 1 of the first auxiliary request the following feature is added at the end:

*and wherein the temporary identification process is executed by a device connected to the motion video camera via a network.*

- VII. Claim 1 of the second auxiliary request differs from that of the main request by specifying that the steps of training a neural network and event identifying use a time series of camera processing data as follows

*training a neural network based event identifying operation using a time series of the stored camera processing data ...*

*...*

*identifying events in the captured scene by inputting a time series of the stored camera processing data..*

- VIII. Claim 1 of the third auxiliary request contains both sets of amendments, those in the first, and those in the second auxiliary request.

- IX. Claim 1 of the fourth auxiliary request differs from that of the main request in that the list of possible

types of camera processing data has been reduced to one element as follows:

*wherein the camera processing data comprises one or more of: an amount of data generated by a temporally compressing video encoder such as a value indicating the data rate of encoded data sent to a communication network.*

- X. Claim 1 of the fifth auxiliary request differs from that of the fourth request in that a feature is inserted as follows:

*... trained neural network based event identifying operation;  
wherein the events relate to vehicles arriving in scene, and  
wherein the temporary identification process is executed ...*

## **Reasons for the Decision**

### *The application*

1. The application relates to identifying events in scenes under surveillance. According to the application, automated systems for this purpose process a substantial amount of data and require a substantial amount of processing power. The objective of the invention is to reduce these amounts (page 1, lines 1 to 21).
- 1.1 The application proposes to use a two step method for event detection. In the first step, called a temporary identification process, "well-known" image data based identification methods are used to detect events and

register corresponding *"camera processing data"*. The registered data and the events are used as input/output to train a neural network. In the second step, called the long-term identification process, events are detected using the trained neural network on the basis of camera processing data alone (paragraph bridging pages 5 and 6).

1.2 According to the application, camera processing data is not *"direct image data"* and may be data indicating the encoding data rate, the contrast, exposure settings and so forth (paragraph bridging pages 4 and 5). So the advantage of *"basing event identification on camera processing data is that processing power required for event detection is decreased as the amount of data that is processed is much lower than for pixel data representing images"* (page 5, lines 7 to 10). For instance (page 12, lines 18 to 29), using the data rate as *"camera processing data"* according to the invention is said to cause a *"radical decrease in the amount of data required"*, e.g. from 6000 Kb/s to 0.12 Kb/s, the former being the amount of data needed to encode the *image data*, and the later being the amount of data needed to encode the *data rate*.

1.3 The considered events are exemplified as follows (page 9, line 30, to page 10, line 2): *"vehicle arriving in scene, type of vehicle arriving in scene, speed of vehicle arriving in scene, number of vehicles arriving in the scene, etc. The event pixel based identifying operations may also or instead relate to person arriving in scene, animal arriving in a scene, a specific object arriving in the scene, etc., also including the variations presented for vehicles above. Further, the pixel based event identifying operations may include identifying that motion occurs in the scene, loitering,*



*large crowd gathering, traffic jam, snow, rain, smog, deteriorated image due to dirt on the housing, lens out of focus, anomalies, etc."*

- 1.4 The neural networks used according to the invention may be of different kinds (page 5, lines 18 to 21). On page 7 of the description it is stated, in the context of what appears to be the most detailed example, that a recursive neural network (RNN) may be employed, and that an RNN may process univariate or multivariate time series: *"For example, processing of a univariate time series may be processing of a sequence of values indicating the data rate of an encoded video stream. An example of processing of a multivariate time series may then be processing of a sequence of vectors where each vector include a value indicating the data rate of an encoded video stream and a value indicating white balance. Any combination of camera processing data are possible and the combinations should not be limited to combining only two different data types but may very well combine a plurality of data types."*

#### *Inventive step*

2. Document D1 relates to surveillance and more specifically to classifying events in images captured using video cameras (paragraph 3). It teaches the use of a two-stage approach (see figure 2). The first stage carries out motion detection (referred to as "motion differencing") and produces so-called "motion boxes" surrounding image areas where there is motion. In a second stage, feature vectors are extracted from these areas and fed to a neural network classifier (paragraphs 32, 42, and 44 to 51).

3. In its analysis the Examining Division equated the motion boxes of D1 with camera processing data, more precisely with the feature "*an amount of data generated by a temporally compressing video encoder*".
- 3.1 The Board does not find this interpretation reasonable. The feature recites an amount, i.e. a quantity, which is not the image data captured, but only a measure of the quantity of data needed to encode the image data. It is this quantity which serves as an input to the neural network. In D1 it is clear that it is the content, i.e. pixel intensity values, of the motion boxes that are processed to obtain feature vectors used as input to the neural network.
- 3.2 This (pixel intensity values) is precisely that with which the application, and the claims, contrast the camera processing data. Thus the Board agrees with the Appellant that D1 does not disclose the use of camera processing data as claimed.
- 3.3 The Board also does not see any hint in D1 towards such use, and has no reason to believe such use to be obvious in view of the common knowledge of the skilled person.

*Sufficiency of disclosure (Article 83 EPC): main request*

*The preliminary opinion*

4. In its preliminary opinion the Board remarked that claim 1 defined a method for identifying events without any definition of the events which are classified. Hence it covered in principle any event in video surveillance, at least those enumerated in the application (see above), and possibly others.

- 4.1 It did not appear credible to the Board that the types of data claimed, separately or in combination, contained the information needed to distinguish between all types of events (e.g. sufficient to distinguish between a person and an animal, and to detect the car type, and to identify loitering, etc.).
- 4.2 This was not credible because the data which served as input to the neural network, contained, as the application explained, no image value information. A change in the data rate, as discussed on pages 11 and 12 of the description, might indicate that something has changed, but not necessarily what has changed (e.g. one large object occurring in the image might change the data rate in the same way as two smaller ones, and the same might be the case for one small, but fast-moving object and a large, but slower object, etc.).
- 4.3 The application also remained very generic as to the implementation of the proposed concept, providing no detailed example, and no results.
- 4.4 Thus the application was limited to presenting a concept, the validity of which was already questionable for theoretical reasons, and which was also not established by any evidence. The application did not contain even a single detailed embodiment. Hence the application did not disclose the claimed invention in a manner sufficiently clear and complete for it to be carried out by the skilled person.

*The Appellant's arguments*

5. During the oral proceedings the Appellant submitted the following arguments in response.

6. The Appellant understood the objection as being one of insufficient disclosure over the full breadth of the claims, and agreed that as such it was not unreasonable. It was true that the claim was based on a limited number of experiments, and that the claim scope included cases which were not covered by these experiments. But the examples were sufficient, because they illustrated how the invention was to be implemented in general and determined the expectations of the skilled person.
  
7. Implementation using a neural network, i.e. choosing a suitable architecture and training the network, was straightforward for the skilled person. If a parameter (a camera processing data type) was not useful for the scene and events considered, it would not be used in real life, because the training would not converge. The temporary identification process claimed provided scene-specific ground truth, so it reduced the claimed scope and training requirements.
  
8. Moreover, the skilled person would have an understanding of which parameters were affected by which of the considered events. For illustration, the Appellant reported the following observations made by the inventors:
  - Auto focus changes if an object enters the scene and moves in a direction towards or away from the camera.
  - Auto-white balance changes if the color composition of a scene changes, which could be caused by an object entering the scene or by a change of weather conditions.
  - Electronic image stabilization changes with conditions in the scene, e.g. the wind picking up or a passing lorry causing vibrations.

- Contrast changes with conditions in the scene or with conditions of imaging optics, caused by someone tampering with the camera.
- Data rate, memory usage, and data generation are indicative of the amount of motion in the scene, i.e. the size of moving objects, the direction of moving objects, or the speed of moving objects.

9. The board in its preliminary opinion had a wrong understanding of what the skilled person in the art would expect. The skilled person had technically reasonable expectations, and would, for instance, not expect that the claimed invention could distinguish car colours or make. The person skilled in the art would not read the claim to cover such cases, i.e. where it is technically unreasonable to expect success.

*The Board's position*

10. The Board agrees that the skilled person would be able to choose an architecture for a neural network and carry out its training, if a set of camera parameters and events to be identified were defined. However, this is insufficient to establish compliance with Article 83 EPC in the present case. That is because the claimed invention is not characterised only by it being a neural network, but also, perhaps primarily, by its purpose, that of being able to (reliably) identify events. This purpose must be achieved in different scenarios, also corresponding to different sets of events.

11. The Board accepts that there are simple cases for which the claimed purpose can be achieved on the basis of the teaching in the application and common general knowledge. For instance, motion can be detected using the encoding data rate. However, it is not at all clear

whether the claimed method can be successful in other, possibly more complex scenarios.

12. The Appellant's argument concerning these other scenarios is that the cases where the method turns out not to work (after trial and error), or "technically unreasonable" cases, where the skilled person does not expect the invention to work, are non-detrimental for the purposes of Article 83 EPC. This should be, and here the argument remains ambiguous,
- (a) because the skilled person would not consider them to be covered by the claimed scope, or
  - (b) possibly, even if covered, because the claimed invention is sufficiently disclosed anyway.

*Disclosure over the full scope of the claim*

13. The success of this argument depends on the validity, and, if valid, the interpretation, of the principle adopted in the jurisprudence of the Boards of Appeal that the disclosure requirement according to Article 83 EPC applies to the full scope of the claims.
- 13.1 The Board agrees with this principle. That for which protection is sought, i.e. the claims, and, if the application is successful, granted, must correspond to the teaching provided by the application. The protection afforded by the claims must correspond to that which the application makes available to the skilled person by way of disclosing it in a manner sufficiently clear and complete for it to be carried out.
- 13.2 It is the applicant who drafts the claims to define the protection it seeks. If it is clear that the claim intentionally covers certain matter, then this matter is part of the claimed invention and the fact that it

cannot be carried out cannot be ignored for the assessment of whether the disclosure is sufficient. Therefore, the Board disagrees with the idea that the claims must, a priori, be construed to exclude instances which the skilled person would "not expect to work", or which only after trial and error turn out not to work.

- 13.3 However, most, if not all, claims may be taken to cover, instances which cannot be carried out or which "do not work", for instance where "technically unreasonable" choices of parameters are made, and specifically ones for which the skilled person would foresee that the claimed invention would not work. The evaluation of sufficiency should take this into account.
- 13.4 In particular, it has been suggested in the jurisprudence of the Boards of Appeal (see also the section below on related decisions of the Boards of Appeal), and argued by the Appellant, that "technically unreasonable" instances of the claimed subject-matter should not be detrimental to sufficiency under Article 83 EPC. The Board agrees that this may sometimes be the case. When the "technically unreasonable" instance is a contrived one, i.e. one which the skilled person would, in view of the provided teaching and of the claimed generalisation, not consider the claim to cover, this instance should not be covered.
- 13.5 However, non-contrived instances, so where it is clear that the claim intends to cover them, should be taken into account for assessing sufficiency of disclosure (see point 13.2 above). That is the case even if they are "technically unreasonable" (for example, a claim to a teleporting machine).

*The present case*

14. The application considers a large range of events to be detected by means of the claimed method, and the claims themselves do not limit the events at all. Also, the application does not provide any guidance to the skilled person as to which events can or cannot be detected, and hence no guidance for the skilled person on how to distinguish a contrived case from a non-contrived one. It is therefore the Board's view that the claims are meant to cover at least the identification of events listed in the application, in their corresponding scenarios.
  
15. For each considered scenario and corresponding set of events, the skilled person needs to define a set of parameters allowing the identification of said events. A large choice of camera parameters used for event detection are disclosed and claimed. But, it is not trivial to see which (or if any) camera parameters contain the information needed for identification. Contrary to the Appellant's statements the application does not provide any clear example of such selection. There is also no guidance provided as to which events may be detected based on which set of parameters.
  
16. Notably, the list of illustrative examples produced by the appellant (see above, point 8), is not part of the application, nor can it be considered to be part of the common general knowledge of the skilled person.
  - 16.1 However, even if the latter were assumed to be the case, it would not lead to a clear understanding as to which events may be identified or how. For instance, a given parameter may be influenced by different events in the same way. The list of examples itself shows this



by mentioning different possible causes for a variation in observed camera processing data (e.g. the wind or a passing truck for electronic image stabilization parameters). So by simply observing that a certain camera parameter varies in some way, one cannot infer the cause of this variation. In mathematical terms, this is an underdetermined (inverse) problem.

17. The same point can be made in view of events mentioned in the application. For instance, it is not clear if or how the claimed method can distinguish between one large object entering the scene and two smaller ones (see point 4.2 above). It is also not at all clear which parameters are related to more complex events such as loitering or the identification of a specific object entering the scene.
18. To carry out the presently claimed invention, the skilled person would thus have to define, without theoretical or practical guidance from the application (e.g. by way of theoretical considerations or concrete examples), in sufficient detail, as a function of the surveillance scenario, the events to be identified, the acquisition setup (e.g. lighting, perspective, resolution, etc.), and test which of the mentioned camera parameters would allow which events to be detected in the given circumstances. The Board considers this to be an undue burden on the skilled person having to carry out the invention. In the Board's judgement, it amounts more to conducting a research program than to carrying out the invention in accordance with the teaching provided.
19. The Board concludes that the application does not disclose a method of identifying events using a neural network trained with camera processing data in a manner

sufficiently clear and complete for it to be carried out by the skilled person and find this to be an unjustified generalisation from the teaching of the application. In the Board's judgement, this is a deficiency under Article 83 EPC.

*Related Board of Appeal decisions*

20. In case T 814/20 (reasons 13.3 to 13.5), which, as the present case, also related to video surveillance, this Board allowed a claimed invention even though it did not work in all conceivable circumstances, because it considered that the skilled person would understand, from the claims and in view of the description, the kind of situations for which the method was designed.
- 20.1 The situation in T 814/20 however was notably different from that in the present case. There, the algorithm was clearly defined (reasons 13.1). Here, without a definition of the camera parameter set, this is not the case. There, the Board found that the theoretical assumption were sound (reasons 13.3), and that the technical effect was proven for a test scenario (reasons 13.2). Here, this is not the case. There, it was clear to the Board what the generalization in the claim was meant to cover, and it held it to be credible that the method would work when so generalized (reasons 13.4 and 13.5).
21. The Board is also aware of recent decisions of the Boards of Appeal discussing the question whether the principle that sufficiency of disclosure is to be assessed over the "full scope of the claims" applies in all fields of technology (see, in particular, T 149/21 on the one hand, and T 2773/18 and T 1983/19 on the other). In this context, the Board offers the following considerations.

- 21.1 Although, as observed in T 2773/18, the principle was introduced and is applied most frequently in chemistry, it has been employed in different fields of technology as pointed out in T 149/21 (point 3.3 of the reasons). This Board agrees with T 149/21 (still reasons 3.3) that the principles applies, and that it applies across all technical fields. There is no basis in the EPC for applying different standards for the compliance with Article 83 EPC depending on the technical field in question. Moreover, the distinction between different fields is not always clear cut, in principle and when technical fields converge, for instance due to the pervasive use of software technology in many technical fields. However, the interpretation of this principle may depend on the nature of the invention (e.g. product or method, structural or functional feature), or the role an individual feature plays in the context of the claimed subject-matter.
- 21.2 At the same time, the Board agrees with T 1983/19 (reasons 2.1.3, sentence 5) that for many claims it will be possible to imagine an arbitrary number of instances which cannot be carried out. Other than suggested in T 1989/19, however, the Board considers this to be the case in all fields.
- 21.3 Moreover, from the above point 13 it should be clear that this Board does not consider this statement of T 1983/19 to contradict the general principle that sufficiency of disclosure is to be assessed over the "full scope of the claims". The decision whether instances literally falling under the letter of the claim are contrived, and therefore to be considered as non-detrimental, and whether the application provides

sufficient disclosure to justify the breadth of the claims will require a case-by-case judgement.

*First to third auxiliary requests*

22. These requests do not restrict the set of events to be identified. Therefore, the same objection still applies.

*The fourth auxiliary request*

23. This request was filed during the oral proceedings before the Board. The restriction of the set of camera processing data considered, without restricting the considered set of events does not overcome the Board's objection. To the contrary: with a smaller choice of parameters, it is likely that even fewer events can be identified. The Board therefore decides to not admit this request (Article 13 RPBA 2020).

*The fifth auxiliary request*

24. This request was also filed during the oral proceedings before the Board. As basis for the amendment, the Appellant referred to page 9, lines 23-37, of the application as filed. The Board does not see that this passage, or the original application as a whole, discloses the now claimed combination of detecting "events relate[d] to vehicles arriving in scene" with the unique camera parameter of "data rate of encoded data sent to a communication network". At least *prima facie*, the amendment does not comply with Article 123(2) EPC. Furthermore, it is still not clear whether, and which precisely, events related to vehicles arriving in the scene can be identified. The Board does not admit this request either (Article 13 RPBA 2020).

**Order**

**For these reasons it is decided that:**

The appeal is dismissed.

The Registrar:

The Chairman:



L. Stridde

Martin Müller

Decision electronically authenticated