**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: IPR2024-00626
U.S. Patent No. 11,805,267

---

# PATENT OWNER'S PRELIMINARY RESPONSE

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

TABLE OF CONTENTS

I. INTRODUCTION .....1
II. BACKGROUND .....2
A. TECHNICAL BACKGROUND .....2
1. Motion Prediction in Video Coding .....3
B. THE '267 PATENT .....13
C. THE '267 PATENT'S PROSECUTION HISTORY .....15
1. Priority Applications .....15
2. Application No. 17/328,750 .....20
3. Prior Art Cited in Prosecution .....20
D. THE ASSERTED PRIOR ART .....26
1. Walker (Ex-1004) .....26
2. Karczewicz-I (Ex-1005) .....27
3. Karczewicz-II (Ex-1006) .....27
E. POSITA .....27
F. CLAIM CONSTRUCTION .....27
G. CLAIM LISTING .....28
III. ARGUMENT .....28
A. THE PETITION SHOULD BE DENIED UNDER *ADVANCED BIONICS* BECAUSE CUMULATIVE REFERENCES HAVE BEEN REPEATEDLY ARGUED TO THE BOARD AND PATENT OFFICE AND PETITIONER HAS NOT DEMONSTRATED EITHER BODY ERRED. .....28
1. Petitioner's art is substantially the same as the art cited in the prosecution of the '267 Patent.....30
2. Petitioner has not established that the Patent Office erred during prosecution.....44
B. PETITIONER HAS NOT ESTABLISHED A REASONABLE LIKELIHOOD THAT THE CHALLENGED CLAIMS ARE UNPATENTABLE IN LIGHT OF GROUND 1. .....46
1. Petitioner fails to demonstrate that the Challenged Claims are obvious over Walker because Walker does not use higher-precision predictions.46

C. PETITIONER HAS NOT ESTABLISHED A REASONABLE LIKELIHOOD THAT THE CHALLENGED CLAIMS ARE UNPATENTABLE IN LIGHT OF GROUND 2. .....57

1. Petitioner fails to demonstrate a motivation to make the modified combination of Karczewicz-I and Karczewicz-II raised in Ground 2 .....57

2. Petitioner cannot show a reasonable likelihood that the Challenged Claims are unpatentable without its hindsight-motivated combination of Karczewicz-I and Karczewicz-II .....62

IV. CONCLUSION.....63

# TABLE OF AUTHORITIES

**Page(s)**

**CASES**

*Advanced Bionics, LLC v. Med-El Elektromedizinishe Geräte GMBH*,
IPR2019-01469, Paper 8 (PTAB Feb. 13, 2020).....29, 44, 45

*Evapco Dry Cooling, Inc., v. SPG Dry Cooling USA*,
IPR2021-00687, Paper 11 (P.T.A.B. Sep. 24, 2021).....31, 35

*Gardner Denver v. Utex Industries, Inc.*,
IPR2020-00333, Paper 12 (P.T.A.B. Aug. 5, 2020).....31, 35

*In re Magnum Oil Tools Int'l, Ltd.*,
829 F.3d 1364 (Fed. Cir. 2016) .....57

*Regents of University of California v. Broad Institute, Inc.*,
903 F.3d 1286 (Fed. Cir. 2018) .....57

*Roku, Inc. v. Universal Elecs., Inc.*,
IPR2019-01619, Paper 11, 16-17 (P.T.A.B. April 2, 2020).....31

# TABLE OF EXHIBITS

| Exhibit | Description |
|---|---|
| 2001 | Declaration of Dr. Iain Richardson |
| 2002 | CV of Dr. Iain Richardson |
| 2003 | File History of App. Ser. No. 15/250,124 ("'124 Application") |
| 2004 | File History of App. Ser. No. 15/490,469 ("'469 Application") |
| 2005 | File History of App. Ser. No. 15/876,495 ("'495 Application") |
| 2006 | File History of App. Ser. No. 16/729,974 ("'974 Application") |
| 2007 | Pat. App. Pub. No. US 2009/0257503 to Yan Ye, Peisong Chen, and Marta Karczewicz ("Ye '503") |
| 2008 | Pat. App. Pub. No. US 2010/0086027 to Rahul Panchal and Marta Karczewicz ("Panchal") |
| 2009 | Pat. App. Pub. No. US 2013/0142262 to Yan Ye and Alexandros Tourapis ("Ye '262") |
| 2010 | VCEG-AI33 by Yan Ye, Peisong Chen, and Marta Karczewicz |
| 2011 | U.S. Pat. No. 9,161,057 to Marta Karczewicz, Peisong Chen, and Yan Ye ("Karczewicz '057") |
| 2012 | Excerpts from "The H.264 Advanced Video Compression Standard", Richardson, 2010 (Cover pages, book pages 28-38, 69, 106, 108, 223-229) |

# I. INTRODUCTION

Petitioner advances two arguments in the petition, both of which were known to, considered by, and rightfully rejected by the Patent Office in issuing the ’267 Patent and each of its five priority applications. The Board should therefore deny institution both under *Advanced Bionics* and on the merits.

Petitioner’s first ground asserts that weighted prediction, where weights are applied to predictions taken from two reference pictures, teaches the challenged claims’ requirement that a first and second prediction with precision greater than the precision of the reference picture be obtained and combined into a combined prediction. An identical argument was presented by the Examiner and overcome in the ’267 Patent’s first priority application. Patent Owner explained that applying a weight to a prediction does not make a higher-precision prediction, as required by the claims, and the pending claims were allowed. The same Examiner who allowed the ’267 Patent’s priority application allowed the ’267 Patent, and he noted in his reasons for allowance that he allowed the claims of the ’267 Patent to issue for the same reasons. Though the specific reference relied upon by Petitioner was not part of the prosecution history, multiple references containing the identical bi-prediction equation relied upon by Petitioner were considered and rejected by the Patent Office. Ground 1 is thus both without merit and duplicative of art and arguments already considered by the Patent Office.

Petitioner's second ground modifies and combines two references, one of which was before the Patent Office during prosecution of the '267 Patent, and both of which are duplicative of numerous references considered by the Patent Examiner. In short, this ground proposes to take two known processes (sub-pixel interpolation using higher-precision intermediate values and bi-prediction using two low-precision prediction results) and materially change both processes so that intermediate values from the first process (selected by Petitioner) are retained, rather than discarded, and are used instead of predictions in subsequent steps rather than the final interpolated prediction values conventionally used. This is the essence of hindsight reconstruction of the Challenged Claims. Neither reference suggests such a combination would be possible, much less desirable, and numerous references teaching both parts of Petitioner's proposed modified combination were considered during prosecution.

Because neither ground is new over art and arguments already considered by the Patent Office, and because neither ground establishes a reasonable likelihood that the Challenged Claims are unpatentable, institution should be denied.

## II. BACKGROUND

### A. Technical Background

The '267 Patent relates to an improvement in the processing of bi-prediction in video coding. As explained in more detail below, motion compensated prediction

in video coding is a computationally intensive aspect of video coding, requiring an encoder to generate and evaluate many possible predictions to obtain a prediction close enough to encode video at an acceptable bit rate. Motion prediction using bi-prediction uses a combination of two predictions, and so may require twice as much computation to generate and evaluate the individual predictions, as well as additional processing to determine which combination of predictions works well enough to use. Ex-2001, ¶¶ 46, 50-53.

### 1. Motion Prediction in Video Coding

#### a) Predictions, Motion Vectors, and Error

Video compression or video coding typically follows two main "phases" (1:34-65). In a first phase, estimates are formed that attempt to match a current picture or a part of a current picture, based on information already included in the video sequence. Ex-2001, ¶¶ 46-48. These estimates are called predictions. *Id.* There are two general categories of prediction in video coding: prediction based on information already coded for the same frame, called intra prediction, and prediction based on information coded in previous frames, called inter prediction. *Id.*, ¶ 48. Inter prediction typically comprises motion compensated prediction. *Id.*, ¶ 49.

In motion compensated prediction, the current frame or picture is estimated from frames or pictures that have already been encoded (at the encoder) or decoded (at the decoder). Ex-2001, ¶ 49. The encoder and decoder each maintain a buffer of

already coded pictures, called reference pictures, which are used in forming predictions. *Id*. These reference pictures are decoded versions of the pictures the encoder coded, rather than the original pictures in the sequence, as it is necessary for the encoder and decoder to construct identical predictions and therefore to use identical reference pictures. *Id*.

Predictions are generally performed for an entire block of pixels. Ex-2001, ¶¶ 47-48. For each current block of pixels to be encoded, the encoder determines a prediction block based on pixels in a reference picture, and identifies this prediction block using a data structure called a motion vector, which is an offset to the position of the prediction block (in the reference picture) relative to the current block (in the current picture). Ex-2001, ¶ 49. The motion vector may indicate an integer-valued offset, i.e. an offset to an actual region of pixels in the reference picture, or a fractional-valued offset, i.e. an offset to a fractional position between actual pixels in the reference picture. *Id.*, ¶ 52. In this second case, prediction block values are generated by interpolating between actual pixel values, so-called sub-pixel interpolation. *Id*.

After identifying a potential motion vector to use for a prediction, an encoder constructs the prediction using the motion vector and the indicated region of reference pixels and subtracts it from the block to be coded, resulting in a prediction error signal, also called a residual signal. Ex-2001, ¶ 55. The residual signal is coded

and transmitted, with the corresponding motion vector and other information necessary to decode the video sequence, in the coded video bitstream. *Id.*

The more similar the prediction block is to the current block, the smaller the residual signal is and the fewer bits that need to be included in the bitstream in order to encode the current block. Ex-2001, ¶ 56. Accurate predictions, meaning predictions that produce a small residual signal, are therefore very important for efficient video coding. *Id.*

As part of the corresponding decoding process, both the motion vector and the prediction error signal are decoded from the coded video bitstream. Ex-2001, ¶ 57. The decoder uses the motion vector to construct the identical prediction the encoder made, and adds the decoded error prediction signal to the prediction to reconstruct the original block. *Id.*

In typical implementations of motion compensated prediction, an encoder can use reference blocks from one reference picture for predictions, or it can use reference blocks from more than one reference picture. Ex-2001, ¶ 49. Pictures that may rely on more than one reference picture to construct each prediction block are called bi-predicted pictures, or B-pictures. Ex-2001, ¶ 51. Predicted pictures that can rely on only one reference picture to construct each prediction block are called P-pictures. *Id.* Pictures that do not rely on references to other pictures for predictions are called I-pictures. *Id.* The example sequence below shows each type:

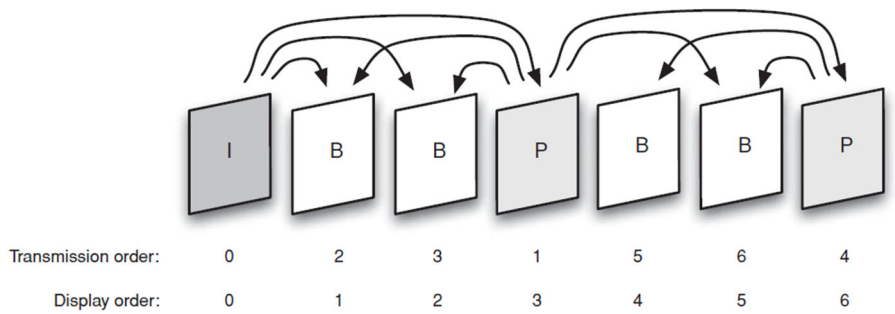



**Figure 5.5** Display order: Type 1 example

EX. 2012, book page 106.

**b) The Role of Encoder and Decoder in Motion Prediction**

While both the encoder and decoder create predictions from blocks of reference pixels based on motion vectors, their tasks differ significantly. Ex-2001, ¶ 58.

The encoder must find and select the prediction to use for a given block, and use the prediction to make the residual value. Ex-2001, ¶ 58. The process of finding an optimal or near-optimal prediction block to use for motion compensated prediction is called motion estimation. *Id.*

Motion estimation and motion compensated prediction has relatively high computation and storage costs at the encoder. Ex-2001, ¶ 59. The encoder must find a closely matching prediction for the block being coded. *Id.* The better the match, the less error data must be sent, and the more efficient the bit stream. *Id.* To do this, the encoder evaluates potential matches in a search area to find an effective match. *Id.* This is illustrated conceptually below.

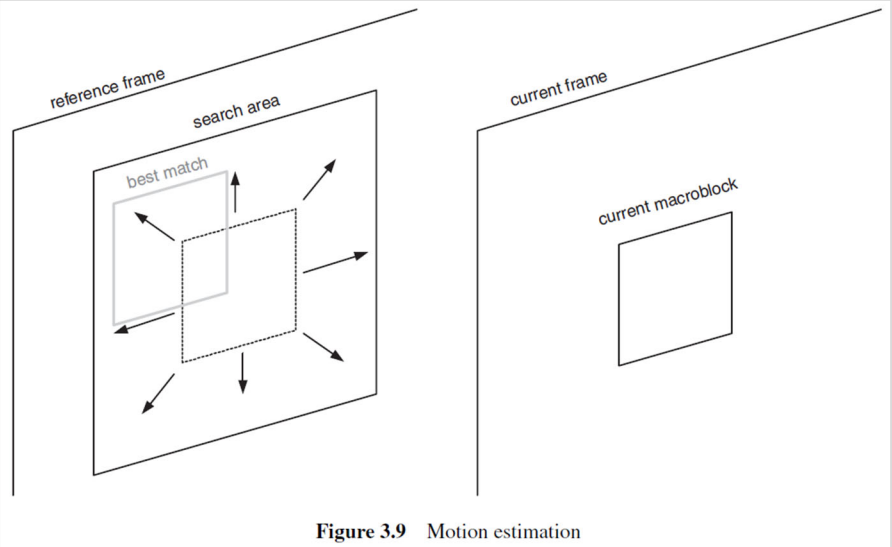



**Figure 3.9** Motion estimation

EX. 2012, Book page 32.

Each potential match in the search area that the encoder evaluates must be generated in order to be evaluated. Ex-2001, ¶ 60. Each potential match in the search space it is identified by a motion vector, which indicates the distance or offset from the upper left corner of the current block to a pixel position that forms the upper left corner of a matching block of pixels. *Id.* Motion vectors can either identify an actual pixel location or a location between the actual pixels. *Id.* Motion vectors identifying actual pixels in the reference picture are called "integer motion vectors" while motion vectors identifying locations between pixels in the reference picture are called "fractional" or "sub-pixel motion vectors." *Id.* When sub-pixel motion vectors are used, generating the potential matching prediction block requires interpolation. Interpolation typically requires a computationally complex series of operations such as multiplications, additions and divisions to arrive at a final value that can be used

as a prediction. *Id.* Generating potential predictions for sub-pixel positions is therefore computationally costly from a processing standpoint, i.e. such processes typically require significant computational resources. *Id.*

Additionally, each current block will generally only use one match from its corresponding search area in a reference frame, but a potential match may be in the search area for many current blocks, and may therefore be evaluated for potential use when carrying out motion estimation for multiple blocks in a current picture or frame. Ex-2001, ¶ 61. Re-interpolating the candidate matching block for each current macroblock can be inefficient, either causing the encoding process to take a long time or curtailing the number of potential matches that can be evaluated. Ex-2001, ¶ 62. As a result, a practical and often efficient strategy is to store interpolation results for later evaluation, for example by calculating all interpolated sub-pixel positions for an entire reference picture. *Id.*

The use of bi-prediction compounds the complexity of the search for potential predictions during motion estimation. Ex-2001, ¶ 62. The encoder must search at least two reference pictures for suitable predictions, match the candidate predictions from each reference picture with candidate predictions from other reference pictures to identify a suitable set of predictions to combine, and then evaluate the residual signal that would result from using that combined prediction to identify the best pair of predictions. *Id.* Video coding standards such as H.264 (Ex. 1014) allow for lists

of reference pictures, meaning that for each "side" of a bi-prediction, an encoder may need to evaluate the candidates in the search spaces of multiple reference pictures. *Id.* The construction and use of reference picture lists is illustrated below:

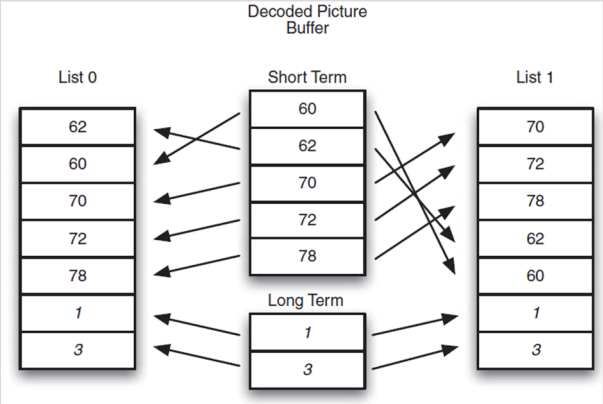



**Figure 5.6** List 0 and List 1 ordering: example

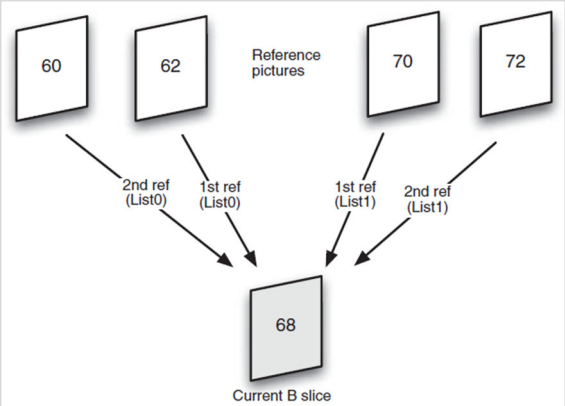



**Figure 5.7** Default reference pictures: example

EX. 2012, Book page 108.

Accordingly, an encoder would potentially need to evaluate the same prediction as a potential match for blocks in multiple pictures, and would potentially need to combine such predictions with potential matches from other reference frames. Ex-2001, ¶ 62. As a result, storing and saving predictions for potential reuse can have significant advantages over re-calculating predictions every time a particular prediction is evaluated. *Id.*

At the decoder, interpolation of predictions that were evaluated by the encoder as part of motion estimation but not selected is not necessary. Ex-2001, ¶ 63. As a result, the decoder can perform fewer interpolations than the encoder, so that motion compensated prediction at the decoder is typically computationally simpler than

motion estimation and motion compensated prediction at the encoder. *Id.* However, the decoder must perform the same set of calculations used to generate the prediction used by the encoder to generate the residual signal. *Id.* As a result, making the process for generating predictions more complex at the encoder will also make the decoder more complex, with corresponding implications for the cost and performance of decoder devices. *Id.*

**c) Precision in Video Coding**

An important parameter in video coding is the bit depth, or precision at which pixels are represented and at which calculations are performed. Ex-2001, ¶ 64. Pixel bit depth defines the range of possible values for the pixels in a video, such that more bit depth allows colors and contrasts in brightness to be more accurately represented. *Id.* For example, representing pixel values in 8-bit precision enables each pixel value (such as a luma or brightness value) to take any one of 256 values ($2^8$=256), for example in a range from 0-255. *Id.* Representing a pixel value with 10-bit precision enables it to take any one of 1024 values, for example in a range from 0-1023. *Id.* When displayed on a suitable display device, a 10-bit image may provide a more faithful representation of the original image, for example with a greater contrast between dark and light areas and more nuanced brightnesses in between. *Id.* In terms of precision and quality of the displayed image, a higher bit depth may therefore be advantageous. *Id.*

However, increasing the bit depth has a number of implications for computational and memory requirements. Moving from 8 bits to 10 bits of bit depth, for example, increases the amount of memory (e.g. the number of stored bits) required to store each sample by 25%. Ex-2001, ¶ 65. Encoding or decoding video requires multiple calculations to be carried out by a processor. *Id.* When processing each pixel sample within an encoder or decoder, all 10 bits need to be made available before a calculation can take place, which requires more connections (wider data buses) and/or more clock cycles when processing each calculation. *Id.* And inside the processor, any digital logic for adding, subtracting, multiplying, dividing and comparing sample values must accommodate two additional bits. *Id.* As a result, representing pixel values or other values at a higher precision in a video codec comes at a cost of increased computation, increased memory requirements and increased rate of access to stored values ("memory bandwidth"). *Id.* Computation, memory and memory bandwidth are at a premium in consumer electronic devices. *Id.*

Computation and memory requirements are an important consideration in the design of video encoders and decoders (Ex. 2012, book pages p223-229). In recognition of this, well-known standards for video coding such as H.264 impose certain limits on computation and memory through the use of Profiles and Levels. Ex-2001, ¶ 66. A Profile is a set of processes or algorithms for video compression that limits the algorithmic capabilities of a video decoder. *Id.* A Level places an

upper limit on the computational and storage requirements of a video decoder. *Id.* For example, H.264's Main Profile and Level 4.1 enables coding and decoding of High Definition video with a precision of 8 bits per luma pixel value. *Id.* H.264's High10 Profile extends this to a precision of 10 bits per luma pixel value. *Id.* Thus the question of how many bits are used to represent pixel values and other values is an important consideration in the specification of video codecs and the design of practical video encoders and decoders. *Id.*

**d) Interpolation**

In video coding, sub-pixel interpolation is used to generate predictions for motion compensated prediction using a set of weights applied to surrounding pixel values. Ex-2001, ¶ 67. Each full-pixel value that is involved in such a calculation, contributes some fractional part of the calculated interpolated value. *Id.* When applied to a block, interpolated values can create prediction blocks with better motion compensated prediction performance than the underlying blocks of full pixel values, due at least in part to the fact that objects in a video scene do not necessarily move by integer-pixel increments between successive video frames. *Id.*

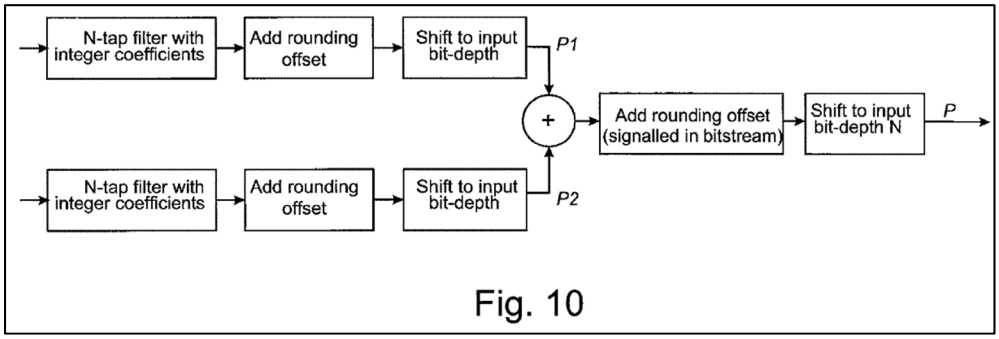
Interpolation in video coding is generally accomplished by an application of integer weights to underlying pixel values followed by a sum of the resulting weighted values, followed by a division by the sum of the weights. Ex-2001, ¶ 68. Interpolation is typically performed in this manner because multiplying by integers

is relatively computationally simple in commonly-available digital hardware and division is relatively computationally expensive. *Id.* Further, choosing filter weights that add to a power of 2 allows the division to be carried out by a right-shift, which is a computationally simple operation in digital logic. *Id.* Before the division or right-shift, such an interpolation method produces an intermediate value at higher precision than the underlying pixel values. *Id.* The H.264 video coding standard specifies a multi-stage interpolation process where certain higher-precision intermediate values are temporarily used in order to interpolate further fractional positions using the intermediate values as inputs. *Id.* However, all the outputs of such an interpolation process are at the same precision as the original input pixels, so that the output values (the interpolated values that may form a prediction) can be used in motion compensated prediction. *Id.*

### B. The ’267 Patent

The Inventors of the ’267 Patent sought to improve bi-prediction in video coding by changing the basis on which predictions from two motion vectors are combined during bi-prediction. Bi-prediction uses prediction values from two different sources, such as prediction blocks formed from two reference frames, and combines these predictions to create a single prediction block. Ex-2001, ¶ 50. In the prior art, bi-prediction used prediction values from each of two reference frames with the same bit depth, or precision, as the pixels of the current block or the

reference blocks. This is true both of the prior art contemplated in the patent itself (Ex-1001 at 3:56-4:25) and of the prior art presented in the Petition. Figure 10 from the ’267 Patent illustrates this approach, with predictions P1 and P2 generated by applying interpolation filters and then returning to the input bit depth before calculating bi-prediction P.




Ex-1001 Fig. 10.

The inventors of the ’267 Patent recognized that this potentially contributed to compounding of rounding error, where sub-pixel interpolation was used, and that the existing solutions of using rounding offsets to counteract rounding error increased complexity at the encoder and required additional signaling, harming efficiency. To resolve this issue, the inventors of the ’267 Patent proposed using higher precision interpolated values as predictions in bi-prediction, and allowing rounding error to accumulate in less significant bits and then be discarded when final values were obtained. Ex-1001 at 4:29-35. Rather than reduce interpolated values to input pixel precision, all predictions are made at the same, higher precision,

requiring some predictions to be made by shifting integer values to the left and other predictions to be made by shifting interpolated values to the right, such that all predictions have the same precision when combined. Ex-1001 at 14:4-10.

### C. The ’267 Patent’s Prosecution History

The ’267 Patent claims priority, through five utility patent applications, to a provisional patent application filed in January 2011. The first utility application (App. No. 13/344,893, Ex. 1007) was subject to multiple Office Actions and responses before being allowed. The four subsequent intervening applications were allowed on the first Office Action, or after an obviousness type double-patenting rejection. During the prosecution of each application, the same art previously considered was cited. The application that issued as the ’267 Patent followed a similar pattern, but additional art was submitted to the Patent Office. The same Examiner and Supervisory Examiner allowed each patent in the family, and the reasons for allowance for each refer back to the original ’893 Application (Ex. 1007). Accordingly, Patent Owner addresses both the file history of the ’267 Patent as well as its related priority applications below.

#### 1. Priority Applications

##### a) Application No. 13/344,893

The originally filed claims in the ’893 Application had many elements in common with the Challenged Claims. *Compare* Ex-1007 at 36 with Ex-1001 claim

1. Both recite a block having a first precision, both recite obtaining a first and second prediction with a higher, second precision, using reference pixel locations, and both recite obtaining a combined prediction using the first and second predictions.

The original claims were rejected over a combination of Ye (US 2013/0142262, Ex-2009) ("Ye '262") and Noda (US 2009/0087111). Ex-1007 at 146. The Examiner wrote that Ye '262 satisfied the requirement that the first and second prediction of the claims have a higher, second precision because the predictions P0(x,y) and P1(x,y) in Ye '262 could be the result of fractional motion vectors, which allow for more precise motion tracking and prediction compared with integer pixel precision. Ex-1007 at 147-148. Patent Owner initially argued that Ye '262 failed to teach the second precision higher than the first precision because the interpolation of the fractional motion vectors is performed on a reference picture, and Ye '262 did not expressly provide the precision of the pixels of the reference picture. Ex-1007 at 188-9. In the subsequent Final Rejection, the Examiner disagreed and re-iterated his Ye '262 based rejection. Ex-1007 at 202-4. Because the predictions P0(x,y) and P1(x,y) were obtained using fractional pixel locations, and fractional motion vectors allow for more accurate motion prediction than integer motion vectors, the Examiner concluded that P0(x,y) and P1(x,y) had the second, higher precision. *Id*.

In its next response, Patent Owner amended the claims to recite that the first precision indicates the number of bits needed to present values of said pixels, and the second precision indicates the number of bits needed to represent values of said first prediction and said second prediction. Ex-1007 at 236. Patent Owner then explained that Ye '262's disclosure did not teach or suggest that the number of bits needed to represent the predictions in Ye '262 was greater than the number of bits needed to represent the values of pixels in the current block. Ex-1007 at 242-243. In other words, the first and second predictions in Ye (P0(x,y) and P1(x,y)) had the same precision at the current block, and so did not have the second precision.

In the following Office Action, the Examiner acknowledged that Ye '262's teaching of weighted predictions ($w*P_0(x,y)$ and $(W-w)*P_1(x,y)$) did not teach a prediction having a second precision, which is higher than said first precision, wherein the second precision indicates the number of bits needed to represent values of said first prediction and values of said second prediction. Ex-1007 at 251-252.

> Ye does not disclose explicitly the following claim limitations (emphasis added):
>
> using said first reference pixel location to obtain a first prediction, said first prediction having a second precision, which is <u>higher than said first precision</u>, wherein the second precision indicates <u>the number of bits needed to represent values of said first prediction and values of said second prediction</u>;

Ex-1007 at 252.

The Examiner then relied on Noda, which used increased bit depth reference frames, to meet this limitation. Ex-1007 at 252. In Noda, the bit depth of each pixel in a reference frame was increased from N to N+M bits, and then predictions were generated at N+M bits. Ex-1007 at 252.

In response, Patent Owner amended the claims to require the reference blocks used to obtain the first and second prediction have the same precision (the first precision) as the current block. Ex-1007 at 272. This excluded Noda, because in Noda the reference blocks used to generate predictions had the same N+M precision as the predictions.

In the remarks, the Patent Owner argued that Ye '262's application of weights to predictions $P_0(x,y)$ or $P_1(x,y)$ did not increase the precision of $P_0(x,y)$ or $P_1(x,y)$, but simply changed their range of values at the existing precision by mapping the original precision values to a new range of values. Ex-1007 at 281.

The claims of the '893 Application were subsequently allowed. Ex-1007 at 294-311. The Primary Examiner on the case was Peter Le, and his Supervisory Examiner was Sath Perungavoor. Ex-1007 at 310-311.

**b) Subsequent Intervening Priority Applications**

Shortly before the '893 Application issued, Patent Owner filed a continuation application, No. 15/250,124. Ex-2003. The Examiner requested a terminal disclaimer over the parent case (the 893 Application), and following entry of a

terminal disclaimer the ’124 Application was allowed without rejection. Ex-2003 at 20-30. The ’124 Application was allowed by the same Examiner who allowed the ’893 Application.

Shortly before the ’124 Application issued, Patent Owner filed a continuation application, No. 15/490,469. Ex-2004. The Examiner allowed the ’469 Application on the first application following the filing of a terminal disclaimer over the ’893 and ’124 Applications. Ex-2004 at 14-24. The ’469 Application was allowed by the same Examiner who allowed the ’893 and ’124 Applications.

Shortly before the ’469 Application issued, Patent Owner filed a continuation application, No. 15/876,495. Ex-2005. The ’495 Application was subsequently rejected for obviousness type double patenting over the claims issued from the ’893 and ’469 Applications. The Examiner subsequently allowed the ’495 Application following the filing of a terminal disclaimer over the ’893, ’124, and ’469 Applications. Ex-2005 at 6-13. The ’495 Application was allowed by the same Examiner who allowed the ’893, ’124, and ’469 Applications.

Shortly before the ’495 Application issued, Patent Owner filed a continuation application, No. 16/729,974. Ex-2006. The Examiner allowed the ’974 Application on the first application following the filing of a terminal disclaimer over the 495 Application. Ex-2006 at 7-14. The ’974 Application was allowed by the same Examiner who allowed the ’893, ’124, ’469, and ’495 Applications.

### 2. Application No. 17/328,750

Shortly before the ’974 Application issued, Patent Owner filed a continuation application, No. 17/328,750. Ex-1002. The ’750 Application ultimately issued as the ’267 Patent. Ex-1002 at 369. During prosecution, the ’750 Application was initially rejected for obviousness type double patenting over the claims issued from the ’974 and ’495 Applications. Ex-1002 at 123-130. The ’750 Application was allowed following Patent Owner’s submission of a terminal disclaimer. Ex-1002 at 181-188. After allowance but before issuance, Patent Owner filed an RCE, requesting consideration of additional prior art references. Ex-1002 at 205-316. The ’750 Application was subsequently allowed again. Ex-1002 at 323-330. Following payment of the issue fee, Patent Owner received an Office Action for a related European application and petitioned the Patent Office to consider that European action as well before allowing the ’750 application to issue. Ex-1002 at 352-355. The ’750 Application was allowed over this additional information as well, and the ’267 Patent issued on October 31, 2023. Ex-1002 at 364-365, 369. The ’750 Application was allowed by the same Examiner who allowed the ’893, ’124, ’469, ’495, and ’974 Applications.

### 3. Prior Art Cited in Prosecution

The chart below summarizes the art of record during the prosecution of the ’267 Patent as well as its priority applications that is relevant to the Petition.

Additional art was also raised and considered, but the art below is either the same art raised by the Petition is not meaningfully different from the art raised in the Petition.

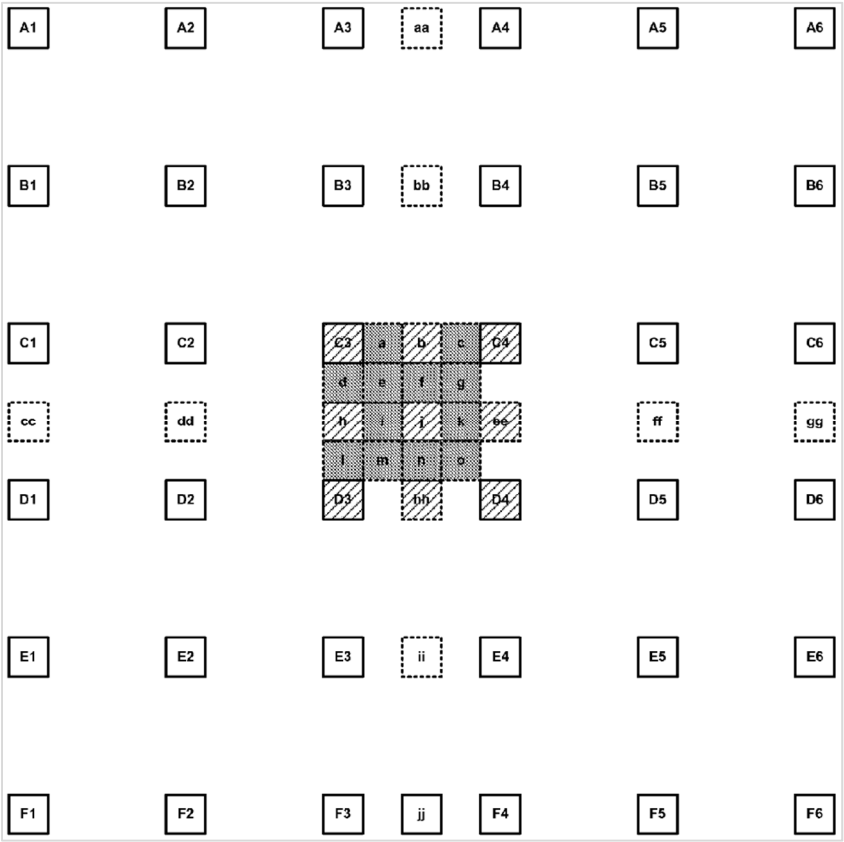
| Reference | Inventors / Authors | ’893 FH | ’124 FH | ’469 FH | ’495 FH | ’974 FH | ’750 FH |
|---|---|---|---|---|---|---|---|
| US 2009/0257503 “Ye ’503” Ex-2007 | Yan Ye, Peisong Chen, Marta Karczewicz | ¥[1] | ¥ | ¥ | ¥ | ¥ | ¥ |
| US 2010/0086027 “Panchal” Ex-2008 | Rahul Panchal, Marta Karczewicz | † | ¥ | ¥ | ¥ | ¥ | ¥ |
| US 2013/0142262 “Ye ’262” Ex-2009 | Yan Ye, Alexandros Tourapis | †‡ | ¥ | ¥ | ¥ | ¥ | ¥ |
| VCEG-AI33 Ex-2010 | Yan Ye, Peisong Chen, Marta Karczewicz | ¥ | ¥ | ¥ | ¥ | ¥ | ¥ |
| U.S. Pat. 9,161,057 “Karczewicz ’057” Ex-2011 | Marta Karczewicz, Peisong Chen, Yan Ye | | | | | | ¥ |
| US 2009/0257499 (Karczewicz-II from petition) Ex-1006 | Marta Karczewicz, Yan Ye, Peisong Chen | | | | | | ¥ |

**a) Ye ’503 (Ex-2007)**

Ye ’503 was cited by Patent Owner to the Examiner during the prosecution of the original application in the family. Ex-1007 at 88. Ye ’503 was also cited by the

[1] Key: † Cited by examiner ‡ Discussed in office actions ¥ Cited by applicant

Patent Owner in each subsequent family member, including the ’750 Application, and is cited on the face of the ’267 Patent. Ex-1002 at 74; Ex-1001 at 2. Ye ’503 teaches the use of an alternative interpolation scheme to that adopted for H.264 which “keep[s] the highest possible precision through intermediate steps” of interpolation by using intermediate values (b1, h1, and j1) to interpolate quarter-pixel subpixel values. Ex-2007 at ¶¶ 98-99. This disclosure is substantively identical to the disclosure Petitioner relies upon from Karczewicz-II, which will be addressed in greater detail below. Ye ’503 also teaches the register mathematics used to accomplish this interpolation of intermediate values without exceeding the 16-bit register limit. Ex-2007 at ¶¶ 100-102. The figure below illustrates both the quarter-pixel positions interpolated (dark grey) and the positions corresponding to the values or intermediate values used in their interpolation.




Ex-2007 at 8.

**b) Panchal (Ex-2008)**

Panchal was cited by the Examiner during the prosecution of the original application in the family and used, in combination with Ye ’262 and Noda, as part of an obviousness rejection. Ex-1007 at 152-156. Panchal was also cited by the Patent Owner in each subsequent family member, including the ’750 Application, and is cited on the face of the ’267 Patent. Ex-1002 at 74; Ex-1001 at 2. Panchal teaches various forms of bi-prediction, as well as strategies for choosing between them, including whether to use default or explicit weights or to signal an offset as

part of the bi-prediction. Ex-2008 at Figs. 4, 5. As part of its disclosure, Panchal teaches default weighted bi-prediction (Equation 1, ¶ 60), weighted bi-prediction with explicit weight and multiple rounding offsets (Equation 2, ¶ 61), and weighted bi-prediction with implicit weights and a fixed rounding offset (Equation 3, ¶ 62).

**c) Ye ’262 (Ex-2009)**

As addressed in the summary of the prosecution history of the ’893 Application above, Ye ’262 was a central focus of the Examiner during the prosecution of the original application in the family, and was the central reference in numerous obviousness rejections. Ye ’262 was also cited by the Patent Owner in each subsequent family member, including the ’750 Application, and is cited on the face of the ’267 Patent. Ex-1002 at 74; Ex-1001 at 2. Ye ’262 teaches block-based motion prediction, including weighted bi-prediction, with predictions from two different reference frames based on motion vectors. Ex-2009 at ¶¶ 24-25. Ye ’262 further teaches that these predictions can made from motion vectors with fractional-pixel precision, and be interpolated from pixel values using a 6-tap filter. Ex-2009 at ¶ 28.

**d) VCEG-AI33 (Ex-2010)**

VCEG-AI33 was cited by Patent Owner to the Examiner during the prosecution of the original application in the family. Ex-1007 at 327. VCEG-AI33 was also cited by the Patent Owner in each subsequent family member, including

the ’750 Application, and is cited on the face of the ’267 Patent. Ex-1002 at 78; Ex-1001 at 3. VCEG-AI33 is a proposal to “increas[e] precision of interpolation and bi-directional prediction during motion compensated prediction” by using intermediate values for half-pixel positions to interpolate values for quarter-pixel positions. Ex-2010 at 1. The disclosed interpolation method from VCEG-AI33 uses the same equations found in Ye ’503 and the Karczewicz-II reference relied upon by the petition. VCEG-AI33 also has the same authors as Ye ’503 and Karczewicz-II.

**e) Karczewicz ’057 (Ex-2011)**

Karczewicz ’057 was cited to the Examiner during the prosecution of the ’750 application, which ultimately issued as the ’267 Patent. Ex-1007 at 75. Karczewicz ’057 is listed as considered by the Examiner on the face of the ’267 Patent. Ex-1001 at 2. Similar to Panchal, Karczewicz ’057 teaches multiple forms of bi-prediction, including default weighted bi-prediction, explicit weighted bi-prediction, and implicit weighted bi-prediction, as well as strategies for selecting between them. Ex-2011 at 1 (abstract), 10 (Fig. 7), and 16-17 (12:11-14:65). Karczewicz ’057 also incorporates VCEG-AI33 by reference as a document related to the techniques described in Karczewicz ’057. Ex-2011 at 14 (8:5-12).

**f) Karczewicz-II (Ex-1006)**

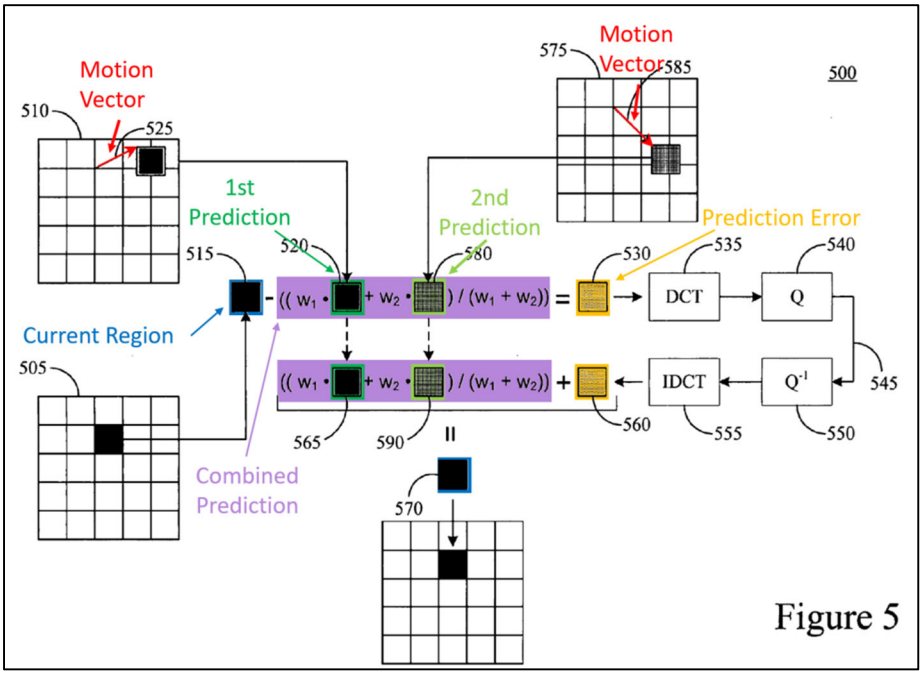
Karczewicz-II was cited to the Examiner during the prosecution of the ’750 application, which ultimately issued as the ’267 Patent. Ex-1007 at 74. Karczewicz-

II is listed as considered by the Examiner on the face of the ’267 Patent. Ex-1001 at 2. The substance of Karczewicz-II will be addressed below in connection with the art raised in the Petition.

### D. The Asserted Prior Art

#### 1. Walker (Ex-1004)

Walker is a weighted prediction reference, and the teachings of Walker relied upon in the Petition are similar to the teachings of Ye ’262, which was addressed above. Specifically, Walker teaches the use of 8-bit predictions in combination with weighting factors applied as part of a bi-prediction process. This is illustrated in the annotated Walker Figure 5, reproduced below.



### 2. Karczewicz-I (Ex-1005)

Karczewicz-I (Ex-1005) teaches bi-prediction both with and without weights, as well as several different techniques for determining weights. Ex-1005 at ¶¶ 60-62. Karczewicz-I also teaches that rounding error, which can accumulate in bi-prediction, can be resolved by using rounding offsets, and presents several strategies for doing so. Ex-1005 at ¶¶ 64-69.

### 3. Karczewicz-II (Ex-1006)

Karczewicz-II (Ex-1006) teaches equations for obtaining 8-bit interpolated values using higher precision intermediate values temporarily stored in 16-bit processor registers. (Ex-1006) ¶¶ 99-106.

## E. POSITA

Patent Owner has applied Petitioner's POSITA for the sake of this Response. Pet. 11. This Response does not depend on this skill level. Patent Owner reserves the right to argue in favor of different level of skill should trial be instituted.

## F. Claim Construction

Petitioner takes a position on the meaning of "precision." *See* Pet. 12. This Response does not depend on this definition so Patent Owner does not address it. Patent Owner reserves the right to address this position should trial instituted.

### G. Claim Listing

For the sake of this analysis, Patent Owner will use Petitioner's limitation labels provided in the petition's claim appendix (Pet. 89-94).

## III. ARGUMENT

### A. The Petition Should Be Denied Under *Advanced Bionics* Because Cumulative References Have Been Repeatedly Argued to The Board and Patent Office and Petitioner Has Not Demonstrated Either Body Erred.

The '267 Patent and its family members have been prosecuted against numerous references teaching bi-prediction, weighted bi-prediction, interpolation, and, in some cases, all three. The Grounds raised by the petition are cumulative of art and arguments considered multiple times, in the context of multiple similar and overlapping references, over the decade of prosecution activity between the filing of the original application and the issuance of the Challenged Claims. Amazingly, over this period, the '267 Patent and each of its priority patents were reviewed and allowed by the same Examiners. There is a strong presumption, then, that the Patent Office considered and approved the Challenged Claims over the art and arguments raised in the Petition, and in the absence of some showing of error in that process institution should be denied.

Under § 325(d), the Board uses a two-part analysis, considering: (1) whether the same or substantially the same art or arguments were presented to the PTO during

original prosecution; and (2) whether the Petitioner has demonstrated that the PTO erred in a manner material to the patentability of Challenged Claims. *Advanced Bionics, LLC v. Med-El Elektromedizinishe Geräte GMBH*, IPR2019-01469, Paper 8 at 10 (PTAB Feb. 13, 2020) ("*Advanced Bionics*") (precedential).

Once the Board determines that the same or substantially the same art or arguments were previously presented to the PTO, it then considers whether the Petitioner has demonstrated error. *Advanced Bionics* at 8-10. "If a condition in the first part of the framework is satisfied and the petitioner fails to make a showing of material error, the Director generally will exercise discretion not to institute *inter partes* review. If reasonable minds can disagree regarding the purported treatment of the art or arguments, it cannot be said that the Office erred in a manner material to patentability. At bottom, this framework reflects a commitment to defer to previous Office evaluations of the evidence of record unless material error is shown." *Id*.

Here, Petitioner has asserted the Walker reference (Ex. 1004) in support of an argument that weighted prediction's application of weights to two low-precision prediction creates two high-precision predictions, which are then combined. The central reference during the prosecution of the '893 Application, the first application of the chain, included essentially the same weighted bi-prediction approach, and Patent Owner secured allowance in that case by arguing that the application of a

weight to a low-precision prediction does not make a high-precision prediction, as argued by the Petition. See Section II.C, *supra*.

For its second Ground, Petitioner asserts it would have been obvious to combine two references, one teaching bi-prediction and one teaching interpolation with higher-precision intermediate values. Both of these concepts were taught by multiple references cited throughout the decade of prosecution of the ’267 patent and its priority applications, including multiple references by the same inventors as the asserted references including the same disclosures relied upon by the Petition in its motivation to modify and combine these references.

The Petition briefly acknowledges that it uses a reference already considered in the prosecution of the Challenged Claims and then ignores that the disclosures it relies upon from the other two references were found in multiple other references cited throughout the prosecution of the ’267 Patent and its family members. Petition at 86.

#### 1. Petitioner’s art is substantially the same as the art cited in the prosecution of the ’267 Patent.

The Grounds presented in the Petition present two arguments: that weighted prediction satisfies the claim (Ground 1) and that it would have been obvious to modify sub-pixel interpolation and bi-prediction to achieve the patented invention (Ground 2). However, the prosecution history of the ’267 Patent shows the Examiner

was well aware of and considered both weighted bi-prediction and sub-pixel interpolation when he allowed the Challenged Claims. The art considered during the prosecution of the Challenged Claims includes the interpolation reference actually relied on in the Petition, several similar references teaching the same interpolation methods, and several references teaching the same bi-prediction and weighted bi-prediction methods in the Petition's relied-upon bi-prediction references.

The Board should find these references are substantially the same as the references the Patent Office has previously reviewed and deny institution. *See e.g.*, *Evapco Dry Cooling, Inc., v. SPG Dry Cooling USA*, IPR2021-00687, Paper 11 at 24-25 (P.T.A.B. Sep. 24, 2021) (finding a newly asserted reference not cited during prosecution as being substantially the same as the references that were before the Examiner during prosecution); *see Gardner Denver v. Utex Industries, Inc.*, IPR2020-00333, Paper 12 at 13-14 (P.T.A.B. Aug. 5, 2020) (denying institution upon finding that newly asserted references Kalsi and Kohl are no more relevant than references that were presented during prosecution); *Roku, Inc. v. Universal Elecs., Inc.,* IPR2019-01619, Paper 11, 16-17 (P.T.A.B. April 2, 2020) (denying institution upon finding that newly applied Zetts was cumulative to previously presented Kim, even though Kim was not considered for the same limitation in reexamination).

**a) Petitioner has failed to show that the Walker reference in the Petition differs in any material way from the Ye '262, Panchal, or Karczewicz 057, references cited in prosecution and on the face of the '267 Patent.**

Petitioner relies on Walker for its disclosure of weighted prediction, as illustrated in Walker Figure 5, equation 13, and Table 2. Pet. 17-25. Each is a different representation of the same weighted bi-prediction process that was already considered and overcome by Patent Owner during prosecution.

According to Petitioner, Walker teaches a first and second prediction at a second precision, higher than a first precision, because in Walker the 8-bit predictions derived from individual reference pictures, pred0 and pred1, are multiplied by corresponding 8-bit weights, w0 and w1, before the resulting products are added together. Petition at 18-21.

Although Patent Owner disputes that this arrangement satisfies the claims, multiple references considered during prosecution taught the same weighted prediction scheme Petitioner relies upon in Walker. Ex-2001, ¶ 73.

For example, the Ye '262 reference that formed the basis for multiple rejections during prosecution of the priority '893 Application taught a nearly identical equation for bi-prediction as the one relied on in the Petition. Ex-2001, ¶ 74. Walker Equation 13 is reproduced below at left next to Ye '262's equation from its paragraph 25.

| Walker (Ex-1004) ¶ 68 | Ye '262 (Ex-2009) ¶ 25 |
|---|---|
| $Final_pred = Clip1[(((pred0)w0 + (pred1)w1 + 2^{\log WD}) >> (\log WD + 1)) + ((o_0 + o_1 + 1) >> 1)]$ | $P(x,y) = (w \cdot P_0(x,y) + (W - w) \cdot P_1(x,y) + W/2) >> S$ |

Compared to Ye '262, Walker adds a Clip1, which confines the Final_pred value to a range between 0 and 255 if the initial result was outside of that range, but otherwise does nothing more of substance to the bi-prediction process. Ex-1004 ¶ 66; Ex-2001, ¶ 75. The $((o_0 + o_1 + 1)>>1)$ term in Walker's equation also adds 1 if offsets $o_0$ and $o_1$ are set to 1, but this addition in Walker is not cited in the Petition to satisfy any claim language. Ex-2001, ¶ 76.

Both equations multiply a first prediction (pred0 in Walker, $P_0(x,y)$ in Ye '262) by a weight (w0 in Walker, w in Ye '262). Ex-2001, ¶ 77. Both equations multiply a second prediction (pred1 in Walker, $P_1(x,y)$ in Ye '262) by a weight (w1 in Walker, W-w in Ye '262). *Id.* Both equations add an offset for rounding ($2^{\log WD}$ in Walker, W/2 in Ye '262). *Id.* Both equations then right-shift (>>(logWD+1) in Walker, >>S in Ye '262). *Id.* And, as W = 1<<S (or $2^S$), the rounding offset values and right-shift amounts are the same when S = logWD+1. *Id.* Accordingly, there are no material differences between the weighted bi-prediction teaching in Walker relied upon in the Petition and the weighted bi-prediction teaching in Ye '262 what was

central to the prosecution of the ’267 Patent’s original priority application. Ex-2001, ¶ 78.

As a further example, Panchal, cited by the Examiner in the priority ’893 Application prosecution and then cited by Patent Owner in every subsequent application, includes the same offsets and the two-stage right-shift found in Walker:

| Walker (Ex-1004) ¶ 68 | Panchal (Ex-2008) ¶ 61 |
|---|---|
| $Final_pred=Clip1[(((pred0)w0+(pred1)w1+2^{\log WD})>>(\log WD+1))+((o_0+o_1+1)>>1)]$ | $\mathrm{pred}(i,j)=(\mathrm{pred0}(i,j)*w0+\mathrm{pred1}(i,j)*w1+2^r)>>(r+1)+((o1+o2+1)>>1)$ |

Ex-2001, ¶ 79.

Although the exact manner in which offsets are rounded and right-shifts are applied does not seem to matter to Petitioner’s theory, even if it did Panchal teaches the same bi-prediction equation taught by Walker, and was cited by the Examiner and used in a rejection during the same prosecution where the Examiner was asserting that the weighted prediction scheme in Ye ’262 taught the claimed predictions at a second precision, greater than the precision of the reference block. Ex-2001, ¶ 80. The only difference between the equations in Walker and Panchal is that the final value in Walker includes a clip function. As noted above, this clip function only serves to keep values within the allowed 0 to 255 bounds, and is not material to any claim limitation at issue. *Id.*

Karczewicz ’057, cited to the Examiner during the prosecution of the application that issued as the ’267 Patent also teaches the same equation for bi-prediction:

| Walker (Ex-1004) ¶ 68 | Karczewicz ’057 (Ex-2011) at 13:11-12 |
|---|---|
| $Final_pred = Clip1[(((pred0)w0+(pred1)w1+2^{\log WD})>>(\log WD+1))+((o_0+o_1+1)>>1)]$ | Bidirectional prediction:pred$(i,j)$=(pred0$(i,j)$*$w0$+pred1$(i,j)$*$w1$+$2^r$)>>$(r+1)$+$((o1+o2+1)>>1)$ |

Ex-2001, ¶ 81.

Karczewicz ’057, like Panchal, teaches the same weighted bi-prediction equation as Walker. Ex-2001, ¶ 82.

Accordingly, the relied-upon disclosure from Walker was included in at least three different references cited to the Patent Office during prosecution of the ’267 Patent, including a reference used as the basis for a rejection for one of the ’267 Patent’s priority applications. Ex-2001, ¶ 83. Walker is nothing more than a cumulative reference of Ye ’262, Panchal, and Karczewicz ’057, and other weighted bi-prediction references that have been considered, and rejected by, the Patent Office when it allowed the Challenged Claims . *See e.g.*, *Evapco Dry Cooling, Inc., v. SPG Dry Cooling USA*, IPR2021-00687, Paper 11 at 24-25 (P.T.A.B. Sep. 24, 2021) (finding a newly asserted reference not cited during prosecution as being substantially the same as the references that were before the Examiner during prosecution); *see Gardner Denver v. Utex Industries, Inc.*, IPR2020-00333, Paper

12 at 13-14 (P.T.A.B. Aug. 5, 2020) (denying institution upon finding that newly asserted references Kalsi and Kohl are no more relevant than references that were presented during prosecution).

**b) Petitioner has failed to show that the Karczewicz-II reference cited in the Petition differs in any material way from the Ye '503 or VCEG-AI33 references cited in prosecution and on the face of the '267 Patent.**

Petitioner relies on Karczewicz-II for its disclosure of sub-pixel interpolation using intermediate values with higher precision, as illustrated in Karczewicz-II tables 3, 4, 5, 6, and 8. These tables include the equations for Karczewicz-II's higher precision interpolation method as well as an illustration for how to implement the method without exceeding the limits of 16 bit registers. Ex-2001, ¶ 84. Petitioner seeks to modify Karczewicz-II by stopping its calculations at an intermediate value and using that intermediate value at as an input to the bi-prediction equations of Karczewicz-I, which also must be modified to accept the new higher-precision values. Pet. 38-53, 59-68. Petitioner argues this modification is motivated by the statement in paragraph 99 of Karczewicz-II to "keep the highest possible precision through the intermediate steps" of the interpolation process. Petition at 42.

Karczewicz-II, however, was before the Patent Office and was considered during the prosecution of the '267 Patent. Petitioner argues that this should be ignored because Karczewicz-II was not expressly addressed by the Examiner in an

Office Action (Petition at 86), however it is not a requirement that a prior art reference be discussed for *Advanced Bionics* to apply, particularly when the teachings of the reference is cumulative to other art considered and rejected during prosecution. Rather, whether and the extent to which a reference was discussed during prosecution is one of the *Becton, Dickinson* factors considered in applying discretion under 325(d) under *Advanced Bionics*. *Advanced Bionics* at p. 9, n 10. Arguments concerning how the patent office treated a reference falls under *Becton, Dickinson* factor (c), "the extent to which the the asserted art was evaluated during examination, including whether the prior art was the basis for rejection." *Id*. Under the two-step framework in Advanced Bionics, Becton, Dickinson factor (c) is considered at the second step, where the petitioner must show the patent office erred in its evaluation of the reference. *Id*. at 10. The Petition does not even argue that the patent office erred here.

Further, it is not surprising that the Examiner did not discuss Karczewicz-II in an Office Action after reviewing it – Karczewicz-II contains substantially the same disclosure as other references the Examiner had already reviewed multiple times in connection with the '267 Patent and its priority applications. Ex-2001, ¶ 85.

Ye '503, which was cited by the Patent Owner during the prosecution of the '267 Patent and each of its five priority applications includes the **same** interpolation equations and tables relied upon by the Petition as well as the same statements relied

upon in the Petition to provide the motivation to combine. Ex-2001, ¶ 86. As can be seen below, the tables in Ye ’503 and Karczewicz-II match exactly:

| Karczewicz-II (Ex-1006) | Ye ’503 (Ex-2007) |
|---|---|
| TABLE 3<br>a = (C3 << 5 + b1 + 32) >> 6<br>c = (C4 << 5 + b1 + 32) >> 6<br>d = (C3 << 5 + h1 + 32) >> 6<br>l = (D3 << 5 + h1 + 32) >> 6<br>f = (j1 >> 5 + b1 + 32) >> 6<br>i = (j1 >> 5 + h1 + 32) >> 6<br>k = (j1 >> 5 + ee1 + 32) >> 6<br>n = (j1 >> 5 + hh1 + 32) >> 6 | TABLE 3<br>a = (C3 << 5 + b1 + 32) >> 6<br>c = (C4 << 5 + b1 + 32) >> 6<br>d = (C3 << 5 + h1 + 32) >> 6<br>l = (D3 << 5 + h1 + 32) >> 6<br>f = (j1 >> 5 + b1 + 32) >> 6<br>i = (j1 >> 5 + h1 + 32) >> 6<br>k = (j1 >> 5 + ee1 + 32) >> 6<br>n = (j1 >> 5 + hh1 + 32) >> 6 |
| TABLE 4<br>e = (b1 + h1 + 32) >> 6<br>g = (b1 + ee1 + 32) >> 6<br>m = (h1 + hh1 + 32) >> 6<br>o = (ee1 + hh1 + 32) >> 6 | TABLE 4<br>e = (b1 + h1 + 32) >> 6<br>g = (b1 + ee1 + 32) >> 6<br>m = (h1 + hh1 + 32) >> 6<br>o = (ee1 + hh1 + 32) >> 6 |

## Karczewicz-II (Ex-1006)

TABLE 5

positions {a, c, d, l} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = x | r1 is integer pixel x | 0 | 255 | 8u |
| r1 = r1 << 5 | r1 is 32 * x | 0 | 8160 | 13u |
| r2 = y0 | r2 is y0 (y0 is a one-dimensional (1-D) half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is 32 * x + y0 | −2550 | 18870 | 16s |
| r1 = r1 + 32 | r1 is 32 * x + y0 + 32 | −2518 | 18902 | 16s |
| r1 = r1 >> 6 | r1 is (32 * x + y0 + 32) >> 6 | −39 | 295 | 11s |
| r1 = max (0, r1) | clip r1 on the low side | 0 | 295 | 10u |
| r1 = min (255, r1) | clip r1 on the high side | 0 | 255 | 8u |

TABLE 6

positions {e, g, m, o} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = y0 | r1 is y0 (y0 is a 1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r2 = y1 | r2 is y1 (y1 is a 1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is y0 + y1 | −5100 | 21420 | 16s |
| r1 = r1 + 32 | r1 is y0 + y1 + 32 | −5068 | 21452 | 16s |
| r1 = r1 >> 6 | r1 is (y0 + y1 + 32) >> 6 | −79 | 335 | 11s |
| r1 = max (0, r1) | clip r1 on the low side | 0 | 335 | 10u |
| r1 = min (255, r1) | clip r1 on the high side | 0 | 255 | 8u |

## Ye '503 (Ex-2007)

TABLE 5

positions {a, c, d, l} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = x | r1 is integer pixel x | 0 | 255 | 8u |
| r1 = r1 << 5 | r1 is 32 * x | 0 | 8160 | 13u |
| r2 = y0 | r2 is y0 (y0 is a one-dimensional (1-D) half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is 32 * x + y0 | −2550 | 18870 | 16s |
| r1 = r1 + 32 | r1 is 32 * x + y0 + 32 | −2518 | 18902 | 16s |
| r1 = r1 >> 6 | r1 is (32 * x + y0 + 32) >> 6 | −39 | 295 | 11s |
| r1 = max(0, r1) | clip r1 on the low side | 0 | 295 | 10u |
| r1 = min(255, r1) | clip r1 on the high side | 0 | 255 | 8u |

TABLE 6

positions {e, g, m, o} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = y0 | r1 is y0 (y0 is a 1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r2 = y1 | r2 is y1 (y1 is a 1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is y0 + y1 | −5100 | 21420 | 16s |
| r1 = r1 + 32 | r1 is y0 + y1 + 32 | −5068 | 21452 | 16s |
| r1 = r1 >> 6 | r1 is (y0 + y1 + 32) >> 6 | −79 | 335 | 11s |
| r1 = max(0, r1) | clip r1 on the low side | 0 | 335 | 10u |
| r1 = min(255, r1) | clip r1 on the high side | 0 | 255 | 8u |

## Karczewicz-II (Ex-1006)

TABLE 7

position {j} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Reg-ister size |
|---|---|---|---|---|
| r1 = y0 | y0 is 1-D half-pixel aa1 before shifting down | -2550 | 10710 | 15s |
| r2 = y5 | y5 is 1-D half-pixel jj1 before shifting down | -2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is y0 + y5 | -5100 | 21420 | 16s |
| r2 = y1 | y1 is 1-D half-pixel bb1 before shifting down | -2550 | 10710 | 15s |
| r3 = y4 | y4 is 1-D half-pixel ii1 before shifting down | -2550 | 10710 | 15s |
| r2 = r2 + r3 | r2 is y1 + y4 | -5100 | 21420 | 16s |
| r1 = r1 - r2 | r1 is (y0 + y5) - (y1 + y4) | -26520 | 26520 | 16s |
| r1 = r1 >> 2 | r1 is ((y0 + y5) - (y1 + y4)) >> 2 | -6630 | 6630 | 14s |
| r3 = y2 | y2 is 1-D half-pixel b1 before shifting down | -2550 | 10710 | 15s |
| r4 = y3 | y3 is 1-D half-pixel hh1 before shifting down | -2550 | 10710 | 15s |
| r3 = r3 + r4 | r3 is y2 + y3 | -5100 | 21420 | 16s |
| r3 = max(-2372, r3) | clip r3 to avoid overflow (low side) | -2372 | 21420 | 16s |
| r3 = min(18640, r3) | clip r3 to avoid overflow (high side) | -2372 | 18640 | 16s |
| r2 = r3 - r4 | r2 is (y2 + y3) - (y1 + y4) | -23792 | 23740 | 16s |
| r1 = r1 + r2 | r1 is ((y0 + y5) - 5* (y1 + y4)) + 4* (y2 + y3)) >> 2 | -30422 | 30370 | 16s |
| r1 = r1 >> 2 | r1 is ((y0 + y5) - 5* (y1 + y4)) + 4* (y2 + y3)) >> 4 | -7605 | 7592 | 14s |
| r1 = r1 + r3 | r1 is ((y0 + y5) - 5* (y1 + y4)) + 20* (y2 + y3)) >> 4 | -9977 | 26232 | 16s |
| j1 = r1 | store result | | | |
| r1 = r1 + 32 | r1 is ((y0 + y5) - 5* (y1 + y4)) + 20* (y2 + y3)) + 512)} >> 4 | -9945 | 26264 | 16s |
| r1 = r1 >> 6 | r1 is ((y0 + y5) - 5* (y1 + y4)) + 20* (y2 + y3)) + 512) >> 10 | -155 | 410 | 10s |
| r1 = max(0, r1) | clip r1 on the low side | 0 | 410 | 9u |
| r1 = min(255, r1) | clip r1 on the high side | 0 | 255 | 8u |

## Ye '503 (Ex-2007)

TABLE 7

position {j} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = y0 | y0 is 1-D half pixel aa1 before shifting down | -2550 | 10710 | 15s |
| r2 = y5 | y5 is 1-D half pixel jj1 before shifting down | -2550 | 10710 | 15s |
| r1 = r1 + r2 | r1 is y0 + y5 | -5100 | 21420 | 16s |
| r2 = y1 | y1 is 1-D half pixel bb1 before shifting down | -2550 | 10710 | 15s |
| r3 = y4 | y4 is 1-D half pixel ii1 before shifting down | -2550 | 10710 | 15s |
| r2 = r2 + r3 | r2 is y1 + y4 | -5100 | 21420 | 16s |
| r1 = r1 - r2 | r1 is (y0 + y5) - (y1 + y4) | -26520 | 26520 | 16s |
| r1 = r1 >> 2 | r1 is ((y0 + y5) - (y1 + y4)) >> 2 | -6630 | 6630 | 14s |
| r3 = y2 | y2 is 1-D half pixel b1 before shifting down | -2550 | 10710 | 15s |
| r4 = y3 | y3 is 1-D half pixel hh1 before shifting down | -2550 | 10710 | 15s |
| r3 = r3 + r4 | r3 is y2 + y3 | -5100 | 21420 | 16s |
| r3 = max (-2372,r3) | clip r3 to avoid overflow (low side) | -2372 | 21420 | 16s |
| r3 = min (18640,r3) | clip r3 to avoid overflow (high side) | -2372 | 18640 | 16s |
| r2 = r3 - r2 | r2 is (y2 + y3) - (y1 + y4) | -23792 | 23740 | 16s |
| r1 = r1 + r2 | r1 is ((y0 + y5) - 5* (y1 + y4) + 4* (y2 + y3)) >> 2 | -30422 | 30370 | 16s |
| r1 = r1 >> 2 | r1 is ((y0 + y5) - 5* (y1 + y4) + 4* (y2 + y3)) >> 4 | -7605 | 7592 | 14s |
| r1 = r1 + r3 | r1 is ((y0 + y5) - 5* (y1 + y4) + 20* (y2 + y3)) >> 4 | -9977 | 26232 | 16s |
| j1 = r1 | store result | | | |
| r1 = r1 + 32 | r1 is ((y0 + y5) - 5* (y1 + y4) + 20* (y2 + y3)) + 512) >> 4 | -9945 | 26264 | 16s |
| r1 = r1 >> 6 | r1 is ((y0 + y5) - 5* (y1 + y4) + 20* (y2 + y3)) + 512) >> 10 | -155 | 410 | 10s |
| r1 = max (0,r1) | clip r1 on the low side | 0 | 410 | 9u |
| r1 = min (255,r1) | clip r1 on the high side | 0 | 255 | 8u |

**Karczewicz-II (Ex-1006)**

TABLE 8

positions {f, i, k, n} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = y0 | r1 is y0 (1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r2 = j1 | r2 is j1 (2-D half-pixel j1 before shifting down) | −9914 | 26232 | 16s |

TABLE 8-continued

positions {f, i, k, n} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r2 = r2 >> 1 | r2 is j1 >> 1 | −4957 | 13116 | 15s |
| r1 = r1 + r2 | r1 is y0 + (j1 >> 1) | −7507 | 23826 | 16s |
| r1 = r1 + 32 | r1 is y0 + (j1 >> 1) + 32 | −7491 | 23842 | 16s |
| r1 = r1 >> 6 | r1 is (y0 + (j1 >> 1) + 32) >> 6 | −235 | 745 | 11s |
| r1 = max (0, r1) | clip r1 on the low side | 0 | 745 | 10u |
| r1 = min (255, r1) | clip r1 on the high side | 0 | 255 | 8u |

**Ye '503 (Ex-2007)**

TABLE 8

positions {f, i, k, n} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = y0 | r1 is y0 (1-D half-pixel such as b1, h1, ee1 and hh1 before shifting down) | −2550 | 10710 | 15s |
| r2 = j1 | r2 is j1 (2-D half-pixel j1 before shifting down) | −9914 | 26232 | 16s |
| r2 = r2 >> 1 | r2 is j1 >> 1 | −4957 | 13116 | 15s |
| r1 = r1 + r2 | r1 is y0 + (j1 >> 1) | −7507 | 23826 | 16s |
| r1 = r1 + 32 | r1 is y0 + (j1 >> 1) + 32 | −7491 | 23842 | 16s |
| r1 = r1 >> 6 | r1 is (y0 + (j1 >> 1) + 32) >> 6 | −235 | 745 | 11s |

TABLE 8-continued

positions {f, i, k, n} of FIGS. 4A-4D

| Operation | Comment | Min value | Max value | Register size |
|---|---|---|---|---|
| r1 = max(0, r1) | clip r1 on the low side | 0 | 745 | 10u |
| r1 = min(255, r1) | clip r1 on the high side | 0 | 255 | 8u |

Likewise, the same statements made in Karczewicz-II and used to support the motivation to modify and combine Karczewicz-I and Karczewicz-II are present in Ye '503. *Compare, e.g.,* Ex-1006, ¶ 99 with Ex-2007, ¶ 98; Ex-2001, ¶ 87.

This overlap is not surprising, as Ye '503 and Karczewicz-II were filed on the same day, naming the same inventors, and claiming priority to the same three provisional patent applications. *Compare,* Ex-1006 at 1 with Ex-2007 at 1.

Ye '503 was also not the only reference the Examiner had seen multiple times containing the essential elements of the relied-upon disclosure of Karczewicz-II. VCEG-AI33, also by Yan Ye, Peisong Chen, and Marta Karczewicz, includes the

same interpolation equations from Karczewicz-II and Ye ’503 tables 3 and 4. Ex-2010 at 1. Ex-2001, ¶ 88. VCEG-AI33 also includes step-by-step instructions for interpolation of the middle-subpixel (j) as well as adjacent quarter-pixel values showing the use of 16 bit registers, as also shown in tables 7 and 8 of Karczewicz-II and Ye ’503. Ex-2010 at 3-4; Ex-2001, ¶ 89. VCEG-AI33 also provides similar language to the motivation to that relied upon by Petitioner as motivation to modify and combine Karczewicz-I and Karczewicz-II, including that “frequent and biased rounding effectively reduces precision of the interpolation filtering” and that, to resolve this issue, “to postpone rounding and clipping until the last step not to compromise the final precision.” Ex-2010 at 1; Ex-2001, ¶ 89.

Accordingly, in addition to having actually been before and considered by the Patent Office during the prosecution of the Challenged Claims, the relied-upon disclosure from Karczewicz-II is was included in at least two different references cited to the Patent Office during prosecution of the ’267 Patent and each of the ’267 Patent’s priority applications. Karczewicz-II is nothing more than the definition of a cumulative reference to Ye ’503 and VCEG-AI33. Ex-2001, ¶ 89. Stated simply, the Board should not allow Petitioner to recycle **related references** filed by **the same inventors**, and **cumulative to the disclosure of references cited to the Patent Office**, in an attempt to challenge the patentability of the claims of the ’267 Patent.

**c) Petitioner has failed to show that the Karczewicz-I reference in the Petition differs in any material way from the Ye '262, Panchal, or Karczewicz '057 references cited in prosecution and on the face of the '267 Patent.**

Petitioner relies on Karczewicz-I for its disclosure of bi-prediction, and argues that a POSITA would have been motivated by Karczewicz-II's teaching of avoiding rounding during interpolation to modify both Karczewicz-I and Karczewicz-II to use Karczewicz-II's intermediate interpolation values as predictions for a modified version of Karczewicz-I's bi-predication equation. Pet. 38-53, 59-68. However, the same bi-prediction teaching relied upon in Karczewicz-I was before the Patent Office in numerous other references during the prosecution of the '267 Patent and its priority applications, including Ye '262, Panchal, and Karczewicz '057. Ex-2001, ¶ 90.

Ye '262 teaches both default weighted bi-prediction and weighted bi-prediction in the same paragraph, and formed the basis for numerous rejections during the prosecution of the '893 Application. Ex-2009 ¶ 25; Ex-2001, ¶ 91. Panchal teaches the same default weighted, explicitly weighted, and implicitly weighted bi-prediction as Karczewicz-I. *Compare* Ex-2008 ¶¶ 60-62 with Ex-1005 ¶¶ 60-62; Ex-2001, ¶ 91. Karczewicz '057 includes substantially the same overall disclosure as Karczewicz-I related to various forms and variations of bi-prediction. *Compare* Ex-2011 at 12:4-16:16 with Ex-1005 ¶¶ 57-76; Ex-2001, ¶ 91. Petitioner

fails to identify this similarity anywhere in its Petition for good reason—as with the Petitions other arguments, the art cited by Petitioner is cumulative to that considered by and rejected by the Patent Office when allowing the claims of the ’267 Patent and its priority applications.

Accordingly, the relied-upon disclosure from Karczewicz-I was included in at least three different references cited to the Patent Office during prosecution of the ’267 Patent, including two references used as the basis for a rejection for one of the ’267 Patent’s priority applications. Ex-2001, ¶ 92. Karczewicz-I is no more than a cumulative reference of Ye ’262, Panchal, and Karczewicz ’057. *Id.*

**2. Petitioner has not established that the Patent Office erred during prosecution.**

Despite acknowledging at least some of the overlap in the references it relies on and the references used in prosecution (Pet. 86) and presenting an *Advanced Bionics* argument, Petitioner has *not* presented an argument (nor can it) that the Patent Office erred in issuing the Challenged Claims included in the ’267 Patent. Pet. 86 (only arguing the first prong). This absence of an argument suggesting error on the part of the Patent Office is in and of itself fatal to Petitioner’s arguments under *Advanced Bionics*. The Patent Office was repeatedly presented with the same or similar, cumulative disclosures to those raised in the Petition during the prosecution of the ’267 Patent and its priority applications. The Examiner applied these and

similar references in a series of rejections of the ’267 Patent’s original priority application, and undoubtedly reviewed the Information Disclosure Statements filed by Patent Owner and the cited references during the prosecution of each application in the ’267 Patent’s family. Patent Owner demonstrated to the Examiner that the art (which is cumulative of the art in this Petition) was lacking and failed to disclose the patented inventions, and the claims of the ’267 Patent, as well as the claims of its priority patents, issued over substantially identical art to that presented in the Petition Grounds.

Even if Petitioner shows some disagreement about the treatment of the art, it will not be enough to overcome the “reasonable mind” standard this Board has established. *Advanced Bionics* at 8-10 (“If reasonable minds can disagree regarding the purported treatment of the art or arguments, it cannot be said that the Office erred in a manner material to patentability. At bottom, this framework reflects a commitment to defer to previous Office evaluations of the evidence of record unless material error is shown.”). The Board should defer to the previous Office evaluations and deny institution.

### B. Petitioner Has Not Established A Reasonable Likelihood That The Challenged Claims Are Unpatentable In Light Of Ground 1.

#### 1. Petitioner fails to demonstrate that the Challenged Claims are obvious over Walker because Walker does not use higher-precision predictions

Petitioner's first Ground, based on Walker, relies on Petitioner's interpretation that a weight, applied to a prediction, creates a higher precision prediction. Pet. 19-23. This interpretation contradicts the specification and prosecution history of the '267 Patent, the understanding of a person of skill in the art and the express teachings of Walker. Ex-2001, ¶ 93, 102-116. Because the weights (w0 and w1) multiplied by the predictions (pred0 and pred1) are not also predictions with higher precision than the reference blocks used to obtain pred0 and pred1, Ground 1 fails to render obvious at least elements 1[c], 1[d], 7[c], 7[d], 13[c], and 13[d]. Each of these elements requires the first/second prediction have the second, higher precision. *Id.*

First, the '267 Patent makes clear that the prediction is the value obtained from the reference frame using the motion vector, and not some intermediate value generated in the process of combining two predictions. Ex-2001, ¶ 94. The '267 Patent plainly states that, in motion compensated prediction, the prediction is "the best match" from a reference frame, which can be "coded into a bit stream as [a] motion vector" at the encoder so the decoder can "reconstruct the exact prediction frame decoding the motion vector data encoded in the bitstream." Ex. 1001 at 2:26-

32; Ex-2001, ¶ 95. The "source" of a prediction is a block of pixels in a previously coded image, and motion vector identifies that source and allows the prediction to be made. Ex. 1001 at 3:14-18; Ex-2001, ¶ 95.

In Columns 12-14, the '267 Patent provides an example of how the first and second predictions are calculated and combined in the preferred embodiment. The first and second prediction values are obtained using the block processor applied to pixel values from the first and second reference blocks based on the first and second motion vectors. Ex-1001 at 12:44-55; Ex-2001, ¶ 96. If a motion vector indicates a fraction of a pixel, or sub-pixel, then an interpolation process will be used to obtain the prediction. Ex-1001 at 12:56-13:18; Ex-2001, ¶ 96. Example applications of a 6-tap filter to underlying pixel values are illustrated at 14:19-22; Ex-2001, ¶ 96.

$$P1=(3*E_1-17*F_1+78*G_1+78*H_1-17*I_1+3*J_1+1)>>1$$

$$P2=(3*E_2-17*F_2+78*G_2+78*H_2-17*I_2+3*J_2+1)>>1$$

Here, P1 and P2 are the predictions because they are the values that result from decoding the motion vector data encoded in the bitstream. Ex-2001, ¶ 97. Intermediate values within that process (e.g., $17*F_1$) are not themselves predictions, even though they are part of the value required to obtain the prediction and include an underlying reference pixel value. Rather, quantities like $17*F_1$ are components of

the prediction $P_1$, the same way that a student's score on an exam is a component of his final grade. Ex-2001, ¶ 97 ("An arbitrary intermediate value calculated within that process (e.g., 17*F1) does not constitute a prediction, even though such a value is calculated during the process of obtaining a prediction such as P1 or P2. Rather, an intermediate value such as for example 17*F1 is a component of the prediction P1, in a similar way that the price of an individual packet of food is a component of a final supermarket bill (grocery bill).").

Applying weights to P1 and P2 before combining them would not change the precision of P1 or P2, and the resulting intermediate values (e.g., w1*P1 and w2*P2) would not be predictions. Ex-2001, ¶ 98. Instead, w1*P1 or w2*P2 are the components or contributions of P1 and P2 to the combined prediction, intermediate values between the predictions P1 and P2 and the combined prediction. *Id.*

Petitioner proposes that applying a weight to a prediction pred0, such as w0*pred0, results in a value that is itself a prediction Pet. 19-23. A value such as w0*pred0 is at most an intermediate value, in a similar way that $17*F_1$ is an intermediate value in the discussion above. In the context of Walker equation 13, w0*pred0 is an intermediate value that contributes to the calculation of an actual prediction, namely Final_pred. Ex-2001, ¶ 99.

Additionally, the '267 Patent never addresses weighted bi-prediction, despite it being a part of multiple prior standards, including the H.264 standard referenced

in the background of the specification. Instead, the ’267 patent refers only to obtaining a combined prediction from two component predictions. Ex-2001, ¶ 100. A POSITA would understand from that that any application of weights during the combination of two predictions is part of the combination process (e.g., the “+” in Fig. 11) and not a separate set of predictions. *Id.*.

This is consistent with the prosecution history for the ’267 Patent’s family as well. As addressed in Sections II.C.1.a and III.A.1.a, weighted prediction references were before the Patent Office during the prosecution of the ’267 Patent and its family members, and the weighted bi-prediction reference Ye ’262 formed the basis for several rejections of the ’267 Patent’s original priority application. However, even while applying a weighted bi-prediction reference, the Examiner recognized that weighted bi-prediction did not teach a first and second prediction with a second precision greater than the first precision (i.e., the precision of full pixel values). Ex-1007 at 252. And the Patent Owner, in response to the Office Actions, made clear that applying a weight to a prediction did not increase the precision of the prediction, it instead mapped that prediction to a different range of values represented by the product of the weight and the prediction. Ex-1007 at 281.

This operates similarly to drawing a picture on a balloon and then inflating it. Ex-2001, ¶ 101. Inflating the balloon makes the original image larger, but does not add precision. *Id.* Scaling a prediction using a weight does not increase the number

of bits needed to represent the value of the prediction, as the total number of possible values for the prediction remains the same before and after scaling. Ex-2001, ¶¶ 102-104. Consider a quantity P that could take values in the range fo 0-9. *Id*. Applying a scaling weight (e.g., 4) to this quantity still only permits ten values. Contrast this with a quantity Q that may take any value from 0 to 36. *Id*. Q has greater precision tha neither P or P×4, as it requires more bits to represent. *Id*. Similarly, scaling a quantity that can take a value 0-255 by a weight w0, as in Walker, simply spaces the 256 allowed values further apart. Ex-2001, ¶ 105.

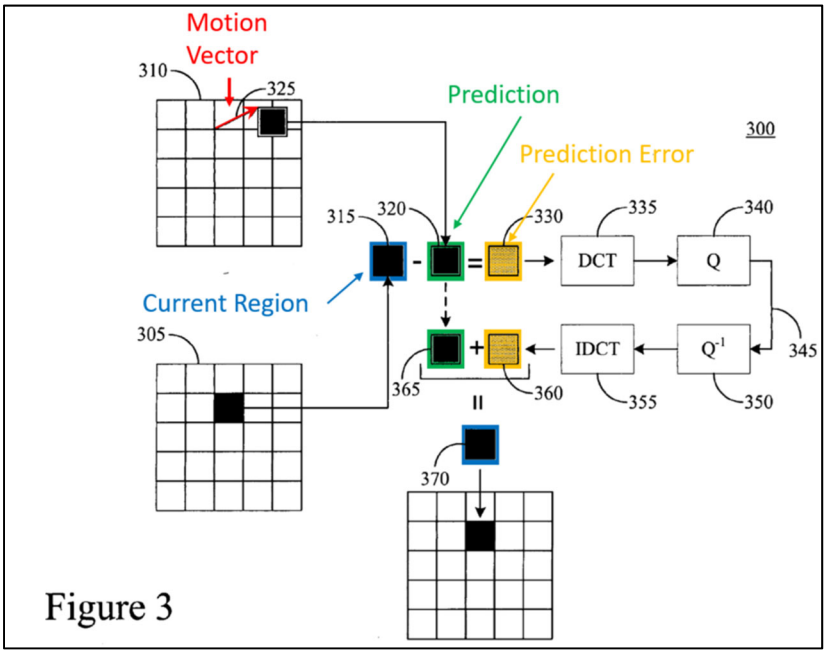
Nor is the Petition's interpretation of "prediction" consistent with the understanding of a POSITA. The Petition points to no contemporary disclosure showing an understanding in the field that a prediction block multiplied by a weight as part of a bi-prediction process is itself a prediction. For example, in H.264, the results of the sub-pixel interpolation processes are called "prediction [luma/chroma] sample values" and are stored in arrays called "PredPart." Ex-1014 at 188 (§ 8.4.2.2); Ex-2001, ¶ 106. The result of applying weights to these prediction samples values, in contrast, is not stored and is not referred to as a prediction sample value. Ex-1014 at 195 (Equation (8-278)); *Id.* The Petition simply takes the (pred0)w0 and (pred1)(w1) values of Walker to be the "predictions" because those values are stored using more than 8 bits. *Id.* However, this is an interpretation driven by hindsight

rather than any faithful adherence to what was understood in the art, or what Walker itself teaches. *Id.*

Finally, Walker itself refers to pred0 and pred1, not w0*pred0 and w1*pred1, as predictions. Walker teaches three kinds of prediction: forward or uni-directional prediction, where predictions are made for a current picture from a single reference picture; bi-predictions, where predictions are made for a current picture from a combination of predictions from two different reference pictures; and weighted bi-prediction, where predictions are made for a current picture from a weighted combination of predictions from two reference pictures. Ex-2001, ¶ 108. Walker's stated purpose is to provide a universal decoder that can use the same prediction pipeline for each of these different forms of prediction found in various different codecs, and add weights to the process when the codec used calls for them. Ex-1004 ¶ 7. Thus the prediction obtained from each reference block is the same in Walker, whether weights are used or not. This is described in reference to the examples provided by Walker below.

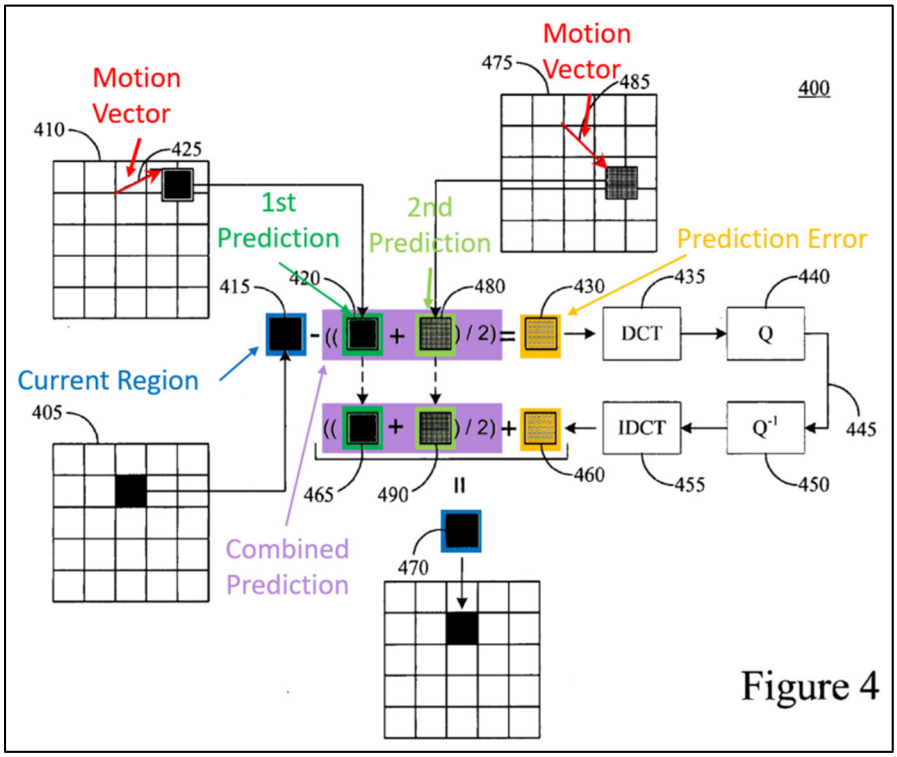
In describing the first form of prediction, Walker describes finding "a best matching prediction region in a reference picture" and taking "the difference between the current region and the best matching reference prediction region" to find the residual error, or prediction error. Ex-1004 ¶ 30; Ex-2001, ¶ 109. Thus, the prediction is obtained from the reference picture, is defined and communicated by

the motion vector, and is used to determine the prediction error. *Id.* This is illustrated in the annotated Walker Figure 3 below. *Id.*



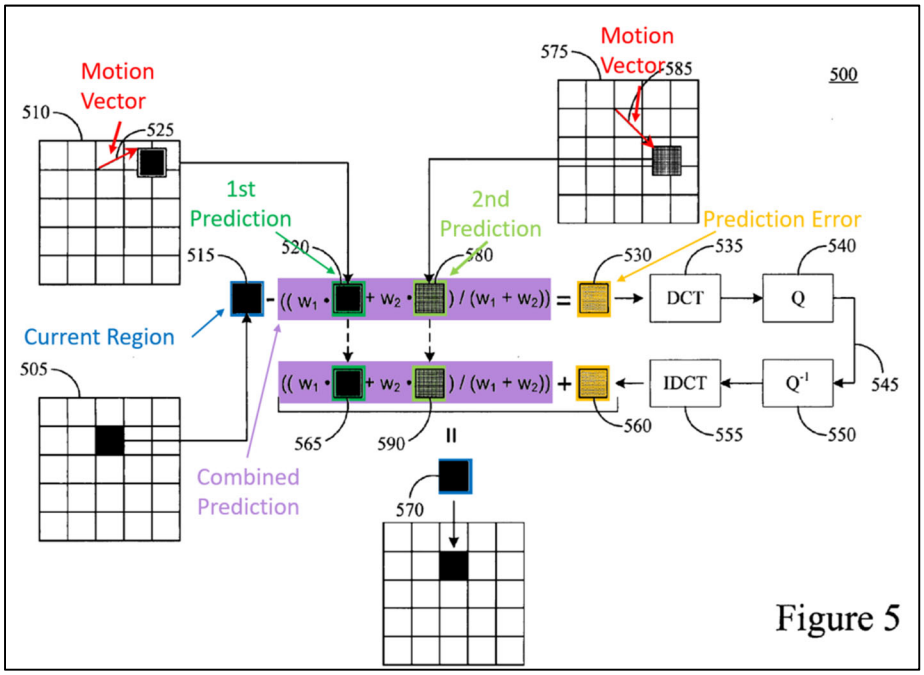

Similarly, Walker describes the use of bi-prediction (without weights) in connection with Figure 4. Ex-2001, ¶ 110. Walker describes identifying "a best matching prediction region in a previous picture" and "a best matching prediction region in a subsequent picture" each with their own corresponding motion vector, and combining those two prediction regions to form a "combined bi-prediction region." Ex-1004 ¶ 32; Ex-2001, ¶ 110. Thus, there is a first prediction, defined by a first motion vector and obtained from a first reference picture based on that motion vector, and a second prediction, defined by a second motion vector and obtained from a second reference picture based on that motion vector, that are combined to

form a combined prediction, which is used to obtain the prediction error. Ex-2001, ¶ 111. This is illustrated in the annotated Walker Figure 4 below. *Id.*




And Walker describes its disclosed weighted bi-prediction in the same way: best matching macroblocks are found in both a previous and a subsequent reference picture, and are combined, this time with weighting factors, to reduce the residual error signal of the combined prediction. Ex-1004 ¶ 34. As with unweighted bi-prediction illustrated with Figure 4, weighted bi-prediction in Walker uses two predictions, one from each of two reference frames, each identified by a motion vector, and then combines them to make a combined prediction. Ex-2001, ¶ 111. The difference between bi-prediction and weighted bi-prediction in Walker is how

these predictions are combined, that is, with our without weight factors. *Id.* This is illustrated below.




Walker's description of prediction and weighted bi-prediction confirms that the predictions combined in bi-prediction are the blocks obtained by the motion vectors, not those prediction blocks after they have had weighting factors applied. Ex-2001, ¶ 112. The simplest description of the bi-prediction process is at Walker ¶¶ 48-49, which describes the bi-prediction equation $pred_{ij}=(pred1_{ij} + pred2_{ij} + 1) >> 1$ as "used to combine predictions from two reference frames," those predictions being pred1ij and pred2ij, the "8-bit luminance and chrominance samples from prediction blocks from two reference frames." Ex-1004 ¶¶ 48-49; Ex-2001, ¶ 114. Walker describes unweighted prediction from H.264 similarly. Ex-1004 ¶¶ 59-60.

For weighted prediction, Walker describes the weighting factors as being "applied to the luma or chroma value" for the corresponding reference picture, and not as a factor or used to obtain a prediction. Ex-2001, ¶ 113.

TABLE 1

| Variables | Description | Range | No. of Bits |
|---|---|---|---|
| logWD | The base 2 logarithm of the denominator for all the luma or all the chroma weighting factors (luma and chroma weighting factors can be derived separately) | 0 to 7 | 3 |
| w0 | The weighting factor applied to the luma or chroma prediction value for the first reference picture (list "0") | −128 to 127 | 8 |
| w1 | The weighting factor applied to the luma or chroma prediction value for the second reference picture (list "1") | −128 to 127 | 8 |

Ex-1004 ¶ 69, Table 1.

Additionally, Walker's detailed step-by-step disclosure of its bi-prediction methods show that both the input predictions and the final values of each bi-prediction process are the same precision: 8 bits. Ex-2001, ¶ 114. This is true of Table 2, 3, and 4, each reproduced below. Ex-1004 ¶¶ 84, 89, 93; Tables 2-4.

TABLE 2

| Operation No. | Operation | Bitwidths Involved | Bitwidth of Operation Result |
|---|---|---|---|
| 1 | (pred0)w0, (pred1)w1 | 8 bits * 8 bits | 16 |
| 2 | $2^{LWD-1}$ | logWD has a maximum value of 7, therefore LWD – 1 has a maximum value of 7 for case 3 | 7 |
| 3 | (pred0)w0 + (pred1)w1 + $2^{LWD-1}$ | 16 bits + 16 bits + 7 bits | 18 |
| 4 | (((pred0)w0 + (pred1)w1 + $2^{logWD}$) >> (LWD + 1)) | (18 bits) >> (8 bits)* *maximum value of LWD + 1 | 10 |
| 5 | $(o_0 + o_1 + 1)$ >> (1) | (8 bits + 8 bits + 1) >> (1) | 8 |
| 6 | Clip1[Op. 4 + Op. 5] (for Clip1, see Eq. (11)) | Clip1[10 bits + 8 bits] | 8 |

TABLE 3

| Operation No. | Operation | Bitwidths Involved | Bitwidth of Operation Result |
|---|---|---|---|
| 1 | (pred0)w0', (pred1)w1' | 8 bits * 9 bits | 17 |
| 2 | (pred0)w0' + (pred1)w1' + $2^7$ | 17 bits + 17 bits + 7 bits | 19 |
| 3 | (((pred0)w0' + (pred1)w1' + $2^7$) >> (8)) | (19 bits) >> (8 bits) | 11 |
| 4 | $(o_0 + o_1 + 1)$ >> (1) | (8 bits + 8 bits + 1) >> (1) | 8 |
| 5 | Clip1[Op. 3 + Op. 4] (for Clip1, see Eq. (11)) | Clip1[11 bits + 8 bits] | 8 |

TABLE 4

| Operation No. | Operation | Bitwidths Involved | Bitwidth of Operation Result |
|---|---|---|---|
| 1 | (pred0) $w_A$, (pred1) $w_B$ | 8 bits * 15 bits | 23 |
| 2 | ((pred0) $w_A$) >> 6 + ((pred1)$w_B$) >> 6 + $2^7$ | 17 bits + 17 bits + 7 bits | 19 |
| 3 | (((pred0) $w_A$) >> 6 + ((pred1)$w_B$) >> 6 + $2^7$) >> (8) | (19 bits) >> (8 bits) | 11 |
| 4 | $(o_0 + o_1 + 1)$ >> (1) | (8 bits + 8 bits + 1) >> (1) | 8 |
| 5 | Clip1[Op. 3 + Op. 4] (for Clip1, see Eq. (11)) | Clip1[11 bits + 8 bits] | 8 |

The first and second predictions of Walker represent only 256 possible values (8 bits), and whether those possible values are separated by 1 (unweighted) or by w0 or w1 (weighted) does not change the precision of the prediction. Ex-2001, ¶ 115.

Accordingly, Walker fails to teach at least elements 1[c], 1[d], 7[c], 7[d], 13[c], and 13[d] and the Petition has failed to show a reasonable likelihood that the Challenged Claims are unpatentable based on this Ground. Ex-2001, ¶ 116.

### C. Petitioner Has Not Established A Reasonable Likelihood That The Challenged Claims Are Unpatentable In Light Of Ground 2.

#### 1. Petitioner fails to demonstrate a motivation to make the modified combination of Karczewicz-I and Karczewicz-II raised in Ground 2

The Petition seeks not just to combine Karczewicz-I and Karczewicz-II, 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 such a combination. Ex-2001, ¶ 117.

"An obviousness determination requires finding that a person of ordinary skill in the art would have been motivated to combine or modify the teachings in the prior art *and would have had a reasonable expectation of success in doing so*." *Regents of University of California v. Broad Institute, Inc.*, 903 F.3d 1286, 1291 (Fed. Cir. 2018) (emphasis added). Petitioner cannot satisfy its burden of proving obviousness by employing "mere conclusory statements," but "must instead articulate specific reasoning, *based on evidence of record*, to support the legal conclusion of obviousness." *In re Magnum Oil Tools Int'l, Ltd.*, 829 F.3d 1364, 1380 (Fed. Cir. 2016) (emphasis added).

As an initial matter, the Petition asserts that Karczewicz-II "teaches improved calculations for averaging interpolated pixel values, where rounding is delayed until

later in the process, thereby maintaining higher precision for intermediate calculations and that those "complimentary teachings" would "improve Karczewicz-I by reducing rounding inaccuracies." Petition at 39. However, Karczewicz-I already includes the higher-precision interpolation equations the Petition relies upon from its incorporation by reference of VCEG-AI33. Ex-1005 at ¶ 37; Ex-2001, ¶ 1118. As noted above, VCEG-AI33 includes the same teachings relied upon in Karczewicz-II, including the same equations for interpolation of quarter-pixel values using intermediate values for half-pixel positions. Ex-2001, ¶ 118; Section III.A.1.b, *supra*. A POSITA would therefore understand Karczewicz-I to already employ the improved interpolation techniques of Karczewicz-II, and would not look to Karczewicz-II for any further combination or modification. *Id.*

Second, the express teachings of both Karczewicz-I and Karczewicz-II call for the output of Karczewicz-II and the input of Karczewicz-I to be at the same precision as the current block of pixels. As a result, even if a POSITA were motivate to combine Karczewicz-I and Karczewicz-II, they would not combine Karczewicz-I and Karczewicz-II in the manner asserted in the Petition. Ex-2001, ¶ 119.

Karczewicz-I refers to the result of its bi-predictions as "prediction data." Ex-1005 at ¶¶ 64, 69. Karczewicz-I further teaches that, once suitable prediction data is found, it is subtracted from the original block being coded to create a residual block. Ex-1005 at ¶ 73. And at the encoder, the received residual block is added to the

prediction data to reproduce the original data. Ex-1005 at ¶ 84. Because they are added and subtracted, the current block pixels and the prediction data must have the same precision. Ex-2001, ¶ 120. The unweighted bi-prediction equation relied upon by the Petition from Karczewicz-I take in predictions pred0(i,j) and pred1(i,j), add them and right-shift by 1, then output prediction data pred(i,j). Ex-1005 ¶ 60; Ex-2001, ¶ 120. This means that the predictions pred0(i,j) and pred1(i,j) and the combined prediction pred(i,j) have the same precision. *Id.* Accordingly, the input predictions in Karczewicz-I (pred0(i,j) and pred1(i,j)) need to have the same precision as the pixels of the current block. *Id.*

Similarly, each of the interpolated values for sub-pixel positions in Karczewicz-II have the same precision as the original input pixels. The 6-tap filter used in Karczewicz-II has weights [1, -5, 20, 20, -5, 1] which add to 32, or $2^5$. Ex-2001, ¶ 121. Ex-1006 ¶ 93. This means that intermediate values such as b1 and h1 that result from a single pass of the 6-tap filter gain five bits of bit depth, and intermediate values such as j1 that result from two passes of the 6-tap filter gain 10 bits of bit depth. Ex-2001, ¶ 121. So when pixels have 8 bits of precision, b1 or h1 have 13 bits of precision, and j1 has 18 bits of precision. *Id.* The equations for sub-pixels a, c, d, l, f, i, k, n, e, g, m, and o in Tables 3 and 4 therefore show the addition of two values with bit depths 5 greater than the original pixel bit depth, followed by a right-shift by 6, which would return a result with the original pixel precision. Ex-

1006 ¶¶ 99-100; Ex-2001, ¶ 121. And the specific examples provided by Karczewicz-II, showing how to implement these interpolation equations using 16-bit math, all use 8-bit pixel precision. Ex-1006 ¶¶ 103-105; Ex-2001, ¶ 121. So Karczewicz-II teaches using 8-bit reference blocks to generate 8-bit predictions, using high-precision intermediate values. . Ex-2001, ¶ 121.

Taken together, then, Karczewicz-II teaches generating 8-bit values for fractional motion vectors by interpolation of 8-bit pixel values, and Karczewicz-I teaches combining those 8-bit values to generate an 8-bit combined prediction. Ex-2001, ¶ 122.

To overcome this shortcoming in its chosen art, Petitioner seeks to modify both references in an unconventional manner: store and use the intermediate values of Karczewicz-II as predictions, and use intermediate values as inputs for bi-prediction in Karczewicz-I. Pet. 38-53, 59-68. To justify this modification, Petitioner points to paragraph 99 of Karczewicz-II as standing for the proposition that Karczewicz-II "teaches improved calculations for averaging interpolated pixel values, where rounding is delayed until later in the process, thereby maintaining higher precision for intermediate calculations." Petition at 39. However, the process addressed by Paragraph 99 of Karczewicz-II is the process of interpolating final 8-bit values to be used as predictions in motion compensated prediction. Ex-2001, ¶ 125.

Petitioner never addresses much less provides a motivation to incur the significant cost of moving to higher precision predictions in its Karczewicz-I/Karczewicz-II combination. Ex-2001, ¶ 126. The advance of Karczewicz-II was the use of a small number of 16-bit registers to calculate the 8-bit final values for each sub-pixel position using higher-precision intermediate values. Ex-1006 ¶¶ 53, 59; Ex-2001, ¶ 126. A POSITA would not be motivated to abandon this advance in order to reduce rounding in bi-prediction, especially considering that Karczewicz-I already provides a solution in the form of rounding offsets. Ex-1005 ¶¶ 64-69; Ex-2001, ¶ 126. In other words, there is no need to modify Karczewicz-I to address the issue of rounding in bi-prediction, so a POSITA would not be motivated to make the change contemplated by the Petitioner. Ex-2001, ¶ 126.

Additionally, a POSITA would understand that the modifications proposed by Petitioner would have considerable costs and disadvantages, and so would not be motivated to make them. Ex-2001, ¶ 127. For example, a POSITA would understand that the proposed modifications would require generating and storing intermediate values for the 16 sub-pixel positions at the encoder in order to generate a suitable search area in each reference frame, requiring a significant increase in the storage and memory bandwidth demands for motion compensated prediction. *Id.* In order to maintain the same search space allowed with 8-bit values, Petitioner's stored intermediate values would need twice the available memory. *Id.* Further, memory

bandwidth would need to be doubled to accommodate the additional precision of the stored intermediate values, or searching a given number of potential predictions would take twice as long. *Id.* This would result either in slower encoding or encoding that searched fewer potential predictions, harming efficiency. *Id.* In the absence of additional storage space, values for sub-pixel positions would need to be recalculated every time they were needed, which would add significant time to each prediction candidate searched, either increasing encode time or decreasing search space, reducing the quality of predictions. *Id.* A POSITA would not be motivated to incur these costs nor reduce the quality of predictions in order to avoid rounding errors. *Id.*

#### 2. Petitioner cannot show a reasonable likelihood that the Challenged Claims are unpatentable without its hindsight-motivated combination of Karczewicz-I and Karczewicz-II

As noted above, Karczewicz-II teaches generating 8-bit values for fractional motion vectors by interpolation of 8-bit pixel values, and Karczewicz-I teaches combining those 8-bit values to generate an 8-bit combined prediction. Absent the modifications to both Karczewicz-I and Karczewicz-II contemplated by the Petition, the predictions combined in Karczewicz-I have the same 8-bit precision as the pixels in the current block and reference blocks. [expert]. The combination therefore fails to render obvious at least elements 1[c], 1[d], 7[c], 7[d], 13[c], and 13[d] and the

Petition has failed to show a reasonable likelihood that the Challenged Claims are unpatentable in view of this Ground. Ex-2001, ¶ 129.

## IV. CONCLUSION

For any or all of the above-described reasons, Patent Owner respectfully requests that the Board deny institution.

Dated: June 27, 2024

Respectfully submitted,

*/s/Scott W. Hejny*
Scott W. Hejny
Reg. No. 65,320
shejny@mckoolsmith.com
Eric S. Hansen
Reg. No. 63095
ehansen@mckoolsmith.com
McKool Smith, P.C.
300 Crescent Court, Suite 1500
Dallas, Texas 75201
Telephone: (214) 978-4000
Facsimile: (214) 978-4044

Counsel for Patent Owner, Nokia Technologies Oy

## CERTIFICATION OF WORD COUNT

The undersigned hereby certifies that the portions of the above-captioned **PATENT OWNER'S PRELIMINARY RESPONSE** specified in 37 C.F.R. §42.24 have 12,081 words, in compliance with the 14,000 word limit set forth in 37 C.F.R. §42.24(b)(1). This word count was prepared using Microsoft Word 2010.

Dated: June 27, 2024

Respectfully submitted,

/s/Scott W. Hejny
Scott W. Hejny
Reg. No. 45,882

## CERTIFICATION OF SERVICE

The undersigned hereby certifies that a copy of this **PATENT OWNER'S PRELIMINARY RESPONSE PURSUANT TO 37 C.F.R. § 42.10(b)** has been served on Petitioner via email on the following counsel of record for Petitioner:

Harper Batts
hbatts@sheppardmullin.com

Chris Ponder
cponder@sheppardmullin.com

Jeffrey Liang
jliang@sheppardmullin.com

Dated: June 27, 2024

Respectfully submitted,

/s/Scott W. Hejny
Scott W. Hejny
Reg. No. 45,882