

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ASUSTEK COMPUTER INC.,

Petitioner

v.

NOKIA CORP. AND NOKIA TECHNOLOGIES OY,

Patent Owner

Case: IPR2025-01279

U.S. Patent No. 8,050,321

DECLARATION OF DR. IAIN RICHARDSON

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	EDUCATIONAL BACKGROUND AND QUALIFICATIONS	2
III.	MATERIALS REVIEWED.....	5
IV.	LEGAL STANDARDS	5
1.	Overview	5
2.	Patent Claims.....	6
3.	Burden of Proof for Invalidity.....	7
4.	Claim Construction	8
5.	Person of Ordinary Skill in the Art	8
6.	Anticipation	8
7.	Obviousness.....	9
V.	PERSON OF ORDINARY SKILL IN THE ART (POSITA).....	12
VI.	CLAIM CONSTRUCTION.....	13
VII.	BACKGROUND	13
1.	The '321 Patent	13
2.	The Prior Art	24
a.	MPEG-1	24
b.	Kim.....	36
VIII.	The Grounds in the Petition do not render obvious the Challenged claims	45
1.	Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 1—MPEG-1.	45
a.	Petitioner fails to demonstrate that limitations 8[a], 10[a], and 11[a] are obvious over MPEG-1	45
b.	Petitioner fails to demonstrate that limitations 8[d], 10[d], and 11[d] are obvious over MPEG-1 because MPEG-1 does not teach or suggest a method for decoding, a video decoder or a computer program product for decoding, resetting the identifier value as required.....	61

2.	Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 2—Kim.....	65
a.	Petitioner fails to demonstrate that limitations 8[a], 10[a], and 11[a] are obvious over Kim.....	65
b.	Petitioner fails to demonstrate that limitations 8[b], 10[b], and 11[b] are obvious over MPEG-1	85
c.	Petitioner fails to demonstrate that limitations 8[d], 10[d], and 11[d] are obvious over MPEG-1	86
3.	Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 3—MPEG-1 In Combination With Yagasaki	90
IX.	SIGNATURE.....	91

I, Dr. Iain Richardson, declare as follows:

I. INTRODUCTION

1. I have been retained on behalf of Nokia Technologies OY and asked to provide opinions in respect of Inter Partes Review (IPR) proceeding IPR2025-01279 which relates to U.S. Patent No. 8,050,321, which I will also refer to as simply the '321 Patent.

2. This Declaration summarises the independent opinions that I have formed to date which are based on the material currently available to me. I reserve the right to modify my opinions, if necessary, based on further review and analysis of information that I receive subsequent to the filing of this Declaration, including in response to any positions taken by Petitioners or their experts with which I have not yet been provided.

3. My findings, as explained below, are based on my study, experience, knowledge and background in the fields discussed below. I am informed by my education and my extensive experience in the fields of video compression and multimedia coding and communications. I have also relied on my review and analysis of the prior art, information provided to me and information I have independently reviewed.

4. I have previously provided five declarations regarding the validity of the '321 Patent: *Amazon.com et al v. Nokia Technologies Oy*, IPR2024-00691, Ex.

2001 (Declaration in Support of Patent Owner Preliminary Response) and Ex 2004 (Declaration in Support of Patent Owner Response); *Snap et al v. Nokia Technologies Oy*, IPR2025-01014, Ex. 2012 (Declaration in Support of Patent Owner Preliminary Response); and two declarations subject to a protective order in the ITC investigation *Certain Video-Capable Laptop, Desktop Computers, Handheld Computers, Tablets, Televisions, Projectors, and Components and Modules Thereof*, Inv. No. 337-TA-1448. I have re-used and adapted text and illustrations from these previous declarations where I feel it is appropriate.

II. EDUCATIONAL BACKGROUND AND QUALIFICATIONS

5. Details of my professional qualifications and background are summarised below and are set out in detail in my Curriculum Vitae, attached to this Declaration.

6. I am Director and co-founder of Vcodex BV (previously Vcodex Ltd).

7. I have over 30 years of experience related to video and image compression and coding. I am qualified by education and experience to testify as an expert with respect to subject matter in the fields of video, image, multimedia and general-purpose data encoding, decoding, compression, decompression, and transmission.

8. I received a Master of Engineering (M.Eng.) degree in Electronic and Electrical Engineering from Heriot-Watt University in Edinburgh, Scotland in 1990.

I received a Doctor of Philosophy (Ph.D.) degree in Video Compression from Robert Gordon University in Aberdeen, Scotland in 1999. I worked as a Digital Signal Processing (“DSP”) Hardware Designer with GEC Avionics Ltd. from 1990 to 1993.

9. In 1993, I assumed a post as a Lecturer, then Reader and eventually a Full Professor in the field of image and video compression in the School of Engineering at Robert Gordon University. In 2009, I was awarded the position of Honorary Professor with Robert Gordon University.

10. At Robert Gordon University, I founded and ran an image communication technology research laboratory. I carried out original research in the field of data, image, and video compression, initiated and managed research projects, and supervised research students.

11. I am the author of four books and over 70 journal and conference papers on video and image coding and communication, including two widely cited books on the H.264 / MPEG-4 industry standards for video compression. In my 2002 book, *Video Codec Design*, I discussed in detail the efficient implementation of image and video compression on software and hardware platforms.

12. In 1995, during my employment at Robert Gordon University, I co-founded and became Technical Director, equivalent of Chief Technology Officer, at 4i2i Communications Limited. I held this position until 2000. In my role at 4i2i, I was responsible for and engaged in the design of software and hardware modules for

compression and communication systems. The designs included video processing algorithms in software and hardware for implementation on processing devices and the design of compression algorithms such as block-based transforms, run-length coding, and variable-length coding in software and hardware. A modified version of one of my video compression hardware designs was successfully used by NASA to stream video from the International Space Station.

13. In 2007, I founded Vcodex Limited (now Vcodex BV) and currently serve as the CEO, through which I provide technical and strategic advice to technology companies particularly relating to video, image, and data compression for streaming.

14. In 2009, I founded Onecodec Limited and served as CEO, leading the company's development of innovative video, image, and multimedia data compression and storage software and systems. Onecodec's technology and business activities were merged with Vcodex Limited in 2015.

15. I have previously prepared expert reports and testified in a number of cases concerning video and image compression, graphics compression, general-purpose data compression, and multimedia streaming.

16. I am a named inventor on 11 patents or patent applications, including U.S. and non-US patents and applications.

17. Because of my background, training, and experience, I am qualified as an expert in the technical areas relevant to the '321 Patent. My company Vcodex BV is being compensated for my independent analysis as an expert with respect to this IPR proceeding, but neither my compensation nor Vcodex BV's compensation is contingent in any way on the content of my analysis or the outcome of this proceeding. I have no financial interest in Petitioners (Snap Inc. or Hisense USA Corporation) or Patent Owner (Nokia Technologies Oy) and I have no financial interest in the challenged patent.

III. MATERIALS REVIEWED

18. The opinions contained in this Declaration are based on the documents that I reviewed or cited here, including the Petition (as well as all materials and expert declaration cited there), my professional judgment, as well as my education, experience and knowledge regarding technologies related to the '321 Patent at and around the time the patent was filed.

IV. LEGAL STANDARDS

1. Overview

19. I am not an attorney or a legal expert, and I do not opine on matters of law. The legal standards, which are listed in this section and which I applied in reaching the conclusions expressed throughout this Declaration, were provided to me and explained to me by counsel. In preparing this Declaration, I have endeavored

to faithfully apply these legal standards. I do not intend to offer opinions on the law. Below, I describe the legal standards that I applied in analyzing the claims of the '321 Patent.

20. I have been informed by counsel that the challenger of a patent's validity, in this case the Petitioners, bears the burden of proving invalidity of the patent.

2. Patent Claims

21. I have been informed by counsel that a patent's claims describe the invention made by the inventors and describe what the patent owner owns and what the owner may prevent others from doing.

22. I have been informed by counsel that claims may describe products, such as machines or chemical compounds, or processes for making or using a product. Claims are usually divided into parts or steps called limitations.

23. I have been informed by counsel that there are two types of claims: independent claims and dependent claims. An independent claim sets forth all the requirements that must be met in order to be covered by that claim. Thus, it is not necessary to look at any other claim to determine what an independent claim covers. A dependent claim does not itself recite all of the requirements of the claim, but refers to another claim for some of its requirements. In this way, a dependent claim depends on another claim. The law considers a dependent claim to incorporate all of

the requirements of the claim to which it refers.

24. I have been informed by counsel that a dependent claim then adds its own additional requirements. To determine what a dependent claim covers, it is necessary to look at both the dependent claim and any other claims to which it refers.

25. I have been informed by counsel that in a case where a dependent claim incorporates the limitations of an independent claim upon which it depends, if a reference fails to disclose any of the limitations of the independent claim, the dependent claims which depend on that independent claim cannot be made invalid by such a reference. As a result, any of the references at issue in this proceeding that fail to disclose all of the limitations of an independent claim necessarily also fail to disclose all of the limitations of claims which depend on that independent claim.

3. Burden of Proof for Invalidity

26. I have been informed by counsel that the party challenging the validity of an issued U.S. patent in an IPR bears the burden of proving invalidity through facts supported by a preponderance of the evidence. I have been informed that a “preponderance of the evidence” means evidence sufficient to prove that a fact is more likely true than it is not. I understand that, for the determination of whether to institute an IPR, the challenger must show a reasonable likelihood of success in invalidating at least one challenged claim.

4. Claim Construction

27. I have been informed by counsel and I understand that the first step in determining validity is to properly construe the claims. I have further been informed by counsel that, in this proceeding, the Board must determine the scope of the claims by using the same claim construction standard that United States District Courts use. Under this standard, the words of a claim are given their “ordinary” and “customary” meaning as would have been understood by a person of ordinary skill in the art at the time of the invention, unless there has been some specific definition or disavowal of subject matter provided in the patent or file history. I have been informed that such a person of ordinary skill in the art looks to the claims, specification, and file history in understanding the meaning of the claim terms. I have applied this understanding in this declaration.

5. Person of Ordinary Skill in the Art

28. I have been informed by counsel and I understand that a person having ordinary skill in the art may be defined by factors such as, when considered at the time of the alleged invention: (a) the levels of education and experience of the inventor and other persons actively involved in the field; (b) the types of problems encountered in the field; (c) prior art solutions to those problems; (d) rapidity with which innovations are made; and (e) the sophistication of the technology.

6. Anticipation

29. I have been informed by counsel and I understand that a patent claim

can be found invalid as anticipated if, and only if, each and every element of the claim, as properly construed, is found either explicitly or inherently in a single prior art reference. If the single prior art reference does not disclose explicitly or inherently any one claim limitation, then that claim is not anticipated. I have further been informed that the single prior art reference cannot merely disclose each element; rather, it must disclose all of the elements as arranged in the claim. I have been informed that for a prior art reference to “inherently” disclose something, the reference must necessarily function in accordance with, or include the claimed limitations, to anticipate. The missing descriptive matter cannot merely probably or possibly include the claimed limitations. I have further been informed by counsel that, with respect to an anticipation analysis, it is improper to combine separate, distinct embodiments of a single prior art reference.

30. I have been informed by counsel that for a reference to be anticipatory, the reference must contain an enabling disclosure. I have been informed that merely naming or describing the subject matter is insufficient if it cannot be produced without undue experimentation.

7. Obviousness

31. I have been informed by counsel that a patent claim can be found invalid as obvious if the differences between the claim and the prior art references would have been obvious to a person of ordinary skill in the art at the time of the

invention, taking into account (1) the scope and content of the prior art, (2) the differences between the prior art and the claims, (3) the level of ordinary skill in the art, and (4) any objective indicia of non-obviousness. I have further been informed by counsel that a claim can be found to be obvious in light of either a single prior art reference or the combination of multiple prior art references. I have been informed that references qualify as prior art for an obviousness determination only when they are analogous to the claimed invention. Two separate tests define the scope of analogous prior art: (1) whether the art is from the same field of endeavor, regardless of the problem addressed; and (2) if the reference is not within the field of the inventor's endeavor, whether the reference still is reasonably pertinent to the particular problem with which the inventor is involved.

32. I have been informed that to be obvious in light of a single prior art reference or the combination of multiple prior art references, there must be a reason either to modify the single prior art reference or to combine the two or more prior art references to achieve the claimed invention. I have been informed by counsel that this reason may come from a teaching, suggestion, or motivation to combine which may come from the reference or references themselves, the knowledge or "common sense" of one skilled in the art, or from the nature of the problem to be solved, and may be explicit or implicit from the prior art as a whole. I have been informed by counsel that it is not sufficient that a person of ordinary skill merely could have

modified or combined references, but that there must be a showing that a person of ordinary skill would have done so.

33. I have been informed by counsel that the combination of familiar elements according to known methods may be obvious when it does no more than yield predictable results.

34. I have further been informed by counsel that there is no suggestion to combine, however, if a reference teaches away from its modification or combination with another source. A reference may be said to teach away when a person of ordinary skill, upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken by the applicant, or if it suggests that the line of development flowing from the reference's disclosure is unlikely to be productive of the result or property sought by the applicant. If, when combined, the references would produce a seemingly inoperative device, then they teach away from their combination. I have been informed that it is improper to combine references where the references teach away from their combination. I have been informed by counsel that the degree of teaching away will depend on the particular facts.

35. I have been informed that if a proposed modification would render the prior art device unsatisfactory or inoperable for its intended purpose, then the proposed modification would generally not be obvious.

36. I have been informed by counsel that it is improper to use hindsight in making the obviousness determination.

37. I have been informed by counsel that some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness must be made. I have been informed that, for example, there may be no rational underpinning to support the legal conclusion of obviousness where systems are already complete and the party challenging validity seeks to mix and match components, which each exist in the separate systems, from these complete systems.

V. PERSON OF ORDINARY SKILL IN THE ART (POSITA)

38. Based on my review of the Petition, I understand that Petitioners have identified that the POSITA “would have had a (1) Bachelor’s degree in electrical engineering, computer engineering, computer science, or a comparable field of study such as physics, and (2) approximately two to three years of practical experience with video encoding/decoding . . . [and] [a]dditional experience can substitute for the level of education, and vice-versa.” For the purposes of my analysis for this Declaration, I have used the definition that the Petitioners have identified.

39. All of my opinions in this Declaration are from the perspective of a person of ordinary skill in the art, as the Petitioners defined it above and during the relevant time frame. During that time frame, I possessed at least the qualifications of such a POSITA.

VI. CLAIM CONSTRUCTION

40. I understand that claim terms are typically given their ordinary and customary meaning, as would have been understood by a person of ordinary skill in the art at the time of the alleged invention. In considering the meaning of the claims, I also understand that one must consider the claim language, specification, prosecution history, and the knowledge of a person of ordinary skill in the art at the relevant time.

41. I understand that the Petition stated that “no constructions are necessary” and therefore has not proposed any terms for construction. Petition at 9.

42. Therefore, at this time, I have applied the plain and ordinary meaning as a POSITA would have understood it at the time of the invention and in accordance with the principles that I state above.

VII. BACKGROUND

1. The '321 Patent

43. The '321 Patent relates to improvements in video encoding and decoding that enables a decoder to identify where to start a decoding process, such as when a user randomly selects a frame to start viewing video, for example when joining a stream part-way through or when rewinding or fast-forwarding. This is known as random access in video coding, which the '321 Patent is directed to improving. '321 Patent at 3:52-56, 3:62-4:4, 4:48-60. The claims of the '321 Patent

provide a solution to specific random-access problems and solutions described in the specification¹.

44. Specifically, “[a] problem in the numbering of image frames arises when the user wishes to start browsing a video file in the middle of a video sequence. Such situations occur for example when the user wishes to browse a locally stored video file backward or forward or to browse a streaming file at a particular point; when the user initiates the playback of a streaming file from a random point, or when a video file that is to be reproduced is detected to contain an error that interrupts the playback or requires the playback to be resumed from a point following the error.” ’321 Patent at 11:11-20. The ’321 Patent identifies this as a problem because “[w]hen the browsing of a video file is resumed from a random point after previous browsing, discontinuity typically occurs in the image numbering. The decoder

¹ I disagree with Petitioner’s assertion that the concept of random access is ‘not reflected in the claims’ of the ’321 Patent (Petition at 6, footnote). The ’321 Patent identifies several challenges that all relate to the concept of random access, including playback of a stream or file from a user-selected position, joining a channel or stream part-way through a transmission and recovering from errors. ’321 Patent at 3:62-4:4, 11:11-25. The claimed invention of the ’321 Patent provides solutions to all of these challenges as I explain herein.

typically interprets this as unintentional loss of image frames and unnecessarily tries to reconstruct the image frames suspected as lost.” ’321 Patent at 11:20-25.

45. The ’321 Patent explains that this is a problem that arises because video coding methods take advantage of a concept called temporal redundancy, which refers to the fact that “information comprised by consecutively displayed image frames is typically largely similar, i.e. successive image frames comprise a considerable amount of redundancy.” ’321 Patent at 1:55-2:4. Video coding methods take advantage of this redundancy by using a technique known as motion-compensated temporal prediction. ’321 Patent at 2:15-17. In this technique, a portion of a current frame to be encoded can be predicted from a portion of another, previously-encoded frame such as a past or future frame in display order. Motion-compensated temporal prediction attempts to identify a matching portion in a previously coded frame, which may be in a different position due to motion of objects or areas, and to use this to predict a portion of a current frame. ’321 Patent at 2:16-20. This technique reduces redundancy in video coding and can greatly improve efficiency in video coding. ’321 Patent at 2:5-33.

46. As the ’321 Patent explains, I-frames (or INTRA-frames) are encoded frames whose encoded image information is not determined using this motion-compensated temporal prediction process, while INTER-frames (also known as P-frames or B-frames) are encoded frames whose image information is, at least in part,

determined based on prediction from one or more previously coded frames in the sequence. Such previously coded frames may be preceding frames in display order and/or subsequent frames in display order. A P-frame is encoded using only one previously encoded frame and a B-frame may be encoded using two previously-encoded frames, for example one preceding frame and one subsequent frame in display order. '321 Patent at 2:20-32, 2:42-47.

47. However, one problem with using motion-compensated temporal prediction is that if a frame that is needed to decode other frames is missing or lost, “frames dependent on it can no longer be correctly decoded” either. '321 Patent at 2:9-33. One situation where this happens is when a user selects a random point of a video to start streaming from. '321 Patent at 11:11-25. The “decoder typically interprets this as unintentional loss of image frames and unnecessarily tries to reconstruct the image frames suspected as lost. '321 Patent at 11:11-25.

48. The '321 Patent explains that “this [problem] can be avoided in the decoder by defining an initiation image in an independently decodable Group of Pictures GOP that is activated at the random point of the video file, and the number of the initiation image is set to zero. . . . Consequently, when activated at a random point, the decoder preferably sets the identifier of the first image frame, preferably an I-frame, of the independently decodable sub-sequence at zero.” '321 Patent at 11:26-40. Similarly, on the encoder side, “identifier values for pictures are encoded

according to a numbering scheme, and the identifier value for the indicated first picture of an independent sequence is reset, preferably to zero.” ’321 Patent at 4:42-45.

49. The ’321 Patent claims such an improved method “that enable[s] detection of image frames from which a decoder can start the decoding process.” ’321 Patent at 4:8-10. “The invention is based on the idea of encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference picture is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order.” ’321 Patent at 4:16-20.

50. Prior to the invention of the ’321 Patent, earlier standards such as MPEG-1 provided limited options for temporal prediction (i.e. for motion compensated inter prediction) and specified a defined prediction structure with limited flexibility. In MPEG-1, each video frame is encoded as an I-picture, P-picture or B-picture. A series of coded pictures are organized into a defined structure known as a Group of Pictures (GOP). Blocks of video data in an I-picture are not predicted from any other coded picture. All blocks of video data in a P-picture are predicted from the most recent previous I- or P-picture in display order. All blocks of video data in a B-picture are predicted from one or both of the previous and next I- or P-pictures in display order. ’321 Patent at 2:34-48. With this somewhat inflexible prediction structure in MPEG-1, random-access could be achieved by

searching for the next GOP and attempting to start decoding. I discuss MPEG-1 in more detail below.

51. The invention of the '321 Patent was created in the context of the emerging H.264/AVC standard. H.264/AVC provided many new concepts compared to the older standards, including greatly increased flexibility for temporal prediction. In H.264, any block in a P-picture can be predicted from any one of a set of previously coded pictures, which may be before or after the current picture in display order. Any block in a B-picture can be predicted from one or two previously coded pictures, each of which may be before or after the current picture in display order. This is achieved using a Decoded Picture Buffer, in which an H.264 decoder stores a number of previously coded pictures. These may include recently-decoded pictures ("short term" reference pictures) and also pictures that could have been coded many tens of frames previously to the current frame ("long term" reference pictures). The concept of flexibly selecting one of a set of reference pictures is known as Reference Picture Selection. *See* Ex. 2020 (Richardson, H.264), 106-111; *see also* '321 Patent at 3:37-4:4.

52. The flexible prediction choices of RPS introduced in the emerging H.264 standard, led to improved compression performance but also to new challenges for random-access. In such a system, the conventional GOP structure of MPEG-1 and MPEG-2 does not necessarily apply and/or does not adequately

support random access. A decoder cannot necessarily begin decoding from an arbitrary Group of Pictures, since subsequent P and/or B-pictures may be predicted from a wide range of potential reference pictures including long-term reference pictures that were decoded much earlier in the video sequence.

53. The inventor of the '321 Patent, Miska Hannuksela, recognised that a new approach to random-access was needed. In contribution JVT-B041.doc, he set out to the Joint Video Team (the team responsible for developing H.264), a number of specific requirements for random-access in the context of the emerging H.264 draft standard. These included:

- The decoder must have received all necessary parameter sets to decode the bitstream (the parameter set concept did not exist in the MPEG-1 or MPEG-2 standards).
- There must be no prediction references to pictures prior to the random-access position.
- The relationship between picture numbers must be 'broken' [sic] by resetting the picture numbering.
- It should be possible to cut and reassemble video sequences at random-access positions.

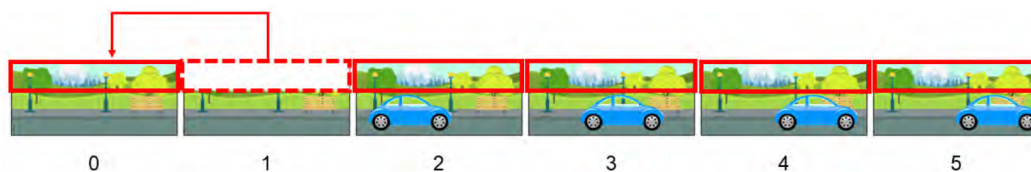
54. See Ex. 2021 ("JVT-B041"); *see also* '321 Patent at 4:16-60. Previous standards and methods were unable to address all these requirements. For example, neither MPEG-1 nor MPEG-2 supported flexible reference picture selection, long-

term reference pictures or parameter sets and so approaches based on either of these standards could not address the requirements for random-access in the emerging H.264 standard.

55. I have provided illustrations below to show the random-access problems addressed by the '321 Patent. As shown below, a video encoder may exploit temporal redundancy, allowing redundant content in a series of video images (*e.g.*, the red boxes below) to be predicted from earlier reference pictures that include the same content, improving compression by avoiding the need to encode redundant content in later pictures.

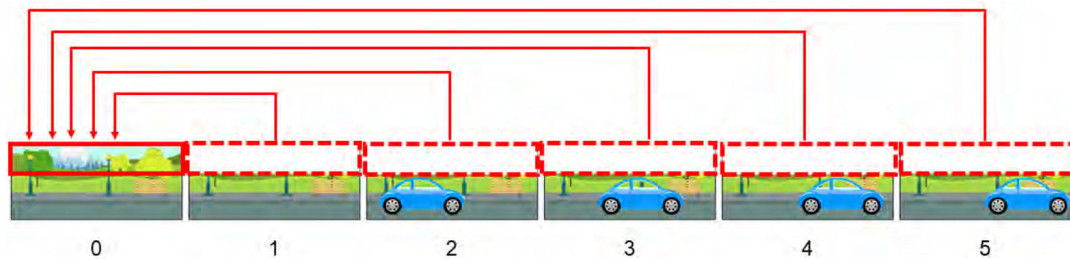


56. As shown below, the top of Picture 1 (dashed rectangle) is predicted from the top of Picture 0 (solid rectangle), exploiting the fact that Picture 0 contains the video content necessary for Picture 1.

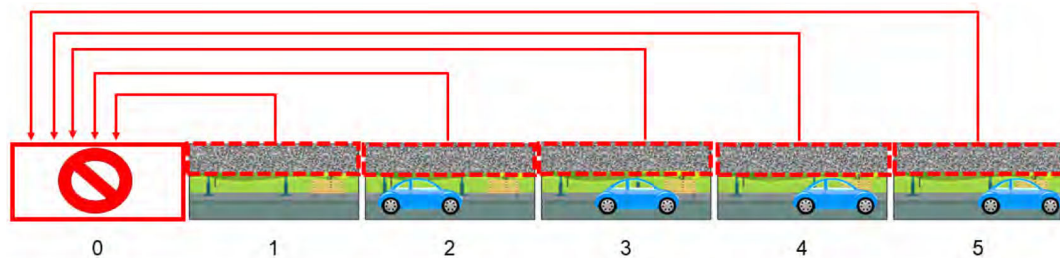


57. In this example, the same information from Picture 0 can be used to improve the compression of a series of pictures. The top of Picture 0 (solid rectangle) is only transmitted once, and Pictures 1-5 predict that portion of their own content

(arrows) from Picture 0.

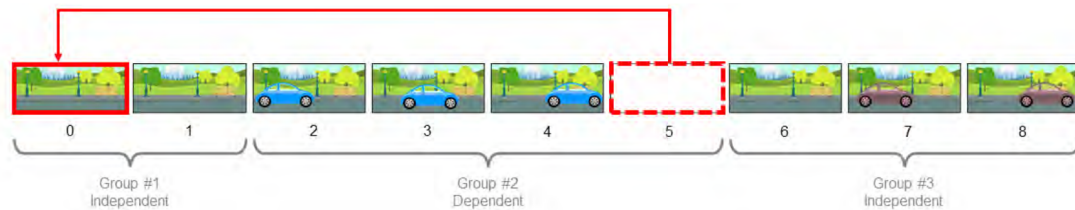


58. Temporal prediction can however lead to unwanted effects if errors occur during transmission. For example, decoding errors occur when the reference picture (Picture 0) is missing, and the predicted pictures (Pictures 1-5) cannot reference that video content, as illustrated below.



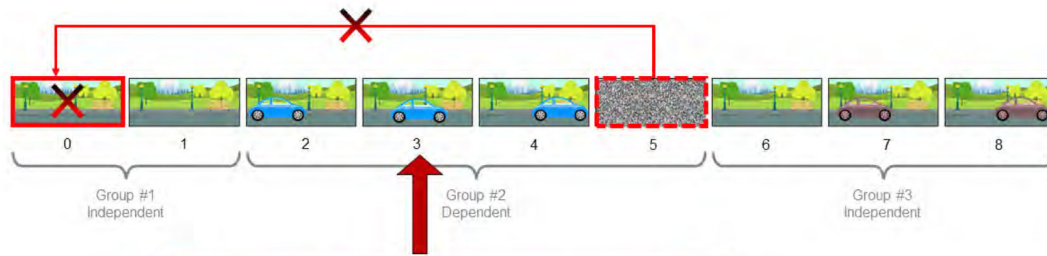
59. The '321 Patent discusses the concept of independent sequences or independently decodable groups of pictures. In the illustrative example below, a video sequence (Pictures 0-8) includes two independent sequences: Group #1 (Pictures 0-1) and Group #3 (Pictures 6-8). It also includes a dependent sequence: Group #2 (Pictures 2-5). An independent sequence makes use of predictions only from pictures within the same independent sequence, whereas a dependent sequence includes at least one picture that refers to another picture outside of that dependent sequence. Because Picture 5 in Group #2 is predicted from Picture 0 in Group #1,

Group #2 is a dependent sequence.



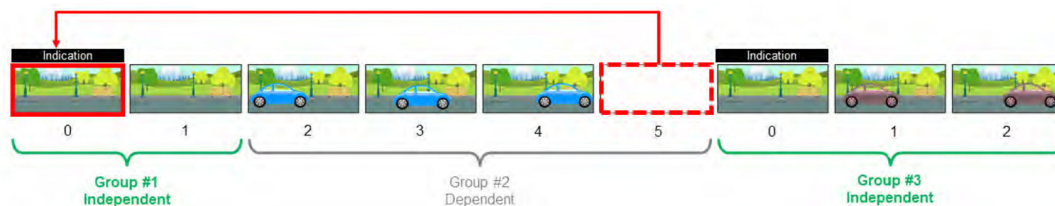
60. In the above example, when a user randomly accesses a specific picture within the middle of a video sequence, the decoder must determine where it can begin decoding so that the decoded pictures can be displayed without any prediction errors. As the '321 Patent explains, prior art methods of identifying sources for prediction (such as reference picture selection) can pose challenges for random access, since certain pictures may be predicted from previous pictures that are not available at the decoder. '321 Patent, 3:52-56, 3:62-4:4. As explained in the '321 Patent, these dependencies present challenges with random access because the decoder does not know from which image frames it can begin decoding. '321 Patent, 3:52-56, 3:62-4:4.

61. The example below illustrates this problem. The user selects Picture 3 as the random-access point. If the decoder starts decoding at Picture 3, an error occurs when decoding Picture 5 because Picture 0 was not previously decoded and is not available at the decoder for predicting Picture 5.



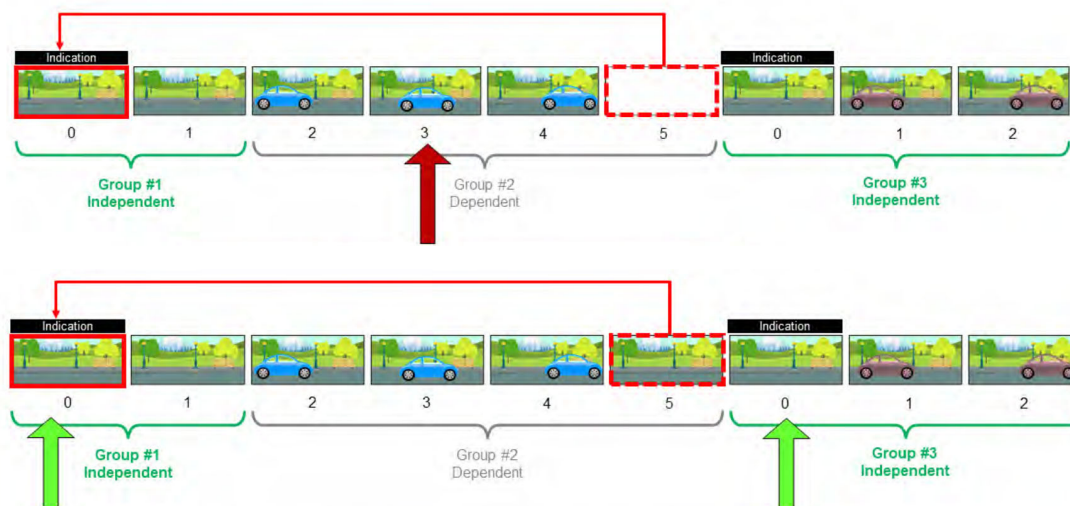
62. The '321 Patent claims a specific solution to address these prior art random-access problems. '321 Patent, 4:48-60. Specifically, the identification of an initiation frame or picture and resetting the numbering of that frame to zero solves the issue so that the decoder does not inadvertently interpret that certain frames are missing. '321 Patent, 11:11-40.

63. According to the '321 Patent claims, a decoder decodes an indication of the first image frame of an independent sequence. As an illustration, below are two independent sequences (Groups #1 and #3), each including an indication (black box) of a first image frame that informs the decoder where decoding can begin to avoid such problems with random access.



64. For example, if the user selects Picture 3 in Group #2 as the random-access point (dark red arrow in the top figure), decoding begins at the first image of one of the independent sequences, either Picture 0 in Group #1 (left-hand green arrow in the bottom figure) or Picture 0 in Group #3 (right-hand green arrow in the

bottom figure). The picture numbering of Group #3 restarts with a Picture 0 and proceeds sequentially because it is an independent sequence.



65. As illustrated in the above example, the '321 Patent claims an improved method “that enable[s] detection of image frames from which a decoder can start the decoding process.” '321 Patent, 4:8-10. By detecting these image frames, the user can start playback of a video sequence from a random-access point and view the video without the errors that were present in the art at the time.

2. The Prior Art

a. MPEG-1

66. The MPEG-1 Part 2 Standard (ISO/IEC 11172-2, Exhibit 1005, hereinafter MPEG-1) defines a bitstream and a decoding process. Encoders can produce compliant bitstreams that are decodable by MPEG-1's specified decoding process. An MPEG-1 decoder must produce identical results to the defined decoding process.

67. The MPEG-1 Standard specifies a decoding process that supports various possible techniques (i.e. encoding options) for encoding MPEG-1 compliant bitstreams. An encoder can choose such options during encoding so long as the bitstream conforms to the standard as discussed above. Examples include picture type (an encoder can choose I, P or B picture types for coding each frame, within the constraints of the standard), macroblock type (an encoder can choose I, P or B macroblock types, within the constraints of the standard), motion vectors, high-level syntax element values, etc.

68. The encoder makes decisions and chooses options whilst encoding a bitstream, as described above. The decoder² processes the received MPEG-1 bitstream and must decode the bitstream according to the syntax element values in the bitstream and according to the specification of the MPEG-1 Part 2 Standard. For example, the temporal_reference syntax element, part of MPEG-1's Picture Layer, is assigned in sequence to pictures in display order (syntax element

² MPEG-1 defines "decoder" as "[a]n embodiment of a decoding process." It further defines "decoding (process)" as "[t]he process defined in ISO/IEC 11172 that reads an input coded bitstream and produces decoded pictures or audio samples. Ex. 1005 at 22.

temporal_reference at the Picture Layer, a 10-bit fixed length binary codeword, Ex 1005 at 38). An encoder chooses a value of Temporal Reference and signals it in the bitstream for an encoded Picture or Frame. The decoder receives the coded bitstream, parses this syntax element (i.e. reads the 10-bit temporal_reference syntax element value from the bitstream), and decodes it according to the standard (Ex 1005 at 44). Thus, as I explain further below, an MPEG-1 decoder cannot *choose* or *reset* a temporal_reference value; rather, an MPEG-1 decoder must only *decode* a temporal_reference value that was assigned by the encoder.

69. An MPEG-1 bitstream is organized into Groups of Pictures (GOPs), each beginning with a syntax element group_start_code (Ex 1005 at 43) and including thereafter (in encoding order) an I-picture followed by zero or more P, B or I-pictures. A typical GOP structure consists of (in display order) an I-picture, optionally preceded by one or more B-pictures, followed by (a) P-pictures spaced at intervals and optionally (b) B-pictures between successive I- or P-pictures and (c) I-pictures spaced at intervals (Ex 1005 at 32-33).

70. In MPEG-1, an I-picture is intra coded, i.e. it can be decoded without reference to any other coded pictures. A P-picture is inter predicted from the previous I- or P- picture in the bitstream. A B-picture is inter predicted from either or both of the previous and following I- or P-pictures in display order.

71. A representative example is shown in Ex. 1005 at p. 71. Note that in coding (bitstream) order, B-pictures are placed after their respective previous and following I- or P-pictures. For example, B-pictures 1 and 2 in the following example are inter predicted from I-picture 0 and P-picture 3. In bitstream order, B-pictures 1 and 2 follow I-picture 0 and P-picture 3. This means that a decoder has to store only two previously coded pictures (0 and 3) prior to decoding and displaying each B-picture (1 and 2 in this example).

Note that because of the picture dependencies, the bitstream order, i.e. the order in which pictures are transmitted, stored, or retrieved, is not the display order, but rather the order which the decoder requires them to decode the bitstream. An example of a sequence of pictures, in display order, might be:

I	B	B	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Figure D.3 -- Typical sequence of pictures in display order

whereas the bitstream order would be as shown below:

I	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B	P	B	B
0	3	1	2	6	4	5	9	7	8	12	10	11	15	13	14	18	16	17

Figure D.4 -- Typical sequence of pictures in bitstream order

72. A GOP may be encoded without the use of B-pictures, i.e. only using I- and P-pictures. An example of such a GOP is shown in Ex. 1005 Fig. D.17:

The original concept of a group of pictures was a set of pictures that could be coded and displayed independently of any other group. In the final version of this part of ISO/IEC 11172 this is not always true, and any B-pictures preceding (in display order) the first I-picture in a group may require the last picture in the previous group in order to be decoded. Nevertheless encoders can still construct groups of pictures which are independent of one another. One way to do this is to omit any B-pictures preceding the first I-picture. Another way is to allow such B-pictures, but to code them using only backward motion compensation.

Property 3. From a coding point of view, a concisely stated property is that a group of pictures begins with a group of pictures header, and either ends at the next group of pictures header or at the next sequence header or at the end of sequence, whichever comes first.

Some examples of groups of pictures are given below:

```

I
I  P  P
I  B  P  B  P
B  B  I  B  P  B  P
B  B  I  B  B  P  B  B  P  B  B  P
B  I  B  B  B  B  P  B  I  B  B  I  I

```

Figure D.17 -- Examples of groups of pictures in display order

These examples illustrate what is possible, and do not constitute a suggestion for structures of groups of pictures.

Ex. 1005, 84-85.

73. This GOP structure begins with an I-picture, followed by a series of P-pictures. Each P-picture is predicted from the immediately preceding I- or P-picture, e.g. consider the following example:

$$I_0 P_1 P_2 P_3 P_4 \dots$$

74. In this example, P-picture P_1 is inter predicted from I-picture I_0 . P-picture P_2 is inter predicted from P-picture P_1 , and so on.

75. One important aspect of MPEG-1 related to Petitioner's arguments is MPEG-1's "closed_gop flag." MPEG-1's closed_gop flag is part of the "Group of Pictures layer," i.e. it is part of the syntax layer that applies to an entire GOP, MPEG-1 at 2.4.3.4 ("Group of pictures layer"), Ex 1005 at 43.

closed_gop -- This is a one-bit flag which may be set to "1" if the group of pictures has been encoded without motion vectors pointing to the previous group of pictures.

76. As specified in MPEG-1, when decoding an MPEG-1 video bitstream, the closed_gop flag is received in the group of pictures header:

2.4.2.4 Group of pictures layer

Syntax	No. of bits	Mnemonic
group_of_pictures() {		
group_start_code	32	bslbf
time_code	25	
closed_gop	1	
broken_link	1	
next_start_code()		
if (nextbits() == extension_start_code) {		
extension_start_code	32	bslbf
while (nextbits() != '0000 0000 0000 0000 0000 0001') {		
group_extension_data	8	
}		
next_start_code()		
}		
if (nextbits() == user_data_start_code) {		
user_data_start_code	32	bslbf
while (nextbits() != '0000 0000 0000 0000 0000 0001') {		
user_data	8	
}		
next_start_code()		
}		
do {		
picture()		
} while (nextbits() == picture_start_code)		
}		

MPEG-1 at 2.4.2.4 (“Group of pictures layer”), Ex 1005 at 37.

77. MPEG-1 defines “Closed GOPs” or “closed groups [of pictures]” as follows³:

³ I disagree with Petitioner’s characterization of Fig. D.18(a)’s example as an ‘exemplary closed GOP’ (Petition at 17 and later). The Oxford English Dictionary

Closed GOP

A one bit flag follows the time code. It denotes whether the group of pictures is open or closed. Closed groups can be decoded without using decoded pictures of the previous group for motion compensation, whereas open groups require such pictures to be available.

A typical example of a closed group is shown in figure D.18a.

I	B	B	P	B	B	P	B	B	P	B	B	P
0	1	2	3	4	5	6	7	8	9	10	11	12

(a) closed group

B	B	I	B	B	P	B	B	P	B	B	P	B	B	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

(b) open or closed group

Figure D.18 -- Example groups of pictures in display order

A less typical example of a closed group is shown in figure D.18b. In this example, the B-pictures which precede the first I-picture must use backward motion compensation only, i.e. any motion compensation must be based only on picture number 2 in the group.

If the closed_gop flag is set to 0 then the group is open. The first B-pictures that precede the first I-picture in the group may have been encoded using the last picture in the previous group for motion compensation.

MPEG-1 at D.5.2 (“Group of pictures”), Ex 1005 at 86.

78. The “one bit flag” referred to in the above extract is the closed_gop flag. As the closed_gop flag is a flag in the group of pictures header, it is a syntax element that applies to an entire GOP and not any one specific coded frame. The closed_gop flag is therefore not an indication of any specific image frame.

79. For example, in Figure D.18 above, the value of the closed_gop flag (0 or 1) defines a property of the entire GOP, such as the 13-picture GOP shown in

defines ‘exemplary (adj)’ as ‘That sets or affords a good example; admirable, commendable; (later also more generally) excellent, outstanding, perfect’. MPEG-1 Fig. D.18(a)’s example of a GOP is just that – an example. It is not ‘exemplary’ or a preferred example in any way.

Fig. D.18a or the 15-picture GOP shown in Fig. D.18b. More specifically, as the cited text above explains, the value of `closed_gop` defines certain properties of the coding of certain B-pictures in the GOP, namely whether or not B-pictures preceding the first I-picture may or may not use forward motion compensation (i.e. motion compensation from the preceding P- or I-picture, which is located in the preceding GOP).

80. In contrast to the `closed_gop` flag, many other syntax elements in an MPEG-1 video bitstream pertain to a specific coded picture, namely syntax elements found at the Picture Layer and lower syntax layers. MPEG-1 specifies that “The picture layer contains all the coded information for one picture”. Ex. 1005 at 87.

81. Thus, the `closed_gop` flag is not an indication of a particular image frame in the GOP and at most, indicates certain properties of certain B-pictures in the GOP.

82. Another aspect of MPEG-1 that is important with respect to Petitioner’s argument is “Temporal Reference,” encoded as a syntax element `temporal_reference`, a fixed-length codeword that is part of the Picture Layer syntax. Unlike the `closed_gop` flag, `temporal_reference` is a syntax element that is associated with a specific coded picture (Ex. 1005 at 38):

2.4.2.5 Picture layer

Syntax	No. of bits	Mnemonic
picture() {		
picture_start_code	32	bslbf
temporal_reference	10	uimsbf
picture_coding_type	3	uimsbf
vbv_delay	16	uimsbf
if ((picture_coding_type == 2) (picture_coding_type == 3)) {		
full_pel_forward_vector	1	
forward_f_code	3	uimsbf
}		

83. MPEG-1 explains that “Temporal Reference” can be used to define “the order in which the pictures must be displayed.” For a given group of pictures (whether open or closed), the first picture in display order is assigned a Temporal Reference of zero:⁴

⁴ Temporal reference is also set to zero when there are more than 1024 pictures in a GOP. Ex. 1005, 87. This is a limitation of the number of bits used to express temporal_reference in the coded video bitstream.

D.5.3.2 Temporal reference

The Temporal Reference is a ten-bit number which can be used to define the order in which the pictures must be displayed. It may be useful since pictures are not transmitted in display order, but rather in the order which the decoder needs to decode them. The first picture, in display order, in each group must have Temporal Reference equal to zero. This is incremented by one for each picture in the group.

Some example groups of pictures with their Temporal Reference numbers are given below:

Example (a) in display order	I	B	P	B	P								
	0	1	2	3	4								
Example (a) in decoding order	I	P	B	P	B								
	0	2	1	4	3								
Example (b) in display order	B	B	I	B	B	P	B	B	P	B	B	P	
	0	1	2	3	4	5	6	7	8	9	10	11	
Example (b) in coded order	I	B	B	P	B	B	P	B	B	P	B	B	
	2	0	1	5	3	4	8	6	7	11	9	10	
Example (c) in display order	B	I	B	B	B	B	P	B	I	B	B	I	I
	0	1	2	3	4	5	6	7	8	9	10	11	12
Example (c) in coded order	I	B	P	B	B	B	B	I	B	I	B	B	I
	1	0	6	2	3	4	5	8	7	11	9	10	12

Figure D.19 -- Examples of groups of pictures and temporal references

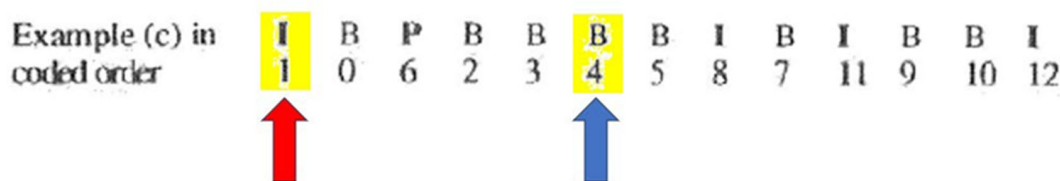
Ex. 1005, 87-88.

84. In an MPEG-1 coded video bitstream, the order in which pictures must be displayed (the display order) can be different from the “coded order” of the pictures, which defines the order that the pictures are decoded. As illustrated in Figure D.19 and discussed earlier, the coded order for a picture can be and typically is different from the temporal order.

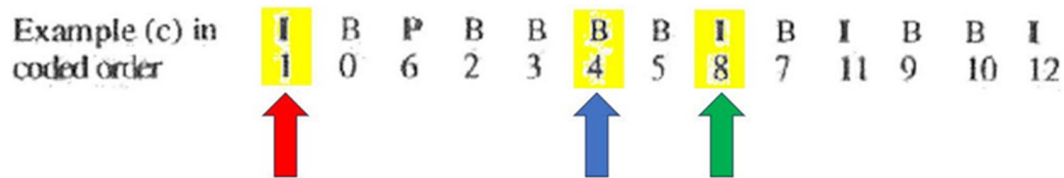
85. Further, MPEG-1 discloses prior art problems (not the claimed solution) because it does not fully enable random access at a user-specified playback point, and therefore does not meet the claim limitations that enable such random access. MPEG-1 describes that after a user specifies a random-access point, the decoder looks for the next sequence start code, group of pictures (GOP) start code, or picture start code which introduces an I-picture (collectively, “start codes”) to

start decoding. Ex. 1005, 43, 87, 120. All frames prior to the next start code are discarded, and decoding and playback starts from that next start code (rather than from a specific identified frame), which is not the user-specified random-access point.

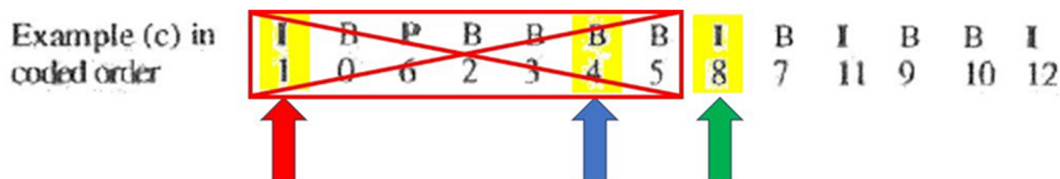
86. Applying the above-described teachings of MPEG-1 to Example (c) of Figure D.19 (which applies equally to all of MPEG-1) highlights the issues with MPEG-1 and confirms that MPEG-1 does not disclose the claimed invention of the '321 Patent. As shown below, the first frame of the GOP in decoding order is I₁ (red arrow) and the user selects, for example, frame B₄ (blue arrow) as the playback point.



87. Thereafter, MPEG-1 starts decoding from frame I₈ (green arrow, below), which is not the first frame I₁ (red arrow) of the independent sequence. MPEG-1 starts decoding from I₈ (green arrow) because MPEG-1 teaches that after a user specifies a random-access point, the decoder looks for the next sequence start code, group of pictures (GOP) start code, or picture start code which introduces an I-picture (collectively, “start codes”) to start decoding. Ex. 1005, 43, 87, 120. In Example (c), the next start code is the picture start code which introduces an I-picture, I₈. *See id.*



88. The decoder discards all frames until frame I_8 and starts decoding and playback from I_8 because MPEG-1 teaches that all prior frames are discarded and decoding and playback starts from that next start code (rather than from a specific identified frame), which is not the user-specified random-access point. Ex. 1005, 43, 87, 120.



89. Further, MPEG-1 does not disclose the decoder resetting any identifier value of the indicated first image frame. In the above example, regardless of whether I_1 or I_8 is the “indicated” first image frame, MPEG-1 never teaches a decoder resetting any identifier value of those frames. An MPEG-1 compliant decoder must follow the decoding process specified in the MPEG-1 Standard, as I explained earlier. The decoding process of MPEG-1 does not instruct a decoder to reset or modify any such identifier and so an MPEG-1 compliant decoder cannot do so.

90. MPEG-1 does not identify decoding any indication of a first image frame of an independent sequence, starting the decoding from the indicated first

image frame, and/or resetting an identifier value of the indicated first image frame. When the user selects a playback point, MPEG-1 teaches to discard all frames and start decoding at a start code (not an indicated first image frame). MPEG-1 suffers from the same random-access problems discussed above and the '321 Patent acknowledged MPEG-1's GOP concept as a problem in random access and sought to address it. '321 Patent, 2:34-55, 3:52-56. Consequently, MPEG-1 fails to disclose several limitations of the asserted claims that enable random access from the user-specified point, as I explain further below.

b. Kim

91. Kim is directed to a Time Lapse recording/reproduction device and method. Ex. 1006 (Kim) at 1:8-12. Specifically, Kim describes recording, compressing and reproducing intermittently recorded image data (*i.e.*, single image frames spaced at intermittent time intervals), which is significantly different from the problems and solutions of the '321 Patent. In particular, Kim was concerned that traditional Time Lapse Video Cassette Recorders using magnetic tapes would suffer from "considerable use and wear" which deteriorates the quality of the images. Ex. 1006 (Kim) at 1:57-67. The solution that Kim suggests is to use an optical disk instead. Ex. 1006 (Kim) at 2:58-61. This is very far from the problems and solutions of the '321 Patent.

92. Further, Kim is not related to video streaming, in contrast to the '321

Patent. Kim's intermittent time lapse technology was directed to security systems for banks or museums. Ex. 1006 (Kim) at 1:14-19. In such implementations, single images may be recorded intermittently and then stored, for example on magnetic tapes or optical disk. These individual images can be individually and incrementally reproduced when needed. However, Kim does not enable random access of streaming video or any playback of video and therefore is not directed to the issues of the '321 patent. Ex. 1006 (Kim) at 1:8-56, 6:17-20 ("only the desired images can be selectively reproduced with high display quality"). In fact, Kim does not discuss independently decodable video sequences anywhere in its disclosure. This makes sense given Kim's emphasis on intermittently stored *individual image frames*. Ex. 1006 (Kim), 1:38-45.

93. More specifically, Kim describes a system that (a) encodes and stores individual images without prediction (I-frame) or with prediction from previously stored images (P-frame) and (b) identifying, decoding and displaying a selected image on a user display. Ex. 1006 (Kim) at 3:33-4:45, Figs. 3-4, 5:8-6:8.

94. According to Kim, in FIG. 5A, each "P-picture frame P1, P2, P3 has a current address ("CAD") indicating the current address of the P-picture, and an I-picture reference address ("IAD") indicating the address of the corresponding (or previous) I-frame." Ex. 1006 (Kim) at 4:46-52 (shown below):

The bit stream generated and transmitted to the recording/
reproducing unit **140** as described above has a logical
structure as shown in, e.g., FIG. 5A. As shown therein, each
P-picture frame P1, P2, P3 has a current address CAD
50 indicating the current address of the P-picture, and an
I-picture reference address IAD indicating the address of the
corresponding (or previous) I-frame. If image signals are

95. The logical address, CAD, is shown in Kim's Fig. 5A as incrementing by one. Ex. 1006 (Kim) at Fig. 5A. The IAD for a P-frame references the logical address of the corresponding or previous I-frame. Ex. 1006 (Kim) at 5:1-7. And the IAD field for an I-frame is either 0000 or absent because it corresponds to nothing and is not used during decoding. Ex. 1006 (Kim) at 4:65-5:1. Kim does not explain why the IAD field for an I-frame can be either 0000 or absent. A POSITA would understand that including the IAD field with an arbitrary value such as 0000 can be used to maintain a constant format or number of bits in the addressing of both I- and P-frames. In some scenarios it would be more convenient from a data formatting or programming perspective to have the same length for both I- and P-frames. Moreover, omission of the IAD field for the I-frame is possible because it has no purpose and is not used in Kim. However, omitting this field may lead to a less consistent (but somewhat more compressed) representation of the data on the optical disc. Thus, a POSITA would understand that Kim's suggestion that the IAD field of an I-frame may be set to an arbitrary value such as 0000 is merely a matter of programming convenience relative to omitting the field completely.

96. Kim's Fig. 3 and specification describes its encoding process.

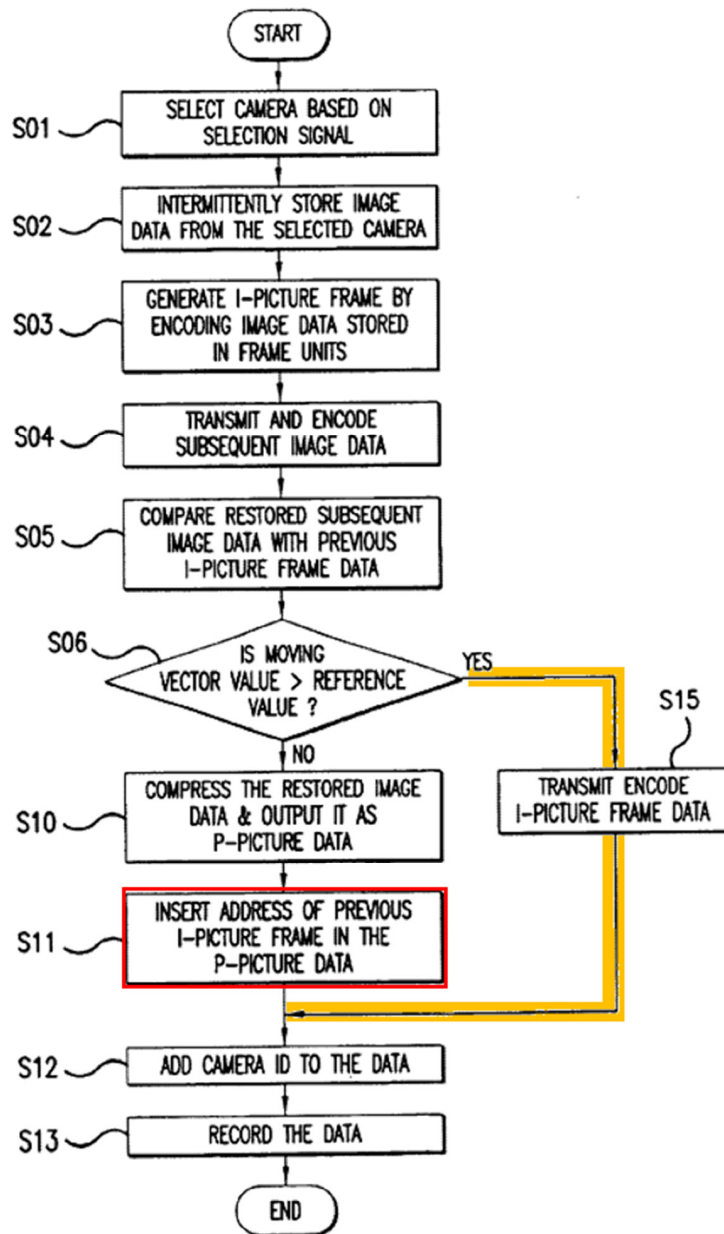


FIG.3

Ex. 1006 (Kim) at Fig. 3 (annotations added).

97. In relevant part, Kim's encoder encodes the first image as an I-frame, which will not depend on any other frame. Ex. 1006 (Kim) at 3:57-65 (Step S03).

For the subsequent image, “intermittently obtained and stored” (Ex. 1006 (Kim) at 3:46-48) at a later time instant, Kim’s encoder will determine whether that image frame’s data is similar enough to the first image that it can use a P-frame to encode the subsequent (second) image. Ex. 1006 (Kim) at 3:65-4:14 (Steps S04-S06). If Kim’s encoder determines that the second image is similar enough to the first image, then it will encode the second image as a P-frame in Step S10 and then, in Step S11, Kim’s encoder will add “the moving vector and the position value of the corresponding (or previous) I-frame, so that the above information is added to the corresponding compressed P-frame output from the encoder.” Ex. 1006 (Kim) at 4:15-33.

98. On the other hand, if Kim’s encoder determines that the second image is not similar enough to the first image in Step S06, then it will encode the second image as another I-frame in Step S15. Ex. 1006 (Kim) at 4:23-27. As I show in the annotated Fig. 3 above, the process for encoding the second image as an I-frame (shown in orange) skips Step S11 (outlined in red): “insert address of previous I-picture frame in the P-picture data.” Thus, Kim’s encoder only adds “the position value of the corresponding (or previous) I-frame” to the P-frames, and not to the I-frames. Ex. 1006 (Kim) at 4:15-33.

99. For the I-frames, Kim teaches only that “the address of the corresponding I-picture frame (3) is zero or absent.” Ex. 1006 (Kim) at 4:65-5:1.

That is, the IAD for the I-frame is not a position value or anything else. The IAD for the I-frame has no meaning or use, which is why it may be entirely absent without impact on the method of Kim. This is confirmed by the claims of Kim, which repeatedly disclose only providing “position information” of the corresponding I-frame for a P-frame but not for an I-frame. Ex. 1006 (Kim) at claims 1-2, 8, 11.

100. Kim’s decoder process is described in Fig. 4 and the associated parts of the specification. More specifically, when the user in Kim selects a picture, Kim’s decoder, shown in Fig. 4, decodes and displays that single picture, based on the user’s request for that one image.

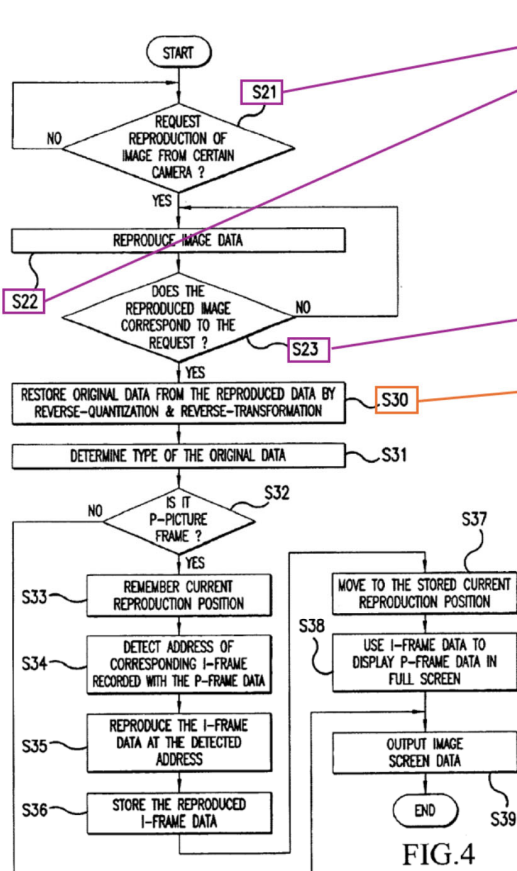


FIG.4

As shown in Steps S21 and S22 of FIG. 4, if the user requests reproduction of an image data from a particular camera or channel among the intermittent image data recorded on the optical disc 1, the controller 150 controls the recording/reproducing unit 140 to reproduce the corresponding image data from the optical disc 1. The image data reproduced from the optical disc 1 is multiplexed data, which is demultiplexed by the de-MUX 131. By demultiplexing, the camera (channel) ID code CHID is separated from the reproduced image data and output to the controller 150. Based on the camera ID code CHID, the controller 150 determines whether or not the reproduced image data corresponds to the selected camera or channel, in Step S23. This ensures that the correct image data is processed to comply with the user's request. If incorrect image data has been reproduced, the controller 150 controls the recording/reproduction unit 140 to reproduce the correct image data from the optical disc 1. In Step S30, the image data corresponding to the user's request is transmitted and decoded in the decoder 111. The decoding process of the decoder 111 involves reverse-quantization and reverse-transformation to restore the original image data prior to data compression carried out during recording. Then the property (type) of the currently restored image data is determined in Step S31 by the reproducing DSP 121, and the determination result is transmitted to the controller 150. More specifically, if the currently restored image data is detected to be P-picture frame data in Step S32, the controller 150 stores the current reproduction position in Step S33. Then in Step S34, the controller 150 detects the address of the corresponding I-picture frame (e.g., data stored in the area (3) of the bit stream B2 shown in FIG. 6) from the reproduced image data output from the de-MUX 131, and reproduces based on the detected address the corresponding (or previous) I-picture frame data recorded on the optical disc 1 by controlling the recording/reproduction unit 140.

In Step S36, the reproduced corresponding I-picture frame data is processed by the reproducing DSP 121, and decoded by the decoder 111 to be stored in the frame memory 101. The controller 150, after Step S36, controls the recording/reproducing unit 140 to move the video heads to the original recording position stored in Step S33.

Ex. 1006 (Kim) at Fig. 4 and 5:1-41 (annotations added).

101. As shown in the annotated figure above (Figure 4 alongside column 5 of Kim), the first image data that is decoded is the image data for the frame the user selected (Step S30). Ex. 1006 (Kim) at 5:29-34. Kim's decoder is unaware until a later point (Step S31) whether that selected frame is an I-frame or a P-frame. Ex. 1006 (Kim) at 5:35-38.

102. According to Kim, if the requested image is a P-frame, the corresponding I-frame address (the IAD of the P-frame) is detected at Step S34, and then the corresponding I-frame data is decoded *in part* at Step S36. Ex. 1006 (Kim)

at 5:39-52. I note that, according to Kim, the I-frame is not completely decoded at this point because the final reverse-transformation and reverse-quantization does not occur until Step S38. In Kim, reverse-transformation and reverse-quantization is a necessary part of the decoding process. Ex. 1006 (Kim) at 5:32-35. At Step S38 of Kim, the partial P-frame⁵ is combined with the partially decoded I-frame data and the moving (motion) vectors, and then the resulting P-frame image is reverse-transformed and reverse-quantized, which completes the decoding of the full P-frame. Ex. 1006 (Kim) at 5:55-65. Thus, according to Kim, when the requested image to decode is a P-frame, only the P-frame is fully decoded because the full I-frame itself is never reverse-transformed and reverse-quantized (only partially decoded I-frame data that are combined with the partial P-frame, prior to reverse transformation and reverse quantization).

103. On the other hand, Kim describes that if the requested image is an I-frame, the corresponding coded image data is simply read from the optical disk, decoded and displayed. Ex. 1006 (Kim) at 5:66-6:3. When the requested image is an I-frame, Step S34 to “detect address of corresponding I-frame recorded with the P-

⁵ I use the expression “partial P-frame” herein to refer to the P-frame data after decoding (Kim Step S30) but before combining with partially decoded I-frame data to ultimately reconstruct the full, displayed P-frame.

frame data” (shown in red below) is skipped (as shown in orange below).

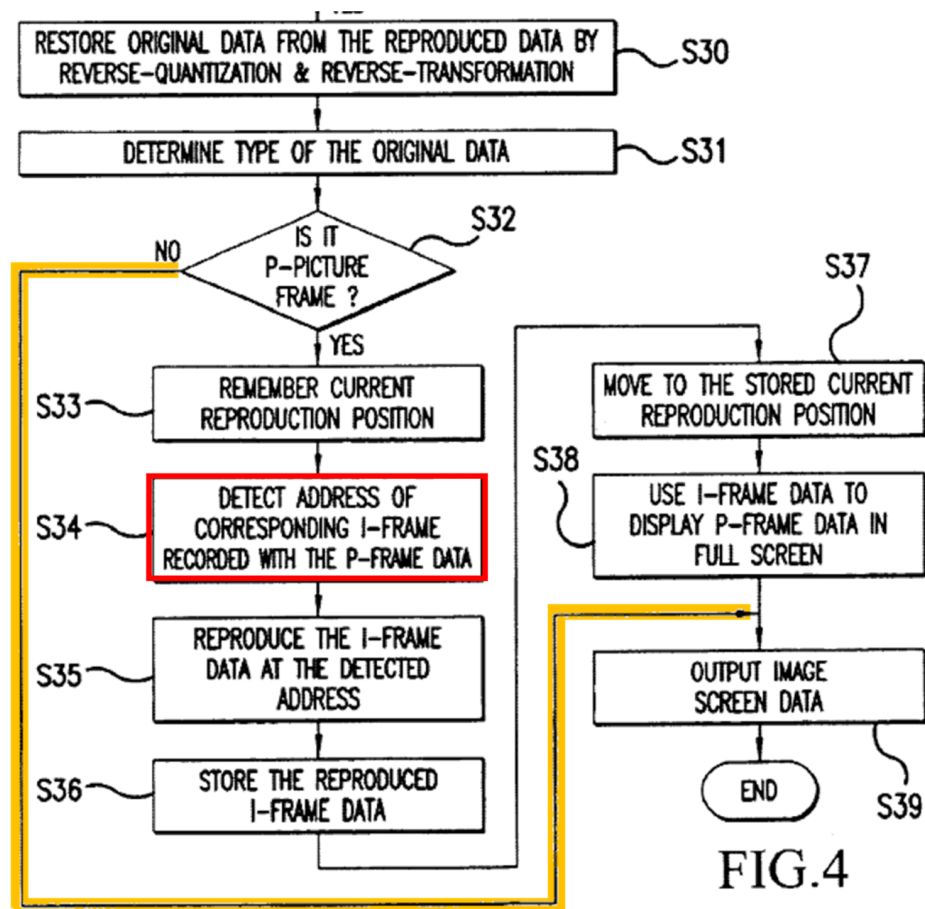


FIG.4

Ex. 1006 (Kim) at Fig. 4 (excerpted and annotations added).

104. Thus, in Kim, the IAD field is only used for the decoding of a P-frame. The IAD of the I-frame is never read or decoded or used for anything, and may be omitted completely, according to Kim. That is why Kim teaches that the IAD of the I-frame may even be absent because it makes no difference to the decoder. Ex. 1006 (Kim) at 4:65-5:1. A POSITA would be familiar with the concept of a “don’t care” value in such a system, i.e., a value that may be zero or absent, because it is simply

not used.

105. This is further confirmed by Kim's claims, which only specify that the step of "detects the address of the corresponding intra-picture frame" occurs "if the restored image data frame is predictive picture frame data" and not "if the original image frame data type is not predictive picture frame data." Ex. 1006 (Kim) at claim 11.

106. As I discuss in greater detail in the following sections of my declaration, Kim does not disclose or teach several limitations of the challenged claims of the '321 Patent, nor does Kim solve or even relate to the '321 Patent's aims of avoiding unnecessary decoding errors and enabling random access from a user-specified point in a video stream.

VIII. THE GROUNDS IN THE PETITION DO NOT RENDER OBVIOUS THE CHALLENGED CLAIMS

1. Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 1—MPEG-1.

a. Petitioner fails to demonstrate that limitations 8[a], 10[a], and 11[a] are obvious over MPEG-1

107. Each of limitations 8[a], 10[a], and 11[a] are written with similar language and require "[decoding/decode] from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, wherein all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said

independent sequence.” In my opinion, Petitioner⁶ has failed to show that MPEG-1 teaches or suggests these limitations.

108. I understand Petitioner asserts two distinct arguments as to how the “[decoding/decode] from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence limitations” are taught or suggested—pointing to MPEG-1’s closed GOP (closed_gop flag) or, alternatively, to MPEG-1’s Temporal Reference (temporal_reference syntax element).

**i. MPEG-1’s Closed GOP Flag Is Not the Claimed
“Indication”**

109. **First**, Petitioner asserts that MPEG-1’s closed_gop flag teaches or suggests “[decoding/decode] from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence.” Petition at 21-24. In my opinion, Petitioner is incorrect for at least two reasons: (1) the closed_gop flag is a single-bit flag that only indicates whether the

⁶ Petitioner put forth an expert declaration that in my view is substantively similar to the limitation-by-limitation analysis. Therefore, references to “Petitioner” as used in these claim sections refers to both Petitioner’s argument and the substantively similar expert declaration substance.

entire GOP is closed or open; it does not identify any image frame and is therefore not an indication according to the claims; and (2) the closed_gop flag does not indicate which frame is the first image frame, in decoding order. In my view it is not sufficient that every closed GOP necessarily has a first image frame because the claims require indicating “which is the first image frame.” I explain this further below.

110. As an initial matter, MPEG-1’s closed_gop flag is only an indication of whether or not the current group of pictures (i.e. the GOP that immediately follows the closed_gop flag) is open or closed; it is not an indication of any image frame (whether or not the image frame is first in decoding order). Ex. 1005, 37; The closed_gop flag is present in a structural level of the coded bitstream data (the GOP Layer) that relates to an entire group of coded pictures and not to any particular frame:

0.5 Structure of the coded video bitstream

This part of ISO/IEC 11172 specifies a syntax for a coded video bitstream. This syntax contains six layers, each of which either supports a signal processing or a system function:

Layers of the syntax	Function
Sequence layer	Random access unit: context
Group of pictures layer	Random access unit: video
Picture layer	Primary coding unit
Slice layer	Resynchronization unit
Macroblock layer	Motion compensation unit
Block layer	DCT unit

Ex. 1005, 16.

111. Syntax elements in the GOP layer apply to an entire GOP, i.e. a

collection of coded pictures, whereas syntax elements in the Picture layer and below apply to a specific coded picture. This means that when a decoder processes the bitstream and decodes the closed_gop flag, it does not decode an indication about any particular coded frame, but instead receives information that pertains to an entire GOP. There is no further information provided about a particular frame, or identifying a particular frame, when the decoder processes the closed_gop flag. In other words, the closed_gop flag is only an indication of existence of a closed GOP, i.e. a GOP that has been encoded without motion vectors pointing to the previous GOP (Ex. 1005 at 43). The closed_gop flag does not tell the decoder anything about an indication of at least one image frame, in decoding order, of an independent sequence. An indication in the context of the '321 Patent is not merely indicating the existence of a particular frame. See e.g. '321 Patent at 4:20-30, 'An indication of at least one image frame is encoded into the video sequence... in the decoding phase, the indication of at least one image frame is decoded from the video sequence, and the decoding of the video sequence is started from said first image frame of the independent sequence'. In other words, the indication is used by the decoder to start decoding from the indicated frame, i.e. the indication needs to identify (indicate) a specific frame. For this reason alone, the closed_gop flag cannot and does not indicate any specific coded frame in an independent sequence, including the first frame in the GOP. Further, when set to '1', MPEG-1's closed_gop flag informs a

decoder that ‘the group of pictures has been encoded without motion vectors pointing to the previous group of pictures’ (EX 1005 at 2.4.3.3, page 43). Thus, if the closed_gop flag indicates anything, it indicates a property of P- and/or B-pictures in the GOP (i.e. a constraint on the use of motion vectors), since I-pictures never contain motion vectors.

112. Further, in my opinion, a POSITA would understand “indication” in the claims to refer to information in an encoded bitstream that necessarily informs decoders of something. Here, the indication is “of at least one image frame, which is the first image frame, in decoding order, of an independent sequence.” A POSITA would understand that this “indication” cannot merely be another element (in this case a particular image frame) that may coincidentally appear alongside certain information in the bitstream (in this case, the closed_gop flag). That is, in video coding, an indication in a bitstream must inform a decoder of something with certainty.

113. Logically, an indication takes the following form in a coded bitstream such as an MPEG-1 bitstream:

If *indication* is present, then *thing indicated* is confirmed
necessarily to the decoder.

114. However, Petitioners’ MPEG-1 Ground relies on a closed_gop flag in

MPEG-1 that does not indicate what Petitioners allege to be the first image frame. That is, in MPEG-1 the fact that a closed_gop flag has a value of “0” or “1” does not inform a decoder that any type of image frame must necessarily follow. Rather, an MPEG-1 decoder that receives a closed_gop flag with a value of “0” or “1” has no certainty about what type of frame will follow, since this is determined by what is actually received in the decoded bitstream, which may for example be affected by transmission errors, including interruptions, bit errors, packet losses, DVD/video-CD read errors, truncated files, etc. These examples are all possible and, in fact, likely and frequent occurrences.

115. The concept of an “indication” that can be decoded from a coded bitstream applies across video coding systems. I will provide an example of an indication from the H.264 Standard.

116. The H.264 Standard defines a “NAL unit” as “A syntax structure containing an indication of the type of data to follow and bytes containing that data in the form of an RBSP interspersed as necessary with emulation prevention bytes.” Ex. 2017, ITU-T Rec. H.264 (03/2005) at 8 (as paginated within the exhibit). In H.264, the NAL unit syntax includes a parameter called nal_unit_type that serves as an indication of the type of data contained in a NAL unit. Table 7-1 (excerpted below) provides possible values of nal_unit_type. Ex. 2017, ITU-T Rec. H.264 (03/2005) at 57 (as paginated within the exhibit).

Table 7-1 – NAL unit type codes

nal_unit_type	Content of NAL unit and RBSP syntax structure	C
0	Unspecified	
1	Coded slice of a non-IDR picture slice_layer_without_partitioning_rbsp()	2, 3, 4
2	Coded slice data partition A slice_data_partition_a_layer_rbsp()	2
3	Coded slice data partition B slice_data_partition_b_layer_rbsp()	3
4	Coded slice data partition C slice_data_partition_c_layer_rbsp()	4
5	Coded slice of an IDR picture slice_layer_without_partitioning_rbsp()	2, 3
6	Supplemental enhancement information (SEI) sei_rbsp()	5
7	Sequence parameter set seq_parameter_set_rbsp()	0
8	Picture parameter set pic_parameter_set_rbsp()	1

117. A POSITA would understand that, for example, nal_unit_type = 7 serves as an indication of a sequence parameter set. In other words, if nal_unit_type=7, then the NAL unit is necessarily a sequence parameter set. If (hypothetically) this were not the case, then a decoder would not know with certainty based on nal_unit_type what type of data was contained in the NAL unit and nal_unit_type would not serve as an indication of the type of data contained in a NAL unit. In such a hypothetical scenario, the decoder would need other information from the bitstream to resolve what type of data was contained in the NAL unit (if the decoder were ever to resolve the uncertainty).

118. Nor can the closed_GOP flag be used inherently to determine an initiation picture, or any specific coded frame or picture, in an independently decodable group of pictures. MPEG-1 specifies that a “group of pictures shall contain one or more I-pictures.” Ex. 1005, 85. That is, MPEG-1 does not constrain the number of I-pictures within a group of pictures and therefore there is no upper bound to that number.

119. The Figures and examples in MPEG-1 illustrate that a wide variation in GOP sizes and structures is possible, such that for example there may be a varying number of I-pictures in GOPs that conform to MPEG-1:

Some examples of groups of pictures are given below:

```

      I
      I  P  P
      I  B  P  B  P
      B  B  I  B  P  B  P
      B  B  I  B  B  P  B  B  P  B  B  P
      B  I  B  B  B  B  P  B  I  B  B  I  I

```

Figure D.17 -- Examples of groups of pictures in display order

Ex. 1005, 85.

120. Thus, while MPEG-1’s closed_gop flag can indicate whether or not the GOP is closed, i.e. whether or not P- and B-pictures in the GOP have been encoded without motion vectors that point to the previous GOP, it cannot and does not identify or indicate an I-picture in the GOP that is the first image, in decoding order, of that group of pictures.

121. This distinction supports my view that MPEG-1’s closed_gop flag fails

to meet the requirement of the claim. The “idea of the [’321 Patent’s] invention” “is to determine an initiation picture in an independently decodable group of pictures, whereby in the decoding phase any picture prior to said initiation picture is defined as unusable for reference.” ’321 Patent, 4:31-38. Conversely, if such a picture is not identified or determined, the scalability and random access benefits described by the ’321 Patent would not be achieved.

122. I understand that in a prior proceeding relating to the same grounds challenging the ’321 Patent, the Board granted institution and, in doing so, the Board relied on the prior petitioner’s representation that the following five factors relating to MPEG-1 were all true:

- (1) a GOP header precedes the first image frame in the GOP in decoding order;
- (2) the GOP header contains the closed_gop flag specifying that the GOP is closed;
- (3) the closed_gop flag in the GOP header is decoded before the first image frame in the closed GOP is decoded;
- (4) the closed_gop flag indicates that the closed GOP is an “independent sequence”; and
- (5) the closed_gop flag also indicates that “the next image frame decoded is the first image frame in decoding order” of the “independent sequence,” i.e., of the closed GOP.

Ex. 2018 (Decision Granting Institution of Inter Partes Review 35 U.S.C. § 314 in IPR2024-00691, hereinafter Amazon Institution Decision).

123. Based on these five factors, the Board determined that the “indication

of at least one image frame, which is the first image frame, in decoding order, of an independent sequence” is met by the closed_gop flag. Ex. 2018 (Amazon Institution Decision), at 26-27. In my view, at least two of these five factors are premised on incorrect interpretations.

124. Considering factor (1), “a GOP header precedes the first image frame in the GOP in decoding order,” this is not always the case in an MPEG-1 bitstream. As Dr. Saber agrees, if playback begins from a random point in the bitstream, an MPEG-1 decoder will “discard all [] bits until it finds a sequence start code, a group of pictures start code, or a picture start code which introduces an I-picture.” Ex. 1003, ¶103.

125. MPEG-1 defines a particular structure for a coded video sequence. A sequence begins with a sequence start code, the first syntax element in a sequence header, which is followed by one or more GOPs (Ex 1005 at 32).

126. A GOP is a series of one or more coded pictures (Ex 1005 at 32). Each GOP starts with a group_start_code (*Id.* at 37) and contains one or more picture() syntax structures.

127. A picture starts with a picture_start_code (*Id.* at 38) and contains one or more slices.

128. The syntax element closed_gop flag (in the GOP header) indicates

whether or not the GOP has been encoded without motion vectors pointing to the previous GOP. Ex. 1005 at 43. It does not indicate the first frame in decoding order.

129. In my view, Petitioner's argument regarding `closed_gop` reads out of the claim, the requirement to indicate one image frame of an independent sequence. Under petitioner's argument, any indication of an independent sequence would of necessity indicate one image frame of an independent sequence, because any sequence of coded frames necessarily begins with a first frame in decoding order. Petitioner's `closed_gop` argument rewrites and broadens the claim, from 'decoding an indication of ... the first image frame in decoding order of an independent sequence' to 'decoding an indication of an independent sequence'.

130. Considering factor (3), "the `closed_gop` flag in the GOP header is decoded before the first image frame in the closed GOP is decoded," the board and petitioner in the Amazon proceeding appeared to assume that the `closed_gop` flag is always decoded before the first image frame in a closed GOP is decoded. Ex. 2018 (Amazon Institution Decision) at 27. This assumption is incorrect. The `closed_gop` flag in the GOP header is not always decoded before a first image frame is decoded. For example, consider a typical (and indeed frequently occurring) random access scenario. When a user selects a frame in the middle of a video sequence, MPEG-1 teaches that all bits are discarded until a start code is reached, as discussed above. In Example (c) of Figure D.19 of MPEG-1 (Ex. 1005, 87), when the user selects, for

example, frame B₄, the first frame that is decoded is frame I₈. In this instance, neither the closed_gop flag nor any syntax elements of the GOP header are decoded or read in the bitstream; however, frame I₈ is decoded first. Thus, factor (3) is not true for every MPEG-1 decoding scenario. Although factor (3) may be true in some circumstances, MPEG-1 teaches that it is not true for others. For at least this reason, MPEG-1 does not teach a decoder that decodes an indication of an image frame, which is the first image frame, in decoding order, of an independent sequence.

131. Considering factor (5), “the closed_gop flag also indicates that ‘the next image frame decoded is the first image frame in decoding order’ of the ‘independent sequence,’ i.e., of the closed GOP,” the prior board and prior petitioner again assumed that the closed_gop flag is always decoded before the first image frame in a closed GOP is decoded. Ex. 2018 (Amazon Institution Decision) at 27. This is not the case for MPEG-1. As I discussed above, the first image frame decoded is not necessarily the image frame following the closed_gop flag. Instead, MPEG-1 teaches that the first decoded image frame is the frame following one of the aforementioned start codes and therefore the closed_gop flag is not always decoded and does not indicate the first image frame in decoding order, as discussed above. Thus, for similar reasons discussed for factors (1) and (3), factor (5) is also not taught by MPEG-1.

**ii. MPEG-1's Temporal Reference Is Not the Claimed
"Indication"**

132. *Second*, Petitioner points to MPEG-1's Temporal Reference of zero as an alternative teaching of "decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of [the] independent sequence." Pet. 24-26. Petitioner is in my view incorrect that MPEG-1's Temporal Reference of zero teaches this claim element. MPEG-1 does not teach that a Temporal Reference of zero is an indication of the first image frame in decoding order of an independent sequence. Rather, MPEG-1 teaches that "[t]he first picture, in display order, in each group must have Temporal Reference equal to zero." Ex. 1005, 87, emphasis added. As I explained earlier, display order and decoding order are not the same in MPEG-1, for any sequence that contains B-pictures.

133. MPEG-1 further teaches that "[t]he Temporal Reference is a ten-bit number which can be used to define the order in which the pictures must be displayed." *Id.* The Temporal Reference is not used only for closed GOPs (what Petitioner alleges to be independent sequences), but also open GOPs (which according to Petitioner's interpretation are not independent sequences). Thus, MPEG-1's Temporal Reference (signalled by the temporal_reference syntax element) does not indicate the first image frame in decoding order of an independent

sequence; it is merely the first picture in display order as taught by MPEG-1. Because of the many possible GOP configurations in MPEG-1, a decoder cannot rely on a Temporal Reference of 0 to indicate a first image frame in decoding order, as I explain further below.

D.5.3.2 Temporal reference

The Temporal Reference is a ten-bit number which can be used to define the order in which the pictures must be displayed. It may be useful since pictures are not transmitted in display order, but rather in the order which the decoder needs to decode them. The first picture, in display order, in each group must have Temporal Reference equal to zero. This is incremented by one for each picture in the group.

Some example groups of pictures with their Temporal Reference numbers are given below:

Example (a) in display order	I	B	P	B	P														
	0	1	2	3	4														
Example (a) in decoding order	I	P	B	P	B														
	0	2	1	4	3														
Example (b) in display order	B	B	I	B	B	P	B	B	P	B	B	P	B	B					
	0	1	2	3	4	5	6	7	8	9	10	11							
Example (b) in coded order	I	B	B	P	B	B	P	B	B	P	B	B							
	2	0	1	5	3	4	8	6	7	11	9	10							
Example (c) in display order	B	I	B	B	B	B	P	B	I	B	B	I	I						
	0	1	2	3	4	5	6	7	8	9	10	11	12						
Example (c) in coded order	I	B	P	B	B	B	B	I	B	I	B	B	I						
	1	0	6	2	3	4	5	8	7	11	9	10	12						

Figure D.19 -- Examples of groups of pictures and temporal references

Ex. 1005, 87.

134. Considering the entirety of MPEG-1's teaching, it is simply untrue that a closed GOP will always have a first picture in decoding order that is an "I-frame with a Temporal Reference of zero" as the Petition suggests (Petition at 25). Neither the closed_gop flag nor the temporal_reference syntax element have any such constraint. For example, Figure D.19 shows multiple instances where an I-picture is not the first image in display order and an I-picture does not have a Temporal

Reference of zero, in contrast to the claims' requirement that the indication be of the first image frame in decoding order (not display order):

D.5.3.2 Temporal reference

The Temporal Reference is a ten-bit number which can be used to define the order in which the pictures must be displayed. It may be useful since pictures are not transmitted in display order, but rather in the order which the decoder needs to decode them. The first picture, in display order, in each group must have Temporal Reference equal to zero. This is incremented by one for each picture in the group.

Some example groups of pictures with their Temporal Reference numbers are given below:

Example (a) in display order	I	B	P	B	P								
	0	1	2	3	4								
Example (a) in decoding order	I	P	B	P	B								
	0	2	1	4	3								
Example (b) in display order	B	B	I	B	B	P	B	B	P	B	B	P	
	0	1	2	3	4	5	6	7	8	9	10	11	
Example (b) in coded order	I	B	B	P	B	B	P	B	B	P	B	B	
	2	0	1	5	3	4	8	6	7	11	9	10	
Example (c) in display order	B	I	B	B	B	B	P	B	I	B	B	I	I
	0	1	2	3	4	5	6	7	8	9	10	11	12
Example (c) in coded order	I	B	P	B	B	B	B	I	B	I	B	B	I
	1	0	6	2	3	4	5	8	7	11	9	10	12

Figure D.19 -- Examples of groups of pictures and temporal references

Ex. 1005, 87.

135. As another example, Figure D.18b shows an example of a GOP that may be closed, where the first I-picture is not the first image in display order and does not have a temporal reference of zero:

B	B	I	B	B	P	B	B	P	B	B	P	B	B	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

(b) open or closed group

Figure D.18 -- Example groups of pictures in display order

Ex. 1005, 86.

136. Petitioner suggests that a Temporal Reference of zero, when present in an I-picture in what Petitioner describes as ‘MPEG-1’s exemplary closed GOP’ (the GOP of MPEG-1 Figure D.18a, Ex. 1005 at 86), indicates the first frame in decoding order of the alleged independent sequence. Petition at 25. First, there is no such thing as an ‘MPEG-1 exemplary closed GOP’, as I explained earlier. The example of MPEG-1 Figure D.18a is not ‘exemplary’, nor is it ‘the exemplary closed GOP’. It is merely one example of a closed GOP in MPEG-1. Second, the fact that even a closed GOP such as the example of MPEG-1 Figure D.18b has a first picture in display order which is a B-picture, not an I-picture, shows that a Temporal Reference of zero does not indicate to the decoder the presence of a first picture in decoding order, even when closed_gop is set to 1. Third, a Temporal Reference of zero can occur in a large group of pictures (more than 1024 pictures), in which case it does not coincide with the first picture in decoding order (Ex. 1005, 87). In other words, an MPEG-1 decoder that decodes a Temporal Reference of zero, cannot assume that the current picture is the first picture in decoding order.

137. To reiterate, there is no ‘exemplary closed GOP’ in MPEG-1. Petitioner has invented the term ‘exemplary closed GOP’. A decoder does not know what form each GOP will take, it only knows that each GOP obeys the constraints of the standard. There is no constraint in MPEG-1 that every GOP must take the form of Fig. D.18a and there is no scenario in which an MPEG-1 decoder ‘knows’ that every

GOP in a bitstream takes this form. This means that decoding a Temporal Reference of 0 does not in itself indicate anything to the decoder about whether the current frame is the first image frame in decoding order of a GOP. The fact that some possible GOP structures, but not others, happen to begin with an I-picture in both decoding and display order with a Temporal Reference of 0, is coincidence. The presence of a Temporal Reference of 0 only informs the decoder that the current frame (which may be an I-, P- or B-picture and may be at the start of a GOP or the 1025th picture in a GOP in display order) has a Temporal Reference of 0.

138. Thus, it is my opinion that the Petitioner has failed to show that an MPEG-1 decoder discloses an indication of an image frame, which is the first image frame, in decoding order, of an independent sequence as required by the [a] limitations.

- b. Petitioner fails to demonstrate that limitations 8[d], 10[d], and 11[d] are obvious over MPEG-1 because MPEG-1 does not teach or suggest a method for decoding, a video decoder or a computer program product for decoding, resetting the identifier value as required**

139. Each of limitations 8[d], 10[d], and 11[d] are written with similar language and require “[resetting/reset] the identifier value for the indicated first image frame of the independent sequence.” Petitioner has failed to show that MPEG-1 teaches or suggests these [d] limitations.

140. Petitioner alleges that the Temporal Reference (temporal_reference

syntax element) being set to zero in an MPEG-1 bitstream discloses the claimed “resetting the identifier value for the indicated first image frame of the independent sequence.” Petitioner does not provide a separate argument for its ‘closed_gop flag’ or ‘temporal_reference syntax element’ arguments discussed above with respect to earlier claim limitations. In my opinion, Petitioner is wrong.

141. First, the challenged claims require the decoder (not the encoder) to reset the identifier value for the indicated first image frame of the independent sequence. The ‘321 patent is clear that this is an important aspect of the invention: “Consequently, when activated at a random point, the **decoder preferably sets the identifier** of the first image frame, preferably an I-frame, of the independently decodable sub-sequence.” ’321 Patent, 11:26-40, emphasis added. In fact, the Temporal Reference is assigned by the encoder by means of the temporal_reference syntax element. Thus, MPEG-1’s temporal reference can only be set or reset by the encoder (not the decoder) for the first image frame in display order. Ex. 1005, 44, 87. Thus, MPEG-1 fails to teach this claim element.

142. The Temporal Reference (temporal_reference syntax element) is a syntax element whose value is determined by the encoder and which informs the decoder of the display order of decoded pictures “since pictures are not transmitted in display order, but rather in the order which the decoder needs to decode them.” Ex. 1005, 87. But MPEG-1 does not teach that the decoder resets the Temporal

Reference to 0. At most, it teaches that the encoder assigns the Temporal Reference of a first picture in display order to zero in certain situations, when encoding an MPEG-1 compliant bitstream, and that an MPEG-1 compliant decoder reads that syntax element from the bitstream, which may be zero-valued. Ex. 1005, 87. Whilst MPEG-1 does not specify an encoder, in order for an encoder to produce a bitstream that conforms to MPEG-1, it must construct the bitstream in a certain way, including assigning a Temporal Reference of a first picture in display order to zero in certain situations according to the MPEG-1 Standard.

143. Petitioner argues that after decoding an exemplary closed GOP, “when an MPEG-1 decoder decodes the next exemplary closed GOP, the Temporal Reference for the first picture in decoding order of that exemplary closed GOP **would be reset to zero**, as that picture would be the first picture in display order for that next GOP.” Pet. 29-30, emphasis added.

144. In the Amazon Institution Decision, the board cited the same passage. Ex. 2018 (Amazon Institution Decision), at 35-36 (citing to Ex. 1005, 44). In my view, this argument fails for two reasons. First, the encoder, not the decoder assigns the Temporal Reference value and therefore there is no system or method for decoding in MPEG-1 that resets the Temporal Reference value. Ex. 1005, 87. Petitioner appears to acknowledge this through the choice of language above, acknowledging that the Temporal Reference ‘would be reset to zero’ but not that an

MPEG-1 decoder itself performs ‘resetting’ this value. As I explained above, a decoder that conforms to MPEG-1 cannot reset the Temporal Reference value.

145. Petitioner suggests that ‘if the [MPEG-1] decoder receives and decodes the value of a Temporal Reference for the first image frame and cannot change the value of that syntax element, it would have been obvious to a POSITA that **an MPEG-1 decoder resets the Temporal Reference to zero** when it decodes the first image frame in decoding order of the exemplary closed GOP’. Petition 30-31, emphasis added. I disagree. As I have explained above, an MPEG-1 decoder does not and cannot reset Temporal Reference or any other such decoded syntax element. Petitioner’s suggestion is completely at odds with a POSITA’s understanding of the well-known engineering term ‘to reset’. A decoder that happens to receive in a bitstream a value of 0, does not reset that value to zero, it merely responds to it.

146. Second, the claim language requires “resetting the identifier value for *the* first image frame of *the* independent sequence.” See ’321 Patent, Claim 8 (and similar for other challenged claims), emphasis added. That is, the resetting of the identifier value is reset for the independent sequence of the prior limitations, which I understand to be the antecedent basis for this limitation.

147. However, Petitioner’s argument relies on the Temporal Reference

being set to zero for the subsequent first image frame of the subsequent independent sequence, not the first image frame of the actual, claimed independent sequence being decoded: “when an MPEG-1 decoder decodes *the next* exemplary closed GOP, the Temporal Reference for the first picture in decoding order of *that exemplary closed GOP* would be reset to zero, as that picture would be the first picture in display order for *that next GOP*.” Pet. 29-30 (emphasis added). Thus, even accepting Petitioner’s argument concerning resetting (which I do not accept, as Petitioner cannot show an MPEG-1 decoder resetting any identifier value), Petitioner has pointed to a resetting of “the next independent sequence” not “the independent sequence” referred to earlier in the same claim.

148. Thus, Petitioner has failed to show that an MPEG-1 decoder discloses the claimed resetting the identifier value for the indicated first image frame of the independent sequence.

2. Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 2—Kim.

149. As I explain in detail below, there are multiple reasons why Kim does not teach or suggest to a POSITA the limitations of independent claims 8, 10, or 11.

a. Petitioner fails to demonstrate that limitations 8[a], 10[a], and 11[a] are obvious over Kim

150. Each of limitations 8[a], 10[a], and 11[a] are written with similar language and require “[decoding/decode] from the video sequence an indication of

at least one image frame, which is the first image frame, in decoding order, of an independent sequence, wherein all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence.” In my opinion, Petitioner has failed to show that Kim teaches or suggests these limitations.

151. The Petition points to the IAD of the I-frame as the claimed “indication.” Pet. at 46.

152. However, there are multiple reasons why a POSITA would have understood that Kim neither teaches nor suggests claims 8[a], 10[a], or 11[a] of the ’321 Patent.

i. Kim’s IAD of the I-frame is not an “indication” of anything.

153. First, Kim does not disclose or render obvious the claimed “indication” to a person of ordinary skill because the IAD of an I-frame is not an indication of anything.

154. As the Petition notes, the IAD of an I-frame “is zero *or absent*.” Ex. 1006 (Kim) at 4:65-5:1 (italics added); Petition at 47. A POSITA reading Kim would have understood that the reason the value of the IAD field of an I-frame can be zero *or absent* is because it has no meaning and will never be used. Unlike the IAD of a P-frame, which refers to “the position value of the corresponding (or previous) I-

frame,” the IAD field of an I-frame has no corresponding meaning or use. Ex. 1006 (Kim) at 4:27-33. This is because the I-frame does not rely on prediction from any other frame and therefore it makes no sense for it to have an associated IAD field with any meaning. In fact, while Kim repeatedly describes the IAD of the P-frame as being a “position value” of the corresponding I-frame, Kim never describes that the IAD field *of the I-frame* has any meaning at all. Ex. 1006 (Kim) at 4:27-33, 4:48-52; claims 1-2, 8, 11. A POSITA would understand that Kim allows that value to be “zero or absent” because the IAD of the I-frame has no meaning or use, and the encoder and decoder do not care about the value of the IAD of the I-frame at all.

155. This is confirmed in Kim’s Fig. 4, which I have excerpted and reproduced below with colored annotations.

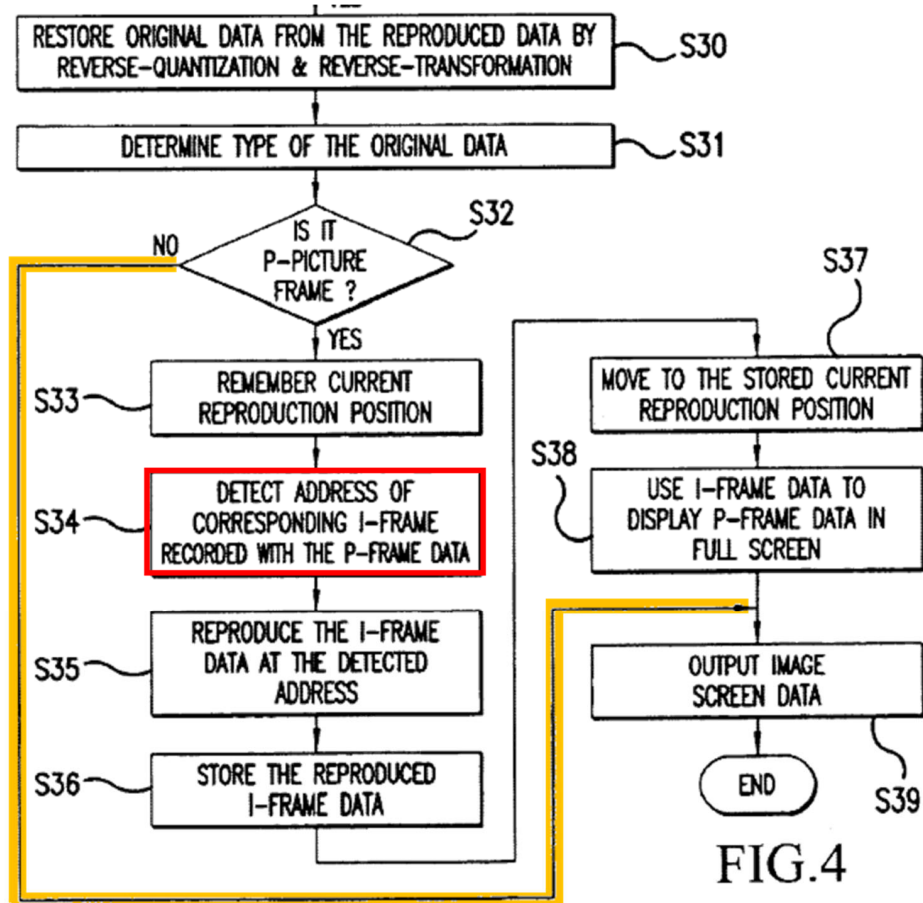


FIG.4

Ex. 1006 (Kim) at Fig. 4 (excerpted and annotations added)

156. As I show in the annotated Fig. 4 above, the process for decoding an I-frame (shown in orange) skips Step S34 (outlined in red): “detect address of corresponding I-frame recorded with the P-frame data.” Thus, Kim’s decoder never detects the IAD for the I-frame. It only detects the IAD for the P-frame. Kim explains that when the requested image is an I-frame, the corresponding coded image data is simply read from the optical disk, decoded and displayed. Ex. 1006 (Kim) at 5:66-6:3.

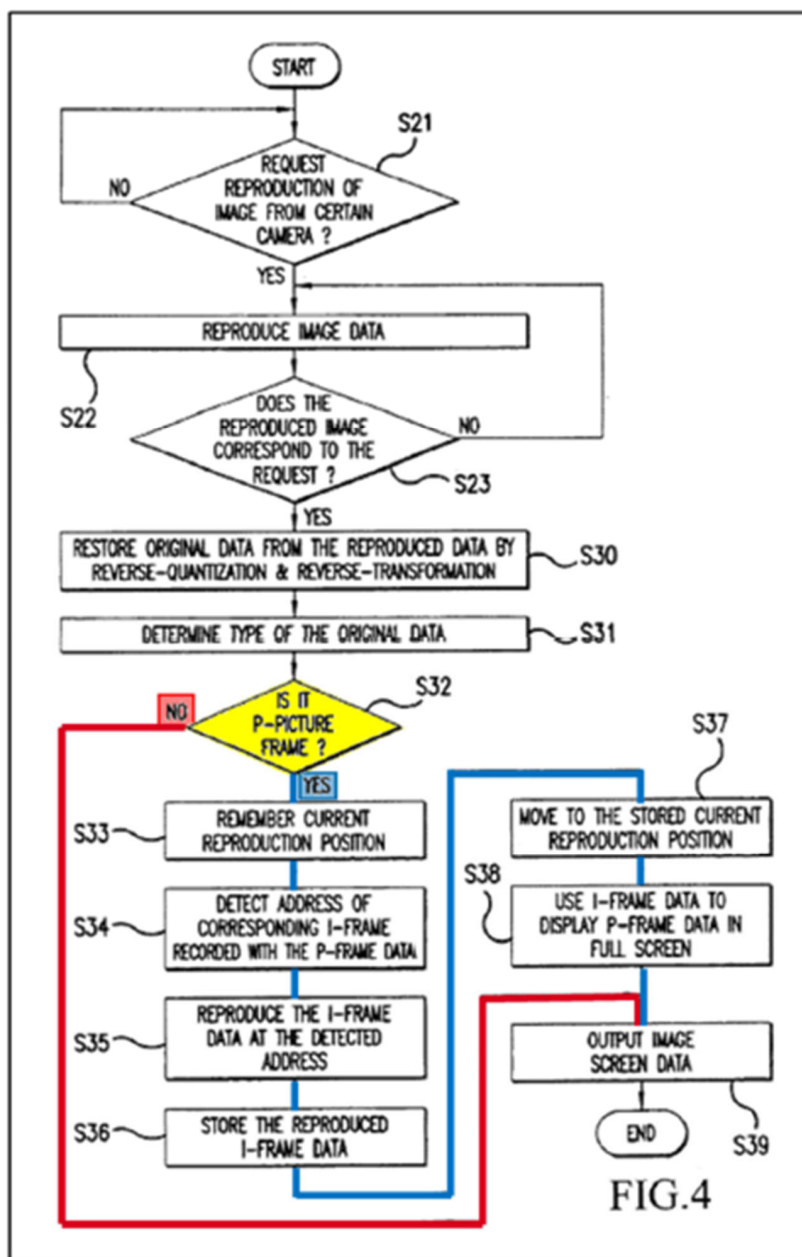
157. Kim never describes any step where the IAD of the I-frame is ever decoded, read, or used for anything. That is why Kim teaches that the IAD of the I-frame may even be absent because it makes no difference to the decoder. Ex. 1006 (Kim) at 4:65-5:1.

158. This is further confirmed by Kim's claims, which only specify that the step of "detect[ing] the address of the corresponding intra-picture frame" occurs "if the restored image data frame is predictive picture frame data" (P-frame) and not "if the original image frame data type is not predictive picture frame data" (I-frame). Ex. 1006 (Kim) at claim 11.

159. Petitioner agrees with me that Step S34 of the decoding process to read the IAD only happens for the P-frame. The I-frame does not perform any such step, as shown from the excerpt from the Petition below:

As shown below in Kim's Figure 4, Kim sets forth two paths for decoding the disclosed bitstream depending on whether the user requests image data that corresponds to an I-picture or a P-picture. (ASUS-1003, ¶126.) When the request corresponds to an I-picture, the picture is decoded and output as an image signal. (ASUS-1006, 5:66-6:3, Fig. 4 ("no" path annotated in red).) When the request is directed to a P-picture, the system "detects the address of the corresponding I-picture frame...and reproduces based on the detected address [of that] I-picture." (ASUS-1006, 5:38-48; Fig. 4 ("yes" path annotated in blue).)

Pet. at 42.



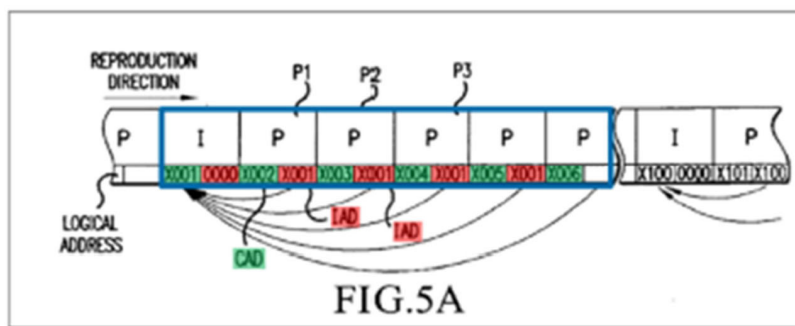
Pet. at 43 (colors and annotations in original).

160. As the Petition highlights in red, when the image is an I-frame, Kim's decoder skips Step S34 and does not detect the IAD of the I-frame. That Step S34 to detect the IAD, only occurs when the image is a P-frame (shown in blue).

161. I note that although Kim's Fig. 6 labels the IAD of the I-frame as the

“corresponding I-picture address,” a POSITA would have understood that this cannot possibly be an indication of the I-frame’s address because every I-frame is given the same “zero or absent” value for the IAD. Ex. 1006 (Kim) at 4:34-45 (similar statements about the “address of the I-picture data,” which a POSITA would know could not actually be the address of I-frame). Indeed, the Petition states that the I-frame’s logical address, as shown in Fig. 5A, is the CAD value, not the IAD value:

(ASUS-1006, 4:46-52.) I-pictures also have a CAD indicating the picture’s current address (ASUS-1006, 4:65-5:1, Figs. 5A, 6 (“physical address”)), but the I-picture’s IAD is zero (ASUS-1006, Figs. 5A (showing “0000”), 6 (showing “00000000” for the “corresponding I-Picture address”)). Figure 5A below is annotated to highlight the IAD values in red and CAD values in green.



Pet. at 41-42 (colors and annotations in original).

162. Therefore, a POSITA would have understood that the IAD of the I-frame has no corresponding meaning at all and is not an indication of anything.

163. I also note that it is not the case that the IAD field of the I-frame has no meaning when it is absent but has some meaning when it is present as 0000. The

IAD of the I-frame, no matter what value it takes, has no meaning or use in either Kim's encoding or decoding processes. I further explain the 0000 value from the perspective of a POSITA in paragraph 95.

164. I understand that expert Dr. Omid Kia testified in an ITC Investigation where the '321 Patent was asserted that "[f]or an I-frame, that IAD does not serve any purpose" and "that value serves no function in the decoding." Ex. 2019 (1380 Public Tr.) at 26 (1733:1-7).

165. I further understand and have been informed by counsel that the same ITC Investigation found that the asserted decoding claims 8 and 10 of the '321 Patent were valid over Kim.

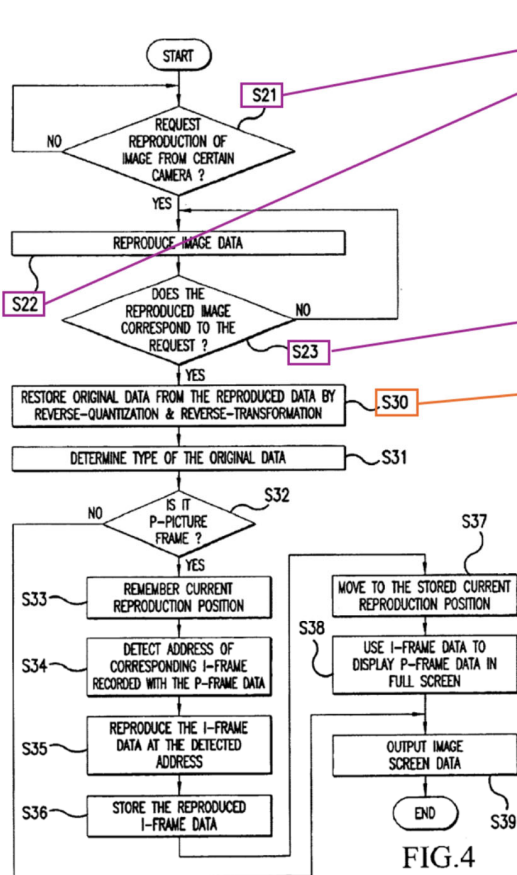
166. Neither the Petition nor Dr. Saber assert that it would have been obvious to modify Kim in any way to make the IAD of the I-frame correspond to an indication of anything or be used for any purpose.

ii. Kim's IAD of the I-frame is not an "indication of at least one image frame, which is the first image frame, in decoding order."

167. Second, Kim does not disclose or render obvious the claims because the IAD of the I-frame is not the required "indication of at least one image frame, which is the first image frame, in decoding order" of the claims.

168. More specifically, an IAD value of 0000 for an I-frame does not indicate the first image frame in decoding order. As Kim explains, the first image

data that is decoded is the image data for the frame the user selected (Step S30 of Fig. 4). Ex. 1006 (Kim) at 5:29-34. There is no consideration of the value of the IAD for the I-frame to indicate which image frame should be decoded first. In fact, as I stated above, the IAD for the I-frame is never read or used at all. The first image frame to be decoded is simply whichever frame the user selects (Steps S21 and S22 of Fig. 4). Ex. 1006 (Kim) at 5:13-32.



As shown in Steps S21 and S22 of FIG. 4, if the user requests reproduction of an image data from a particular camera or channel among the intermittent image data recorded on the optical disc 1, the controller 150 controls the recording/reproducing unit 140 to reproduce the corresponding image data from the optical disc 1. The image data reproduced from the optical disc 1 is multiplexed data, which is demultiplexed by the de-MUX 131. By demultiplexing, the camera (channel) ID code CHID is separated from the reproduced image data and output to the controller 150. Based on the camera ID code CHID, the controller 150 determines whether or not the reproduced image data corresponds to the selected camera or channel, in Step S23. This ensures that the correct image data is processed to comply with the user's request. If incorrect image data has been reproduced, the controller 150 controls the recording/reproduction unit 140 to reproduce the correct image data from the optical disc 1. In Step S30, the image data corresponding to the user's request is transmitted and decoded in the decoder 111. The decoding process of the decoder 111 involves reverse-quantization and reverse-transformation to restore the original image data prior to data compression carried out during recording. Then the property (type) of the currently restored image data is determined in Step S31 by the reproducing DSP 121, and the determination result is transmitted to the controller 150. More specifically, if the currently restored image data is detected to be P-picture frame data in Step S32, the controller 150 stores the current reproduction position in Step S33. Then in Step S34, the controller 150 detects the address of the corresponding I-picture frame (e.g., data stored in the area (3) of the bit stream B2 shown in FIG. 6) from the reproduced image data output from the de-MUX 131, and reproduces based on the detected address the corresponding (or previous) I-picture frame data recorded on the optical disc 1 by controlling the recording/reproduction unit 140.

In Step S36, the reproduced corresponding I-picture frame data is processed by the reproducing DSP 121, and decoded by the decoder 111 to be stored in the frame memory 101. The controller 150, after Step S36, controls the recording/reproducing unit 140 to move the video heads to the original recording position stored in Step S33.

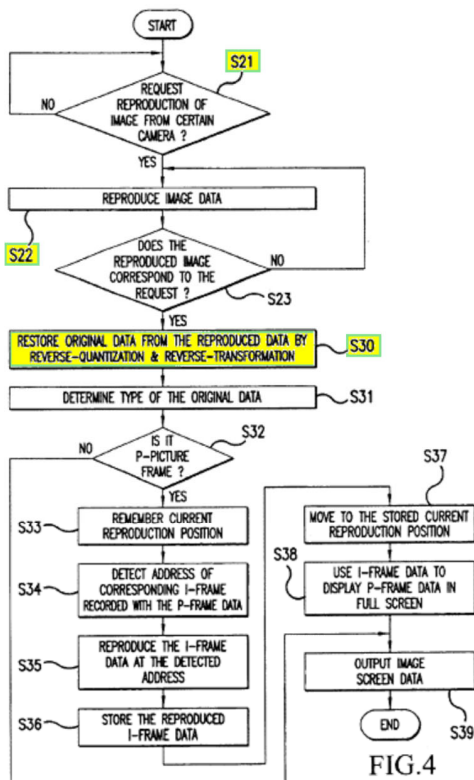
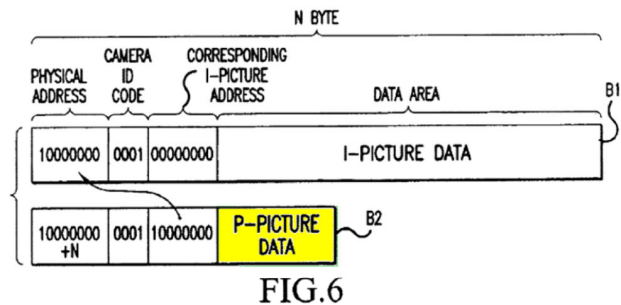
FIG.4

Ex. 1006 (Kim) at Fig. 4 and 5:1-41 (annotations added).

169. In fact, Kim's decoder fully decodes all of the image data in Step S30 before determining at a later point (Step S31) whether that selected frame is an I-

frame or a P-frame. Ex. 1006 (Kim) at 5:30-38.

170. The example where a user selects a P-frame to decode is instructive. If the requested image is a P-frame, Kim's decoder fully decodes all of the P-frame's image data (which is only part of the full P-frame) in Step S30, including reverse-quantization and reverse-transformation. Ex. 1006 (Kim) at 5:30-35 ("In Step S30, the image data corresponding to the user's request is transmitted and decoded in the decoder 111. The decoding process of the decoder 111 involves reverse-quantization and reverse-transformation to restore the original image data prior to data compression carried out during recording."). Thus, the IAD of the I-frame does not indicate the first image frame in decoding order because the first image frame decoded is whatever picture the user selects, as shown by the example of when the user selects a P-frame.



Ex. 1006 (Kim), Fig. 4 (highlighting added to Steps S20, S21, S30) and Fig. 6 (highlighting added to the corresponding partial P-frame that is decoded in Step S30).

171. I understand that Petitioner asserts that because the P-frame is a predicted frame, there will be portions of the image that rely on image data from a corresponding I-frame. (Petition at 49). While it is true that a P-frame in Kim is at least partly predicted from an I-frame, the coded P-frame data is already fully decoded (comprising only a partial P-frame at this point) before Kim's decoder even knows whether the frame is a P-frame or I-frame. Ex. 1006 (Kim) at 5:30-38, explaining that the partial P-frame is decoded, including reverse quantization and

reverse transform, to restore the original image data prior to compression: 'The decoding process of the decoder 111 involves reverse-quantization and reverse-transformation to restore the original image data prior to data compression carried out during recording. Then the property (type) of the currently restored image data is determined in Step S31'. Moreover, Kim explains that the P-frame is still the first image in decoding order when the P-frame is selected.

172. Specifically, Kim explains that if the decoder determines that the image selected by the user is a P-frame in Step S31, the corresponding I-frame address (the IAD of the P-frame) is detected at Step S34 (shown in red below), and then the corresponding I-frame data is decoded *in part* at Step S36. Ex. 1006 (Kim) at 5:35-52. I note that, according to Kim, the I-frame is not completely decoded at this point because the final reverse-transformation and reverse-quantization does not occur until Step S38. In Kim, reverse-transformation and reverse-quantization is a necessary part of the decoding process. Ex. 1006 (Kim) at 5:32-35.

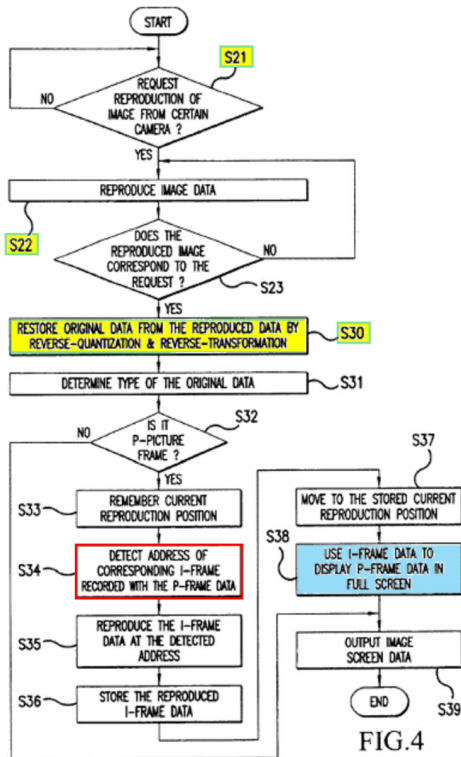
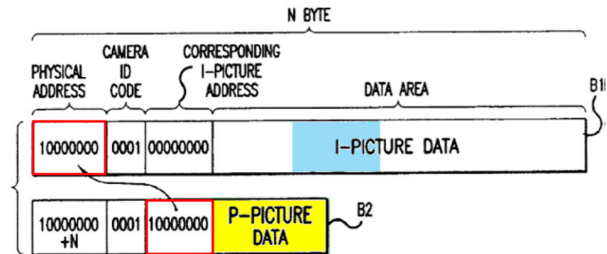


FIG. 6

Ex. 1006 (Kim), Fig. 4 and Fig. 6 (highlighting added).

173. According to Kim, the I-frame data is at most partially decoded when the requested image to decode is a P-frame because the full I-frame itself is never reverse-transformed and reverse-quantized. That is, at Step S38, the partial P-frame is combined with the partially decoded I-frame data and the moving (motion) vectors, and then the complete P-frame image is reverse-transformed and reverse-quantized, which completes the decoding of the full P-frame. Ex. 1006 (Kim) at 5:55-65. Thus, when the requested image to decode is a P-frame, only the P-frame is fully decoded because the full I-frame itself is never reverse-transformed and reverse-quantized (only the parts that are used to predict the P-frame). Thus Petitioner is incorrect to state that in Kim, ‘the I-picture is the first image frame in

decoding order, even when a P-picture is selected'. Petition at 49.

174. This example of where a user selects a P-frame to decode, illustrates how the IAD field of an I-frame does not indicate the first image frame in decoding order. In Kim's method, the IAD of an I-frame never indicates the first image frame in decoding order. In other words, even if an I-frame was selected by the user such that the I-frame is decoded first, the IAD of the I-frame still indicates nothing and has no purpose during decoding, just as it indicated nothing and had no purpose in the case when the P-frame was selected first. This is why the IAD of the I-frame can be entirely absent. Ex. 1006 (Kim) at 4:65-5:1. As I explain in further detail below, in Kim, the IAD of the I-frame is never decoded and is never used. Therefore, it does not indicate the first image frame in decoding order, as the claim requires.

175. I understand that expert Dr. Omid Kia testified in an ITC Investigation where the '321 Patent was asserted that "[i]t is the user that selects the picture to decode" and not the value of the IAD of the I-frame. Ex. 2019 (1380 Public Tr.) at 23-24 (1730:10-1731:2).

176. Neither the Petition nor Dr. Saber address this fundamental problem with Kim and that there is no "decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence."

177. Neither the Petition nor Dr. Saber assert that it would have been obvious to modify Kim in any way to make the IAD field of the I-frame correspond to the claimed “indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence.”

iii. Kim’s IAD of the I-frame is never “decod[ed] from the video sequence”

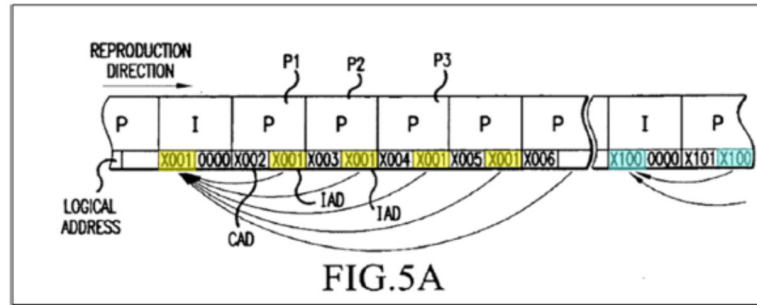
178. Third, even if the IAD field of the I-frame was the claimed indication (and I have explained why it is not), Kim’s decoder never discloses or renders obvious “decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence.”

179. Kim’s decoding process, shown in Fig. 4, teaches that Step S34 to detect the IAD is only performed for a P-frame. That step is skipped for an I-frame. Ex. 1006 (Kim) at 5:66-6:3. The Petition appears to agree with me that Kim’s decoding process only detects the IAD for the P-frame and not for the I-frame. Pet. at 42-43.

180. Dr. Saber states that a “A POSITA would have [] understood that Kim’s decoder decodes the IAD of image frames.” Ex. 1003 (Saber Decl.) at ¶ 132. While he cites to Fig. 4 and the same passages that I refer to above, Dr. Saber makes no distinction or recognition that Kim teaches that only the IAD field *of the P-frames*

is ever read and that the critical Step S34 is skipped for I-frames. He makes no attempt to explain why a POSITA would have understood Kim's decoder to decode the IAD of the I-frames when he admits that Kim expressly teaches *not* to decode the IAD of the I-frames. Ex. 1003 (Saber Decl.) at ¶ 126. Neither the Petition nor Dr. Saber address this fundamental problem with Kim.

181. Dr. Saber also cites to Fig. 5A, stating that a "a POSITA would have understood that [decoding the bitstream in Figure 5A] include[s] decoding the components of the bitstream, which would include IAD values..." Ex. 1003 (Saber Decl.) at ¶ 132. However, his own annotated figure shows that there are no arrows pointing to or from the IAD field of the I-frame. There is simply no evidence or description in Kim that the IAD field of the I-frame is ever used for anything. There is nothing to support Dr. Saber's opinion that a POSITA would have understood the IAD of the I-frame is decoded. In fact, all of the evidence and disclosures in Kim expressly state the opposite. A POSITA would not be motivated to decode or use the IAD field of the I-frame, first because it is "zero or absent" (i.e. a "don't care" value) and second because this would involve an unnecessary decoding and processing step that has no use or benefit within Kim's method.



Ex. 1003 (Saber Decl.), ¶ 131.

182. Even the Board acknowledged in a previous proceeding that “Kim discloses that ‘the image data corresponding to the user’s request is transmitted and decoded’ by a decoder” and then “[a]fter decoding the image data, Kim’s controller uses the IAD in the image data, e.g. to ‘detect[] the address of the corresponding I-picture frame’ *if the image data is a P-picture.*” IPR2025-01068, Paper No. 10 at 58 (emphasis added).

183. Neither the Petition nor Dr. Saber assert that it would have been obvious to modify Kim in any way to make the decoder perform Step S34 for an I-frame or otherwise decode the IAD of the I-frame as the claim requires. Given that the IAD field of the I-frame serves no functional purpose, it would be a waste of processing power and design effort to make such a modification, i.e. to perform a decoding process for a data field that has no purpose and is not required for successful decoding and display of image frames.

iv. Kim does not have a “video sequence.”

184. Fourth, Kim does not disclose or render obvious “decoding from the

video sequence” because there is no teaching or suggestion of a video sequence in Kim.

185. Kim describes intermittent time lapse recording that requires “mov[ing] and stop[ping] the tape repeatedly” to capture time lapsed single images rather than a video sequence. Ex. 1006 (Kim) at 1:57-60. Critically, the purpose of Kim’s described technique is that “only the desired images can be selectively reproduced with high display quality.” Ex. 1006 (Kim) at 6:17-20. In other words, Kim only teaches the select encoding and decoding of individual images at a time, and not the video sequence of the ’321 Patent.

186. I understand that expert Dr. Omid Kia similarly testified in an ITC Investigation where the ’321 Patent was asserted that Kim does not disclose or suggest recording or playback of continuous video. Ex. 2019 (1380 Public Tr.) at 1726:1-23. Rather, Kim only describes decoding a single image instead of a video sequence. Ex. 2019 (1380 Public Tr.) at 1727:5-20.

187. Dr. Saber agrees that Kim teaches only time lapse recording and reproduction. Ex. 1003 (Saber Decl.) at ¶ 122. Critically, neither the Petition nor Dr. Saber ever explain how the encoding and decoding of a single image can satisfy a video sequence of claim 8[a], 10[a], or 11[a]. Petition at 44-51, 66, 69; Ex. 1003 (Saber Decl.) at ¶¶ 127-137, 159-169.

188. Dr. Saber asserts that it would have been obvious to modify Kim to

decode a video sequence. Ex. 1003 (Saber Decl.) at ¶¶ 144. To support this, he points to Figure 4, which he claims contemplates a video sequence merely by the mention of I- and P-frames. *Id.* As I explain above, Kim describes reproducing individual images from intermittent time lapse recordings, and it uses I-frame and P-frame terminology. Kim does not contemplate, nor does it suggest, decoding video sequences.

189. Furthermore, modifying Kim to repeatedly decode images into a video sequence would have required more effort, processing power, memory, and cost. Such a modification would have been less efficient and unnecessary for the applications Kim targeted, which only require decoding of individual, intermittently captured time lapse images rather than a video sequence. Ex. 1006 (Kim) at 1:14-20.

190. Kim's Fig. 5B further confirms that Kim does not teach or suggest a video sequence. Fig. 5B shows a data stream of images takes from three different cameras (CHID: camera (channel) ID code 0, 1, 2). Ex. 1006 (Kim) at 4:56-58.

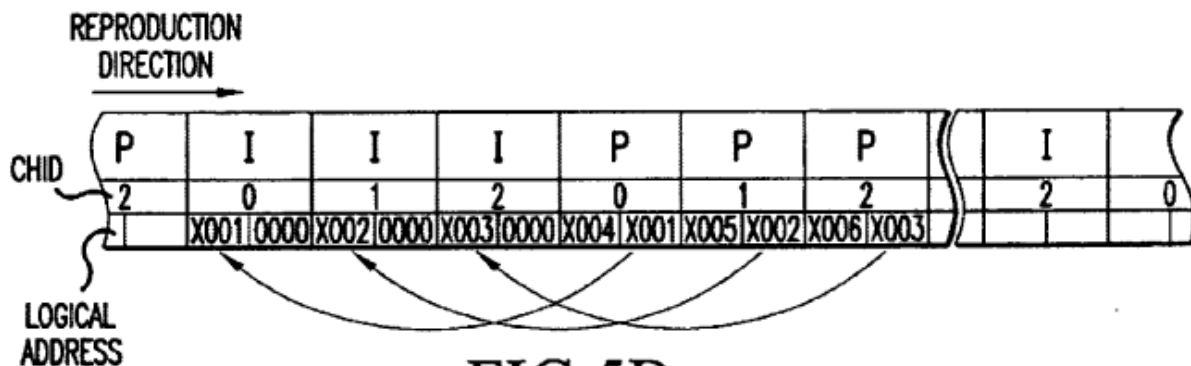


FIG.5B

Ex. 1006 (Kim) at Fig. 5B.

191. Given that the “Reproduction Direction” in Fig. 5B is from left to right, if the stream was treated as a video sequence, every frame would hop to a different camera and cycle through each of the three different cameras before returning to the next frame for a particular camera. Such a disjointed set of images would not make sense to a viewer, and a POSITA would have understood that Kim did not teach or suggest that any such disjointed set of images would comprise a video sequence.

192. I understand that expert Dr. Omid Kia similarly testified in an ITC Investigation in which the ’321 Patent was asserted, that modifying Kim to decode a video sequence would result in a “very time-consuming” implementation that is “very process intensive” as compared to Kim’s actual teaching of “single image retrieval.” Ex. 2019 (1380 Public Tr.) at 1727:21-1728:21.

b. Petitioner fails to demonstrate that limitations 8[b], 10[b], and 11[b] are obvious over MPEG-1

193. Each of limitations 8[b], 10[b], and 11[b] are written with similar language and require an “starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame” where “from said first image frame of the independent sequence” is the first image frame identified from the “indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence” from 8[a], 10[a], and 11[a].

194. Kim does not disclose or render obvious “*starting the decoding of the video sequence from said first image frame of the independent sequence*” because Kim teaches starting decoding from the user-selected frame. Ex. 1006 (Kim) at 5:12-38, Fig. 4.

195. A POSITA would understand that “said first image frame” is the frame that the claimed “indication” of the previous element refers to (for example, 8[a]: “indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence”). Therefore, this element requires starting the decoding from that first frame that was indicated by the “indication” of the previous element.

196. As I explained above in Section VIII(2)(a)(i)-(ii), any IAD field of an

I-frame does not indicate the first image frame in decoding order because the first image frame to be decoded is simply whichever frame the user selects. Ex. 1006 (Kim) at 5:13-32, Fig. 4 (Steps S21 and S22).

197. While there may be certain circumstances where the decoder will start decoding from an I-frame because the user happens to select an I-frame, it would never be because the alleged indication (IAD of the I-frame) indicated to do so. Kim's decoder always starts decoding from the user-selected picture. Section VIII(2)(a)(i)-(ii).

198. I understand that expert Dr. Omid Kia similarly testified in an ITC Investigation in which the '321 Patent was asserted that "[i]t is the user that selects the picture to decode" and not the value of the IAD of the I-frame. Ex. 2019 (1380 Public Tr.) at 1730:10-1731:2.

199. Thus, for the same reasons as I explained previously in Section VIII(2)(a)(i)-(ii), Kim does not disclose or render obvious "*starting the decoding of the video sequence from said first image frame of the independent sequence.*"

c. Petitioner fails to demonstrate that limitations 8[d], 10[d], and 11[d] are obvious over MPEG-1

200. Kim does not disclose or render obvious "[resetting/reset] the identifier value for the indicated first image frame of the independent sequence" for multiple reasons.

201. First, for the same reasons I explained above, there is no "indicated

first image frame of the independent sequence.”

202. Second, the claims require that the *decoder* reset the identifier value, but Kim teaches at most that the *encoder* sets the alleged identifier value. Kim never teaches a decoder resetting any identifier values, as the claims require.

203. Specifically, the Petition states that “Kim’s Fig. 5A shows that the IAD component of each logical address is reset to zero for each I-picture.” Petition at 60. However, Fig. 5A shows “a digital data stream *compressed and recorded* on the optical disk” by Kim’s *encoder*, not its decoder. Ex. 1006 (Kim) at 2:49-51. Kim makes clear that the IAD values are generated only by the encoder. Ex. 1006 (Kim) at 4:34-45, Fig. 3. There is no disclosure or suggestion in Kim that the *decoder* ever sets or resets the IAD field of the I-frame.

204. A decoder simply reading a value that was set or reset by the encoder does not satisfy the claims, which clearly require the *decoder* to perform the resetting. As the ’321 Patent specification states, “when activated at a random point, the *decoder preferably sets* the identifier of the first image frame, preferably an I-frame, of the independently decodable sub-sequence.” ’321 Patent at 11:26-40 (bold and italics added). Both the Petition and Dr. Saber quote this same passage. Petition at 7; Ex. 1003 (Saber Decl.) at ¶ 59.

205. In the aforementioned previous ITC Investigation, the expert for the patent challenger, Dr. Eli Saber, testified at trial that he agreed it “would not meet

the claim limitation to just read a value that was reset by an encoder, because the claim requires that the resetting be done by the decoder.” Ex. 2019 (1380 Public Tr.) at 1065:19-1066:3.

206. In that same trial, after Dr. Saber presented the same arguments as this Petition regarding Fig. 5A and resetting the IAD of the I-frame, expert Dr. Omid Kia explained (just as I did above) that the Kim teachings are related to the encoder, not the decoder, in contrast to the claims. Ex. 2019 (1380 Public Tr.) at 1038:3-18, 1732:11-1733:9.

207. Considering the same arguments presented in this Petition, the judge in that previous ITC Investigation concluded that “[t]here is really no dispute that ... the alleged ‘identifier value’ in Kim is set/reset by the encoder and the decoder simply reads or derives it from the bitstream.” Ex. 2018 (1380 ID Public) at 56-57. I agree with the judge’s conclusion that “Kim does not teach a decoder resetting the IAD value.” Ex. 2018 (1380 ID Public) at 57.

208. I understand that expert Dr. Omid Kia similarly testified in an ITC Investigation in which the ’321 Patent was asserted that “[f]or an I-frame, that IAD does not serve any purpose ... So it would not be obvious to add functionality that was not there.” Ex. 2019 (1380 Public Tr.) at 1732:19-1733:9.

209. Therefore, it would not have been obvious to a POSITA for a decoder to reset the “0000” value of an IAD of an I-frame repeatedly for the same reasons

that I described above, including at least because Kim's decoder never reads, decodes, or uses the IAD of an I-frame for any purpose. Resetting the value would be a waste of resources, processing power, and memory and would achieve no purpose within Kim.

210. Considering the same arguments presented in this Petition, including "that the decoder reading a 'reset' value from a bitstream and storing it locally meets the resetting limitation," the judge in the previous ITC Investigation rejected that the claims were anticipated or rendered obvious, in view of the testimony from Dr. Kia that I refer to above. Ex. 2018 (1380 ID Public) at 57.

211. In the '321 Patent, the decoder resets the identifier in order to solve the problems with random access that I described previously in the Background section of this declaration. The solution of the '321 Patent explains that "this [random access problem] can be avoided in the decoder by defining an initiation image in an independently decodable Group of Pictures GOP that is activated at the random point of the video file, and the number of the initiation image is set to zero Consequently, when activated at a random point, the *decoder preferably sets* the identifier of the first image frame, preferably an I-frame, of the independently decodable sub-sequence." '321 Patent at 11:26-40 (emphasis added).

212. Kim has no reason to teach or suggest the solution of the '321 Patent because Kim has nothing to do with random access during streaming video. Kim's

single image reproduction for a time lapse recording on an optical disk is not related to the same problem or solution of the '321 Patent.

213. As with other arguments Petitioners raise in its Petition, the judge in the previous ITC Investigation already considered and rejected this exact argument in view of the same evidence and testimony above, and concluded that “claims 8 and 10 require the decoder, not the encoder, to perform the resetting step.” Ex. 2018 (1380 ID Public) at 56.

214. I agree with the judge’s determination that Kim does not teach or render obvious the claims of the '321 Patent for all of the reasons I explain above.

3. Petitioner Has Not Established That The Challenged Claims Are Unpatentable In Light Of Ground 3—MPEG-1 In Combination With Yagasaki

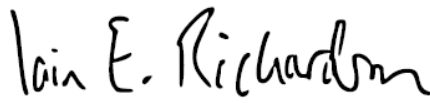
215. It is my opinion that Petitioner cannot show it has met its burden for Ground 3. Claim 9 depends on Claim 8 and Petitioner relies on its Ground 1 to show that independent claim 8 is taught or suggested. As I have explained above, MPEG-1 fails to teach or suggest at least limitations 8[a], 8[b] and 8[d]. For at least that reason, Ground 3 fails.

IX. SIGNATURE

This Declaration represents the opinions that I have formed up to date based on the material currently available to me. I reserve the right to revise, supplement, and/or amend my opinions stated in this Declaration based on new information that becomes available to me and on my continuing analysis of the materials available to me.

I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine, imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: October 29, 2025

A handwritten signature in black ink that reads "Iain E. Richardson". The signature is written in a cursive, slightly slanted style. Below the signature is a horizontal line.

Dr. Iain Richardson