

**Internal distribution code:**

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

**Datasheet for the decision  
of 17 February 2026**

**Case Number:** T 1272/24 - 3.5.04

**Application Number:** 14807897.5

**Publication Number:** 3007444

**IPC:** H04N19/30, H04N19/31,  
H04N19/187, H04N19/156,  
H04N19/132

**Language of the proceedings:** EN

**Title of invention:**

IMAGE ENCODING METHOD, IMAGE DECODING METHOD, IMAGE ENCODING  
APPARATUS, AND IMAGE DECODING APPARATUS

**Applicant:**

Sun Patent Trust

**Headword:**

**Relevant legal provisions:**

EPC Art. 56

**Keyword:**

Main request and auxiliary requests 1 to 3 - inventive step  
(no)

**Decisions cited:**

**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

Case Number: T 1272/24 - 3.5.04

**D E C I S I O N**  
**of Technical Board of Appeal 3.5.04**  
**of 17 February 2026**

**Appellant:** Sun Patent Trust  
(Applicant) 450 Lexington Avenue, 38th Floor  
New York, NY 10017 (US)

**Representative:** Grünecker Patent- und Rechtsanwälte  
PartG mbB  
Leopoldstraße 4  
80802 München (DE)

**Decision under appeal:** **Decision of the Examining Division of the  
European Patent Office posted on 26 June 2024  
refusing European patent application  
No. 14807897.5 pursuant to Article 97(2) EPC.**

**Composition of the Board:**

**Chair** B. Willems  
**Members:** B. Le Guen  
W. Ungler

## Summary of Facts and Submissions

- I. The appeal is against the decision to refuse European patent application No. 14 807 897.5.
- II. The prior-art documents cited in the decision included the following:
- D5: A. Leontaris and P. C. Cosman, "*Compression Efficiency and Delay Tradeoffs for Hierarchical B-Pictures and Pulsed-Quality Frames*", IEEE Transactions on Image Processing, vol. 16, No. 7, July 2007, pp. 1726-1740, XP011185441
- III. The decision was based on the following grounds, *inter alia*.
- The subject-matter of claims 1, 8, 15 and 16 of the main request and the second auxiliary request then on file did not involve an inventive step in view of the disclosure of document D5 (Articles 52(1) and 56 EPC).
  - Claims 1 and 8 of the first auxiliary request then on file contained subject-matter extending beyond the content of the application as filed (Article 123(2) EPC).
- IV. The applicant (appellant) filed notice of appeal. With the statement setting out the grounds of appeal, the appellant filed claims of a main request and auxiliary requests 1 to 3. It submitted that the claims of the main request and auxiliary requests 1 and 2 corresponded to the claims of the main request and the

second and first auxiliary requests underlying the decision under appeal, respectively. Auxiliary request 3 was a new request on appeal. The appellant also explained why, in its view, the examining division had erred in its findings.

- V. The appellant was summoned to oral proceedings. In a communication under Article 15(1) RPBA the board gave its preliminary opinion that the subject-matter of claim 1 of all requests on file lacked inventive step starting from the disclosure of document D5 (Articles 52(1) and 56 EPC).
- VI. In its reply dated 16 January 2026, the appellant provided further arguments as to why the claimed subject-matter involved an inventive step.
- VII. The oral proceedings before the board took place on 17 February 2026.

The appellant's final requests were that the decision under appeal be set aside and that a European patent be granted on the basis of the claims of the main request filed with the statement of grounds of appeal or, alternatively, one of auxiliary requests 1 to 3 filed with the statement of grounds of appeal.

At the end of the oral proceedings, the chair announced the board's decision.

- VIII. Claim 1 of the **main request** reads as follows:

"A video encoding method of performing temporally scalable encoding on video, the video encoding method comprising:

performing the temporally scalable encoding on pictures included in the video and associated with a layer structure, to generate a bitstream, wherein a picture type of the pictures includes I-picture, P-picture, and B-picture, the I-picture being an intra-frame prediction picture, the P-picture being a forward reference prediction picture, and the B-picture being a bi-directional reference prediction picture,

characterized by

determining a total number of layers of the temporally scalable encoding to be less than or equal to a maximum layer count determined according to a frame rate of the video; and

the layer structure having as many layers as the total number of layers determined,

wherein a shallowest layer of the layers includes an I-picture and a P-picture, the shallowest layer includes no B-picture, at least one B-picture is included in another layer than the shallowest layer of the layers, and any layer other than the shallowest layer of layers includes no P picture."

IX. Claim 1 of **auxiliary request 1** reads as follows (additions with respect to claim 1 of the main request are shown by underlining):

"A video encoding method of performing temporally scalable encoding on video, the video encoding method comprising:

performing the temporally scalable encoding on pictures included in the video and associated with a layer

structure, to generate a bitstream, wherein a picture type of the pictures includes I-picture, P-picture, and B-picture, the I-picture being an intra-frame prediction picture, the P-picture being a forward reference prediction picture, and the B-picture being a bi-directional reference prediction picture,

characterized by

determining a total number of layers of the temporally scalable encoding to be less than or equal to a maximum layer count determined according to a frame rate of the video and an encoder latency limit; and

the layer structure having as many layers as the total number of layers determined,

wherein a shallowest layer of the layers includes an I-picture and a P-picture, the shallowest layer includes no B-picture, at least one B-picture is included in another layer than the shallowest layer of the layers, and any layer other than the shallowest layer of layers includes no P picture."

- X. Claim 1 of **auxiliary request 2** reads as follows (additions with respect to claim 1 of the main request are shown by underlining):

"A video encoding method of performing temporally scalable encoding on video, the video encoding method comprising:

performing the temporally scalable encoding on pictures included in the video and associated with a layer structure, to generate a bitstream, wherein a picture type of the pictures includes I-picture, P-picture, and

B-picture, the I-picture being an intra-frame prediction picture, the P-picture being a forward reference prediction picture, and the B-picture being a bi-directional reference prediction picture,

characterized by

determining a total number of layers of the temporally scalable encoding to be less than or equal to a maximum layer count determined according to a frame rate of the video and an output latency limit, wherein the output latency limit is a limit for a lag between input of a first picture of the input video and a start of an output of the bitstream, and the maximum layer count is defined by:

—  
Maximum layer count =  $\text{int}(\log_2(\text{encoder output latency [s]} \times \text{frame rate [fps]}) + 1)$  ; and

the layer structure having as many layers as the total number of layers determined,

wherein a shallowest layer of the layers includes an I-picture and a P-picture, the shallowest layer includes no B-picture, at least one B-picture is included in another layer than the shallowest layer of the layers, and any layer other than the shallowest layer of layers includes no P picture."

XI. Claim 1 of **auxiliary request 3** reads as follows (additions with respect to claim 1 of auxiliary request 2 are shown by underlining):

"A video encoding method of performing temporally scalable encoding on video, the video encoding method comprising:

performing the temporally scalable encoding on pictures included in the video and associated with a layer structure, to generate a bitstream, wherein a picture type of the pictures includes I-picture, P-picture, and B-picture, the I-picture being an intra-frame prediction picture, the P-picture being a forward reference prediction picture, and the B-picture being a bi-directional reference prediction picture,

characterized by

determining a total number of layers of the temporally scalable encoding to be less than or equal to a maximum layer count determined according to a frame rate of the video and an output latency limit, wherein the output latency limit is a limit for an encoder output latency, the encoder output latency being a lag between input of a first picture of the input video and a start of an output of the bitstream, and the maximum layer count is defined by:

Maximum layer count =  $\text{int}(\log_2(\text{encoder output latency [s]} \times \text{frame rate [fps]})) + 1$  ; and

the layer structure having as many layers as the total number of layers determined,

wherein a shallowest layer of the layers includes an I-picture and a P-picture, the shallowest layer includes no B-picture, at least one B-picture is included in another layer than the shallowest layer of the layers, and any layer other than the shallowest layer of layers includes no P picture."

## **Reasons for the Decision**

1. The appeal is admissible.

*All requests - inventive step (Articles 52(1) and 56 EPC)*

2. An invention is to be considered to involve an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art (Article 56 EPC).
3. At the oral proceedings before the board, the appellant contended that document D5 was not the best starting point for assessing inventive step; however, it did not dispute that document D5 represented a suitable starting point for assessing inventive step.

The board adheres to the case law according to which if the skilled person has a choice of several workable routes, i.e. routes starting from different documents, which might lead to the invention, the rationale of the problem-and-solution approach requires that the invention be assessed relative to all these possible routes, before an inventive step can be acknowledged (see Case Law of the Boards of Appeal of the European Patent Office, 11th edition, 2025 ("Case Law"), I.D.3.3).

4. Document D5 discloses a method of performing temporally scalable encoding on video pictures associated with a layer structure to generate a bitstream, wherein a picture type of the pictures includes I-picture, P-picture, and B-picture, the I-picture being an intra-frame prediction picture, the P-picture being a forward reference prediction picture, and the B-picture being a bi-directional reference prediction picture,

wherein a shallowest layer of the layers includes an I-picture and a P-picture, the shallowest layer includes no B-picture, at least one B-picture is included in another layer than the shallowest layer of the layers, and any layer other than the shallowest layer of layers includes no P picture (see Figures 3 and 4 and the paragraph bridging pages 1728 and 1729 describing an encoding method with hierarchical bi-predictive motion-compensated ("HIER") structures; see also the second paragraph of section III: "*temporal scalability*").

Document D5 also discloses that the number of hierarchical temporal levels of the HIER structure is given by  $\log_2(N_{\text{GOP}}) + 1$  (see paragraph bridging pages 1728 and 1729), where  $N_{\text{GOP}}$  denotes the number of frames in a group of pictures consisting of an I-picture or P-picture and the following pictures up to (but excluding) the next P-picture ("GOP") (see Figure 3 and the paragraph below that figure).

Document D5 further discloses that frames are encoded at a given frequency (number of frames per second (fps)) which is the inverse of the display time duration of a frame ( $t_{\text{fr}}$ ) (see D5, page 1729, second and fourth paragraphs).

5. In the case of a HIER structure with a GOP of size  $N_{\text{GOP}}$ , the P-picture of the shallowest layer must be encoded immediately after encoding the I-picture of that layer because it is used for encoding the pictures of the remaining layers (see D5, Figure 4). This order is the same at the decoder. Therefore, the P-picture must immediately follow the I-picture in the bitstream. To achieve this, the encoder must await the arrival of the P-picture before starting to output the bits of the

I-picture. This translates into an encoder output latency of  $N_{\text{GOP}} \times t_{\text{fr}}$  seconds between the time at which the I-picture starts being input into the encoder and the time at which the bits of the I-picture start being output by the encoder. It is undisputed that this latency is inherent to the HIER coder disclosed in document D5.

6. It is common ground that the method of claim 1 of **auxiliary request 2** differs from the HIER coding method disclosed in document D5 in that it comprises the following features:

*"determining a total number of layers of the temporally scalable encoding to be less than or equal to a maximum layer count determined according to a frame rate of the video and an output latency limit, wherein the output latency limit is a limit for a lag between input of a first picture of the input video and a start of an output of the bitstream, and the maximum layer count is defined by:*

*Maximum layer count =  $\text{int}(\log_2(\text{encoder output latency [s]} \times \text{frame rate [fps]})) + 1$ ; and*

*the layer structure having as many layers as the total number of layers determined"*

7. In the light of paragraph [0053] of the description, the expression "encoder output latency" is to be understood as referring to the encoder output latency limit.
8. On page 12 of its statement of grounds of appeal, the appellant argued that the present invention provided the advantage that the display latency could be

controlled and thus limited. The appellant formulated the objective technical problem as that of enhancing the control of display timing.

The display latency relates to the time from the start of decoding a picture to the start of displaying that picture (see paragraph [0058] of the application in hand).

On page 4 of its reply dated 16 January 2026 and at the oral proceedings before the board, the appellant clarified that the term "display timing" in the objective technical problem was meant to reflect the cumulative delay resulting from both encoding and decoding, not only the display latency.

9. Insofar as "display timing" is understood to relate to the end-to-end delay encompassing both the encoder output latency and the display latency, the board agrees with the appellant's formulation of the objective technical problem. This is consistent with the second sentence of paragraph [0089] of the application as filed, according to which the invention guarantees that both the output latency at the encoder and the display latency at the decoder do not exceed limits adopted by the Association of Radio Industries and Businesses (ARIB) (see paragraph [0003] of the application as filed and the reference to "NPL 2" in paragraph [0089]).
  
10. At the oral proceedings before the board, the appellant argued that, although it disclosed formulas related to the end-to-end delay and the number of hierarchical temporal levels, document D5 did not teach the use of these formulas to enhance the control of display timing. Instead, document D5 proposed reducing the

delay by removing some prediction. Therefore, the distinguishing features identified in point 6. above would not have been obvious to the skilled person.

11. The board disagrees with the appellant for the following reasons.
  - 11.1 The problem-and-solution approach does not require document D5 to hint at the distinguishing features. The motivation to modify the parameters of the HIER coder and the teaching towards the distinguishing features may come from the skilled person's common general knowledge and abilities or from another item of prior art.
  - 11.2 Making sure that a delay remains under a preset limit is an obvious way of controlling that delay. By definition, the end-to-end delay is the addition of the encoder output latency and the decoder output latency. Therefore, the skilled person would have thought to configure the HIER coder such that both latencies remained under preset limits. This has not been disputed by the appellant.
  - 11.3 As indicated in point 5. above, the encoder output latency in the case of the HIER coder amounts to  $N_{\text{GOP}} \times t_{\text{fr}}$  seconds. Therefore, using basic mathematics, the skilled person would have realised that  $N_{\text{GOP}}$  must be set to a value less than or equal to the encoder output latency limit multiplied by the inverse of  $t_{\text{fr}}$ , i.e. to a value less than or equal to the encoder output latency limit multiplied by the frame rate. Since the number of hierarchical temporal levels for the HIER structure is given by  $\log_2(N_{\text{GOP}}) + 1$  (see point 4. above), this would have led the skilled person to a maximum number of temporal

layers equal to  $\text{int}(\log_2(\text{encoder output latency limit [s]} \times \text{frame rate [fps]})) + 1$ , the  $\text{int}()$  function obviously being necessary to account for maximum values of  $N_{\text{GOP}}$  not being powers of two.

- 11.4 On page 4 of its reply dated 16 January 2026, the appellant submitted that a skilled person setting the GOP did not correspond to the claimed method step of determining a maximum layer count according to the frame rate and determining the total number of layers to be less than or equal to the determined maximum layer count. With this determination step being provided by an apparatus performing the method, the present invention went beyond merely converting units and mathematical terms.

The board does not find these arguments persuasive. The maximum value of  $N_{\text{GOP}}$  must necessarily be calculated before calculating the maximum number of temporal layers. For the reasons given in point 11.3 above, these calculations would have anticipated the formula specified in claim 1 of auxiliary request 2 ("*Maximum layer count = int(log2(encoder output latency [s] × frame rate [fps])) + 1*"). Since that maximum value is, by definition, the limit not to be exceeded, the skilled person would evidently have chosen a number of layers that is less than or equal to that maximum value. Claim 1 of auxiliary request 2 does not require the maximum layer count and the total number of layers to be determined by the encoder, i.e. does not rule out the maximum layer count and the total number of layers being determined by a person and then input into the encoder.

12. For the above reasons, the board finds that the subject-matter of claim 1 of **auxiliary request 2** lacks an inventive step within the meaning of Article 56 EPC.
13. It follows that the subject-matter of claim 1 of the **main request** and **auxiliary request 1**, being broader than the subject-matter of claim 1 of auxiliary request 2, also lacks an inventive step.
14. Claim 1 of **auxiliary request 3** clarifies that the output latency limit is a limit for an encoder output latency. This is the meaning the board ascribed to the term "output latency limit" in claim 1 of auxiliary request 2 anyway (see point 6. above, last sentence). Therefore, the subject-matter of claim 1 of auxiliary request 3 lacks an inventive step for the reasons given with respect to claim 1 of auxiliary request 2.

*Conclusion*

15. Since none of the requests in hand is allowable, the appeal must be dismissed.

**Order**

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

The appeal is dismissed.

The Registrar:

The Chair:



K. Boelicke

B. Willems

Decision electronically authenticated