UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM, INC., AMAZON.COM SERVICES, LLC,

Petitioner

v.

NOKIA TECHNOLOGIES OY,

Patent Owner

Case: IPR2024-01505

U.S. Patent No. 9,571,833

DECLARATION OF DR. IAIN RICHARDSON

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

4918-0884-1756

TABLE OF CONTENTS

I. INTRODUCTION 1
II. EDUCATIONAL BACKGROUND AND QUALIFICATIONS 2
III. MATERIALS REVIEWED AND TIME OF INVENTION 5
IV. LEGAL STANDARDS 5
A. Overview 5
B. Patent Claims 6
C. Burden of Proof for Invalidity 7
D. Claim Construction 7
E. Person of Ordinary Skill in the Art 8
F. Anticipation 8
G. Obviousness 9
V. PERSON OF ORDINARY SKILL IN THE ART (POSITA) 12
VI. CLAIM CONSTRUCTION 13
VII. BACKGROUND OF THE RELEVANT TECHNOLOGY 18
A. Overview of Video Coding and Video Coding Standards 18
B. Overview of Motion Prediction in Video Coding 22
C. Technical Challenges in Motion Prediction 31
D. Overview of the ’833 Patent 32
E. Overview of Rusert 36
F. Overview of Nakamura 40
G. Comparison of Rusert and Nakamura 43
1. In Nakamura and Rusert, the Alleged Subsets are Determined Before the Alleged First Spatial Motion Vector Prediction Candidates are Known or Confirmed 44
2. In Nakamura and Rusert, Comparisons are Performed with All Possible Candidate Pairs 49

Nokia Exhibit 2001, p. 2
Amazon.com v. Nokia
IPR2024-01505

VIII. The Grounds in the Petition do not anticipate or render obvious the Challenged claims 53

A. Ground 1: Rusert in Combination with Zheng.....53

1. Rusert does not teach "determining a set of spatial motion vector prediction candidates for the encoded block of pixels" as required by 1[b], 15[b], and 17[b] 53

2. Rusert does not teach "determining a subset of spatial motion vector prediction candidates based on a location of the block associated with the first spatial motion vector prediction candidate" as required by 1[d], 15[d], and 17[d].....57

3. Rusert does not teach "comparing . . . without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates," as required by 1[e], 15[e], and 17[e].....67

B. Ground 2: Nakamura and WD4 - Nakamura and WD4 do not teach the challenged claims.....74

1. Nakamura and WD4 do not teach "determining a subset of spatial motion vector prediction candidates based on a location of the block associated with the first spatial motion vector prediction candidate," as required by 1[d], 15[d], and 17[d] 74

2. Nakamura and WD4 do not teach "comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate of the subset of spatial motion vector prediction candidates without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates," as required by 1[e], 15[e], and 17[e].....77

3. A POSITA would not have been motivated to combine Nakamura with WD4.....79

Nokia Exhibit 2001, p. 3
Amazon.com v. Nokia
IPR2024-01505

IX. CONCLUSION 80

Nokia Exhibit 2001, p. 4
Amazon.com v. Nokia
IPR2024-01505

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 IPR2024-01505 which relates to U.S. Patent No. 9,571,833 the ’833 Patent.

2. This Declaration summarises the independent opinions that I have formed to date, 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 that I have not yet been provided with. I further understand that an institution decision is “all-or-nothing” and so I have focused my opinions on the independent claims at issue in this Petition. Should the Petition be instituted, I anticipate offering additional opinions regarding both the dependent and independent claims.

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.

II. EDUCATIONAL BACKGROUND AND QUALIFICATIONS

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

5. I am Director and co-founder of Vcodex BV.

6. 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.

7. 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.

8. 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.

Nokia Exhibit 2001, p. 6
Amazon.com v. Nokia
IPR2024-01505

9. 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.

10. I am the author of five 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.

11. 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.

Nokia Exhibit 2001, p. 7
Amazon.com v. Nokia
IPR2024-01505

12. In 2007, I founded Vcodex Limited 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.

13. 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.

14. 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.

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

16. Because of my background, training, and experience, I am qualified as an expert in the technical areas relevant to the '833 patent.

17. My company Vcodex BV is being compensated for my independent analysis as an expert with respect to this inter partes review (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 Petitioner (Amazon.com, Inc., Amazon.com Services LLC) or Patent

Owner (Nokia Technologies Oy) and I have no financial interest in the challenged patent.

III. MATERIALS REVIEWED AND TIME OF INVENTION

18. The opinions contained in this Declaration are based on the documents that I have reviewed, including the Petition and all materials cited therein, the Institution Decisions (Paper 11) of IPR2024-00604 and IPR2024-00605 (the "'714 IPRs"), the deposition testimonies of Dr. Creusere, Dr. Sze, and Dr. Reader from the '714 IPRs, my professional judgment, as well as my education, experience, and my knowledge regarding technologies related to the '833 Patent. I have also reviewed and considered Rusert, Zheng, Nakamura, and WD4, which I understand have been designated Exhibits Ex-1004, Ex-1005, Ex-1007-Ex-1009, and Ex-1010, respectively. Nakamura consists of three documents and has been given three exhibit numbers by Petitioner.

IV. LEGAL STANDARDS

A. Overview

19. I am not an attorney or a legal expert, and I do not opine on matters of law. All 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 my analysis of the claims of the ’833 Patent.

20. I have been informed by counsel and I understand that an issued patent is presumed valid and that the challenger of a patent’s validity bears the burden of proving invalidity of the patent.

B. Patent Claims

21. I have been informed by counsel and I understand 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 claims 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 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.

C. Burden of Proof for Invalidity

26. I have been informed by counsel and I understand 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.

D. 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. I have been informed that such a person looks to the claims, specification, and file history in understanding the meaning of the claim terms. I have applied this understanding in this declaration.

E. 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.

F. Anticipation

29. I have been informed by counsel and I understand that a patent claim can be found invalid as anticipated 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 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.

G. Obviousness

31. I have been informed by counsel and I understand 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 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 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.

33. I have been informed by counsel that the combination of familiar elements according to known methods is likely to 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 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 inventor, or if it suggests that the line of development flowing from the reference's disclosure is unlikely to be productive of the result sought by the inventor. 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.

35. I have been informed by counsel that the degree of teaching away will depend on the particular facts; in general, a reference will teach away if it suggests that the line of development flowing from the reference's disclosure is unlikely to be productive of the result sought by the inventor. I have been informed that a reference teaches away, for example, if the references leave the impression that the product would not have the property sought by the inventor.

36. I have been informed by counsel that if the teachings of a prior art reference would lead one skilled in the art to make a modification that would render another prior art device inoperable, then such a modification would generally not be obvious. I have been informed that if a proposed modification would render the prior art device unsatisfactory for its intended purpose, then the proposed modification would generally not be obvious.

Nokia Exhibit 2001, p. 15
Amazon.com v. Nokia
IPR2024-01505

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

38. 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)

39. Based on my review of the ’833 Patent, the types of problems encountered in the art at and prior to the time of the alleged invention, prior solutions to those problems, the pace with which innovations were made during the relevant time period, the sophistication of the technology, and the educational level of active professionals in the field, I believe that a person of ordinary skill in the art, at the time of the invention, would have had at least a Bachelor’s degree in electrical engineering, computer science, or computer engineering; familiarity with the relevant video coding standards as of this time; and at least two years of work experience in research, design, development, and/or testing of video coding hardware and/or software. More education could have substituted for less work experience and vice versa.

40. All of my opinions in this Declaration are from the perspective of a person of ordinary skill in the art, as I have defined it above and during the relevant time frame (e.g., 2011). During that time frame, I possessed at least the qualifications of a POSITA, as defined above.

VI. CLAIM CONSTRUCTION

41. 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.

42. In the institution decisions for Amazon's petitions challenging the related U.S. Pat. No. 10,536,714 (IPR2024-00604 and -00605), the Board interpreted the language "determine a subset of spatial motion vector prediction candidates ***based on the location of the block associated with the first spatial motion vector prediction candidate***" to mean that a subset is determined based on the location of the block with which the first spatial motion vector prediction candidate is associated. (-604 IPR Institution Decision[1] at 16-18; -605 IPR

[1] IPR2024-00604, Paper 11.

Institution Decision[2] at 15-17). In my opinion, this interpretation is correct and a POSITA would readily understand this interpretation. For example, in Figure 9 of the '833 Patent, block "901" is associated with the first spatial motion vector prediction candidate "A0."

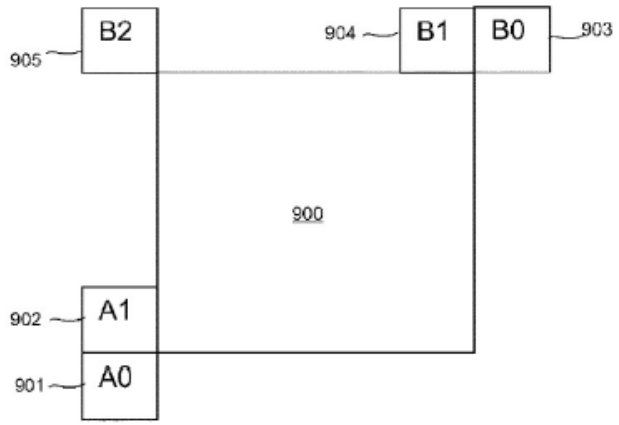


Fig. 9

(Ex-1001, Figure 9).

43. Petitioner argues that the term "the block" from the phrase "based on the location of the block associated with the first spatial motion vector prediction candidate" in limitation [1d] should be interpreted to mean based on the location of the encoded block of pixels in limitation [1a]. (Petition, 20). In other words, Petitioner believes the claim language "the block associated with the first spatial motion vector prediction candidate" should be interpreted to refer to block 900 in Figure 9 above rather than to one of the blocks that correspond to the spatial motion

[2] IPR2024-00605, Paper 11.

vector prediction candidates A0, A1, B0, B1, and B2.

44. In my opinion, Petitioner's argument is inconsistent with how POSITAs would understand the plain language of the claims. A POSITA would readily understand that "the block associated with the first spatial motion vector prediction candidate" refers to the blocks labeled 901 through 905 or A0 through B2 in Figure 9 above. Petitioner's Petition effectively erases "associated with the ***first spatial motion vector prediction candidate***" from the claim language.

45. Petitioner and Dr. Creusere's allegations about the Examiner's interpretation of the claims are in my view incorrect. For example, Dr. Creusere alleges that the Examiner interpreted "based on the location of the block associated with the first spatial motion vector prediction candidate" as "based on the location of the 'current block.'" The principal support Dr. Creusere provides for his opinion is the Examiner's Non-Final Office Action. (Ex-1003, ¶ 46 (citing Ex-1002, 000198)). However, the Examiner's Non-Final Office Action supports Nokia and the Board's interpretation of this claim limitation rather than Petitioner's.

46. The Petition cites to page 198 of the prosecution history, an excerpt of which is highlighted below. (Ex. 1002, 000198). At this part of the prosecution history, the Examiner is considering the phrase "based on the location of the block associated with the first spatial motion vector prediction candidate." The Examiner points to part of a reference called "Tai" (Ex. 2007) and notes that they believe a

subset is defined "by selecting ***one of the*** $\overrightarrow{\boldsymbol{PMV}}$ ***from MVs of neighboring blocks.***" (Ex. 1002, 000198; Emphasis added.)

> determining a subset *(i.e.* $CS^i(\vec{X},t) = \{\vec{V} | \vec{V} = \overrightarrow{PMV}^i(\vec{X}) + \vec{U}, \vec{U} \in US^i\}$*- equation 2)* of spatial motion vector predictions based on the location of the block associated with the first spatial motion vector prediction candidate *(i.e. by selecting one of the* $\overrightarrow{PMV}$ *from MVs of neighboring blocks, a subset of* $CS^i(\vec{X},t)$ *is defined for different linear combination of that PMV and* $\vec{U} \in US^i$*);*

(*Id.*; annotated).

47. I will now examine the Tai reference. Below is a highlighted excerpt from Tai containing the equation relied on by the Examiner above. I note that "$\overrightarrow{\boldsymbol{PMV}}$" is "chosen from the MVs ***of the neighboring blocks***" (*see, e.g.*, Ex-2007, Figure 3 of Tai for exemplary patterns of neighbouring blocks). In other words, each neighbouring block is associated with an MV that may be chosen as "$\overrightarrow{\boldsymbol{PMV}}$." This is what the Examiner was referring to when stating "by selecting one of the PMV from MVs of neighboring blocks." (Ex-1002, 000198). In my opinion, the Examiner (contrary to Petitioner and Dr. Creusere's apparent allegations) did not point to a PMV selected from the MV of the current block. That would not make sense because the point of Tai's exercise is that the current block does not have a MV and thus needs a predicted motion vector.

The candidate vector set $CS^i(\overrightarrow{X}, t)$ is comprised by a single PMV and a set of update vectors. Let $\overrightarrow{X}$ represent the block location and let t be the frame number. The candidate vector set for this particular block in the ith pass are then defined as

$$CS^i(\overrightarrow{X}, t) = \left\{ \overrightarrow{V} \,|\, \overrightarrow{V} = \overrightarrow{\text{PMV}}^i(\overrightarrow{X}) + \overrightarrow{\text{U}}, \overrightarrow{\text{U}} \in US^i \right\} \quad (2)$$

where $\overrightarrow{\text{PMV}}^i(\overrightarrow{X})$ is the PMV of block $\overrightarrow{X}$ and is chosen from the MVs of the neighboring blocks and US^i is a set of update vectors. Every motion vector $\overrightarrow{V}$ in the candidate vector set is comprised by a single PMV and a update vector $\overrightarrow{U}$ selected from update vector set US^i.

(Ex-2007, 3; annotated).

48. The Petition and Dr. Creusere also rely on the applicant's supposed failure to dispute the examiner's interpretation in a subsequent office action response, but as I point out above the Petition and Dr. Creusere's interpretation of this prosecution history do not make sense in view of the actual prosecution history language and the disclosure of the relevant reference. Additionally, the office action the Petition relies upon is 20 pages long, and the applicant responded with both amendment and argument where the applicant argued that the examiner failed to show that the cited references collectively failed to teach or suggest "selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates as a potential spatial motion vector prediction candidate to be included in a merge list for the prediction unit, where the merge list is selected based

on the motion information; **determining a subset of spatial motion vector prediction candidates based on the location of the block associated with the first spatial motion vector prediction candidate**; comparing motion information of the first spatial motion vector prediction candidate with motion information of ~~the~~ spatial motion vector prediction candidates in the determined subset of spatial motion vector prediction candidates; if at least one of the comparisons indicates that the motion vector information of the spatial motion vector prediction candidates correspond with each other, excluding the first spatial motion vector prediction candidate from the merge list[.]" (Ex-1002, 000787 (emphasis added, strikethrough in original)). Accordingly, the applicant contested the examiner's rejection and makes no statement adopting the interpretation advanced by Petitioner here.

VII. BACKGROUND OF THE RELEVANT TECHNOLOGY

A. Overview of Video Coding and Video Coding Standards

49. Modern-day video coding and compression schemes involve a delicate balance of competing concerns. On one hand, video technicians and consumers demand high fidelity and large resolution of the video images to make use of large high-definition displays and wider availability of high-speed Internet connections. However, the competing concern becomes the efficient use of storage space on hard drives and memory cards, as well as available bandwidth of Internet connections and computer processing demands. The use of modern video coding and compression

Nokia Exhibit 2001, p. 22
Amazon.com v. Nokia
IPR2024-01505

schemes aims to represent the video images in a compressed form with as little data as possible, while maintaining as much quality of the original source images as possible after decoding and ensuring that the video images can be decoded and displayed in real time. Present-day applications of video coding often involve a trade-off between computation, compression and video quality. Increasing the amount by which data is compressed can save storage space and transmission bandwidth, but may result in degraded video quality to the point where the visual quality of decoded video is unacceptably poor when viewed by the consumer. More compression for the same video quality, or better video quality for the same level of compression, may be achieved at the expense of increased computational processing and/or increased memory requirements. However, an increase in computational processing may require more sophisticated and perhaps more expensive processors and may lead to problems such as increased battery drain on mobile devices or reduced video frame rates.

Excerpt from Richardson 2010 book (Ex-2011[3]) at page 69:

---

[3] I refer here to Exhibit 2011 which contains excerpts from my 2010 book, "The H.264 Advanced Video Compression Standard."

Nokia Exhibit 2001, p. 23
Amazon.com v. Nokia
IPR2024-01505

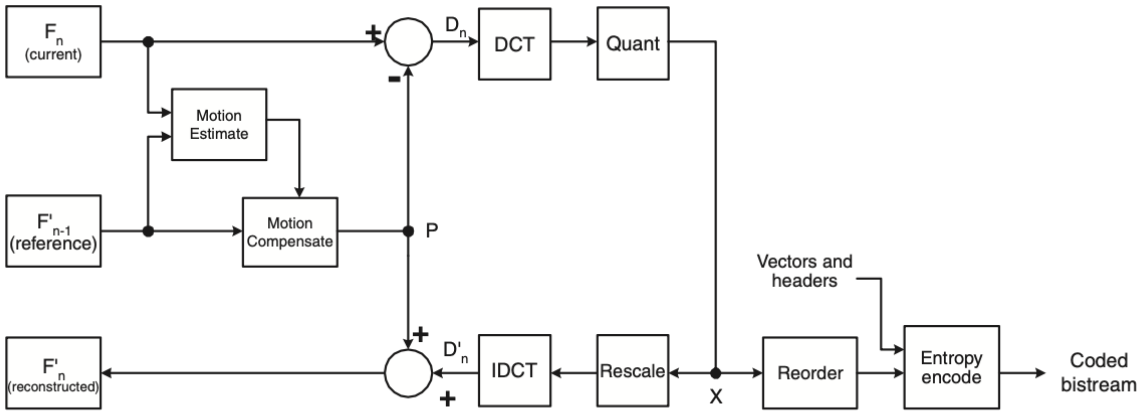


Figure 3.51 DPCM/DCT video encoder

……

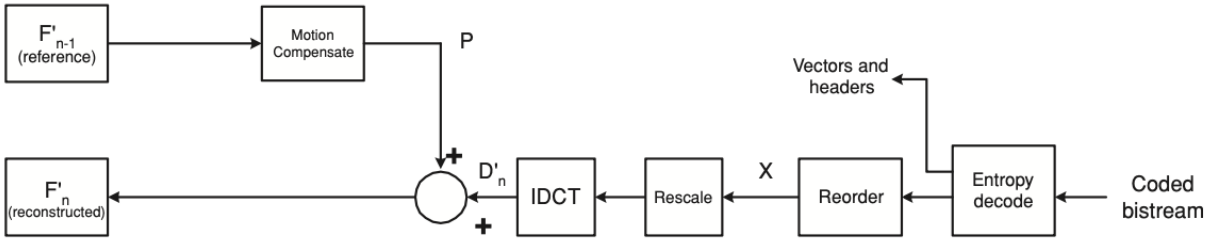


Figure 3.52 DPCM/DCT video decoder

50. A video codec comprises a video encoder and a video decoder, which I illustrate in the above block diagrams from my 2010 book. A video encoder converts a sequence of input video frames (such as frame F_n in the above diagram, video encoder) into a coded bitstream, which is a compressed representation suitable for storage or transmission and a video decoder converts such a compressed representation into a series of decoded video frames (such as frames F'_n in the above diagram, video decoder). (Ex-2011, 69; Ex-1001, 1:21-28). As the ’833 patent explains, a video encoder forms a prediction for each block of pixels to be coded,

Nokia Exhibit 2001, p. 24
Amazon.com v. Nokia
IPR2024-01505

labelled P in the above diagrams, using motion compensated prediction from an area in a previously coded video frame or spatial prediction from previously coded pixels in the same frame. The prediction is subtracted from the block of pixels to be coded to produce a residual or error, labelled D_n in the above encoder diagram. The residual or error is transformed, quantized and entropy encoded. (Ex-1001, 1:29-53). A video decoder entropy decodes the coded bitstream and rescales and inverse transforms block data in order to produce a decoded residual or error, labelled D'_n in the above decoder diagram. The decoder also decodes information that enables it to form the same prediction as the encoder, labelled P above, which it adds to the decoded residual in order to construct a decoded block of pixels and thus a decoded frame F'_n.

51. By 2011, a POSITA would have been familiar with video coding standards such as the ITU-T H.263 and H.264 standards discussed in the '833 patent (Ex-1001, 1:29-32). In common with other video coding standards, these standards each specify a format for a compressed bitstream and a method of decoding such a bitstream to produce a decoded video sequence but do not fully specify an encoding process. A video decoder conforming to a standard such as H.263 or H.264 decodes a bitstream that conforms to the specified bitstream format, in order to produce the same output as the specified decoding process.

Nokia Exhibit 2001, p. 25
Amazon.com v. Nokia
IPR2024-01505

B. Overview of Motion Prediction in Video Coding

52. Most video coding techniques exploit pixel redundancy in the spatial and temporal domains, and involve finding ways to communicate new information in a bitstream by referencing similar, already communicated information. The references to already coded information are used to generate approximations (known as predictions) of each block of a video frame being encoded. At the decoder, predictions are combined with additional information (known as prediction error or residual information) in order to reproduce a reconstructed image. The encoder communicates to the decoder how to make predictions for each block as well as the residual information in a compressed form. The decoder creates the same prediction, decodes the residual information, and combines residual and prediction to reconstruct a block of video information (pixel values) from the coded block.

53. There are two main ways of exploiting redundancy in video coding to create such predictions, namely—intra coding and inter coding. In an intra-coded block in a frame of video, previously coded pixels in the same frame are used to create a prediction and to reduce the number of bits necessary to encode and represent the current block. In an inter-coded block, a prediction block is constructed with reference to pixels in one or more previously coded frames, which may be frames in the past or the future in display order. The choice of one or more previously coded frames (known as reference pictures) and an offset (known as a motion vector,

Nokia Exhibit 2001, p. 26
Amazon.com v. Nokia
IPR2024-01505

i.e., an offset indicating the location of a prediction source in another frame) identifying each similar region of pixels in a reference picture, are used to create a prediction block. The instructions for creating the prediction block, including the choice of reference picture and the motion vector if present, are encoded along with any prediction error or residual, i.e., motion vector information is provided in the bitstream.

54. A widely-used technique in video compression is the encoding of the source video through motion compensated prediction, such as inter prediction which I introduced in the previous paragraph. (Ex-2011, 28-38). This process can be carried out by dividing the individual picture elements (pixels) of a video frame into blocks of pixels. Motion compensated prediction attempts to find a prediction or match for a block of pixels in a current video frame, in a correspondingly sized region of a previously coded frame. Such a match may be found in the same position as the current block or in an offset or shifted position in the previously coded frame. Such a shift in position may indicate, for example, that motion or other change is occurring in a particular area of the video scene. Motion compensated prediction has a very significant contribution to compression performance, since encoding the motion compensated differences between video frames typically results in much less data than encoding each video frame without such motion compensated prediction.

55. In video coding, motion may be approximated as being block-wise

Nokia Exhibit 2001, p. 27
Amazon.com v. Nokia
IPR2024-01505

constant, i.e. blocks of pixels in a frame are assumed to undergo translational motion represented by a motion vector with a horizontal and a vertical component. An example block size for prediction is 16x16 pixels, however modern video coding techniques allow for motion compensated blocks of various sizes, including blocks as small as 4x4 or as large as 64x64.

56. A strategy for reducing the amount of encoded data in the bitstream related to motion compensated prediction is to reduce the amount of data needed to represent the motion vectors themselves. Motion vector information can make up a substantial portion of a compressed video frame, particularly when small block sizes are used. However, the motion vectors of adjacent blocks (in space and time) are generally correlated, and this correlation or redundancy can be exploited through efficient encoding of motion vector information, e.g., by forming a prediction for the motion vector of a current block by taking into account the motion vectors of its neighbours.

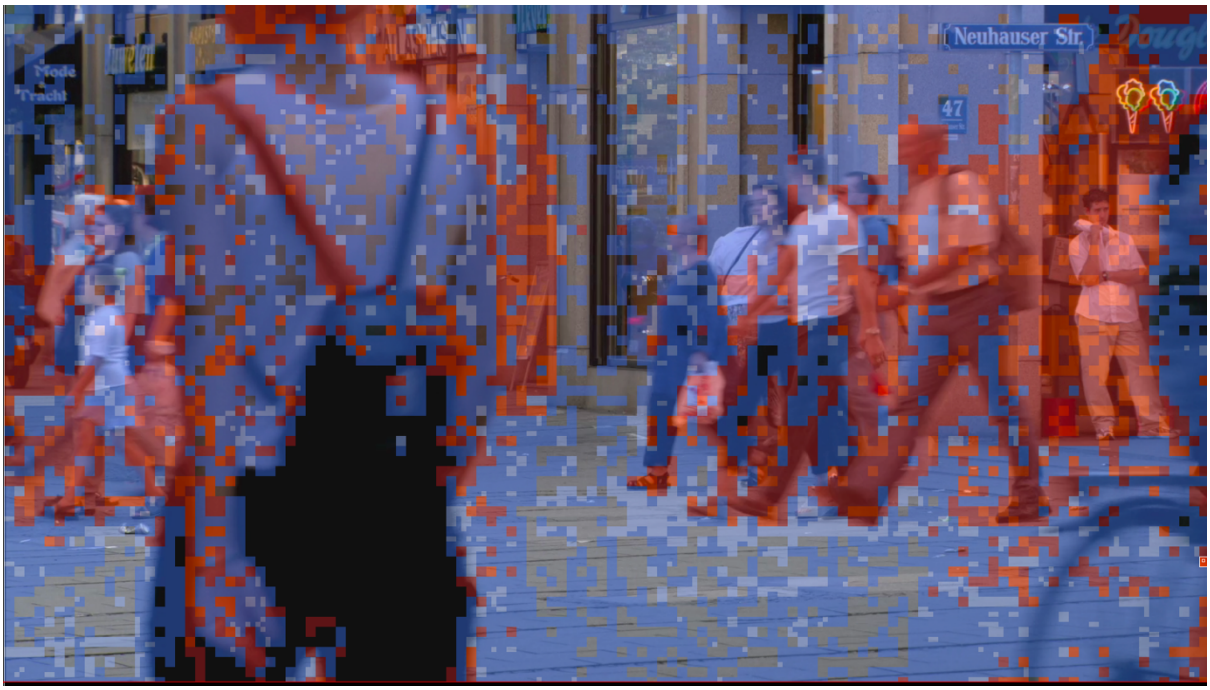
57. An H.264-complaint decoder can use the median value of three already-decoded spatially adjacent blocks to form a prediction for the current block of pixels being decoded. Any difference between the current motion vector and its prediction can be losslessly encoded in the bitstream and decoded by the decoder. Thus, an H.264-compliant decoder can predict a motion vector based on previously decoded motion vectors and add any motion vector difference to reconstruct the

motion vector.

58. An H.264/Advanced Video Coding (AVC) inter-coded video frame (Ex-1004, ¶ 3), includes blocks (macroblocks) that may be coded in different ways, including: inter predicted with a non-zero motion vector, inter predicted without sending a motion vector (skipped/Direct mode) or intra predicted. Consider the following sample frame from a 1080p video sequence, 'Pedestrian Area'. When encoded using H.264, the frame is processed in units of a macroblock, 16x16 pixels, shown as the larger-scale grid:

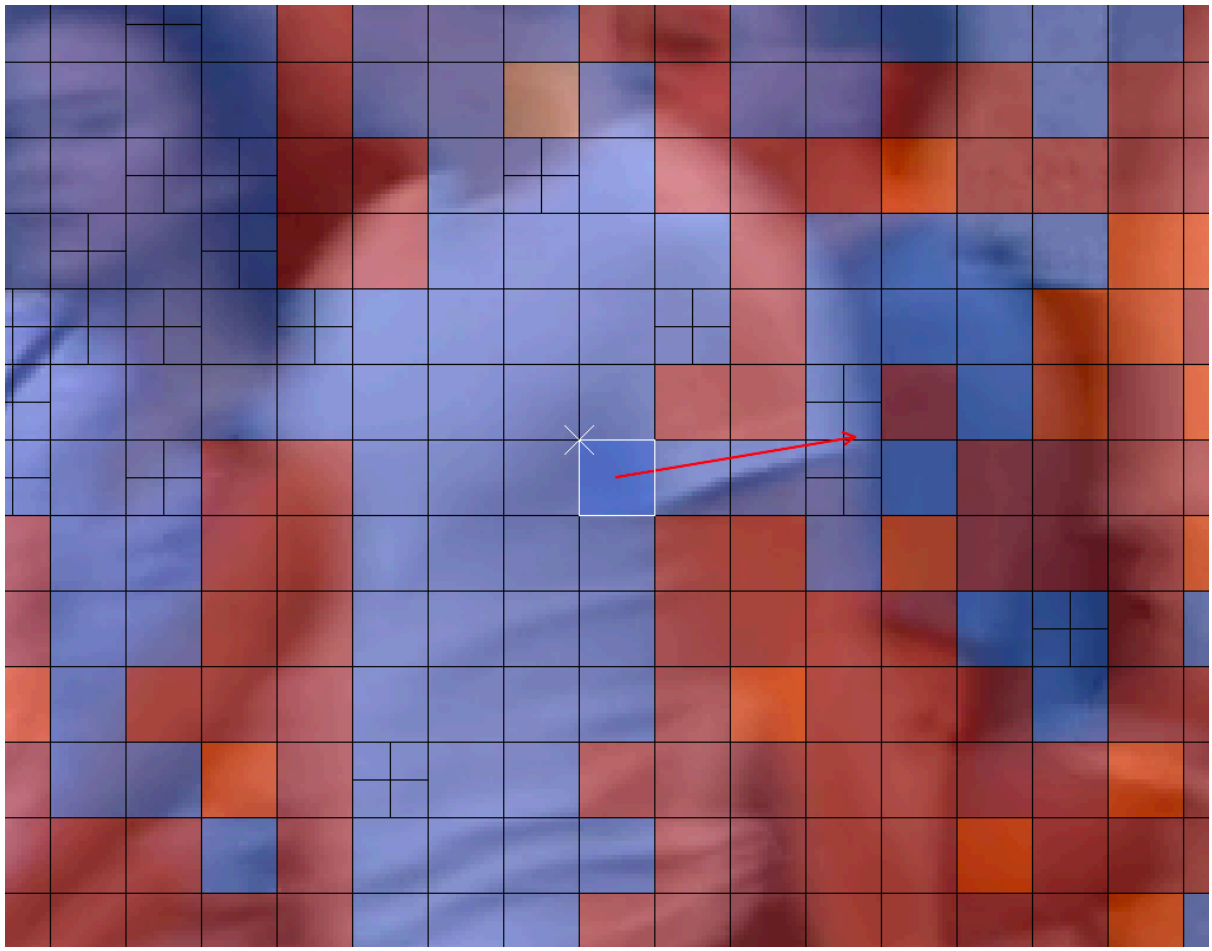
59. In an H.264 inter-coded frame, each macroblock is encoded using inter prediction with a non-zero motion vector(blue highlighting), inter prediction without sending a motion vector (for example, skipped) (no highlighting) or intra prediction

(red highlighting)[4]:



60. Consider an inter-coded macroblock such as the one highlighted in the center of this close-up:

[4] Screenshots generated using video codec stream analysis tool, Elecard Streameye.



Some of the neighbouring macroblocks may also be inter-coded with motion vectors (blue highlighted macroblocks). Some of the neighbouring macroblocks, such as the upper-right macroblock in this example, may be intra-coded (red highlighting). Intra-coded neighbours have no associated motion vector, since they do not use prediction from another frame.

61. H.265-compliant decoders can use more sophisticated ways of predicting motion vectors. Encoders that encode video into an H.265-compliant bitstream can reduce the number of bits necessary to encode motion vectors. Two options in H.265 for signaling each motion vector are known as the Merge mode and

Nokia Exhibit 2001, p. 31
Amazon.com v. Nokia
IPR2024-01505

the Advanced Motion Vector Prediction (AMVP) mode. The Merge mode allows spatiotemporally adjacent blocks to share the same motion vectors and other motion information. (Ex-2014, § 8.5.3.2.3). The AMVP mode provides an efficient way to predict the motion vector for the current block to be encoded from spatiotemporally adjacent blocks and to encode a difference or delta between the predicted motion vector and the actual motion vector (Ex-2014, § 8.5.3.2.6).

62. Nokia contributed to the Merge mode process specified in the H.265 Standard. The inventors of the '833 Patent proposed their solution in the JCTVC-G593 and JCTVC-H0252 contribution documents. (Ex-2015; *see also* Ex-2016). The topic of Merge mode was the subject of a "core experiment" by the JCTVC, and Nokia's JCTVC-H0252 proposal was tested and ultimately adopted by the JCTVC, the committee responsible for developing the H.265 Standard.[5] The '833 Patent corresponds with the inventors' contributions.

63. Merge mode exploits the property that motion vectors of neighbouring blocks are often similar. In Merge mode, the encoder and decoder form lists of

---

[5] The Report on the Merge mode core experiment is JCTVC-H0039, and the Meeting Report adopting Nokia's contribution by way of JCTVC-H0039 is JCTVC-H1000. (Ex-2017; *see also* Ex-2018). Nokia's approach was adopted at the eighth meeting of the JCTVC in February 2012.

motion vector prediction candidates called merge lists, and then—through an indication in a bitstream—the encoder communicates to the decoder which candidate from the list to use.[6] Merge mode can save bits by avoiding the need to send the motion vector explicitly and instead sending an index value to identify a motion vector prediction included in the merge list, but this potentially reduces prediction accuracy if the motion vectors available on the list do not produce accurate predictions since the decoder needs to form a merge list corresponding to that used at the encoder to correctly interpret the received index value. Further, if a transmission error occurs, selection from the candidate list at the decoder may be incorrect.

64. In Merge mode, upon completion of the merging process, the resulting merged region, i.e. the selected candidate block and the current block, share the same motion vector. The following Figure from Ex-2012 conceptually illustrates five spatial neighbour blocks a0, a1, b0, b1, and b2 that are spatially adjacent to the current block to be encoded. The availability of these blocks for merging is evaluated

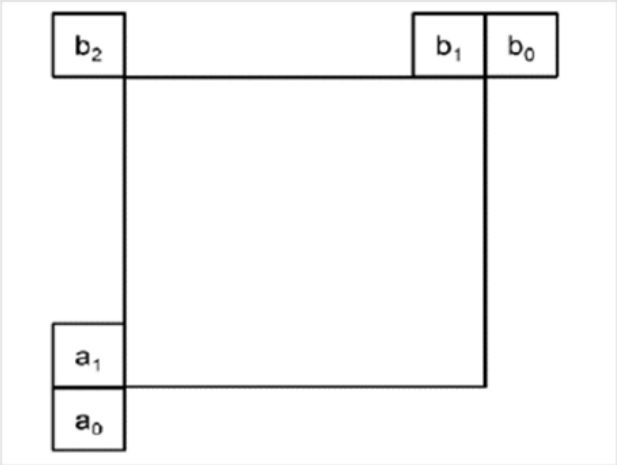
---

[6] In Merge mode, a motion vector predictor is selected from a list of candidates, each of which is a motion vector of a previously coded block, and the index of the selected candidate in the list is encoded and sent to the decoder. The list can be constructed based on the five spatial neighbours, as well as a temporal candidate.

Nokia Exhibit 2001, p. 33
Amazon.com v. Nokia
IPR2024-01505

in the following order: a1, b1, b0, a0, b2. A temporal candidate for block merging at the same location in a reference frame, i.e. a previously coded frame that is not the current frame, is also evaluated for block merging. A candidate block is selected and the motion information for the current block is "inherited" from the candidate block, such that the motion information of the two blocks is merged.

Merge Mode Overview (Ex-2012, Figure 8):



65. It is frequently the case that two or more candidates in the list have the same motion vector. It is beneficial to eliminate such duplicates to shorten the list and reduce the number of bits needed to specify the index in the list. Comparison of candidates for potential elimination can be a computationally costly process. As mentioned, the '833 Patent correlates with the inventors' Merge mode contribution to the H.265 Standard development. One goal of the '833 Patent is to provide an enhanced Merge list process that "reduce[s] the complexity of the implementation [of Merge mode]." (Ex-1001, 4:7-19).

Nokia Exhibit 2001, p. 34
Amazon.com v. Nokia
IPR2024-01505

C. Technical Challenges in Motion Prediction

66. A computationally challenging part of motion-compensated prediction is identifying the motion vector to use for the prediction. This is typically done by sub-dividing the current frame into blocks and finding matching regions in previously coded reference frames to use to predict the blocks in the current frame. While the blocks in the current frame are generally formed based on some type of grid pattern, the best matching blocks in reference frames may not, and typically do not, align with the same grid. As a result, the search for a best matching block typically involves many comparisons between the current block and many corresponding regions from reference frames, with differing motion vectors relative to the current block, to find the best available matching block. This search for a matching block is referred to as motion estimation.

67. The encoder may select a motion vector that points to an integer pixel offset value, which means that the corresponding block is a block of pixels present in the reference frame. The encoder may also choose a motion vector that points to a sub-pixel location, i.e. a location between full pixel values in the reference frame. The advantage of using sub-pixel locations is greater accuracy of motion prediction. However, the use of sub-pixel locations generally results in a greater number of calculations because pixel values must be generated for sub-pixel locations.

68. The encoder encodes the instructions necessary to recreate the block

Nokia Exhibit 2001, p. 35
Amazon.com v. Nokia
IPR2024-01505

prediction. These instructions may include motion vectors and whether they are to be used as part of a uni- or bi-prediction, and in some cases, an indication of one or more chosen reference pictures. Additionally, the encoder may compress and encode the difference (prediction error/residual) between the predicted values and the actual values, i.e., the original pixel values of the current from being encoded. The prediction error/residual data is typically compressed using a block transform such as the Discrete Cosine Transform, and the resulting transform coefficients representing the residual are subsequently quantized and entropy encoded, to limit the amount of data needed to represent the residual in the encoded video bitstream.

69. In inter prediction, a decoder uses the encoder's instructions including any motion vectors and the chosen prediction method (e.g., inter or intra prediction) to recreate the prediction. decoder then decodes any compressed prediction error to add it to the prediction to obtain the decoded video. Generally, encode processes use as reference frames previously coded frames that the encoder has decoded, so that the decoder can reconstruct a prediction for a current block using the prediction information communicated to the decoder and already decoded.

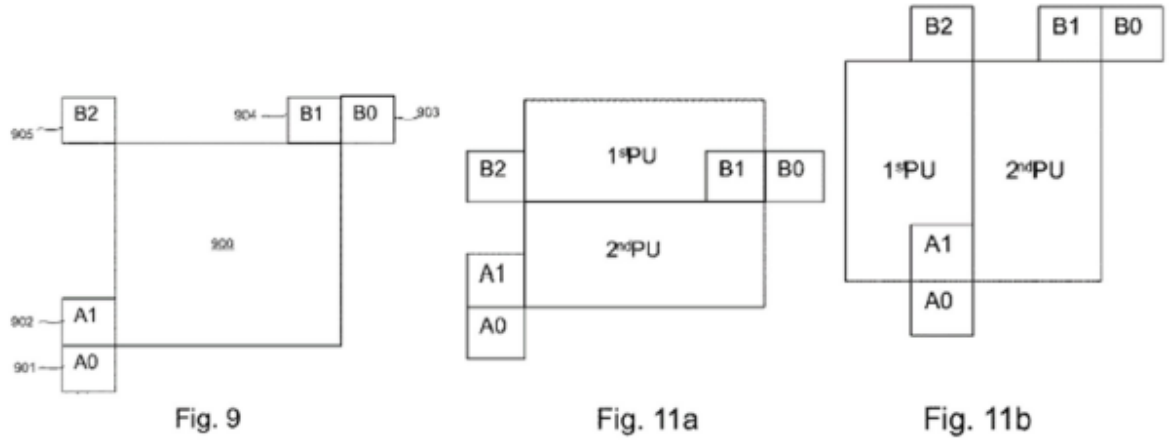
D. Overview of the '833 Patent

70. The '833 Patent correlates with the inventors' efforts to improve the Merge mode of the then-draft H.265 Standard. One goal of the '833 Patent is to provide an enhanced Merge list process that "reduce[s] the complexity of the

Nokia Exhibit 2001, p. 36
Amazon.com v. Nokia
IPR2024-01505

implementation [of Merge mode]." (Ex-1001 at 4:7-19). One benefit of the '833 Patent's approach is a reduced number of comparisons between spatial motion vector prediction candidates. (*Id*., 4:11-27).

71. Figure 5 of the '833 Patent, with reference to Figures 9 and 11, illustrates an improved merge list construction embodiment developed by the inventors. Figures 9 and 11, reproduced below, label the locations of merge list candidates relative to a current block, where A0, A1, B0, B1, and B2 are part of a set of spatial motion vector prediction candidates and block 900 is a block of pixels. (Ex-1001, cl. 1, cl. 15, cl. 17; *see also id.*, 14:31-34, 16:62-17:3). As Claim 9 of the '833 Patent discloses, a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates is a "potential spatial motion vector prediction candidate to be included in a motion vector prediction list for a prediction unit" of the "encoded block of pixels." (*Id.*)



Ex-1001 at Figures 9, 11a, and 11b

72. Each motion vector prediction candidate A0, A1, B0, B1, and B2

Nokia Exhibit 2001, p. 37
Amazon.com v. Nokia
IPR2024-01505

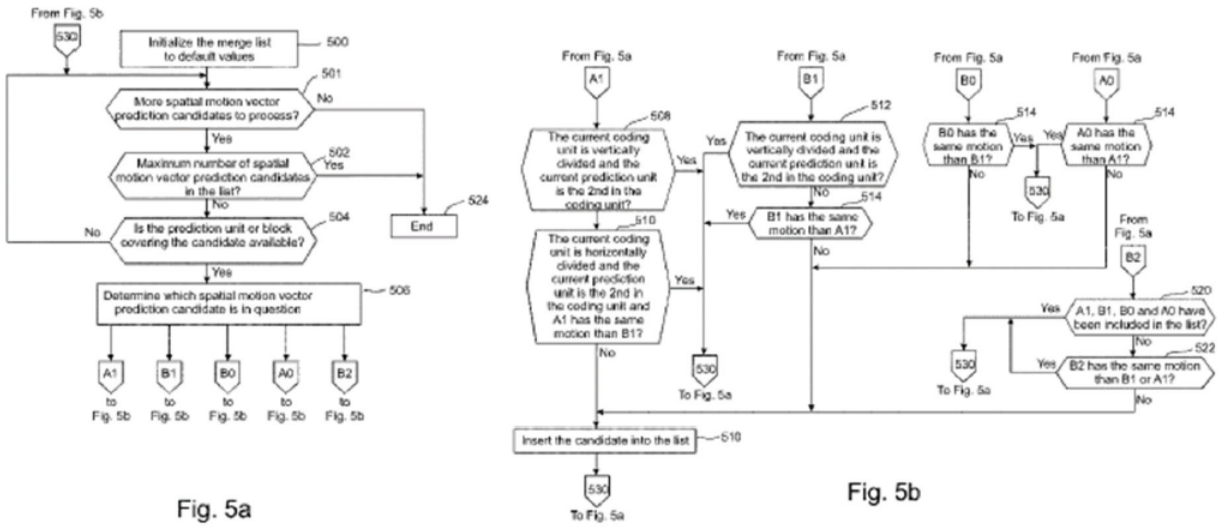
(these labels represent the locations of the blocks with which the motion prediction candidates are associated)[7] is subject to a limited set of comparisons and other conditions to determine whether it can be included on the merge list. For instance, suppose candidate A0 is the first spatial motion vector prediction candidate. A subset of spatial motion vector candidates is determined based on a location of the block associated with the first spatial motion vector prediction candidate. (Ex-1001, cl. 1; *see also id.*, 14:52-15:5).

73. Using candidate A0 as the example first spatial motion vector prediction candidate, the subset determined for A0 is {A1}. The subset for A0 in the embodiment of Figure 5 does not include B1 or B0 even though A0/B1 and A0/B0 are possible candidate pairs. After the subset is determined based on a location of the block associated with the first spatial motion vector prediction candidate, the "motion information of the first spatial motion vector prediction candidate" is compared "with motion information of the spatial motion vector prediction

---

[7] A POSITA would understand that the labels can refer to both block locations and the motion vectors associated with the blocks at those locations. For example, "A0" can refer to the block location A0 in Figure 9 and—depending on the context—also refer to the motion vector associated with the block at location A0, which can be called candidate A0.

candidate in the determined subset of spatial motion vector prediction candidates." (Ex-1001, cl. 1). Again using A0 as an example, A0 is compared to A1 based on the conditions described in Figures 5a and 5b. (*Id.*, 15:38-16:39).



Ex-1001 at Figures 5a and 5b

74. Since A0 is only being compared to A1 (i.e., A0 is not being compared to B0, B1, and B2), the remainder of the "comparing motion information" proceeds "without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates." (Ex-1001, cl. 1). This is because A0's motion vector is only compared to A1's motion vector. The determination of whether to include or exclude the first spatial motion vector prediction candidate is based on the comparing step. (Ex-1001, cl. 1).

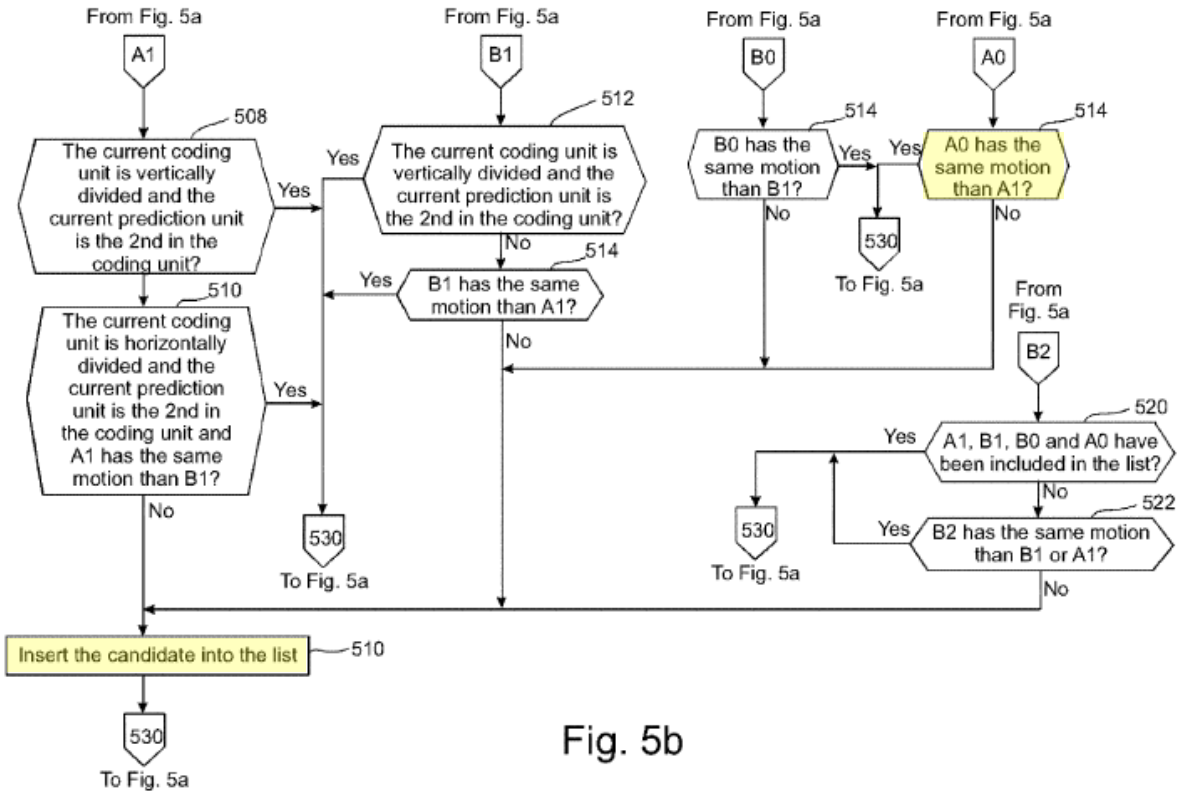


Fig. 5b

Ex-1001 at Figure 5b (annotating the example comparison)

75. If it is determined that A0 does not have the same motion as A1, then A0 is added to the motion vector prediction list based on the comparing step. This same iterative process can be carried out for each candidate in the set by selecting another one of the candidates as the "first spatial motion vector".

E. Overview of Rusert

76. Rusert discloses a method for generating a potential motion vector ("PMV") candidate list, PMV_CANDS, for a current block of pixels that is being encoded or decoded. (Ex-1004, ¶¶ 11-12, 36-37). The PMV candidates can be obtained from previously coded neighbouring blocks within an allowed distance from the current block. (*Id.*, ¶¶ 11, 40-44, 73). Dr. Creusere identifies the previously

coded blocks within Rusert's scan order as "spatial motion vector prediction candidates." (Ex-2010, 72:11-76:19).

77. Rusert teaches search patterns for "an outwards going scan … around the current block." (Ex-1004, ¶¶ 44, 51-66; *see id.*, Figures 3a-3n (illustrating numerical scan order around the current block, labelled ".", from block 1 onwards)).

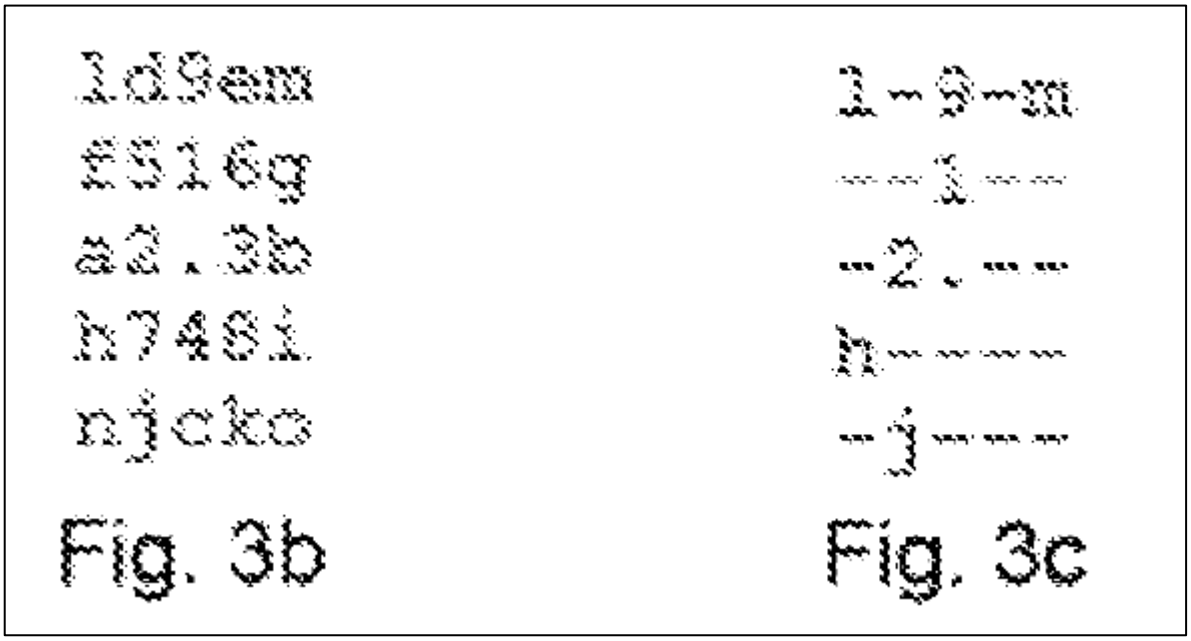
Fig. 3n

(*Id.*, Figure 3n (annotated)). Using Figure 3n as an example, the list of PMV_CANDS can ultimately include the blocks that are available in the scan order, i.e., some or all of 1, 2, 3, …, a, b, c, …. (Ex-2010, 72:11-76:19; *see also id.*, 75:6-76:1). In the case of Fig. 3n, the position of the current block is indicated by a dot '.'. The previously coded block immediately above the current block and labelled '1' is evaluated first, if available. The block to the left of the current block labelled '2' is evaluated next, if it is available, and so on.

78. Some of the numbered blocks in Rusert's scan order may not be

Nokia Exhibit 2001, p. 41
Amazon.com v. Nokia
IPR2024-01505

available, for example if the current block is along a frame boundary, or if some of the blocks have not yet been encoded when the current block is processed, or if the numbered block(s) do not have a motion vector present. Rusert describes exemplary scenarios of block unavailability, including for example blocks that have not yet been encoded, or blocks without a motion vector present such as an intra-coded block. (Ex-1004, ¶ 54, Figure 3c). Rusert provides Figure 3c as an example where some blocks in the scan order of Figure 3b were unavailable, showing that only blocks 1, 2, 9, h, j, l and m are available for inclusion in PMV_CANDS.



(*Id.*, Figures 3b and 3c).

79. Rusert performs an outward scan of selected and available previously coded PMV candidates, starting from the current block and remaining within an allowed distance. (*Id.*, ¶¶ 4-5, 36-39, 43-44, 51-66).

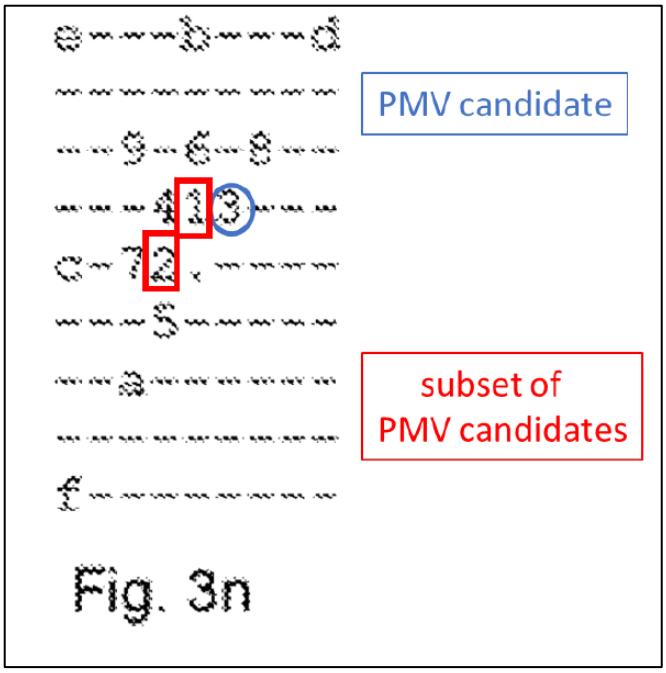
80. As the scan expands outward, what Petitioner alleges is "the subset,"

Nokia Exhibit 2001, p. 42
Amazon.com v. Nokia
IPR2024-01505

(Petition, 35-41), is determined based on which blocks have available candidates and whether those candidates satisfy comparisons with candidates already in PMV_CANDS (Ex-1004, ¶¶ 44, 51-66; *see also id.*, Figures 3a-3n).

81. For example, when the motion vector for the block with location 3 ("block 3") is evaluated during Rusert's scan sequence, the subset of candidates PMV_CANDS already comprises candidates evaluated from earlier scan locations, which in this example may include blocks with locations 1 and 2, "block 1 and block 2", assuming those blocks were available and that block 2 is not sufficiently similar to block 1 to have been excluded. In other words, Rusert's subset PMV_CANDS is determined prior to knowing the location or value of the next spatial motion prediction candidate, which in this example is block 3 but need not be 3 (e.g., block 3 may be unavailable). Petitioner argues that the "subset" (here, blocks 1 and/or 2) is determined based on the location of the block with location 3, which Petitioner alleges is the claimed "first spatial motion vector prediction candidate." (Petition, 40). As I will explain, I disagree.

Nokia Exhibit 2001, p. 43
Amazon.com v. Nokia
IPR2024-01505



(Ex-1003, ¶ 132 (annotated by Dr. Creusere)).

82. The motion information of the block being evaluated (block 3 in this example) is then compared to the alleged subset, blocks 1 and 2 in this example, to determine whether the block being evaluated will be included in PMV_CANDS. (Ex-1004, ¶¶ 62, 71). When Rusert evaluates block 3 in this example, Rusert compares all possible pairs of candidates,[8] as I will explain.

F. Overview of Nakamura

83. Nakamura is a proposal to the H.265 development effort that attempted to provide methods for deriving candidates for Merge mode and the motion vector predictor (MVP). (Ex-1007, § 1). Nakamura suggested that the

[8] That is, comparisons are done with the pairs 1/3 and 2/3.

Nokia Exhibit 2001, p. 44
Amazon.com v. Nokia
IPR2024-01505

positions of the spatial neighbours eligible as merging candidates in merge mode should be the same as those of the spatial MVP candidates, thereby unifying the location of the spatial neighbours for both Merge mode and MVP. (*Id.*) Additionally, the proposed techniques sought to unify the derivation processes for Merge mode and MVP, as illustrated in Figure 1:

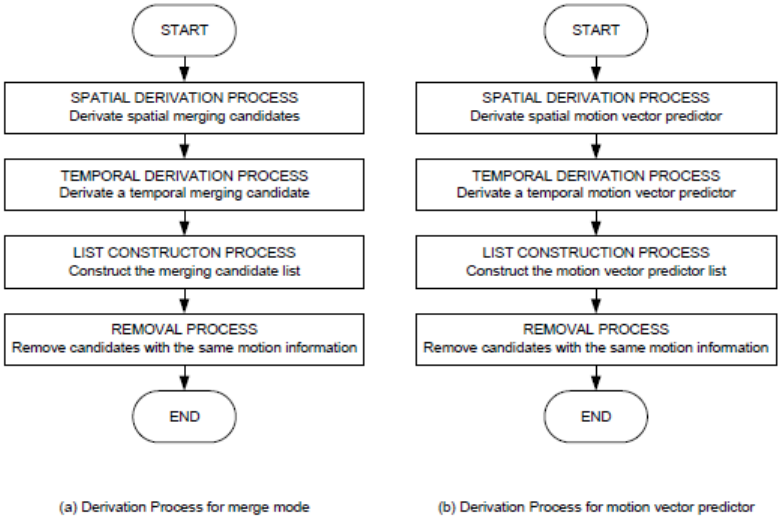


Figure 1 Derivation Process of merge mode and MVP in WD3

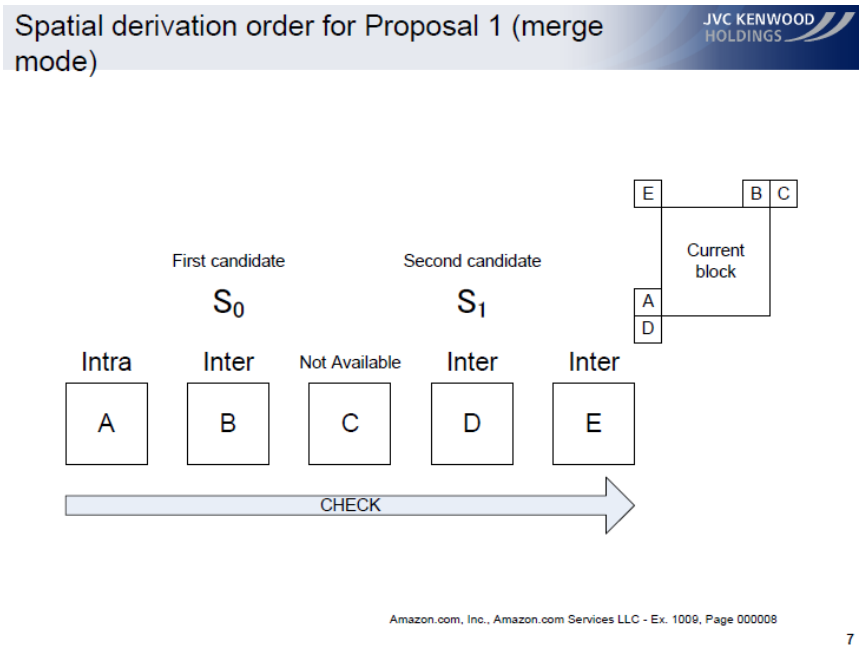
(*Id.*; *see also id.*, § 2.1).

84. For Merge mode, Nakamura teaches selecting two spatial motion vector prediction candidates (i.e., spatial candidates S0 and S1) for a merge list selected in a spatial derivation order from blocks A, B, C, D, and E (i.e., blocks in the current frame) relative to the current prediction unit.[9] (*Id.*, Abstract, § 1, § 2.2.1,

---

[9] The JCTVC committee adopted Nokia's merge mode proposal over Nakamura's proposal as part of the draft HEVC/H.265 standard. (Ex-2018)

Figures 1, 2(b), Tables 2-4). Following this spatial derivation order, Nakamura checks blocks A-E to identify the first available motion vector. When this motion vector is found, it is denoted as spatial candidate S0. (Ex-1009, 8). Once S0 is determined (i.e. S0 is assigned to the motion vector associated with one of the neighbouring block locations), S0 is automatically added to Nakamura's merge list. (*Id.*) Then Nakamura continues its availability check to identify a second motion vector, denoted as candidate S1 (i.e. to assign S1 to the motion vector of one of the neighbouring block locations), which, like S0, is added automatically to the merge list once identified. (*Id.*)



(*Id.*)

85. Petitioner alleges that Nakamura's S0 is a "subset" according to the claims of the '833 Patent. In my view this is incorrect as I will explain. Nakamura's

S0 (analogous to Rusert's alleged subset) is determined before Nakamura has knowledge of which block's motion vector will be assigned to S1 (what Petitioner alleges is the "first spatial motion vector prediction candidate"). Thus, S0 (the alleged "subset") cannot be determined based on the location of S1 (the alleged block associated with the "first spatial motion vector prediction candidate") because that information (i.e., the location of S1, which Petitioner alleges to be the location of the block associated with the "first spatial motion vector prediction candidate") is not known or confirmed at the time Nakamura determines S0.

86. In other words, S0 is not and cannot be determined based on the location of S1. (Ex-1007, § 2.2.1; Ex-1009, 1, 8). The alleged subset S0 does not change depending on the later choice of S1, since S0 is already selected regardless of what the selection of S1 thereafter might be.

G. Comparison of Rusert and Nakamura

87. Rusert and Nakamura are similar in at least two important respects. First, both Rusert and Nakamura determine what Petitioner alleges to be the "subset of spatial motion vector candidates" before ever determining or selecting what Petitioner alleges to be the "first spatial motion vector prediction candidate." Second, both Rusert and Nakamura compare what Petitioner alleges to be the "first spatial motion vector prediction candidate" with each "possible candidate pair." (*See, e.g.*, Ex-1001, cl. 1, cl. 15, cl. 17). I will discuss these two similarities in turn.

Nokia Exhibit 2001, p. 47
Amazon.com v. Nokia
IPR2024-01505

1. In Nakamura and Rusert, the Alleged Subsets are Determined Before the Alleged First Spatial Motion Vector Prediction Candidates are Known or Confirmed

88. I will first discuss how Rusert and Nakamura both determine what Petitioner alleges to be the "subset of spatial motion vector candidates" before determining or selecting what Petitioner alleges to be the "first spatial motion vector prediction candidate." Dr. Creusere agreed with this similarity in his deposition in the IPR proceedings instituted for the related '714 Patent:

> **Q.** I think I agree with you, that's certainly true in Rusert, Doctor, that what will be the subset will be based on prior processing done at -- to the point where you reach the -- the first spatial motion vector prediction candidate.
>
> And so to clarify with an example, by the time we reach the -- candidate 3 in Rusert using Figure 3N, we would have done some processing on 1 and 2.
>
> And depending on 1 and 2 and whether they're the same or sufficiently similar or available, that's going to help us inform our choice of subset 3.
>
> Do you agree with that?
>
> **A.** Oh, for Rusert?
>
> **Q.** Yes.
>
> **A.** Yeah, I agree with that on -- with respect to Rusert.
>
> **Q.** Okay. And to some extent -- now I'm going to refer to the Nakamura reference.
>
> That's -- that's also true of the Nakamura reference, where we have S0 and S1; and if we're identifying S1 as the first spatial motion vector candidate, what exists in the subset of S0 will depend on some prior processing before we get to S1 in Nakamura's order.
>
> **A.** I -- I believe it's accurate, yeah. S0 will be filled first, yeah, and then

Nokia Exhibit 2001, p. 48
Amazon.com v. Nokia
IPR2024-01505

-- and then -- and we don't know what's going to be a priori, right?

So then you'll -- you'll have -- you'll evaluate candidates for S1.

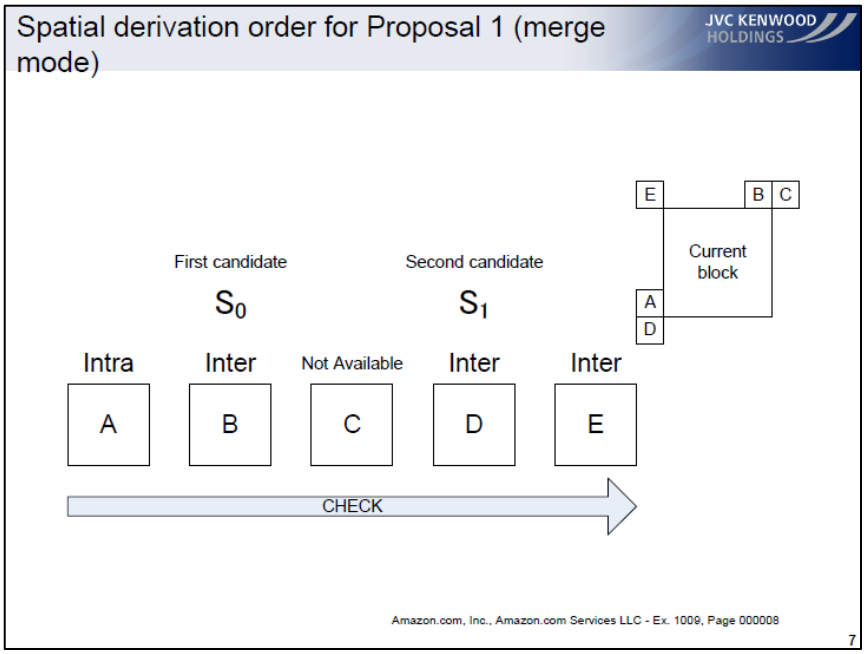
(Ex-2010, 103:20-105:9)

89. In my opinion, Dr. Creusere's statement is correct. In both Rusert and Nakamura, the content of the alleged "subset" has already been determined before any selection, processing, or evaluation of what Petitioner alleges is the "first spatial motion vector prediction candidate."

90. Considering Nakamura, Nakamura determines S0 (the alleged subset) first. Nakamura does this without knowledge or confirmation of what S1 will be. For example, in the scenario shown below from Nakamura's proposal (Ex-1009, 8), the motion vector associated with neighbouring block "B" is selected as the S0 value.[10]

[10] The "check" arrow in Nakamura shows the order in which A to E are considered.

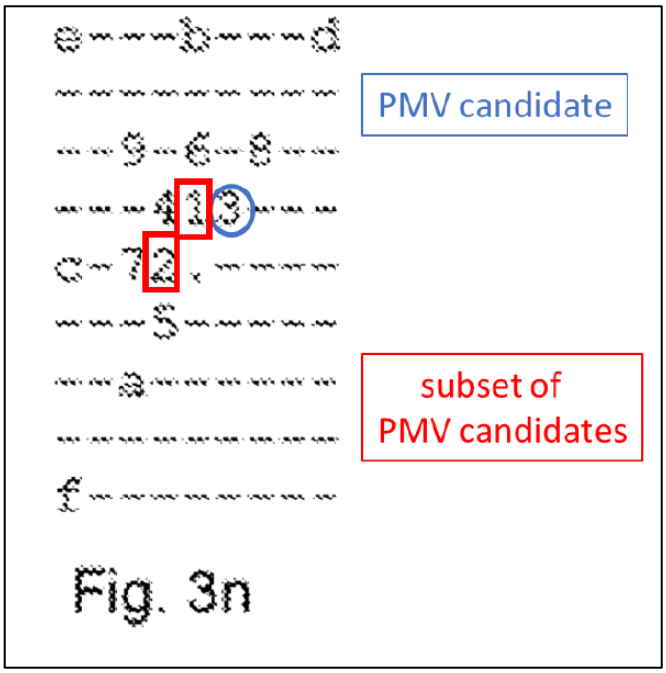
Nokia Exhibit 2001, p. 49
Amazon.com v. Nokia
IPR2024-01505



91. In this example from Nakamura, the selection of the motion vector of B as the value of S0 does not depend on the identity or block location of S1, which—at the time S0 is determined—could be any of C, D, E, or nothing, based on availability. This is true for any possible arrangement in Nakamura. In other words, Nakamura's alleged subset (the single value S0) is determined without knowledge of the alleged first spatial motion vector prediction candidate (the single value S1) or the location of the block associated with S1. Logically, a determination of the subset (alleged to be S0) cannot be based on the location of the first spatial motion vector prediction candidate (alleged to be S1) if the first spatial motion vector prediction candidate (S1) is not known or confirmed at the time the determination of the subset (S0) is made. A POSITA would understand that the first spatial motion vector prediction candidate must be known and confirmed for any determination of

a subset to be "based on" the location of the first spatial motion vector prediction candidate. The Board correctly noted this in its Institution Decisions in the IPR proceedings relating to the '714 Patent. (IPR2024-00604 Institution Decision (paper 11) at 56 ("As Patent Owner notes, S0 is already selected regardless of what the selection of S1 thereafter might be. The subset S0 does not change depending on the later choice of S1.")).

92. Considering Rusert, I have provided below an annotated version of Rusert's Figure 3n from Paragraph 132 of Dr. Creusere's declaration (Ex-1003). In Dr. Creusere's example, blocks '1' and '2' are the alleged subset while '3' is the alleged first spatial motion vector prediction candidate.



93. The content of Rusert's alleged subset depends on block availability. For example, if '1' is unavailable, (Ex-1004, ¶ 54 (giving exemplary scenarios where

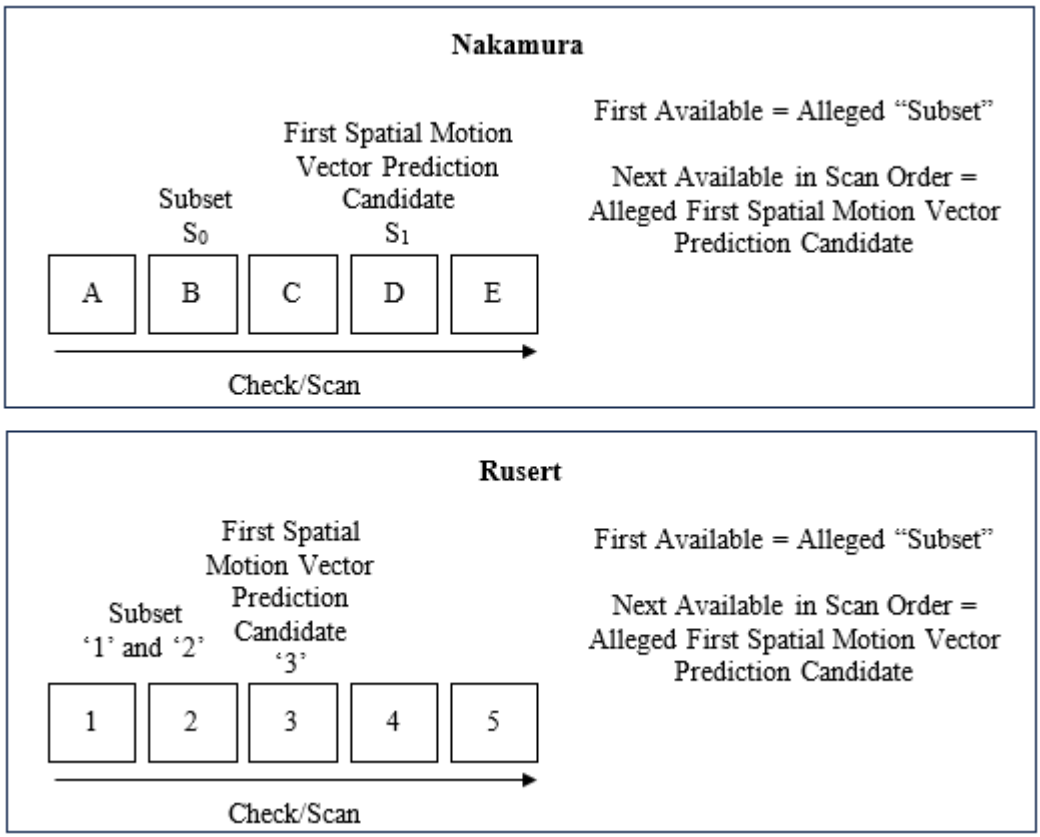
blocks are unavailable due to coding order or lack of motion vector)), the alleged subset will be just '2.' The alleged subset further depends on a comparison between the motion vectors of blocks '1' and '2' assuming both are available. If '1' and '2' are both available and their motion vectors are sufficiently similar or identical, the alleged subset will be just '1.'

94. By the time Rusert begins its processing to determine whether '3' is available and whether to add '3' to PMV_CANDS, the alleged subset has already been determined. Whether '3' is available is unknown to Rusert at the time the alleged subset is determined. If '3' is not available (for example, if block 3 is intra coded), then—assuming both '1' and '2' are in the alleged subset—the next available candidate, which may be '4', '5' or a later block, will be compared to '1' and '2.' Thus in Rusert, the alleged subset ('1' or '1, 2' in this example) has already been selected regardless of what the alleged first spatial motion vector prediction candidate ('3', '4' or a later block depending on which blocks are available) thereafter might be. The alleged subset does not change depending on the later choice of candidate for evaluation.

95. Furthermore, the Figure below illustrate that ***both*** Nakamura ***and*** Rusert's alleged subsets are determined before the alleged first spatial motion vector prediction candidate is known or confirmed. In Nakamura, S_0 is assigned the motion vector of block B, prior to determining a suitable motion vector for candidate S_1. In

Nokia Exhibit 2001, p. 52
Amazon.com v. Nokia
IPR2024-01505

Rusert, PMV_CANDS is determined to be '1', '1, 2' or '2', prior to determining whether to add the next available block motion vector to PMV_CANDS.



2. In Nakamura and Rusert, Comparisons are Performed with All Possible Candidate Pairs

96. I will now discuss how both Rusert and Nakamura compare what Petitioner alleges to be the "first spatial motion vector prediction candidate" with ***each*** "possible candidate pair." (*See, e.g.*, Ex-1001, cl. 1, cl. 15, cl. 17).

97. In Nakamura, the first available motion vector prediction is added to the list as S0. Nakamura then scans for the next available candidate, which, when found, is labelled S1. S1 will always be compared to S0. Given Nakamura's process,

Nokia Exhibit 2001, p. 53
Amazon.com v. Nokia
IPR2024-01505

the only possible pair of spatial motion vector prediction candidates is S1 and S0.[11] (*See, e.g.*, Ex-1009, 8). In other words, Nakamura does not reduce computation by comparing less than all of the possible candidate pairs, but instead proposes to reduce the number of possible candidate pairs to 3 and then subsequently comparing every possible candidate pair.

98. I provide below a highlighted excerpt from Nakamura's JCTVC-F419 proposal explaining that the only possible pair of spatial motion vector prediction candidates is S1 and S0. (Ex-1007, 2). The abbreviation "Col" refers to "collocated" block, which is a temporal motion vector prediction candidate and thus not a "spatial motion vector prediction candidate" as considered by the '833 Patent.

---

[11] This is true in any given scenario regardless of whether one considers S0 and S1 as the set of "candidates" or A, B, C, D, and E as the set of "candidates."

Nokia Exhibit 2001, p. 54
Amazon.com v. Nokia
IPR2024-01505

Table 2 Comparison between HM3.0 and proposed technique for merge mode

	HM3.0	Proposal 1
The number of spatial candidates	4 in 4 [positions]	2 in 5 [positions]
Spatial derivation order	A, B, C, D	A, B, C, D, E
The number of times of comparison of redundant candidates in the spatial derivation process	0 [time]	0 [time]
The number of temporal candidates	1	1
Merging candidate list order	A, B, Col, C, D	S_0, S_1, Col
The number of times of comparison in the removal process	10 [times] (A vs B, Col, C, D, B vs Col, C, D, Col vs C, D, and C vs D)	3 [times] (S_0 vs S_1, S_0 vs Col, and S_1 vs Col)

Notes:

S_0: The first spatial candidate found in the spatial derivation process

S_1: The second spatial candidate found in the spatial derivation process

99. Rusert operates in an analogous way. With reference to Figure 3n, the first available candidate in scan order will automatically be added to the PMV_CANDS list. (Ex. 1004, Figure 3n). For example, '1' is added to the PMV_CANDS list if block 1 is available. Rusert will then scan for the next available candidate in scan order. For example, '2' will be the next candidate if block 2 is available. Then, '2' is compared with '1.' This is a comparison of all possible candidate pairs at this point in Rusert's process to determine whether '2' will be added to the PMV_CANDS list.

Nokia Exhibit 2001, p. 55
Amazon.com v. Nokia
IPR2024-01505

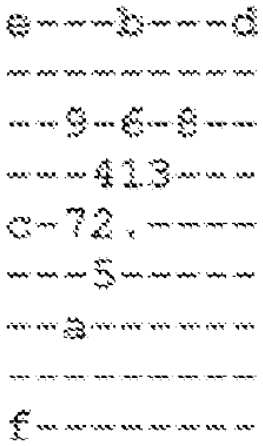


Fig. 3n

(*Id.*)

100. Rusert continues this iterative process until the scan order is complete or until it reaches a maximum list size.[12] For example, '3' may be the next candidate if available. At this point, if the candidate list PMV_CANDS comprises '2' and '1', '3' is compared to '2' and '1.'[13] This is a comparison of all possible candidate pairs

---

[12] Under Petitioner's interpretation of Rusert, each iteration along Rusert's scan requires selecting a new first spatial motion prediction candidate. In Rusert, for each subsequent iteration, the next "first spatial motion predication candidate" is the next available candidate in the scan order.

[13] If available, '3' may be compared to only '1', only '2' or both, depending on the availability and/or similarity of '1' and '2.' Regardless of the scenario, Rusert specifies a comparison of all possible candidate pairs to determine whether to add a

Nokia Exhibit 2001, p. 56
Amazon.com v. Nokia
IPR2024-01505

to determine whether '3' will be added to the list.

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

A. Ground 1: Rusert in Combined with Zheng

101. The Petition relies on Rusert to satisfy the limitations of the independent claims for its Ground 1. Petition at 21. I understand that, at the institution stage, if the Board determines institution is warranted as to a single claim then the proceeding is instituted as to all claims. I therefore focus my analysis on the independent Challenged Claims. For the independent Challenged Claims, Zheng is only relied upon for its use of the terms "prediction unit" and "merge list," and is not relied upon for any of the limitations I address below. (Petition, 28, 30).

1. Rusert does not teach "determining a set of spatial motion vector prediction candidates for the encoded block of pixels" as required by 1[b], 15[b], and 17[b]

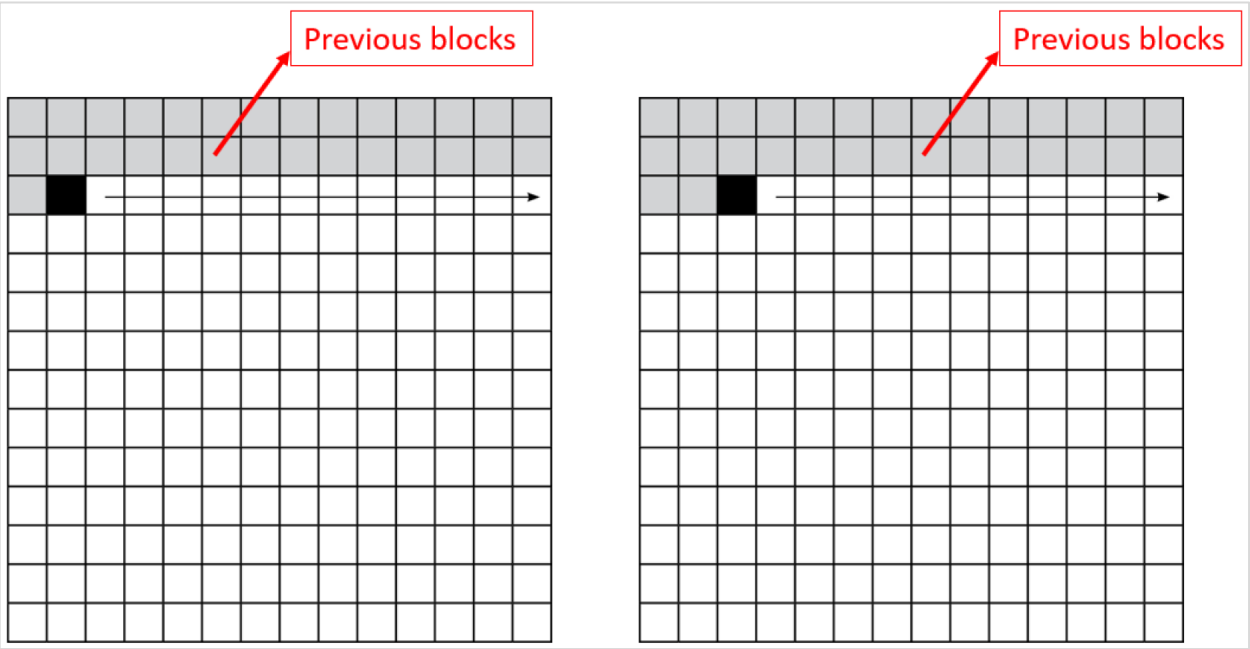
102. For the "determining a set of spatial motion vector prediction candidates for the encoded block of pixels" portion of limitations 1[b], 15[b], and 17[b], the Petition asserts that in its asserted combination, "the set of spatial motion vector prediction candidates encompasses all previously encoded blocks," i.e., all

---

candidate to the PMV_CANDS list. For example, if PMV_CANDS contains just '2' at the point when '3' is evaluated, then the only possible candidates are '2' and '3' and the comparison of '3' with '2' is a comparison of all possible candidate pairs.

Nokia Exhibit 2001, p. 57
Amazon.com v. Nokia
IPR2024-01505

previously coded vectors for the current frame. (Petition, 29). The Petition illustrates its theory of what makes up the "set of spatial motion vector prediction candidates: in its Ground 1 showing a current block in black and previously coded blocks in grey:



(*Id.*).

103. Contrary to the Petition's assertion, a POSITA would understand this to show a *lack* of any "determining a set of spatial motion vector prediction candidates for the encoded block of pixels" as required by the challenged claims. Because all previously coded motion vectors are included in the alleged "set of spatial motion vector prediction candidates", the asserted combination does not perform any "determining" of the alleged "set", "for the encoded block of pixels."

104. The Petition's treatment of this portion of limitations 1[b], 15[b], and

17[b] reads "determining a set of spatial motion vector prediction candidates for the block of pixels" out of the claim and renders it superfluous, as the same set previously coded motion vectors (all previously coded motion vectors for the current frame) exists, in the same form, regardless of whether a current block is received, and regardless of how it is encoded. The existence of any spatial motion vector candidates associated with previously coded blocks in the frame, no matter how obtained, would satisfy the limitation as interpreted by Petitioner. As a result, the claim would have the same scope whether limitation 1[b] were included in Claim 1 or not. I understand that a POSITA in interpreting a claim is required to give meaning to all of its terms. Here, the Petition ignores the "determining a set" step and focuses only on the existence of spatial motion vector prediction candidates.

105. Further, the plain language of the challenged claims requires the "determining a set of spatial motion vector prediction candidates" step (1[b]) to be performed before the "selecting a first spatial motion vector prediction candidate from the set of spatial motion vector prediction candidates" step (1[c]) and the "comparing motion information of the first spatial motion vector prediction candidate … without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates" step (1[e]). Rusert does not teach this. A vector cannot be "selected … from the set" until it is part of the set, and "without making a comparison of each possible candidate pair from the set" requires

the set to have first been determined.

106. However, in Rusert the set of spatial motion vector prediction candidates is not determined in advance, but as part of the scan order. (Ex-1004, ¶ 54). As Rusert paragraph [0054] explains, its contemplated list is constructed according to scan order, and at each position in the scan order the availability of a PMV for the block is determined and then, if available, is compared and potentially added to the list. (*Id*.) At a particular scan position, the first step in Rusert is to determine whether the position has an available MV. (*Id*.) The set of spatial motion vector prediction candidates is not closed (i.e., is not determined) until the list limit is reached (*Id.*, ¶ 48) or no more positions in the scan order remain (*Id.*, ¶ 49).

107. When the set is closed because the list limit is reached, this is not known (and the set therefore determined) until after the "comparing" step, so Rusert would not complete the "determining a set of spatial motion vector candidates" step until after completing the "select a subset" or "comparing" steps, which would not satisfy the challenged claims.

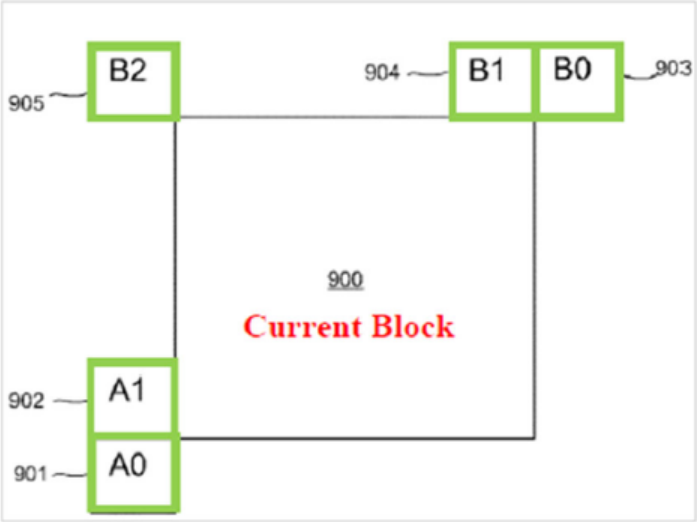
108. In the case where the set is closed because no more positions remain in the scan order, then the motion information of the last candidate is compared against the motion information for every possible candidate in the set and cannot satisfy the "comparing" limitation. Rusert therefore cannot satisfy the "determining a set" limitation under the Petition's theory while also satisfying the "comparing"

limitation.

2. Rusert does not teach "determining a subset of spatial motion vector prediction candidates based on a location of the block associated with the first spatial motion vector prediction candidate" as required by 1[d], 15[d], and 17[d]

109. As discussed above, Rusert's subset of spatial motion vector candidates is not based on the location of the block associated with the alleged first spatial motion vector candidate, nor is it in any way influenced by the alleged first spatial motion vector candidate.

110. The '833 Patent discloses that a set of spatial motion vector prediction candidates is obtained from neighbouring blocks located around a current block of pixels being either encoded or decoded. (Ex-1001, 13:13-19). This relationship is illustrated in the following Figure. I have annotated the Figure to show the neighbouring blocks from which the set of spatial motion vector prediction candidates is obtained.



Ex-1001, Figure 9 (annotated to show the set of motion vector prediction candidates A0, A1, B0, B1 and B2)

111. As a POSITA would understand, and as the ’833 Patent makes clear, when generating the claimed merge list, a spatial motion vector prediction candidate is selected for consideration and is subsequently compared with a subset of the other spatial motion vector prediction candidates, by checking whether each member of the subset has the same motion information as the selected candidate. (Ex-1001, 15:38-16:45). The purpose of this comparison is to avoid adding a spatial motion vector prediction candidate to the motion vector prediction merge list if the new candidate would be redundant. (*Id.*, 4:7-15).

112. The subset that the ’833 Patent describes is determined “based on the location of the block associated with the first spatial motion vector prediction candidate,” so that for each candidate in question, a subset is determined which is less than the full candidate set. Thus, instead of comparing a candidate with all the other four blocks (in the set of five candidates), the candidate is compared with a subset in order to avoid redundancy, while also reducing the computational work required to perform comparisons. (*Id.*) Columns 15-16 of the ’833 Patent provide example of comparing a candidate with a subset. Referencing the Figure above, for the candidate located to the left (A1), the ’833 Patent teaches that a subset containing only B1 can be used for comparison. (*Id.*, 15:38-62). Similarly, for the candidate located at the top left (B2), the ’833 Patent teaches that the subset containing only

Nokia Exhibit 2001, p. 62
Amazon.com v. Nokia
IPR2024-01505

A1 and B1 can be used for comparison. (*Id.*, 16:29-39). In the embodiments of the '833 Patent, the location of a selected spatial motion vector prediction candidate must first be known in order to then determine a subset based on this location. The fact that the location of a selected spatial motion vector prediction candidate (A1 or B2 in this example) is determinative of the members of the subset used for comparison (B1 or {A1, B1} in this example), in my opinion directly contradicts Petitioner's construction of "the block" in element 1[d]. (Petition, 20; arguing that "the block" in element 1[d] means the "block of pixels" (i.e., the "current block" being encoded or decoded); *see also* Section VI).

113. In its Institution Decision in the IPRs for the related '714 Patent, the Board disagreed with Petitioner's proposed construction of "the block" to mean "the block of pixels" in limitations 1[a] and 9[a] of the '714 Patent. (-00604 Institution Decision (Paper 11) at 18 -00605 Institution Decision (Paper 11) at 17). Instead, the Board agreed with Patent Owner that "the block" in '714 limitation 1[b]/9[b] means "the block associated with the first spatial motion vector prediction candidate," as '714 limitation 1[b]/9[b] states. (*Id.*) I agree with both Patent Owner and the Board. Petitioner makes a similar argument regarding "the block" and "the block associated with the first spatial motion vector prediction candidate" in the challenged claims of the '833 Patent, and in my opinion the same reasoning the Board applied in the institution decisions of the '714 Patent applies here.

114. The Petition approaches the construction of a motion vector prediction list differently to the claimed technique of the ’833 Patent. In Rusert, each spatial motion vector candidates is obtained from neighbouring blocks in a scanning order based on each neighbouring block’s distance from the current block. For each spatial motion vector candidate, Rusert determines whether, if available, it should be added to the list, and then moves on to the next PMV candidate in the scan order until the process is completed. (Ex-1004, ¶ 44). As I show above, Figure 3n of Rusert provides an exemplary depiction of this distance-based scanning method, with the closest PMV candidates being evaluated first and the farthest last (i.e., in the order 1, 2, 3 . . . 9, a, b . . . depending on availability), in a counter-clockwise spiral starting from the current block, represented by the ‘.’. (*Id.*, ¶ 58).

115. When determining whether to add a PMV candidate to the motion vector prediction (PMV_CANDS) list, Rusert compares the candidate to each previously added PMV candidate, i.e. to all members of PMV_CANDS at this point. For example, if candidate 1 has previously been added to the list, candidate 2 will be compared to candidate 1; if candidate 2 has also previously been added, then candidate 3 will be compared to candidates 1 and 2, and so on. (*Id.*, ¶¶ 71-72). If the comparison determines that the candidate under consideration is too similar to one already added, then it can be excluded, i.e., not added to the PMV_CANDS list. (*Id.*)

116. In my opinion, Petitioner is incorrect in asserting that the alleged

"subset" is determined based on the location of a selected spatial motion vector prediction candidate for at least two reasons. First, the content of PMV_CANDS at the time that a PMV candidate is considered has already been determined without regard for the PMV candidate under consideration. Second, the content of the PMV_CANDS list is based on the availability and values of the previous candidates, not on the location of the current candidate under consideration. (Petition, 39-40).

117. Considering the first point, Rusert's PMV_CANDS list is constructed by scanning neighbouring blocks in an order determined by their distance to the current block, where each motion vector block that is on the scan path is a PMV candidate.[14] (Ex-1004, ¶ 61, Fig. 3n, reproduced below). Previously coded blocks that are outside of the allowed distance or are unavailable cannot be included in PMV_CANDS. (*Id.*, ¶ 54; EX-2010, 125:9-127:20).

---

[14] Rusert explains that if a block has no motion vector present (such as the red highlighted, intra coded macroblocks above in paragraphs 58-60), they are not available as motion vector candidate. In Rusert, only blocks with valid motion vectors may be added to the subset of motion vectors in PMV_CANDS. (Ex-1004, ¶ 54). Thus a block that is unavailable, for example because it does not have an associated motion vector, is not included in Rusert's scan order and is not evaluated for adding to PMV_CANDS.

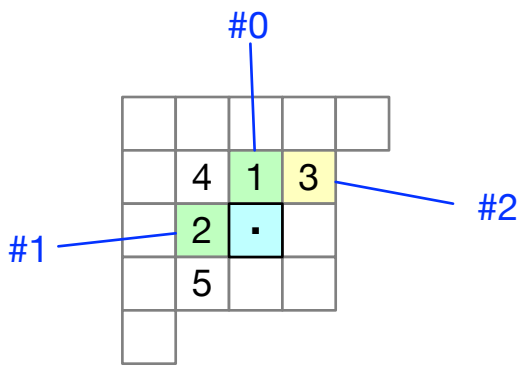
Nokia Exhibit 2001, p. 65
Amazon.com v. Nokia
IPR2024-01505

(Ex-1004, Figure 3n).

118. Petitioner argues that the subset is "based on" the current PMV block location, because for example if the current block is being evaluated (which Petitioner describes as the first spatial motion vector prediction candidate) 3, the subset is at most 1 and 2, i.e., according to Petitioner, the subset {1, 2} depends on the fact that block 3 is currently being evaluated. (Petition, 41-42). In my opinion, Petitioner is incorrect. Consider an example. At iteration 0 of Rusert's scan, PMV_CANDS is empty and block 1 is evaluated, then added to PMV_CANDS. At iteration 1, PMV_CANDS contains {1} and block 2 is evaluated, then added to PMV_CANDS. At the next iteration 2, regardless of the location of block 3, PMV_CANDS (the alleged subset) can only contain (at most) blocks 1 and 2. The

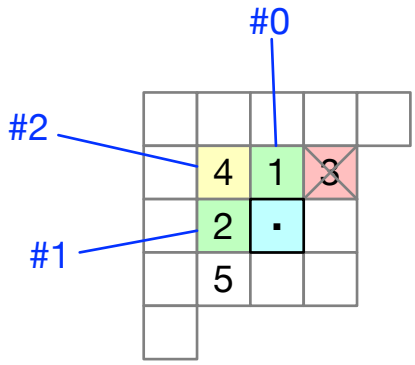
Nokia Exhibit 2001, p. 66
Amazon.com v. Nokia
IPR2024-01505

location of the block being evaluated is not taken into account in any way when constructing the previous content of PMV_CANDS (the alleged subset). For example, if block 3 is available, then iteration number 2 (highlighted in yellow and labelled #2 in the Figure below) compares block 3 (assuming the scan order of Figure 3n above) with a subset (PMV_CANDS) containing 1 and 2.



Example 1: Block 3 is Available

119. If block 3 is not available, iteration number 2 would evaluate block 4 (which has a different position) with the same subset containing 1 and 2.



Example 2: Block 3 is Not Available

120. Thus, the subset does not depend on the location of the block currently being evaluated.

Nokia Exhibit 2001, p. 67
Amazon.com v. Nokia
IPR2024-01505

121. At the point at which Rusert has finished its comparison of neighbouring block 2, the content of PMV_CANDS for the next block to be evaluated, is already known. PMV_CANDS will contain the set of candidates 1 and 2, assuming that 1 and 2 are available and dissimilar. The location of, and information related to, block 3 (or block 4, if block 3 is unavailable has no effect on the contents of PMV_CANDS at the time block 3 is under consideration. PMV_CANDS is determined (i.e. all its members are assigned) prior to considering a next block position. This is very different to the method of the '833 Patent. In Rusert, no information about the next candidate (including the position of the next candidate) is taken into account when determining PMV_CANDS prior to evaluating the next candidate. Therefore, PMV_CANDS (the alleged "subset") is not determined based on the location or any other aspect of the next candidate to be evaluated (the alleged "first spatial MVP candidate"). For example, assuming that blocks 1, 2, 3 and 4 are available, Rusert's method proceeds as followed:

Iteration	PMV_CANDS before	Block being evaluated	PMV_CANDS after
0	{}	1	{1}
1	{1}	2	{1, 2}
2	{1, 2}	3	{1, 2, 3}
3	{1, 2, 3}	4	{1, 2, 3, 4}

122. Now assume that block 3 is not available, Rusert's method proceeds as follows:

Nokia Exhibit 2001, p. 68
Amazon.com v. Nokia
IPR2024-01505

Iteration	PMV_CANDS before	Block being evaluated	PMV_CANDS after
0	{}	1	{1}
1	{1}	2	{1, 2}
2	{1, 2}	4	{1, 2, 4}
3	{1, 2, 4}	5	{1, 2, 4, 5}

123. In the examples of these two tables, at iteration 2, PMV_CANDS always contains at most blocks 1 and 2. Depending on whether block 3 is available, the current block being evaluated may be block 3 or block 4, or a later-numbered block. Which block is being evaluated at iteration 2 (block 3, block 4 or a later block) has no bearing on the content of PMV_CANDS and therefore the location of the block being evaluated (the alleged "first spatial motion vector prediction candidate") has no bearing on the content of PMV_CANDS (the alleged "subset"). The only thing that determines the candidates in the PMV_CANDS list, at any iteration, is the availability of the blocks that are considered before the current block that is being evaluated.

124. As a further example, consider block 4. PMV_CANDS, prior to processing block 4, cannot depend on the location of block 4. The contents of PMV_CANDS depend only on the previously processed blocks in the scan order, i.e., on the values and processing of the motion vectors for blocks 1, 2, and 3 in this example. Block 4 in this example is the alleged "first spatial MVP candidate" and PMV_CANDS prior to evaluating block 4 is the alleged "subset of spatial MVP

candidates," according to Petitioner. Once again, PMV_CANDS is fully determined prior to the processing block 4, i.e., prior to selecting block 4 for processing. Consequently, the "determining a subset" (determining of PMV_CANDS prior to a comparison involving block 4) cannot be based on the location of block 4 or any other aspect of block 4.

125. Rather than being based on the location of a PMV candidate under consideration, the set of PMV candidates used for comparison in Rusert is determined based on the availability and value of each PMV candidate that has already been considered and added to the list. (Ex-1004). As I explain above, if candidate 3 is compared to candidates 1 and 2, this is not because candidate 3 is located at the top right, but because candidates 1 and 2 are available and dissimilar and thus both have previously been added to PMV_CANDS. (*Id.*, ¶¶ 70-72).

126. In reaching its institution decision in the IPRs related to the '714 Patent (IPR2024-00604 and -00605, Paper 11 in each), the Board considered the following example: "if the first selected candidate is at the location of block 5, then the subset would be the four members 1, 2, 3, and 4, and if the first selected candidate is at the location of block 3, then the subset would be the two members 1 and 2" (-00604 Institution Decision at 35; -00605 Institution Decision at 34). In my opinion, this example is an incorrect characterization of Rusert. As I explained above, if the first selected candidate is at the location of block 5, then Rusert's PMV_CANDS *may*

consist of {1,2,3,4}, but may (and often will) consist of a smaller set such as {1,2}, {1,3} etc, depending on availability of blocks and similarity of motion vector values. If the first selected candidate is at the location of block 3, then PMV_CANDS *may* consist of {1,2} but may alternatively consist of {1}, {2} or no entries at all, depending on availability and similarity as above. Stated another way, the alleged subset PMV_CANDS may consist of {1,2} for the first selected candidate at block 3, block 4, block 5 or any numbered block in Rusert's scan, depending on availability of blocks. Thus I disagree with the Board's conclusion that "The members of the determined subset [PMV_CANDS] depend on the location of the block associated with the first selected spatial motion vector prediction candidate" ((-00604 Institution Decision at 35; -00605 Institution Decision at 34).

3. Rusert does not teach "comparing . . . without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates," as required by 1[e], 15[e], and 17[e]

127. The Petitioner's first ground relies on the interpretation that the subset of Rusert is compared without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates, suggesting that Rusert compares potential candidates with preceding candidates already in PMV_CANDS, which are a subset of candidates in the scan order. (Petition, 41-42). I disagree with this interpretation because it is inconsistent with the specification of the '833 Patent, the understanding of a person of skill in the art and the express teachings of Rusert.

Nokia Exhibit 2001, p. 71
Amazon.com v. Nokia
IPR2024-01505

128. As I explained above, in Section VIII.A.1, Columns 15-16 of the ’833 Patent provide an example of comparing candidates in the subset. Referencing Figure 9 above, for the candidate located to the left (A1), assuming A1 is the first spatial motion prediction candidate, the ’833 Patent teaches that a subset containing only B1 can be used for comparison. (Ex-1001, 15:38-16:62). Similarly, for the candidate located on the top left (B2), assuming B2 is the first spatial motion prediction candidate, the ’833 Patent teaches that the subset containing A1 and B1 can be used for comparison. (*Id.*, 16:29-39).

129. As required by claim limitation 1[e], the motion information of the first spatial motion vector prediction candidate (left column of the below table) is compared to the motion information of spatial motion vector prediction candidates in the determined subset of spatial motion vector prediction candidates (right column of the below table). Each first spatial motion vector prediction candidate and its respective subset is also illustrated in the ’833 Patent’s Figure 5b’s flow diagram.

First Spatial Motion Vector Prediction Candidate	Motion Information of Spatial Motion Vector Prediction Candidates in the Determined Subset of Spatial Motion Vector Prediction Candidates
A1	No subset. It’s automatically included in the merge list if available.
A0	{A1}
B0	{B1}
B1	{A1}
B2	{A1, B1}

130. As the ’833 Patent makes clear, and as a POSITA would understand,

the first spatial motion vector prediction candidate is compared to the determined subset without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates. For example, when A0 is the first spatial motion vector prediction candidate, A0 is not compared to B0, B1, or B2; when B0 is the first spatial motion vector prediction candidate, B0 is not compared to A1, A0, or B2; when B1 is the first spatial motion vector prediction candidate, B1 is not compared to A0, B0, or B2; and when B2 is the first spatial motion vector prediction candidate, B2 is not compared to A0 or B0.

131. The Petition approaches comparing in a different way to the claimed technique of the '833 Patent. Petitioner quotes various portions of Rusert to support its argument. First, the Petitioner states, "the selected set of PMV candidates comprises a subset of the set of previously coded motion vectors." (Petition, 35; discussing PMV_CANDS and citing Ex-1004, ¶ [0011]). But Rusert here is discussing ***all*** previously coded MVs, ***not*** previously coded MVs that are candidates for predicting the current MV. Rusert's "subset" is a "subset of the set of previously coded motion vectors" (Ex-1004, ¶ [0011]) that are in the scan order (what Petitioner and Dr. Creusere call the "set"), not the claimed "subset of spatial motion vector candidates" for comparison purposes.

132. Second, the Petitioner suggests that, "Rusert's … comparisons … compare … without making a comparison of each pair from the set of motion vectors

used for all previous blocks." (Petition, 43-47). However, in Rusert, the candidates for predicting the motion vector of the current block are not the entire set of motion vectors for all previously coded blocks. Dr. Creusere agreed in his deposition that the "candidates" are those in Rusert's scan order, not all previously coded blocks. (Ex-2010, 72:11-76:19; *see also id.*, 75:6-76:1). Moreover, this statement in the Petition does not address the fact that Rusert compares each *possible candidate pair* when Rusert compares the alleged first spatial motion vector prediction candidate (e.g., '3') to the alleged subset (e.g., '1' and '2').

133. Third, the Petitioner states that, "[e]ven within the scan order, Rusert compares potential candidates with preceding candidates already in PMV_CANDS, which are a subset of the candidates in the scan order; therefore, Rusert avoids comparing each pair of candidates from the scan order." (Petition, 44). In my view, by making this statement, the Petitioner concedes that Rusert compares each possible pair of candidates from the scan order. As stated by Petitioner, "within the scan order, Rusert compares potential candidates with preceding candidates that are already in PMV_CANDS, which are a subset of the candidates in the scan order." (*Id.*) However, the subset of candidates in Rusert is also the entire set of candidates available for at the current iteration, i.e. at the point when a potential candidate (the alleged "first spatial motion vector prediction candidate") is compared. In other words, the higher scan-order locations do not include "possible candidate pairs"

because they have not been evaluated yet to determine whether they include a PMV. Since the higher-scan order locations are not part of the set at that point, they cannot be considered as part of a possible candidate pair. Therefore, Petitioner is incorrect to state that "Rusert compares the selected candidate with a subset of candidates, without comparing each pair from the set of previously-coded motion vectors or the entire scan order" because the Petitioner does not address the fact that Rusert does indeed compare each possible candidate pair that is available at the time a particular current block is being evaluated for inclusion on the PMV_CANDS list. (*Id.*).

134. For example, if the maximum number of candidates in Rusert's scan is four, and if PMV_CANDS contains {1, 2, 3, 4} after the third iteration number 3, assuming that and blocks 1-4 are available, then the motion vector of the fifth block in Rusert's scan is not available as a candidate. At minimum, the MV of the fifth block is not a candidate available for comparison such that the MV of the fifth block would be part of any "possible candidate pair" at the time a first spatial motion vector prediction candidate (e.g., '3' or '4') is compared to its corresponding subset and the determination to include or exclude based on comparing is made. For example, when block 4 is been evaluated and compared to '1,' '2,' and '3,' all the possible pairs of candidates have been compared. This is because, during the processing scan for a given block, Rusert compares a current block with all blocks previously stored in PMV_CANDS, which means that all the available candidates for the current block

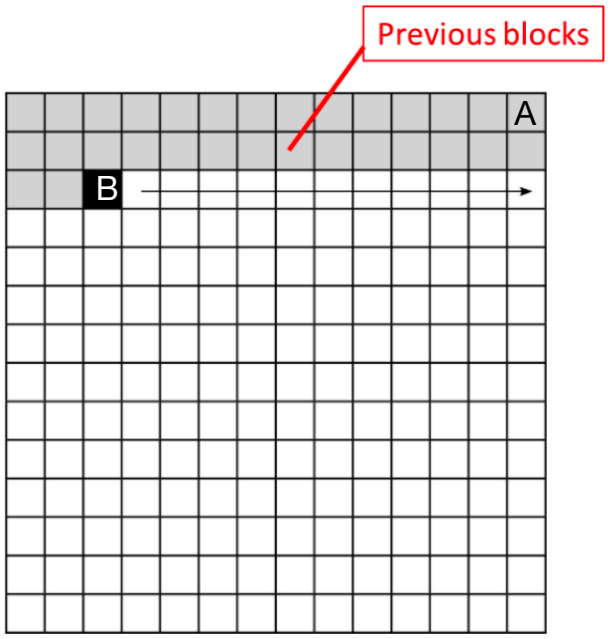
are compared during the current iteration. In other words, all possible (available) candidate pairs are compared.

135. If on the other hand the "set" comprises all the previously coded motion vectors in the current frame, as Petitioner appears to allege in the alternative,[15] then the previously coded motion vectors outside of the blocks in what Petitioner calls the "subset" are not available candidates (or possible candidate pairs) for the current block. For example, the top-right block MV (which I have labelled A below) is not available for consideration (or part of a possible candidate pair) for the current block towards the left of the picture (which I have labelled B below), because in Rusert's method A is too far away from B and can therefore never be considered as a candidate or used in any comparison.

---

[15] Dr. Creusere identifies the previously coded blocks within Rusert's scan order as spatial motion vector prediction candidates. (Ex-2010, 72:11-76:19).

Nokia Exhibit 2001, p. 76
Amazon.com v. Nokia
IPR2024-01505



(Petition, 29; annotated).

136. Further, as discussed in Section VIII.A.1 above, the method employed by Rusert to build its list means that, to the extent a "set of spatial motion vector prediction candidates for the block of pixels" is determined, it is determined as the Rusert method moves through the scan order. As a result, motion vectors corresponding to block locations at future positions in the scan order cannot be considered part of the "set" at the point in time when the "comparing" step occurs. As a result, future potential motion vector candidates cannot make up "possible candidate pair from the set of spatial motion vector prediction candidates" as required by 1[e], 15[e], or 17[e] of the challenged claims. Whether comparisons of these potential future set members are made is irrelevant to the claim.

137. Additionally, at each position in the scan order if a motion vector is

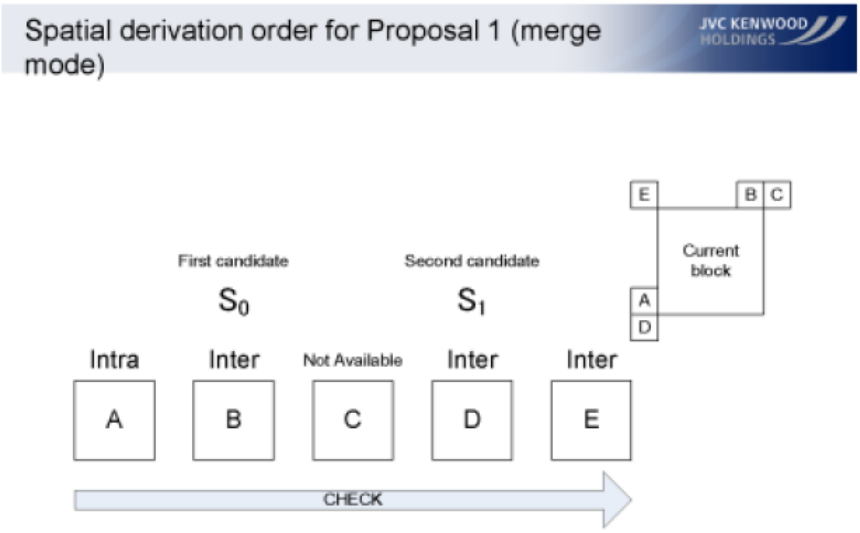
present its motion information is compared to the motion information of every previously identified motion vector and is added to the list if its motion information is different. Prior scan order candidates with unique motion information are included in the PMV_CANDS list and have their motion information compared directly. Prior scan order candidates with redundant motion information that are excluded based on a comparison of motion information have the same motion information as candidates on the PMV_CANDS list. As a result, the motion information of each new candidate is compared to the motion information of all previous candidates in the scan order. As a result, in Rusert each new candidate has its motion information compared against the motion information of candidates in the set in an earlier scan order position. As candidates from earlier scan order positions make up the entirety of the set at the time of the comparison, Rusert's PMV_CANDS cannot satisfy 1[e], 15[e], or 17[e] of the challenged claims.

B. Ground 2: Nakamura and WD4 - Nakamura and WD4 do not teach the challenged claims

1. Nakamura and WD4 do not teach "determining a subset of spatial motion vector prediction candidates based on a location of the block associated with the first spatial motion vector prediction candidate," as required by 1[d], 15[d], and 17[d]

138. Petitioner's argument fails because the Nakamura and WD4 combination does not teach element 1[d], "determining a subset of spatial motion vector prediction candidates based on the location of the block associated with the

first spatial motion vector prediction." First, Petitioner suggests that S0 is a subset of the spatial motion vector prediction candidates. (Petition, 73). Second, Petitioner indicates that S1 is the first spatial motion vector prediction candidate, even though Nakamura discloses that S1 is derived after S0, as illustrated in the Figure below, showing the check order (arrow) generating S0 (first candidate) then S1 (second candidate). (Petition, 68-69; see also Section VII.F).



Ex-1009, 8

139. Elements 1[c] and 1[d] in the '833 Patent have a logical order, since one cannot "determine a subset … based on a location … associated with the first spatial motion vector prediction candidate" until one has "selected a first spatial motion vector candidate." (Ex. 1001, cl. 1[c]). Unlike the '833 Patent, Nakamura adds two candidates, S0 and S1, to its merge list based on availability. As illustrated in the Figure above, S0 (the "first candidate" added to the list in Nakamura's

Nokia Exhibit 2001, p. 79
Amazon.com v. Nokia
IPR2024-01505

example) corresponds to the "B" block because it is the first inter-coded block encountered. S1 (the "second candidate" added to the list) corresponds to block "D," skipping over C because it is not available. Petitioner's arguments concerning S0 and S1 are flawed at least for the following reasons:

140. S0 (the alleged "subset") is not determined based on the location and S1 (the alleged "first spatial motion vector prediction candidate"). There is no connection between the selection of block B as the "subset" and, for example, block D's location in the bottom non-adjacent corner. Rather, block B is the alleged "subset" in Nakamura because it happens to be the first inter-block with an available motion vector. Additionally, by the time S0 is under consideration as the alleged "subset," relative to S1, S0 has already been determined. Therefore, the alleged subset determination has nothing to do with any characteristic of S1.

141. Petitioner relies on S1 as the "first spatial motion vector prediction candidate" presumably because S0 is added to the list without any comparison, and so S0 in any case cannot possibly meet the requirements of the claims. Furthermore, neither S0 nor S1 can satisfy the claimed "first spatial motion vector prediction candidate." S0 is determined based on neighbour availability and is included automatically on the merge list. S1 is then selected as the next available candidate following S0. Further, when S0 and S1 are compared (the only possible comparison), only S1 can be removed from the merge list if S0 and S1 have the same motion

information.

142. Accordingly, Petitioner's alleged S0 subset cannot be determined based on anything associated with S1 at the point in time when Nakamura derives S0, since S1 has not been derived and does not exist at this point.

143. The combination therefore fails to render obvious at least elements 1[d], 1[e], 15[d], 15[e], 17[d], 17[e], and the Petitioner has failed to show a reasonable likelihood that the challenged claims are invalid on this ground.

2. Nakamura and WD4 do not teach "comparing motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate of the subset of spatial motion vector prediction candidates without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates," as required by 1[e], 15[e], and 17[e]

144. With respect to the limitations of the challenged claims which describe "compar[ing] motion information of the first spatial motion vector prediction candidate with motion information of the spatial motion vector prediction candidate in the determined subset of spatial motion vector prediction candidates without making a comparison of each possible candidate pair from the set of spatial motion vector prediction candidates," it is my opinion that the Nakamura/WD4 combination again fails. The Petition cites to the table below wherein Nakamura contrasts its proposal with HM3.0.

Nokia Exhibit 2001, p. 81
Amazon.com v. Nokia
IPR2024-01505

Table 2 presents the comparison between HM3.0 and proposed technique for merge mode.

Table 2 Comparison between HM3.0 and proposed technique for merge mode

	HM3.0	Proposal 1
The number of spatial candidates	4 in 4 [positions]	2 in 5 [positions]
Spatial derivation order	A, B, C, D	A, B, C, D, E
The number of times of comparison of redundant candidates in the spatial derivation process	0 [time]	0 [time]
The number of temporal candidates	1	1
Merging candidate list order	A, B, Col, C, D	S_0, S_1, Col
The number of times of comparison in the removal process	10 [times] (A vs B, Col, C, D, B vs Col, C, D, Col vs C, D, and C vs D)	3 [times] (S_0 vs S_1, S_0 vs Col, and S_1 vs Col)

Notes:

S_0: The first spatial candidate found in the spatial derivation process

S_1: The second spatial candidate found in the spatial derivation process

Ex-1007 at § 2.2.1, Table 2

145. In my opinion, with respect to Nakamura's Proposal 1, Petitioner is incorrect to suggest that less than each possible candidate pair is compared. (Petition, 75). This erroneous conclusion appears to be based on Nakamura's lack of comparison between each of the A-E neighbours, which results from only "2 in 5" neighbours being used as "spatial candidates." (*Id.*.; Ex-1007, Table 2). However, this does not address whether Nakamura compares each "possible candidate pair" of spatial motion vector prediction candidates. Nakamura does. Specifically, Nakamura can only compare S0 and S1 and always does so. Thus, for the alleged first spatial motion vector prediction candidate 'S1,' Nakamura always compares all possible

candidate pairs.[16]

3. A POSITA would not have been motivated to combine Nakamura with WD4

146. The petition seeks not only to combine Nakamura and WD4, but to modify them both in the process. The petition fails to show that a POSITA would be motivated to make this modified combination and fails to show that a POSITA would have had a reasonable expectation of success in making it.

147. First, the petition suggests that once Nakamura was proposed, it would have been obvious to a POSITA to combine it with WD4 or any working draft of the HEVC standard. (Petition, 63-65). I disagree, as demonstrated by WD4 itself. Petitioner agrees that Nakamura was originally proposed as an improvement to Working Draft 3 of the HEVC standard but fails to address that Nakamura's proposed simplification to the derivation method for merge mode candidate was expressly evaluated by experts in the field and then not included in Working Draft 4. (*Id.*).

148. In my opinion, rather than following the suggestion of Nakamura, which "tries to reduce the number of candidates in the spatial derivation process," WD4 reflects a decision not to reduce the number of candidates during the spatial derivation process, instead adopting a technique that more strongly resembles the

---

[16] Note "Col" is not a spatial motion vector prediction candidate.

same HM3.0 technique with which Nakamura contrasts its proposed technique. (Ex-1007, Abstract; *compare* Ex-1007, § 2.2.1 *with* Ex-1010, § 8.4.2.1.1).

149. Despite Petitioner's assertion to the contrary, the unification of the spatial neighbour locations used for merge mode and Motion Vector Prediction ("MVP") does not reflect an adoption of Nakamura's proposal that would motivate a POSITA to combine Nakamura with WD4. (Petition, 65-66). In JCTVC-F744-v3 ("F744"), where the JCTVC evaluated the Nakamura proposal, the JCT-VC adopted only an extremely limited aspect of Nakamura: the "candidate search order and TMVP refidx derivation." (Ex-2009, 8).

150. The principal technique of Nakamura, in which only "two spatial candidates and a temporal candidate are derived," was also not adopted, with WD4 making use of 5 spatial candidates where WD3 and Nakamura suggested using 4 and 2 spatial candidates respectively. (Ex-1010, Figure 8-5; Ex-1007, Table 2). Given that the JCTVC considered Nakamura for integration into the Working Document and adopted only a narrow portion thereof, it is my opinion that a POSITA would not be compelled to reopen the JCTVC's consideration of Nakamura and adopt further aspects of the proposal that were previously rejected by the JCTVC.

IX. CONCLUSION

151. This Declaration represents the opinions that I have formed to date based on the material currently available to me. I reserve the right to revise,

Nokia Exhibit 2001, p. 84
Amazon.com v. Nokia
IPR2024-01505

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.

152. 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: February 21, 2025

Iain E. Richardson

Dr. Iain Richardson

Nokia Exhibit 2001, p. 85
Amazon.com v. Nokia
IPR2024-01505