IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:)
U.S. Patent No. 10,218,995)
Issued: February 26, 2019)
Application No.: 14/692,138)
Filing Date: April 21, 2015)

For: **Moving Picture Encoding System, Moving Picture Encoding Method, Moving Picture Encoding Program, Moving Picture Decoding System, Moving Picture Decoding Method, Moving Picture Decoding Program, Moving Picture Reencoding System, Moving Picture Reencoding Method, Moving Picture Reencoding Program**

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 10,218,995

TABLE OF CONTENTS

I. INTRODUCTION 1

II. IDENTIFICATION OF CHALLENGES (37 C.F.R. §42.104(b)) 1

III. BACKGROUND 2

A. '995 Patent (Ex.1001) 2

1. Technical Background 2

2. Brief Description 4

B. Prosecution History 8

C. POSITA 10

D. Claim Construction 10

IV. Ground 1: Phek, Segall, Martins, and He Render Obvious Claims 1-6 and 8-11 10

A. Prior Art 10

1. Phek 10

2. Segall 12

3. Martins 14

4. He 15

5. Combining Phek, Segall, Martins, and He 18

a. Obvious based on Segall to Encode/Decode Phek's Base and (First) Enhancement Layer Streams at the Same Spatial Resolution 19

b. Obvious based on Segall to Implement a Second Enhancement Layer in Phek's System for Spatial Scalability 23

c. Obvious based on Segall to Multiplex/Demultiplex Phek's Base and Enhancement Layer Bitstreams.....26

d. Obvious based on Martins to Downscale Phek's Super-Resolution Enlarged Reference Pictures 29

e. Obvious based on He to Selectively Apply Phek's/Martin's Super-Resolution (and Downscaling) Loop Filter.....34

B. Claim 2 39

1. [2.P] 39

2. [2.1] 40

3. [2.2] 44

4. [2.3] 46

5. [2.4] 51

6. [2.5.i] 56

7. [2.5.ii] 61

8. [2.6] 68

9. [2.7] 72

C. Claims 3-4 Limitations Correspondence to Claim 2 Limitations 72

D. Claim 11 73

E. Claim 1 74

1. [1.P] 74

2. [1.1] 76

3. [1.2] 77

4. [1.3] 81

5. [1.4] 82

6. [1.5] 86
7. [1.6.i] 87
8. [1.6.ii] 87
9. [1.6.iii] 88
10. [1.6.iv] 90
11. [1.7] 91

F. Claim 5 91
1. [5.P] 91
2. [5.1] 91
3. [5.2] 92
4. [5.3] 94
5. [5.4] 94

G. Claim 6 95
1. [6.P] 95
2. [6.1] 95

H. Claims 8-9 Limitations Correspondence to Claim 1 Limitations 96

I. Claim 10 97

V. GROUND 2: Phek, Segall, Martins, He, and Wang Render Obvious Claim 7 98

A. Prior Art 98
1. Wang 98
2. Phek-Segall-Martins-He-Wang 101

B. Claim 7 102

1. [7.P] 102
2. [7.1] 102
3. [7.2] 103
4. [7.3] 104
5. [7.4] 104

VI. THE BOARD SHOULD REACH THE MERITS 105
A. *Advanced Bionics* Favors Institution 105
B. *General Plastic* Favors Institution 105
C. *Fintiv* Favors Institution 106

VII. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8(a)(1) 107
A. Real Party-In-Interest 107
B. Related Matters 107
C. Grounds for Standing 108
D. Counsel and Service Information 108
E. IPR Fee 109

VIII. CONCLUSION 109

TABLE OF ABBREVIATIONS

Term	Abbreviation
super-resolution and downscaling loop filter	SR-D filter
signal-to-noise ratio	SNR
bitstream	BTS
output bitstream	OBTS
input bitstream	IBTS
minimum sum of absolute differences	MINSAD
power saving factor	PSF
scalable video coding	SVC

TABLE OF AUTHORITIES

Page(s)

CASES

Advanced Coding Technologies LLC v. Apple Inc.,
No. 2:24-cv-00572-JRG (E.D. Tex.) 107

Advanced Coding Technologies LLC v. Google LLC,
No. 2:24-cv-00353-JRG (E.D. Tex.) 107

Ericsson Inc. v. Intellectual Ventures I LLC,
IPR2014-00527, Paper 41, 11 (PTAB May 18, 2015) 1

Ford Motor Co. v. Neo Wireless LLC,
IPR2023-00763, Paper 28 (PTAB Mar. 22, 2024) 106

KSR Int'l Co. v. Teleflex Inc.,
550 U.S. 398 (2007).....*passim*

Samsung Elecs. Co., Ltd. v. Advanced Coding Techs. LLC,
IPR2024-00372, Paper 9 (PTAB Jul. 18, 2024) 106

Vivendum Prod. Sols., Inc. v. Rotolight Ltd.,
IPR2023-01218, Paper 12 (PTAB Apr. 19, 2024) 106

EXHIBIT LIST

Exhibit	Description
1001	U.S. Patent No. 10,218,995 to Sakazume ("the '995 Patent")
1002	Excerpts from the Prosecution History of the '995 Patent ("the '995 Patent File History")
1003	Declaration and Curriculum Vitae of Dr. Clifford Reader
1004	**[reserved]**
1005	English Translation of Japanese Patent Publication No. 2007316161 A with Translation Certificate ("Phek")
1006	Segall et al., *Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension*; IEEE Transactions on Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 ("Segall")
1007	Martins et al., *A Unified Approach to Restoration, Deinterlacing and Resolution Enhancement in Decoding MPEG-2 Video*, IEEE Transactions on Circuits & Systems for Video Tech., Vol. 12, No. 9, September 2002 ("Martins")
1008	U.S. Patent Application Publication No. 2008/0137753 ("He")
1009	U.S. Patent No. 5,886,736 ("Chen")
1010	Schwarz et al., *Overview of the Scalable Video Coding Extension of the H.264/AVC Standard*, IEEE Transactions on Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 ("Schwarz")
1011	U.S. Patent Application Publication No. 2009/0154567 ("Lei")
1012	Park et al., *Super-Resolution Image Reconstruction: A Technical Overview*, IEEE Signal Processing Magazine, pp. 21-36, May 2003 ("Park")
1013	English Machine Translation of Japanese Patent Publication No. 2008053848 ("Hatanaka")
1014	U.S. Patent Application Publication No. 2010/0165077 ("Yin")
1015	U.S. Patent Application Publication No. 2009/0257664 ("Kao")
1016	U.S. Patent Application Publication No. 2009/0046995 ("Kanumuri")
1017	U.S. Pat. 6,470,051 ("Campisano")
1018	U.S. Patent Application Publication No. 2006/0013306 ("Kim")

1019	Lu et al., Mechanisms of MPEG Stream Synchronization, ACG SIGCOMM Computer Communication Review, Vol. 24, Issue 1, pp. 57-67 (January 1994) ("Lu")
1020	U.S. Patent Publication No. 2003/0021345 ("Brusewitz")
1021	U.S. Patent Publication No. 2003/0021347 ("Lan")
1022	EE|Times, How Video Compression Works (Aug. 6, 2007), available at https://www.eetimes.com/how-video-compressionworks (retrieved Dec. 26, 2023) ("EETimes")
1023	Bier, Introduction to Video Compression, Berkeley Design Tech., Inc. (October 2005), available at https://www.bdti.com/MyBDTI/pubs/20051024_GSPx05_Video_Intro.pdf (retrieved Dec. 30, 2023) ("Bier")
1024	U.S. Patent No. 7,379,496 ("Holcomb")
1025	U.S. Patent Publication No. 2008/0043832 ("Barkley")
1026	U.S. Patent No. 6,173,013 ("Suzuki")
1027	**[reserved]**
1028	**[reserved]**
1029	Declaration of June Ann Munford
1030	Ely, MPEG Video Coding: A Simple Introduction, EBU Technical Review (Winter 1995) ("Ely")
1031	Japanese Patent Publication No. 2007316161 A with English Machine Translation
1032	ITU-T Recommendation H.264 (11/07)
1033	Karczewicz et al., The SP- and SI-Frames Design for H.264/AVC, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, v.13 (July 2003)
1034	ISO/IEC 11172-2: 1993 (E) (MPEG1 Video)
1035	ISO/IEC 13818-2: 2000 (E) (MPEG2/H.262 Video)
1036	ITU-T Recommendation H.263 (02/98)
1037	Reader, Intraframe and Interframe Adaptive Transform Coding, Efficient Transmission of Pictorial Information, SPIE v.66, 108 (1975)
1038	Reader, MPEG4: Coding for Content, Interactivity, and Universal Accessibility, Optical Engineering 35(1), 104 (January 1996)
1039	Reader, MPEG Patents, in MPEG Video Compression Standard,

	Chapter 16, at 357-362 (Mitchell et al ed.) (1996)
1040	Radha et al., The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP, IEEE Transactions on Multimedia, v.3, 53 (March 2001)
1041	H.264 / MPEG-4 Part 10 White Paper, January 31, 2003 (downloaded from https://github.com/yistLin/H264-Encoder/blob/1cdaf090b63642932ed726b102ed2cf80908edbb/doc/H.264%20%3A%20MPEG-4%20Part%2010%20White%20Paper.pdf on Jan. 5, 2024)
1042	ISO/IEC 14496-2: 2004 (MPEG4)
1043	ISO/IEC 13818-2: 1996 (MPEG2)
1044	EP 1811786 to Wang ("Wang")
1045	Third Amended Docket Control Order

LIST OF CHALLENGED CLAIM ELEMENTS

Claim 1	
[1.P]	A moving picture encoding system comprising:
[1.1]	a first encoder configured to work on a subsequence of a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution;
[1.2]	a first super-resolution enlarger configured to work on the subsequence of the sequence of moving pictures with the standard resolution to implement an interpolation of pixels with a first enlargement to create a set of super-resolution enlarged pictures with a first resolution higher than the standard resolution;
[1.3]	a first resolution converter configured to work on the set of super-resolution enlarged pictures to implement a process for a first resolution conversion to create a set of super-resolution enlarged and converted pictures with a standard resolution;
[1.4]	a second super-resolution enlarger configured to acquire the set of decoded pictures with the standard resolution from the first encoder to work on the sequence of decoded pictures to implement an interpolation of pixels with a second enlargement to create a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution;
[1.5]	a second resolution converter configured to work on the set of super-resolution enlarged decoded pictures to implement a process for a second resolution conversion to create a set of super-resolution enlarged and converted decoded pictures with a standard resolution; and
[1.6.i]	a second encoder configured to:
[1.6.ii]	have the set of super-resolution enlarged and converted pictures from the first resolution converter as a set of encoding target pictures, the set of decoded pictures from the first encoder as a set of first reference pictures, and the set of super-resolution

	enlarged and converted decoded pictures from the second resolution converter as a set of second reference pictures,
[1.6.iii]	select one of the set of first reference pictures and the set of second reference pictures to create reference picture selection information to identify the set of selected reference pictures to implement a second process for encoding to create a second sequence of encoded bits based on the set of encoding target pictures and the set of selected reference pictures, and
[1.6.iv]	implement a third process for encoding for the reference picture selection information to create a sequence of encoded bits of the reference picture selection information,
[1.7]	wherein the set of encoding target pictures, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 2	
[2.P]	A moving picture decoding system comprising:
[2.1]	a demultiplexer configured to work on a sequence of input encoded bits to implement a process for a prescribed demultiplexing to out- put at least a first and a second sequence of encoded bits;
[2.2]	a first decoder configured to acquire the first sequence of encoded bits obtained with a standard resolution at the demultiplexer to implement thereon a process for a prescribed first decoding to create a sequence of decoded pictures with a standard resolution;
[2.3]	a first super-resolution enlarger configured to acquire the sequence of decoded pictures created with a standard resolution at the first decoder to work on the sequence of decoded pictures to implement an interpolation of pixels with a first enlargement to create a sequence of super-resolution enlarged decoded pictures with a first resolution higher than a standard resolution;
[2.4]	a first resolution converter configured to acquire the sequence of super-resolution enlarged decoded pictures created at the first super-resolution enlarger to work on the sequence of super-resolution enlarged decoded pictures to implement a process for a prescribed resolution conversion to create a sequence of super-resolution decoded pictures with a standard resolution;

[2.5.i]	a second decoder configured to acquire the second sequence of encoded bits obtained with a standard resolution at the demultiplexer as a set of decoding targets, the sequence of decoded pictures created with the standard resolution at the first decoder as a set of first reference pictures, and the sequence of super-resolution decoded pictures created with the standard resolution at the first resolution converter as a set of second reference pictures, and
[2.5.ii]	select one of the set of first reference pictures and the set of second reference pictures based on reference picture selection information to implement a combination of processes for a prescribed prediction and a prescribed second decoding being a decoding with an extension of the standard resolution, to create a sequence of super-resolution pictures decoded with the standard resolution based on the set of decoding targets and the set of selected reference pictures; and
[2.6]	a second resolution converter configured to acquire the sequence of decoded pictures with the standard resolution from the first decoder to work on the sequence of decoded pictures to implement an interpolation of pixels with the second enlargement to create a sequence of enlarged decoded pictures with a high resolution as a second resolution higher than the standard resolution,
[2.7]	wherein the set of decoding targets, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 3	
[3.P]	A moving picture decoding method comprising:
[3.1]	a step of implementing a process for a prescribed demultiplexing on a sequence of input encoded bits, outputting at least a first and a second sequence of encoded bits;
[3.2]	a step of acquiring the first sequence of encoded bits obtained with a standard resolution through the process for the prescribed demultiplexing, implementing thereon a process for a prescribed first decoding, creating a sequence of decoded pictures with the standard resolution;
[3.3]	a step of acquiring the sequence of decoded pictures created with the standard resolution through the process for the

	prescribed first decoding, working on the sequence of decoded pictures to implement an interpolation of pixels with a first enlargement, creating a sequence of super-resolution enlarged decoded pictures with a first resolution higher than the standard resolution;
[3.4]	a step of acquiring the sequence of super-resolution enlarged decoded pictures created through the process for the prescribed super-resolution enlargement working on the sequence of super-resolution enlarged decoded pictures to implement a process for a first resolution conversion, creating a sequence of super-resolution decoded pictures with a standard resolution;
[3.5.i]	a step of acquiring the second sequence of encoded bits obtained with a standard resolution through the process for the prescribed demultiplexing as a set of decoding targets, the sequence of decoded pictures created with the standard resolution through the process for the prescribed first decoding as a set of first reference pictures, and the sequence of super-resolution decoded pictures created with the standard resolution through the process for the prescribed resolution conversion as a set of second reference pictures, and
[3.5.ii]	selecting one of the set of first reference pictures and the set of second reference pictures based on reference picture selection information to implement a combination of processes for a prescribed prediction and a prescribed second decoding being a decoding with an extension of the standard resolution, based on the set of decoding targets and the set of selected reference pictures to create a sequence of super-resolution pictures decoded with the standard resolution; and
[3.6]	a step of acquiring the sequence of decoded pictures created with the standard resolution through the process for the prescribed first decoding working on the sequence of decoded pictures to implement an interpolation of pixels with the second enlargement, creating a sequence of enlarged decoded pictures with a high resolution as a second resolution higher than the standard resolution,

[3.7]	wherein the set of decoding targets, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 4	
[4.P]	A recording medium storing a moving picture decoding program comprising a non-transitory computer-readable medium configured to have a computer execute:
[4.1]	a step of implementing a process for a prescribed demultiplexing on a sequence of input encoded bits, outputting at least a first and a second sequence of encoded bits;
[4.2]	a step of acquiring the first sequence of encoded bits obtained with a standard resolution through the process for the prescribed demultiplexing, implementing thereon a process for a prescribed first decoding, creating a sequence of decoded pictures with the standard resolution;
[4.3]	a step of acquiring the sequence of decoded pictures created with the standard resolution through the process for the prescribed first decoding, working on the sequence of decoded pictures to implement an interpolation of pixels with a first enlargement, creating a sequence of super-resolution enlarged decoded pictures with a first resolution higher than the standard resolution;
[4.4]	a step of acquiring the sequence of super-resolution enlarged decoded pictures created through the process for the prescribed super- resolution enlargement working on the sequence of super-resolution enlarged decoded pictures to implement a process for a first resolution conversion, creating a sequence of super-resolution decoded pictures with a standard resolution;
[4.5.i]	a step of acquiring the second sequence of encoded bits obtained with a standard resolution through the process for the prescribed demultiplexing as a set of decoding targets, the sequence of decoded pictures created with the standard resolution through the process for the prescribed first decoding as a set of first reference pictures, and the sequence of super-resolution decoded pictures created with the standard resolution through the process for the prescribed resolution conversion as a set of second reference pictures, and

[4.5.ii]	selecting one of the set of first reference pictures and the set of second reference pictures based on reference picture selection information to implement a combination of processes for a prescribed prediction and a prescribed second decoding being a decoding with an extension of the standard resolution, based on the set of decoding targets and the set of selected reference pictures to create a sequence of super-resolution pictures decoded with the standard resolution; and
[4.6]	a step of acquiring the sequence of decoded pictures created with the standard resolution through the process for the prescribed first decoding, working on the sequence of decoded pictures to implement an interpolation of pixels with the second enlargement creating a sequence of enlarged decoded pictures with a high resolution as a second resolution higher than the standard resolution,
[4.7]	wherein the first resolution is different from the second resolution,
[4.8]	wherein the standard resolution of the second sequence of encoded bits is the same as the standard resolution of the super-resolution decoded pictures,
[4.9]	and wherein the set of decoding targets, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 5	
[5.P]	The moving picture encoding system according to claim 1, wherein the first super-resolution enlarger comprises:
[5.1]	a positioner configured to work in a processing for super-resolution enlargement, on a combination of a base picture constituting a base therein and one or more observation pictures to be based on, to make a positioning to pixel positions of a desirable high resolution, for an interval-unequal sampling to create nonhomogeneous high resolution pictures;
[5.2]	an interpolator configured to work on nonhomogeneous high resolution pictures created at the positioner, to implement a process for a prescribed nonhomogeneous interpolation to create interpolated pictures with a desirable high resolution;

[5.3]	an estimated picture creator configured to acquire interpolated pictures created at the interpolator, to implement thereon a process for a prescribed reconstruction to create estimated pictures with a desirable resolution; and
[5.4]	a repetition determiner configured to acquire a nonhomogeneous high resolution picture from the positioner and a nonhomogeneous estimated picture from the estimated picture creator, for employment of the acquired pictures to follow a prescribed determination method to determine whether or not a repetition of the processing for super-resolution enlargement is necessary, and work depending on a result thereof, to operate as the repetition is necessary, to provide the interpolator with information for a control to continue the processing, and operate as the repetition is unnecessary, to provide the estimated picture creator with information for a control to output a set of estimated pictures with a high resolution after super-resolution enlargement.
Claim 6	
[6.P]	The moving picture encoding system according to claim 1, further comprising
[6.1]	a multiplexer configured to operate complying with a prescribed syntax structure to multiplex the first sequence of encoded bits from the first encoder, the second sequence of encoded bits from the second encoder, and information on encoding parameters of respective types used in the encoding at the first encoder and the encoding at the second encoder.
Claim 7	
[7.P]	The moving picture encoding system according to claim 1, further comprising:
[7.1]	a first accumulation buffer configured to work in a course before the first encoder and the first super-resolution enlarger, to accumulate subsequences of the sequence of moving pictures with the standard resolution;

[7.2]	a second accumulation buffer configured to work in a course between the first encoder and the second encoder, to accumulate sets of decoded pictures from the first encoder;
[7.3]	an accumulation controller configured to work for detections of buffer accumulation amounts at the first accumulation buffer, the second accumulation buffer, the first encoder, and the second encoder, to control the buffer accumulation amounts; and
[7.4]	a code rate controller configured to work on bases of the detections of buffer accumulation amounts at the accumulation controller, to control code rates at the first encoder and the second encoder.
Claim 8	
[8.P]	A moving picture encoding method comprising:
[8.1]	a step of implementing a first combination of processes for an encoding and a decoding on a sequence of moving pictures with a standard resolution, creating a first sequence of encoded bits and a set of decoded pictures with the standard resolution;
[8.2]	a step of implementing an interpolation of pixels with a first enlargement on the sequence of moving pictures with the standard resolution, creating a set of super-resolution enlarged pictures with a first resolution higher than a standard resolution;
[8.3]	a step of implementing a process for a first resolution conversion on the set of super-resolution enlarged pictures, creating a set of super-resolution enlarged and converted pictures with a standard resolution;
[8.4]	a step of acquiring the set of decoded pictures with the standard resolution to work on the sequence of decoded pictures to implement an interpolation of pixels with a second enlargement to create a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution;
[8.5]	a step of working on the set of super-resolution enlarged decoded pictures to implement process for a second resolution conversion to create a set of super-resolution enlarged and converted decoded pictures with a standard resolution; and

[8.6.i]	a step of:
[8.6.ii]	having the set of super-resolution enlarged and converted pictures as a set of encoding target pictures, the set of decoded pictures as a set of first reference pictures, and the set of super-resolution enlarged and converted decoded pictures as a set of second reference pictures,
[8.6.iii]	selecting one of the set of first reference pictures and the set of second reference pictures to create reference picture selection information to identify the set of selected reference pictures to implement a second process for encoding to create a second sequence of encoded bits based on the set of encoding target pictures and the set of selected reference pictures, and
[8.6.iv]	implementing a third process for encoding for the reference picture selection information to create a sequence of encoded bits of the reference picture selection information,
[8.7]	wherein the set of encoding target pictures, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 9	
[9.P]	A recording medium storing a moving picture encoding program comprising a non-transitory computer-readable medium configured to have a computer execute:
[9.1]	a step of implementing a first combination of processes for an encoding and a decoding on a sequence of moving pictures with a standard resolution, creating a first sequence of encoded bits and a set of decoded pictures with the standard resolution;
[9.2]	a step of implementing an interpolation of pixels with a first enlargement on the sequence of moving pictures with the standard resolution, creating a set of super-resolution enlarged pictures with a first resolution higher than the standard resolution;
[9.3]	a step of implementing process for a first resolution conversion on the set of super-resolution enlarged pictures, creating a set of super-resolution enlarged and converted pictures with a standard resolution;

[9.4]	a step of acquiring the set of decoded pictures with the standard resolution from the first encoder to work on the sequence of decoded pictures to implement an interpolation of pixels with a second enlargement to create a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution;
[9.5]	a step of working on the set of super-resolution enlarged decoded pictures to implement a process for a second resolution conversion to create a set of super-resolution enlarged and converted decoded pictures with a standard resolution; and
[9.6.i]	a step of:
[9.6.ii]	having the set of super-resolution enlarged and converted pictures as a set of encoding target pictures, the set of decoded pictures as a set of first reference pictures, and the set of super-resolution enlarged and converted decoded pictures as a set of second reference pictures,
[9.6.iii]	selecting one of the set of first reference pictures and the set of second reference pictures to create reference picture selection information to identify the set of selected reference pictures to implement a second process for encoding to create a second sequence of encoded bits based on the set of encoding target pictures and the set of selected reference pictures, and
[9.6.iv]	implementing a third process for encoding for the reference picture selection information to create a sequence of encoded bits of the reference picture selection information,
[9.7]	wherein the set of encoding target pictures, the set of first reference pictures, and the set of second reference pictures have the same value in spatial resolution.
Claim 10	
[10]	The moving picture encoding system according to claim 1, wherein the second encoder divides an encoding target picture into a plurality of regions to select one of the set of first reference pictures and the set of second reference pictures for each divided region.

Claim 11	
[11]	The moving picture decoding system according to claim 2, wherein the second decoder divides a decoding target into a plurality of regions to select one of the set of first reference pictures and the set of second reference pictures for each divided region.

I. INTRODUCTION

Google LLC ("Petitioner") requests *Inter Partes* Review of claims 1-11 ("Challenged Claims") of U.S. Patent 10,218,995 ("'995 patent") (Ex.1001).

II. IDENTIFICATION OF CHALLENGES (37 C.F.R. §42.104(b))

- **Ground 1:** Claims 1-6 and 8-11 are unpatentable under 35 U.S.C. §103 over Phek (Ex.1005), Segall (Ex.1006), Martins (Ex.1007), and He (Ex.1008)
- **Ground 2:** Claim 7 is unpatentable under §103 over Phek, Martins, He, and Wang (Ex.1043)

The '995 patent claims priority to Japanese Application 2008-142433, filed May 30, 2008. Ex.1001, Face. Petitioner does not concede the Challenged Claims are entitled to the benefit of any claimed priority date, but nonetheless treats May 30, 2008, as the Challenged Claims' Critical Date. Each reference qualifies as prior art:[1]

[1] Segall and Martins were published in IEEE journals before the Critical Date. IPR2014-00527, Paper 41, 11 (PTAB May 18, 2015) ("IEEE is a well-known, reputable compiler and publisher of scientific and technical publications"). Expert

Reference	Filing Date	Publication Date	Pre-AIA Prior Art
Phek	---	Dec. 6, 2007	§102(a)
Segall	---	Sept. 2007	§102(a)
Martins	---	Sept. 2002	§102(a)-(b)
He	Dec. 6, 2006	June 12, 2008	§102(e)
Wang	June 9, 2006	July 25, 2007	§102(a)

III. BACKGROUND

A. '995 Patent (Ex.1001)

1. Technical Background

The '995 Patent relates to video coding technology, and to using "super-resolution" techniques to enhance "reference pictures" in a video coding system.

The subject matter amounts to nothing more than an obvious collection of well-known video coding concepts (Ex.1003, §II), including:

(i) **motion-compensated coding** — compressing pictures (or picture regions) in a video stream using a previously-decoded picture (or a region thereof)—a "reference picture"—to predict other picture(s) (or region(s)) in

librarian June Munford presents evidence that each was publicly-accessible. Ex.1029, ¶¶9-12, 17-20.

the video stream (Ex.1003, ¶¶62-73);

(ii) **loop filtering** — enhancing "reference pictures" before being used to predict other picture(s) (Ex.1003, ¶¶74-78);

(iii) **super-resolution processing** — multiple lower-resolution pictures of the same scene are registered and used to form a super-resolution picture of the scene with higher spatial frequencies (e.g., more detail) than any individual lower-resolution input picture and higher spatial resolution (e.g., higher pixel count[2]) than input pictures (Ex.1003, ¶¶44-54); and

(iv) **layered, hierarchical coding** — two or more related video streams are coded together to provide a base layer that can be decoded independent of other layers and one or more enhancement layers decoded in dependence on the base layer (Ex.1003, ¶¶95-103; *see also id.*, ¶¶103-129).[3]

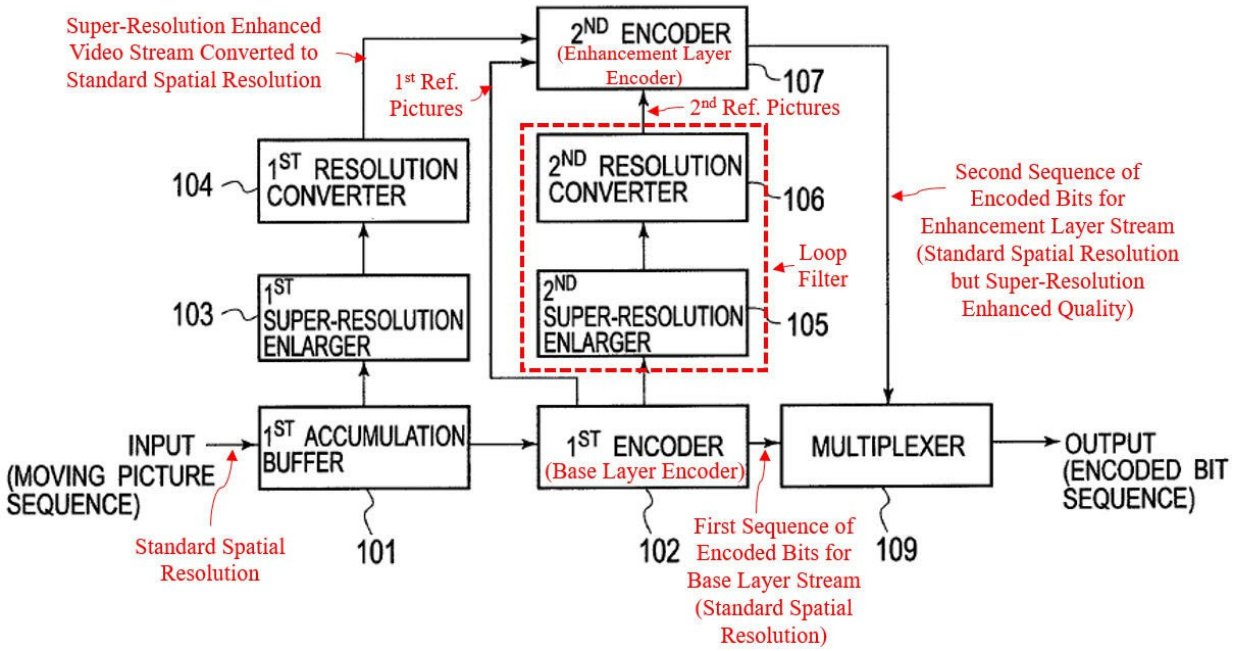
[2] The '995 Patent sometimes uses "resolution" to refer to a picture's detail (e.g., spatial frequency or information content) and elsewhere uses "resolution" to refer to the picture's size (e.g., pixel number). Ex.1003, ¶¶172-174.

[3] Common forms of layered, hierarchical coding include multi-view coding (e.g., for stereoscopic video) and scalable coding for transmitting multiple streams

2. Brief Description

The patent describes a layered video coding system using conventional super-resolution algorithms to enhance the quality of decoded base-layer pictures used as reference pictures for subsequent decoding of enhancement layer pictures. Ex.1003, ¶¶170-190. FIG. 1's encoding apparatus is instructive:

FIG. 1



Ex.1001, FIG. 1[4]

differentiated by SNR, frame rate, and/or spatial resolution. Ex.1003, ¶¶99; 93-102, 216.

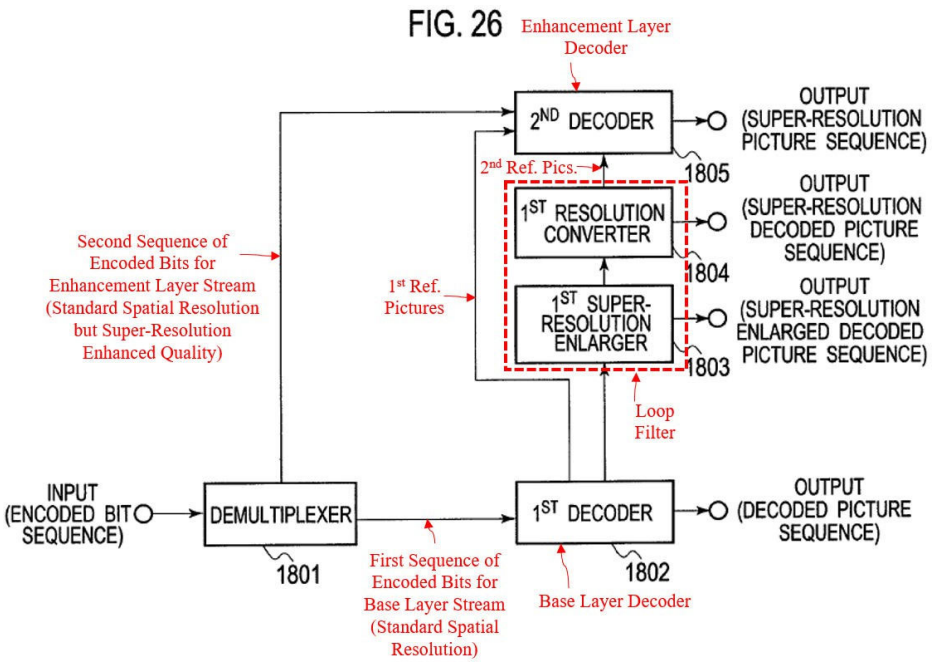
[4] Figures annotated and emphasis added unless otherwise specified.

The apparatus passes input video pictures through an accumulation buffer 101 to two encoding paths. Ex.1001, 13:40-17:45. In the first, a base version of input video is encoded by first encoder 102 to produce a first coded bitstream for the base version. *Id.*, 13:46-14:9. In the second, a super-resolution enhanced version of input video is encoded by second encoder 107 to produce a second coded bitstream for the super-resolution enhanced version. *Id.*, 15:30-42. The system generates the super-resolution enhanced version using (i) first super-resolution enlarger 103 to increase spatial resolution and information content (e.g., quality) of input video pictures, and using (ii) first resolution converter 104 to reduce spatial resolution of the super-resolution enlarged video. *Id.*, 14:10-48. For input video pictures subject to super-resolution processing, first accumulation buffer 101 accumulates subsequences of pictures from the neighborhood of the input video picture to be super-resolution enhanced; pictures in the subsequences are registered with sub-pixel accuracy to provide a corresponding super-resolution image. *Id.*, 4:34-49.

To remove redundancies in information coded between base- and enhancement-layer streams, second encoder 107 implements inter-layer predictive processes to predict enhancement-layer stream pictures from base-layer stream decoded pictures. Ex.1001, 15:30-64. Base-layer stream decoded pictures can also be enhanced before being used as reference by second encoder 107 by subjecting

base-layer pictures that first encoder 102 encoded and then decoded to processing by second-resolution enlarger 105 and second-resolution converter 106. *Id.*, 14:49-16:18. Elements 105 and 106 form a type of loop filter to enhance reference picture quality and revert reference pictures to standard resolution. Ex.1003, ¶185. Second encoder 107 selects either non-super resolution enhanced base layer pictures or super-resolution enhanced base layer pictures as reference pictures for enhanced stream encoding. Ex.1001, 15:30-16:45. Base- and enhancement-layer streams are multiplexed (109) into a single output signal, with a signal indicating which base-layer picture was selected; the multiplexed signal is transmitted to a decoder. *Id.*, 16:64-17:20, 16:46-63, 29:67-30:5, 59:62-60:20.

FIG. 26 illustrates decoding techniques for the multiplexed signal from FIG.1's apparatus. Ex.1001, 59:4-41, 50:39-54:6.



Ex.1001, FIG. 26

Demultiplexer 1801 parses the input encoded bit sequence to produce first and second bitstreams corresponding to base- and enhancement-layer video streams, respectively. Ex.1001, 50:52-51:18. First decoder 1802 decodes the base layer stream; second decoder 1805 decodes the enhancement layer stream. *Id.*, 51:16-37, 52:27-53:46. FIG. 26's apparatus includes a loop filter formed by first super-resolution enlarger 1803 and first resolution converter 1804 that enhances decoded base layer pictures from first decoder 1802. *Id.*, 51:38-52:26.

The first decoder's output is also passed directly to the second decoder. *Id.* Second decoder 1805 selects either non-super-resolution enhanced decoded base layer pictures or super-resolution enhanced decoded base layer pictures for inter-

layer predicting pictures in the enhancement-layer video stream; selection of which picture to use is based on reference picture selection information in the signal. *Id.*, 52:40-53:46.

In additional embodiments, non-super-resolution enhanced decoded base layer pictures can be upscaled to higher spatial resolution by second resolution converter 1806.

B. Prosecution History

In each Office Action during prosecution, the claims were rejected as obvious over combinations based on Phek (Ex.1005) and additional references. The Examiner repeatedly found Phek disclosed most limitations but relied on additional references to disclose the "first" and "second" "resolution converter" limitations. Ex.1002, 85-86, 88-91, 161-162, 165-167, 234-235. For example, although the Examiner stated Phek "does not explicitly disclose a first resolution converter," the Examiner nonetheless observed "use of resolution converters…was well known…as evidence[d] from the teaching of Holcomb." Ex.1002, 89; *see also id.*, 85. The applicant never disputed the Examiner's finding in this regard. The Examiner likewise acknowledged Phek and Holcomb "do not explicitly disclose a second resolution converter" and relied on Barkley's "spatial enhancement layers."

Ex.1002, 90-91. The applicant also never disputed Barkley's teachings in this regard.

In the final rejection before allowance, the Examiner asserted Barkley taught using two sets of reference pictures, as the then-pending language of the "second decoder" limitation recited (Ex.1002, 91), asserting Barkley's decoder "selectively decode[s] video information from the base layer and one or more enhancement layers to playback or reproduce the video information at a desired level of quality" and met the reference picture selection requirement. *Id.* In response, applicant amended claim 23 to recite that the "second decoder" must further "select one of the set of first reference pictures and the second set of reference pictures ***based on reference picture selection information***." Ex.1002, 65-66. Applicant argued "*Barkley* (newly-cited) only encodes information on a plurality of layers" and "*Barkley's* device fails to disclose…selecting one of two kinds of reference pictures and fails to suggest using any such reference picture selection information." Ex.1002, 62 (emphasis original). The Examiner subsequently allowed the application and identified the "second decoder" limitation (including the "reference picture selection information") in reasons for allowance. *Id.*, 17-18.

The Examiner did not consider Phek's teachings in combination with Segall, Martins, He, and Wang.

C. POSITA

A person of ordinary skill in the art ("POSITA") would have had a bachelor's degree in computer science, electrical engineering, computer engineering, or related field, with two years' experience in digital video processing, or related field. Ex.1003 ¶¶137-139.

D. Claim Construction

The Challenged Claims' subject matter is rendered obvious under any reasonable construction, including the plain meaning.[5] Petitioner reserves the right to respond to constructions offered by Patent Owner or adopted by the Board.

IV. Ground 1: Phek, Segall, Martins, and He Render Obvious Claims 1-6 and 8-11

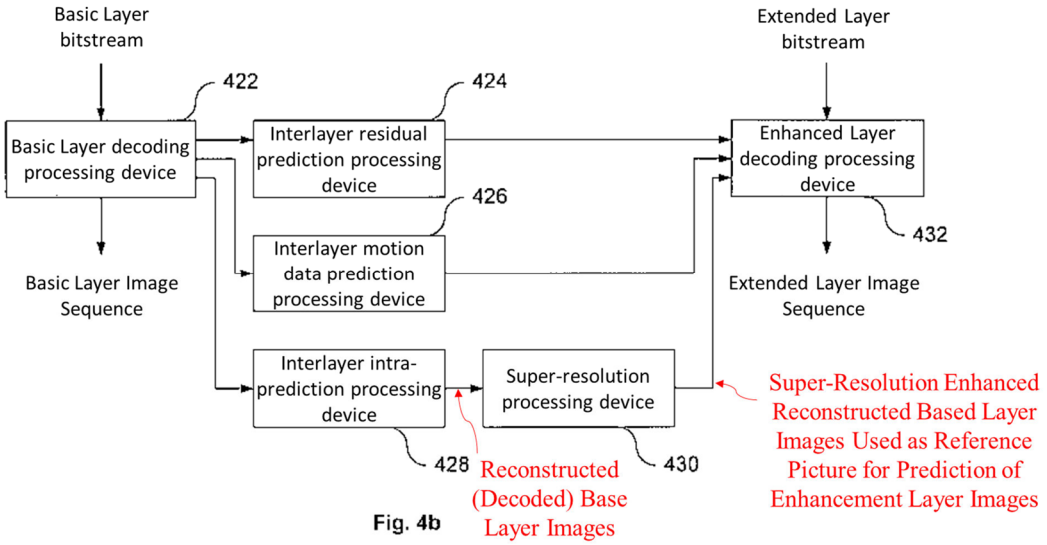
A. Prior Art

1. Phek

Phek's video coding system uses "[i]nterlayer prediction…for hierarchical encoding in order to improve encoding efficiency." Ex.1005, [0014]. "One example of hierarchical encoding that uses interlayer prediction is…scalable video encoding." *Id.* Phek's interlayer coding system includes basic and enhanced

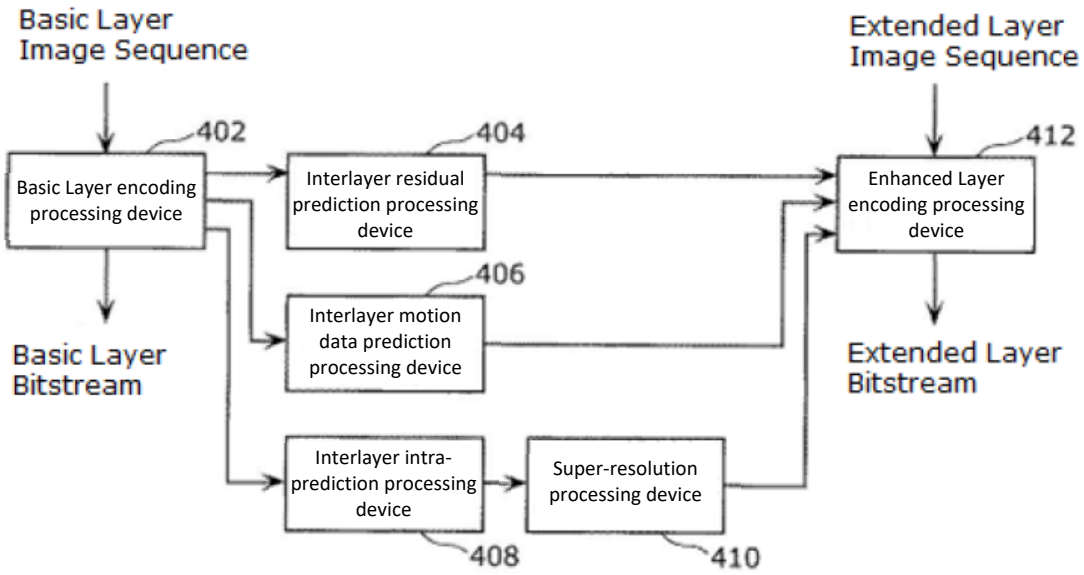
[5] Petitioner reserves the right to argue alternative constructions in other proceedings, and where such a defense is available, that the claims are indefinite.

layers, where "an intra image of the basic layer or a reconstructed image of the basic layer" is used "as an additional reference image for encoding the image of the enhanced layer." *Id.* "Super-resolution processing is used to improve these additional reference images" from the basic layer. *Id.*



Ex.1005, FIG. 4b (decoding)[6]; Ex.1003, ¶204.

[6] Blank boxes in Phek's figures are identified using the certified translation's labels provided external to the boxes. *See* Ex.1005. Translations (Ex.1005, Ex.1031) use "base" and "basic" layers (or streams) interchangeably, and use "extended," "enhanced," and "enhancement" layers (or streams) interchangeably. Petitioner cites/quotes from the certified translation that uses different terms, but typically



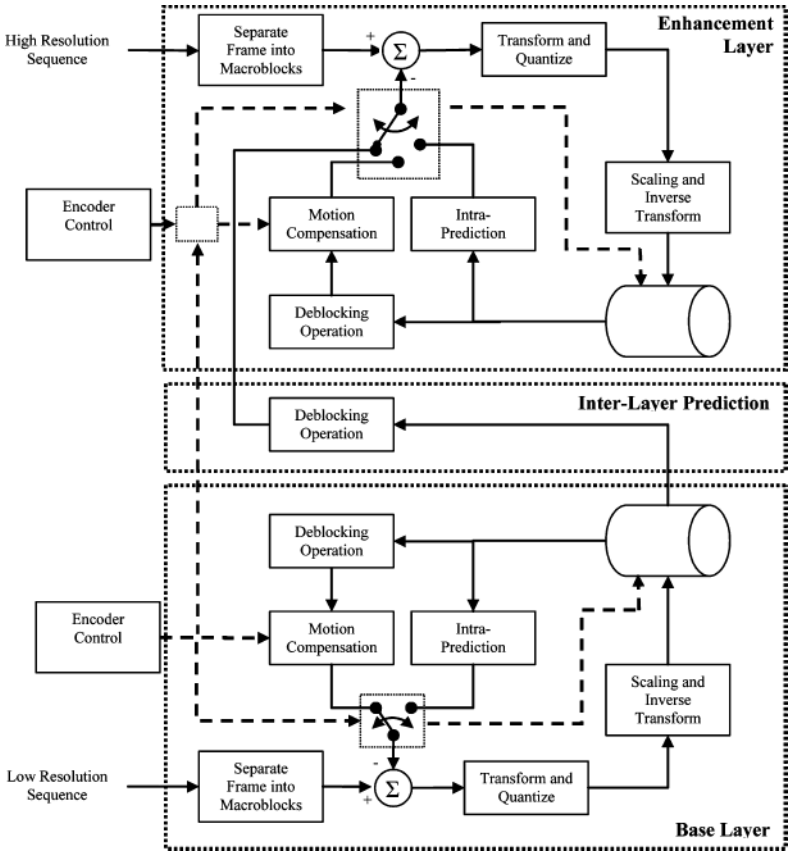
Ex.1005, FIG. 4a (encoding); Ex.1003, ¶305.

2. Segall

Segall surveys the scalable video coding (SVC) extension that "provide[] support for multiple display resolutions within a single compressed bit stream (or in hierarchically related bit streams), which is referred to [] as *spatial scalability*." Ex.1006, 1 (emphasis original). Segall describes spatial scalability with reference to a layered coding system like Phek's:

---

refers simply to "base" and "enhancement" layers (or streams) consistent with common usage of these terms. Ex.1003, FN12.

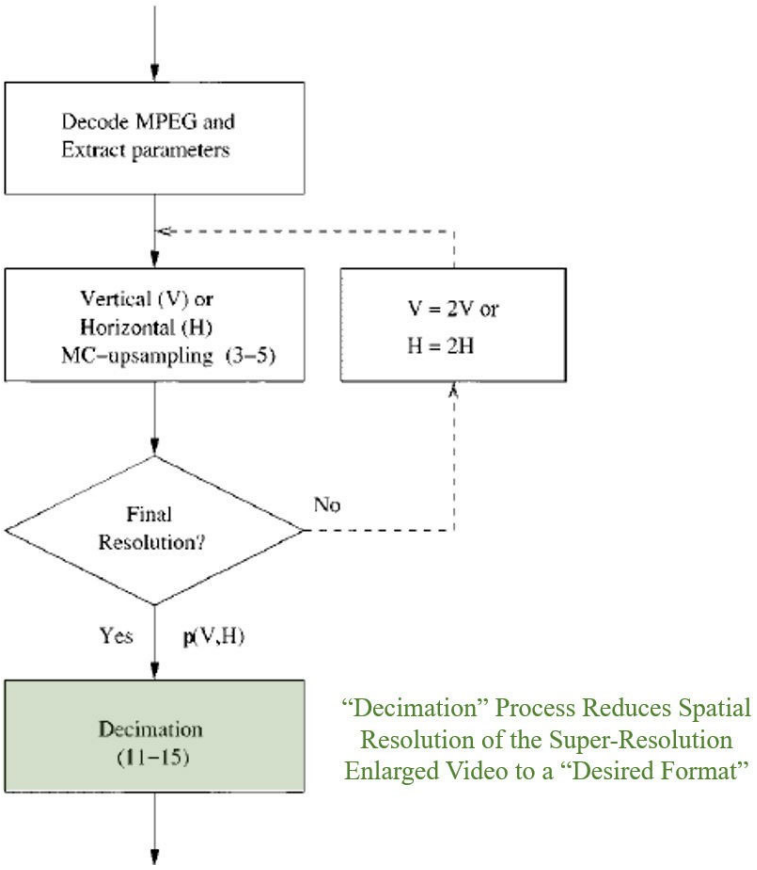


Ex.1006, FIG. 1.

Segall's "'base layer' encoder takes a lower resolution video sequence as input and encodes it." Ex.1006, 2. "[E]nhancement layer encoder takes a higher resolution sequence as input" and encodes it. *Id.* "[I]nter-layer prediction can be used to provide additional coding choices" and "[f]or the case of intra-picture coded blocks in the base layer, reconstructed intensities provide a prediction for the enhancement layer." *Id.* "For the case of inter-picture coded blocks in the base layer, enhancement layer motion vectors and residual difference information can be predicted from the base layer." *Id.*; Ex.1003, ¶¶205-206.

3. Martins

Martins teaches "quality and spatial resolution of video can be improved by combining multiple pictures to form a single super-resolution picture." Ex.1007, 1. "[S]uper-resolution pictures obtained by the algorithm are of much higher visual quality and have lower [mean-squared error] than super-resolution pictures obtained by simple spatial interpolation." *Id.* "The algorithm is based on motion-compensated spatial upsampling from multiple images and decimation to the desired format." *Id.* FIG. 2 illustrates Martins' algorithm, including a "decimation" process downscaling super-resolution enlarged images to a desired lower spatial resolution:



Ex.1007, FIG. 2; *see also* Ex.1003, ¶207.

4. He

He discloses selective in-loop filtering in SVC systems and selective in-loop deblock filtering controlling deblock filtering parameters at each layer of scalable coding hierarchy. Ex.1008, [0002]-[0006], FIGS. 1-4; Ex.1003, ¶¶208-211.

"The deblocking filter enhances the decoded picture quality and increases the encoding efficiency…," but "also consumes a significant amount of power and processing cycles." Ex.1008, [0006]. "Disabling the entire deblocking filter [] provides a significant reduction of computational complexity and power

consumption." *Id.*, [0028]. He examines "content of the video information…to determine whether to disable the deblock filtering," e.g., based on "a characteristic indicative of a relative amount of content complexity and/or a relative amount of motion" and "power savings considerations." *Id.*, [0028], [0037].

Decoder 200 implements selective deblock filtering. Ex.1008, [0032], FIG. 2. Decoder 200 receives coded bitstream (BTS) and performs prediction and reconstruction to decode at least a picture portion, generating "unfiltered reconstructed frame information URF'." *Id.*, [0032]. URF' is provided to deblocking filter 211, controlled by deblock control circuit 221. *Id.*, [0033]-[0034]. Circuit 221 issues "decoder filter control DFC signal" that selectively enables/disables filter 211. *Id.*, [0034]. When disabled, deblock filtering is bypassed and "URF' is passed as the 'filtered' reconstructed information RF' to reference frame buffer 213 for use as a reference picture for motion-compensated prediction processes. *Id.* When enabled, URF' is filtered to generate RF' for use as a reference picture in the motion compensated prediction processes. *Id.*

"[E]ncoder 100 provides control information within the bitstream BTS to cause the deblock control circuit 221 of the video decoder 200 to provide the DFC signal to adaptively disable deblock filtering for a corresponding slice or frame of URF' information being processed in the video decoder 200." *Id.*

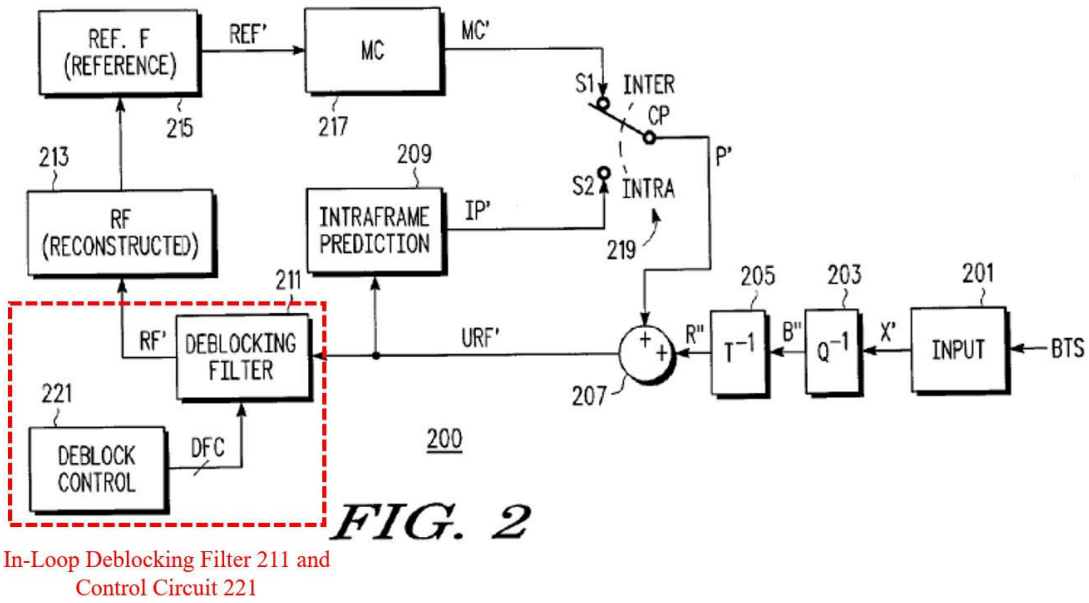


FIG. 2

Ex.1008, FIG. 2

Encoder 100 implements selective deblock filtering. Ex.1008, [0023], FIG. 1. Encoder 100 encodes at least a picture portion and outputs coded bitstream BTS containing that picture's "video information" and includes frame buffer 109 providing reference information (REF), "includ[ing] information from one or more previously encoded and decoded frames that have been further filtered, such as by deblocking filter 131." *Id.* URF is provided to deblocking filter 131 which, when enabled, "filters the unfiltered reconstructed frame information URF and provides filtered reconstructed information RF to an RF buffer 133." *Id.*, [0026]. If disabled, "the unfiltered reconstructed frame information URF is passed as the reconstructed information RF in which deblock filtering is bypassed." *Id.*

Disabling filter 131 reduces computational complexity and power consumption. *Id.*, [0028]. "If the video information content exhibits a relatively low amount of motion or is relatively simple content," "the deblocking filter 131 may be fully disabled without significant degradation in visual quality." *Id.*

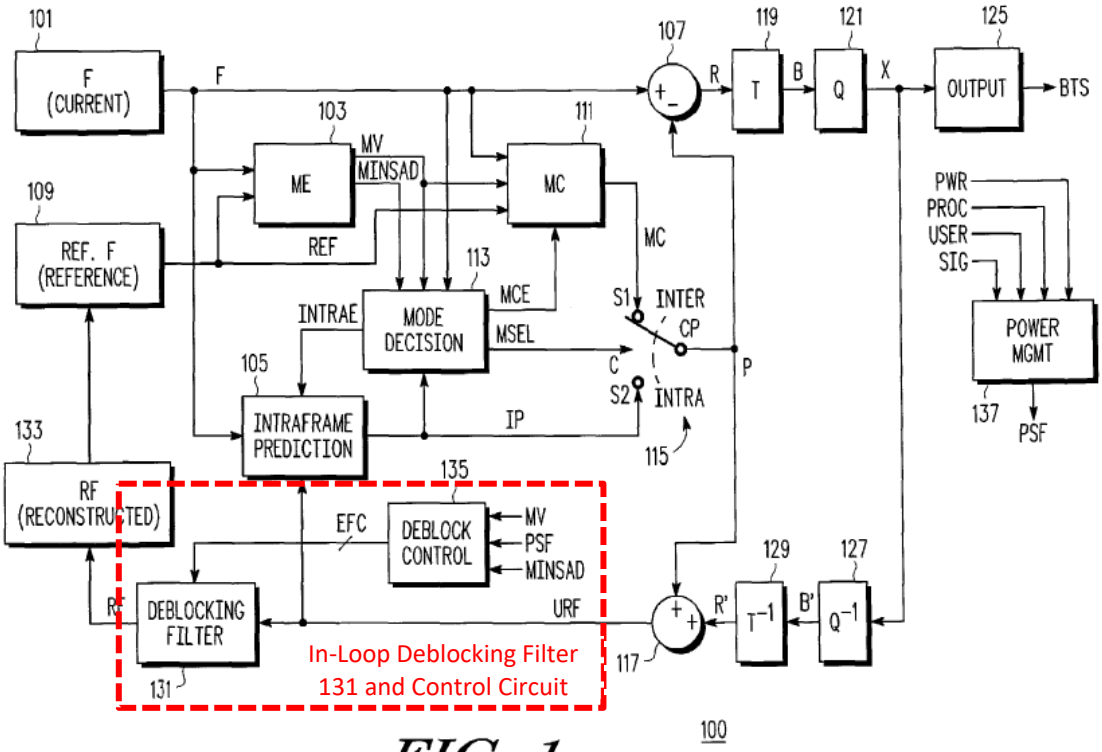


FIG. 1

Ex.1008, FIG. 1

5. Combining Phek, Segall, Martins, and He

Like the '995 patent, Phek discloses a layered video coding system that enhances reference picture quality in an enhancement layer by super-resolution processing. *Supra*, §IV.A.1. Phek details its super-resolution algorithm, but Phek's layered coding system description itself (which uses the proposed algorithm) is less

detailed. Phek's encoder/decoder descriptions are "simplified" and "do not have to be performed as shown" (Ex.1005, [0013]); thus its system is ripe for improvement. Ex.1003, ¶213.

The following subsections address straightforward obvious modifications to Phek's system in view of Segall, Martins, and He.[7] Ex.1003, ¶¶213-214.

a. *Obvious based on Segall to Encode/Decode Phek's Base and (First) Enhancement Layer Streams at the Same Spatial Resolution*

Phek's "hierarchical encoding" system "uses interlayer prediction" between video streams coded at base and enhancement layers. Ex.1005, [0014]-[0015], FIGs. 4a-4b. Phek does not expressly disclose (nor restrict) the type(s) of hierarchical coding that its system can implement or relative spatial resolutions of base- and enhancement-layer streams. *Id.* "One example of hierarchical encoding that uses interlayer prediction is…scalable video encoding." Ex.1005, [0014].

Segall teaches that SVC systems can implement "combinations of temporal scalability (frame rate enhancement) and quality scalability (fidelity enhancement for pictures of the same resolution)" as well as "spatial scalability" for spatial resolution enhancement. Ex.1006, 1. SVC allows "mixing" of layers including

[7] This overview is referenced and incorporated in the analysis of each claim element.

"layers that provide temporal or quality scalability." *Id.* Hierarchical coders can also be used for multi-view coding to encode/decode video streams corresponding to different viewpoints of a scene with the same spatial resolution. *Supra*, §III.A.1.

Given Segall, it would have been obvious to implement Phek's system to encode/decode base- and enhancement-layer streams at the same spatial resolution.[8] Ex.1003, ¶¶215-219. For example, it would have been obvious for Phek's first enhancement-layer[9][10] to implement quality scalability, temporal scalability, and/or

---

[8] Video streams encoded in base- and enhancement layers can be obtained, and optionally pre-processed, in any suitable manner, e.g., it would have been obvious for a stream captured by a low-resolution camera to be super-resolution enhanced before encoding (compression) to increase quality and detail of the encoded video. Ex.1003, FN14; Ex.1005, [0002] ("resolution and quality of images…can be improved"); Ex.1012, 1-2. The enhanced video stream can be encoded in the enhancement layer and a temporally or quality-decimated version can be encoded in the basic layer, consistent with common techniques. Ex.1003, FN14.

[9] A second enhancement layer is described in §IV.A.5.b.

[10] Regarding claims 1 and 5-10, the first enhancement-layer refers to the combination's single enhancement-layer.

multi-view coding—all of which can use hierarchical coding and interlayer prediction between streams of the same spatial resolution. *Id.* Quality scalability would have been beneficial to allow bitstream receivers to decode different quality versions depending on receiver capabilities, channel bandwidth differences, and transmission rates, which impact whether the receiver successfully receives a quality enhancement layer stream along with the base layer. Ex.1003, ¶217. Temporal scalability would have been beneficial to efficiently encode streams with different frame rates and to allow for higher frame rate decoding when receiving an enhanced stream. *Id.* And multi-view coding would have been beneficial to allow receivers to decode video from multiple viewpoints (e.g., allowing stereoscopic or 3D reproduction). *Id.*

Further, whether implementing quality scalability, temporal scalability, multi-view coding, or a combination thereof, it would have been obvious to code base- and enhancement-layer streams at the same spatial resolution—particularly when the spatial resolution of the base layer is adequate for an intended display that does not support higher resolutions. Ex.1003, ¶218. Segall teaches that "lower resolution displays…are [] popular in applications constrained by size, power and weight." Ex.1006, 1. Upscaling Phek's first enhancement-layer video stream in time and/or quality—but not spatially—relative to the base-layer stream would have allowed

lower-resolution displays to present a temporally- or quality-enhanced video despite the display's native/maximum spatial resolution limitations. Ex.1003, ¶218. Segall acknowledges "it is rarely justifiable to design a device with low display resolution with the capacity for decoding and down-sampling high resolution video material" since "[s]uch a requirement could increase the cost and power of the device" and results in "waste of [the device's] receiving channel bit rate." *Id.* Similarly, in multi-view coding, different viewpoints can be encoded at the same spatial resolution in different layers to allow for reproduction of the video stream in each layer on the same display device. Ex.1003, ¶218.

Coding Phek's base- and enhancement-layer video streams at the same spatial resolution also would have been obvious as a predictable application of known layered coding techniques to achieve predictable results. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) ("arrang[ing] old elements with each performing the same function it had been known to perform" is obvious). A POSITA would have had a reasonable expectation of success implementing Phek's system as described since layered coding systems commonly coded streams with the same spatial resolution. Ex.1003, ¶219.

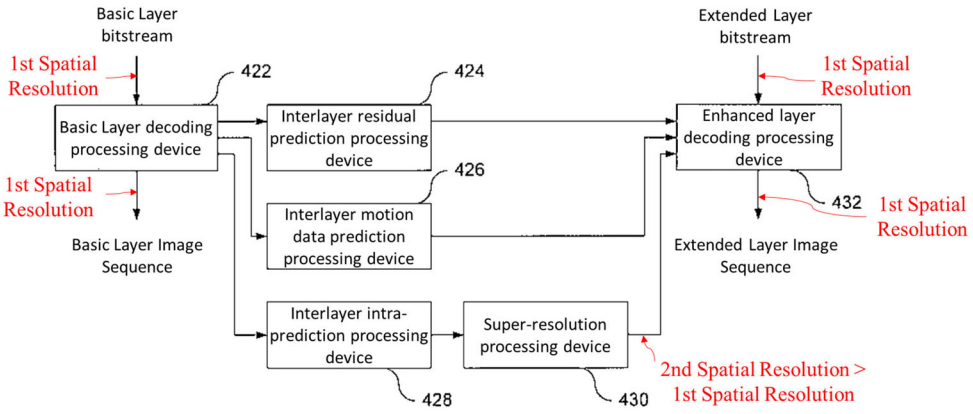


Fig. 4b

First Intermediate Illustration of Proposed Combination (Phek-Segall)

b. *Obvious based on Segall to Implement a Second Enhancement Layer in Phek's System for Spatial Scalability*

Segall teaches that SVC allows "mixing of spatial scalability layers with other layers that provide temporal or quality scalability." Ex.1006, 1-2; *supra*, §IV.A.2. It would have been obvious to implement a second-enhancement layer in Phek's system that codes a video stream with higher spatial resolution than video streams in the base- and first-enhancement layers. Ex.1003, ¶220. The second-enhancement layer can be a spatially scalable layer that encodes video whose frames are predicted from frame(s) in the base layer, first enhancement layer, or both. *Id.*

Encoding a stream with higher spatial resolution in the second-enhancement layer would have been obvious to support displays that can render video at higher resolutions than video in the base- and first-enhancement layers. Ex.1003, ¶221. Segall recognizes "resolution diversity of current display devices motivates the need

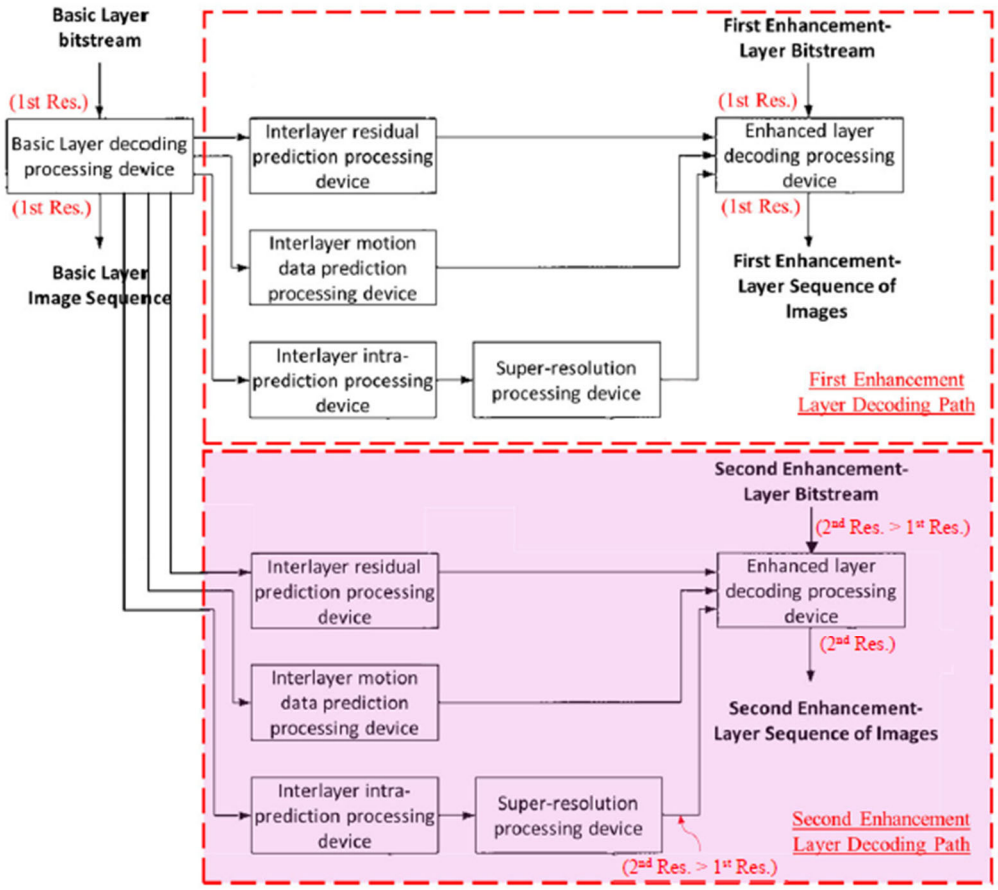
for spatial scalability" and notes "larger format, high definition displays are becoming common." Ex.1006, 1. Providing coded video streams at lower-resolution in base- and first-enhancement layers and a higher-resolution stream in a second enhancement layer furthers Segall's goal of supporting broader ranges of display devices through "resolution diversity." *Id.*

Adding a second-enhancement layer in Phek's encoder/decoder (including a second-enhancement layer decoder processor) would have been beneficial to allow higher-definition display devices to decode and display higher-resolution video. Ex.1003, ¶222. For example, a higher-definition display's decoder could switch from decoding a lower-layer stream to a higher-layer stream at higher spatial resolution when transmission rate of a streaming video increases sufficiently to allow the decoder to successfully receive and decode additional video data in the second enhancement layer. Ex.1006, 1.

Segall underscores this, teaching "[d]iverse, limited, and time-varying channel capacity provides an [additional] motivation for spatial scalability." *Id.* "Spatial scalability allows for the rapid bit rate adaptation" needed to address time-varying channel conditions. *Id.* Increased viewer satisfaction would be realized by allowing decoders to decode and display video at higher spatial resolutions during

favorable channel conditions, while allowing decoders to continue playing video from a lower-layer stream under degraded conditions. Ex.1003, ¶222.

Adding a second, spatial enhancement layer to Phek also would have been obvious as a predictable application of known spatial scalability techniques to achieve predictable results. *KSR*, 550 U.S., 417. A POSITA would have had a reasonable expectation of success given common usage of multiple enhancement layers. Ex.1003, ¶223. For example, Segall teaches that standard SVC techniques "support[ed] multilayer scenarios including special multiple spatial scalability layers and the mixing of spatial scalability layers with other layers that provide temporal or quality scalability." Ex.1006, 3. Further, "[c]onsiderable flexibility is [] provided in regard to the selection of the reference layer for each enhancement layer[]…." *Id.*; *see also* Ex.1011, 2 ("different types of scalability" can "be combined").



Second Intermediate Illustration of Proposed Combination (Phek-Segall)

c. *Obvious based on Segall to Multiplex/Demultiplex Phek's Base and Enhancement Layer Bitstreams*

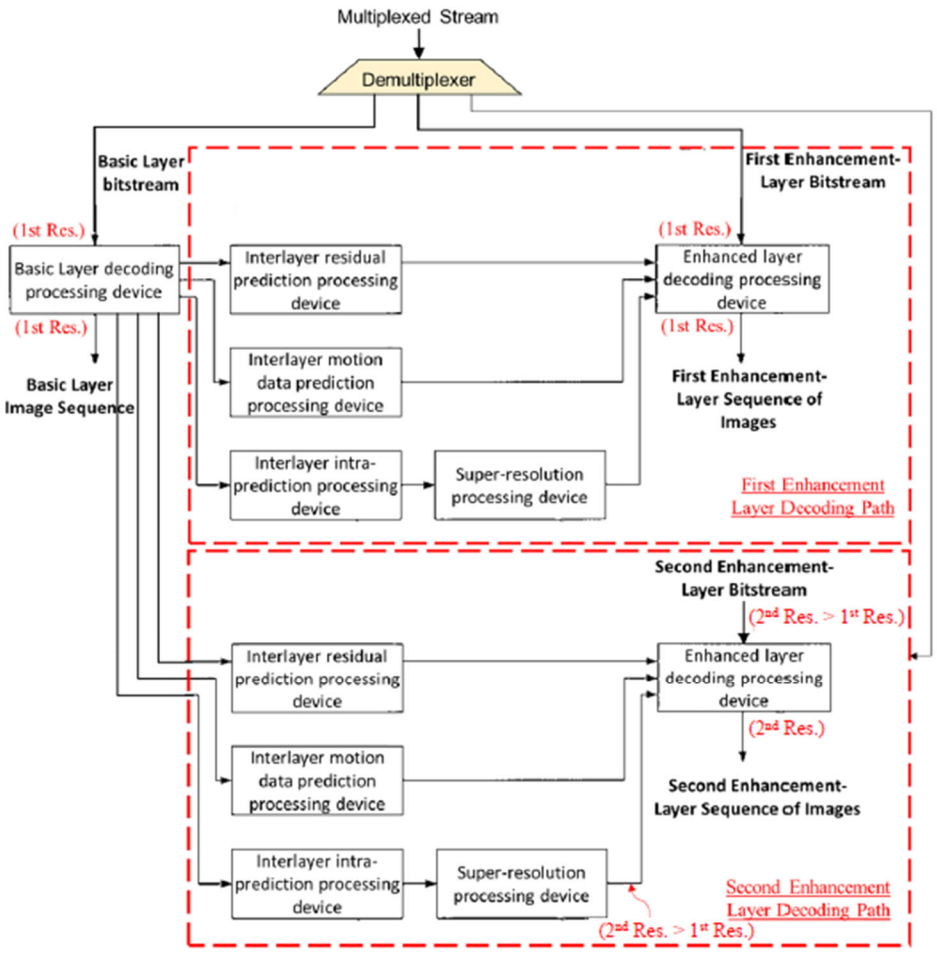
Phek does not expressly disclose how base- and enhancement-layer bitstreams are transmitted from encoder to decoder or vice-versa. It nonetheless would have been obvious to multiplex respective bitstreams from the base-, first-enhancement, and/or second-enhancement layers of the combination's encoder into a single, composite bitstream for transmission. Ex.1003, ¶¶224-226. The combination's hierarchical decoder would then predictably include a demultiplexer to parse

separate basic and enhancement-layer bitstreams. *Id.* Likewise, the combination's hierarchical encoding device would include a multiplexer to create a single bitstream comprising separate basic and enhancement-layer bitstreams. Multiplexing and demultiplexing bitstreams in the combination's layered coding system would have been obvious based on Phek's and Segall's teachings. *Id.*

Phek describes a "video source [that] is separated from the other sources using an inverse multiplexer 202." Ex.1005, [0010], FIG. 2. Phek then discloses that its hierarchical video encoder (FIG. 4a) produces separate bitstreams for base- and enhancement-layers, and that its hierarchical video decoder (FIG. 4b) separately processes base-layer and enhancement-layer bitstreams. *Id.*, [0014]-[0015], FIGs. 4a-4b. Segall teaches scalable video streams are treated "from a systems ***multiplex***…perspective" where "data might often be handled…as distinct hierarchically related streams of content that are coordinated using decoding timestamps or other such mechanisms." Ex.1006, 3. A POSITA would have understood Segall as teaching the known technique of multiplexing/demultiplexing of scalable or otherwise hierarchically coded bitstreams. Ex.1003,¶225 (citing Ex.1010, FIG. 12; Ex.1018, [0079], FIG. 10).

It would have been obvious and a POSITA would have been motivated to multiplex/demultiplex base- and enhancement-layer streams to support transmission

of respective streams over communication channels, while also facilitating stream synchronization by referencing common timing information in the multiplexed signal. Ex.1003, ¶226. A decoder-side demultiplexer also would have been obvious to separate streams to process each layer by portions of the decoder. *Id.* (citing Ex.1019, 1; Ex.1018, [0079]). Multiplexing/demultiplexing layered streams also would have been a predictable application of known techniques to yield predictable results for which a POSITA would have had a reasonable expectation of success, especially since multiplexing and demultiplexing layered bitstreams was commonly performed before the '995 patent. *KSR*, 550 U.S. at 417; Ex.1003, ¶226.



Third Intermediate Illustration of Proposed Combination (Phek-Segall)

d. *Obvious based on Martins to Downscale Phek's Super-Resolution Enlarged Reference Pictures*

It next would have been obvious given Martins to downscale super-resolution enhanced reference pictures that Phek's (first) enhancement layer uses for prediction to the first (standard) resolution at which images in the base- and first-enhancement layers are encoded/decoded. Ex.1003, ¶¶227-235. Phek's decoding path for the (first) enhancement layer includes "super-resolution processing device 430…used to

improve the reference picture from the inter-layer intra-prediction processing device 428." Ex.1005, [0015]. Using "super-resolution processing…improve[s] encoding efficiency or picture quality by improving the reference frames." *Id.*, [0013].

Phek's super-resolution enhancement involves processing lower-resolution images to generate a super-resolution enlarged image of higher quality and spatial resolution than individual lower-resolution images. Ex.1005, [0002]. The combination's base- and (first) enhancement-layer streams are both coded at the same spatial resolution. *Supra*, §IV.A.5.a. Subjecting reference pictures to super-resolution processing yields enlarged images whose spatial resolution is greater than the resolution of the decoding targets in the first enhancement layer stream. Ex.1005, FIG. 4b. In this context, it would have been obvious to apply Martins' decimation technique to revert Phek's super-resolution enlarged reference picture to the smaller spatial resolution of the base- and first enhancement-layer streams (while retaining SNR or detail gains from super-resolution enhancement). Ex.1003, ¶¶227-229; Ex.1007, §IV; *supra*, §IV.A.5.a.

Multiple reasons would have prompted a POSITA to follow Martins' suggestion to downscale super-resolution images back to the original (standard) resolution such that reference pictures that Phek's (first) enhancement layer uses for

interlayer prediction would be downscaled super-resolution images.[11] Ex.1003, ¶¶231-232.

First, a POSITA would have sought to downscale Phek's super-resolution reference images to reduce computational expense and complexity of interlayer predictions of (first) enhancement layer images. Ex.1003, ¶232. A POSITA would have recognized decoding higher-resolution reference pictures generally requires more processing time/resources and power than decoding lower-resolution reference images—a fact that explains multiple proposals for reducing reference image size. A POSITA would have sought to downscale Phek's super-resolution reference images to obtain similar benefits. Ex.1003, ¶232 (citing Ex.1021; Ex.1020; Ex.1013).

---

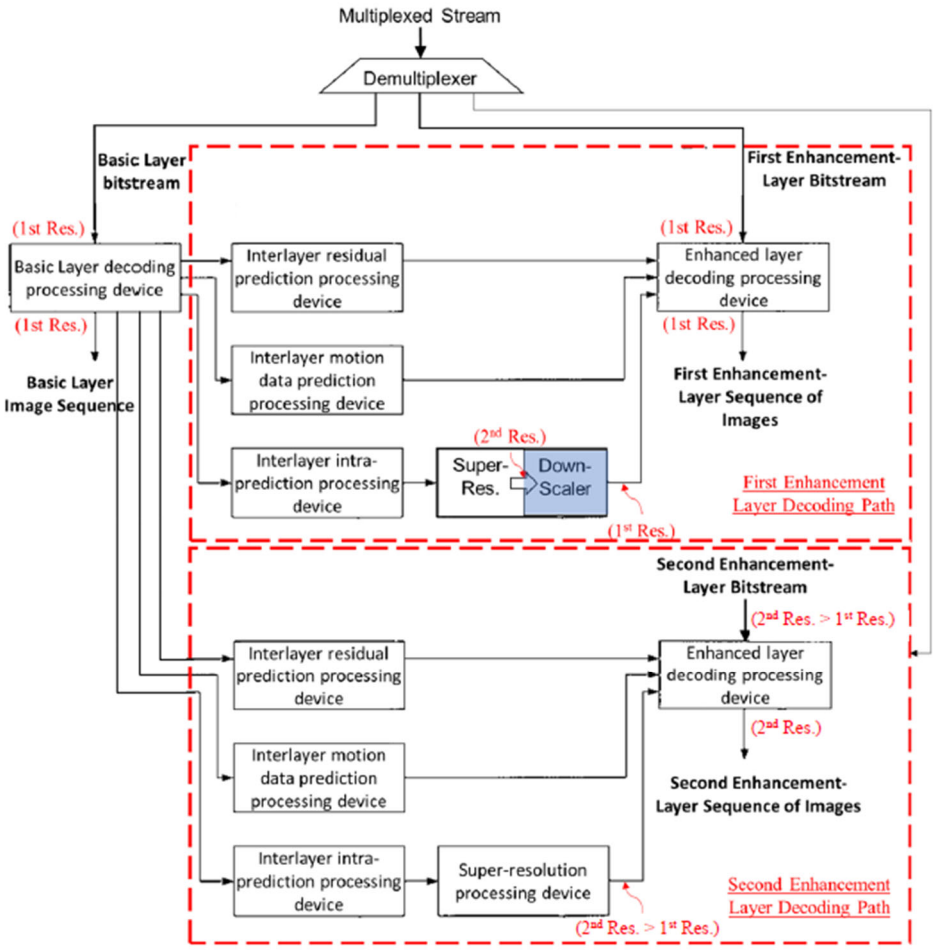
[11] In combination, it would have been obvious for super-resolution processor 430 in the (first) enhancement layer to generate super-resolution images according to the Phek or Martins techniques. Ex.1003, FN16. Both references' super-resolution processes produce higher quality and higher resolution images than the lower-resolution input images. *Id.* Each process likewise meets the claimed requirements for super-resolution enlargement. Post super-resolution downscaling can then be performed in any suitable manner including, e.g., Martins' decimation process. *Id.*

Second, a POSITA would have sought to downscale Phek's super-resolution reference images to reduce memory amounts required to buffer or store reference images for decoding. Ex.1003, ¶233. Reference images are commonly stored in a frame buffer where they are made available to the decoder for downstream image prediction. Ex.1003, ¶233; Ex.1009, FIG. 2 (buffers 213, 215). Frame buffer size is proportional to reference image size and it would have been obvious to reduce buffer size by reducing reference images' resolution before storing. Ex.1003, ¶234. Frame buffer size was a well-known concern, especially since larger buffers generally require more expensive hardware, processing, and power consumption to maintain. Ex.1003, ¶233 (citing Ex.1009, 3:21-28).

Third, a POSITA would have sought to downscale Phek's super-resolution reference images to separately allow the base-layer stream to be output for display with improved picture quality at the first (standard) spatial resolution that does not require a higher resolution, more expensive display device. Ex.1003, ¶234. Phek expressly notes that super-resolution images are usable as reference images in a decoding loop and can be displayed. Ex.1005, [0002], [0012]-[0013]. Martins likewise describes how decimation can be applied to "obtain a sequence of higher fidelity…in the same format as the directly decoded one." Ex.1007, 1. Downscaling Phek's super-resolution images would have allowed rendering those

images at a lower resolution display or even on devices that support higher resolutions during time periods when the decoder is unable to obtain enhancement layer data, e.g., due to degraded channel conditions. Ex.1003, ¶234; Ex.1006, 1.

Fourth, downscaling super-resolution enlarged pictures would have been obvious as merely applying known techniques to yield predictable results, and a POSITA would have had a reasonable expectation of success, given the known use of decimation and downscaling techniques on super-resolution images. *KSR*, 550 U.S., 417; Ex.1003, ¶235.



Fourth Intermediate Illustration of Proposed Combination (Phek-Segall-Martins)

e. *Obvious based on He to Selectively Apply Phek's/Martin's Super-Resolution (and Downscaling) Loop Filter*

A POSITA would have next found it obvious based on He to implement Phek's (first) enhancement layer decoder to selectively enable or disable super-resolution processing and downscaling reference pictures in the combination's (first) enhancement layer. Ex.1003, ¶¶236-244.

As discussed, the Phek-Martins teachings render obvious performing super-resolution processing and downscaling reference images used for inter-layer predicting images in the (first) enhancement layer. Ex.1005, [0011]-[0015], FIGS. 3a-3b, 4a-4b; Ex.1007, 1-2, 6; *supra*, §IV.A.5.d.

Super-resolution and downscaling operations performed on reference pictures in the (first) enhancement layer decoding path form a type of "loop filter," referred to herein as a "super-resolution and downscaling loop filter" ("SR-D filter"). Ex.1003, ¶238. Like other known loop filters, the Phek-Segall-Martins-He SR-D loop filter enhances decoded frames before being used as references for predicting frames in the first enhancement layer. *Supra*, §IV.B.6-7; Ex.1022, 10-11; Ex.1008, [0006]. As Dr. Reader explains, many loop filters were known and the combination's SR-D filter similarly serves to enhance reference picture quality consistent with other known filters. Ex.1003, ¶238 (citing Ex.1008, [0006]; Ex.1015, [0068]; Ex.1011, [0022]-[0023], FIGS. 2-5); *supra*, §IV.B.6-7.

Phek does not explicitly disclose criteria or conditions for using super-resolution processing on reference pictures, nor does Martins explain when SR-D filtering should be applied. As discussed above, He describes known costs of loop filtering, which may or may not be justified depending on factors, e.g., degree of decoded image quality and compression improvements that would be realized by

using the loop filter, power consumption, etc. *Supra*, §IV.A.4; Ex.1003, ¶¶239-241. He explains that an in-loop "deblocking filter…consumes a significant amount of power and processing cycles" and "[g]enerally…the entire deblocking filter…consumes approximately 8% of the computations at the encoder and approximately 35% at the decoder for H.264/AVC." Ex.1008, [0006]. "Disabling the entire deblocking filter, therefore, provides a significant reduction of power and processing complexity, especially in the decoder." *Id.* To maintain a "desired tradeoff between video quality and complexity reduction," He discloses control signals that allow decoders to selectively enable/disable the loop filter. *Id.*, [0069].

A POSITA would have found it obvious and would have been motivated to apply He's suggestion for selective in-loop filtering to the Phek-Segall-Martins combination to achieve benefits like those He expressly describes—e.g., reduce power consumption and computational complexity of decoding frames in at least the (first) enhancement layer. Ex.1008, [0028], [0031], [0037], [0076], [0069]; Ex.1003, ¶241. In the combination, based on He, control flags in frame or slice header(s) can predictably indicate whether in-loop SR-D filtering should be applied when decoding a given frame or slice in the (first) enhancement layer. *Id*.

Although He describes selective loop filtering in the context of "deblocking" filters, He's teachings are readily applicable to other loop filter types including SR-

D filtering. Ex.1003, ¶242. Indeed, the same concerns that motivated He's use of selective "deblocking"-type loop-filters—e.g., computational complexity and power consumption—are similarly, if not more, applicable to SR-D filtering since SR-D filtering involves processing of more images and imposes an even greater computational burden than traditional deblocking filters. Ex.1003, ¶242; Ex.1007, 6 (describing need to "keep the algorithmic complexity down"); Ex.1012, 13 (describing importance implementing "efficient [super-resolution] algorithm that reduces the computational cost").

A POSITA also would have known that super-resolution processing is most effective to improve quality of an image when "some relative scene motions…exist from frame to frame." Ex.1011, 2; Ex.1003, ¶243. But in the absence of motion, or if lower-resolution images from which the super-resolution image is constructed are merely "shifted by integer [pixel] units," then "no new information can be used to reconstruct" a super-resolution image and super-resolution processing will be less effective. Ex.1011, 2; Ex.1003, ¶243. Accordingly, when super-resolution processing does not improve the reference picture significantly, imposes an unacceptable computational load increase, and/or when power consumption required for super-resolution processing is substantial, it would have been obvious to bypass SR-D filtering based on He's selective in-loop filtering teachings. Ex.1003, ¶243.

For example, consistent with He's use of unfiltered reference pictures when loop filtering is disabled, it would have been obvious for the combination's (first) enhancement layer decoder/encoder to use a decoded base layer-frame that has not been SR-D filtered for inter-layer prediction of a corresponding frame in the first enhancement layer when the SR-D filter is disabled. *Id*.; Ex.1008, [0026] ("unfiltered reconstructed frame information URF is passed as the reconstructed information RF" when loop filtering disabled), FIG. 2.

The fact that selective loop filtering was standard in video coders and that He describes such techniques in a layered coding system like Phek's underscores the obviousness of the combination and that a POSITA would have had a reasonable expectation of success implementing selective SR-D loop filtering. Ex.1003, ¶244; *KSR*, 550 U.S., 417.

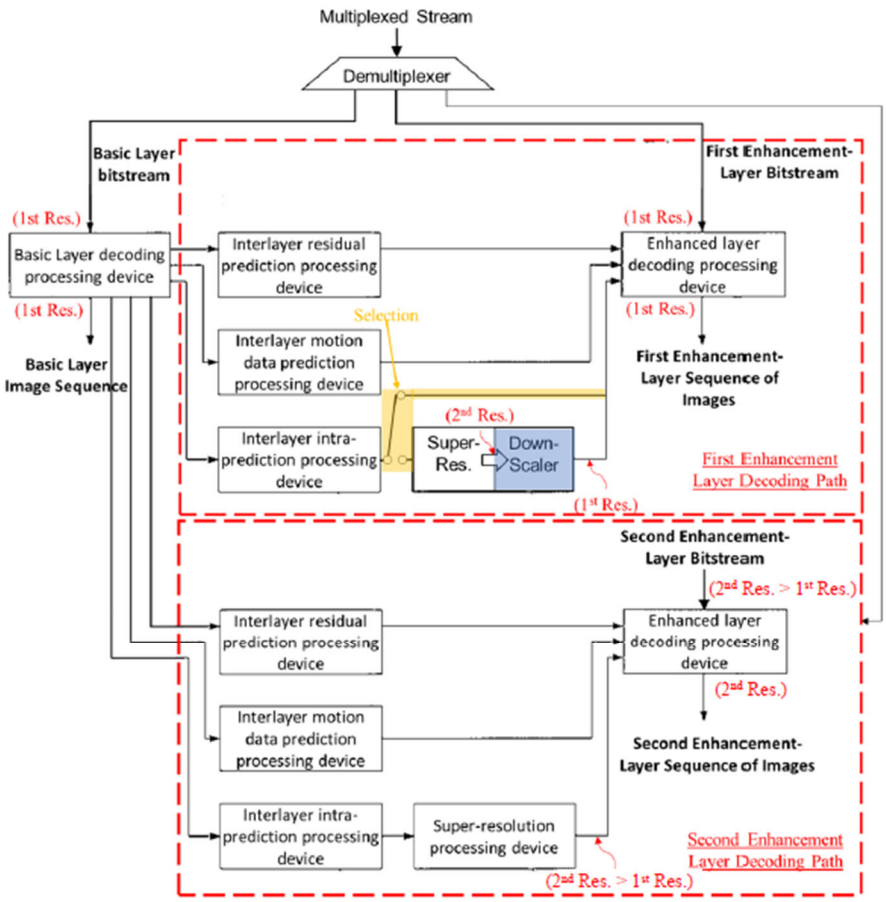
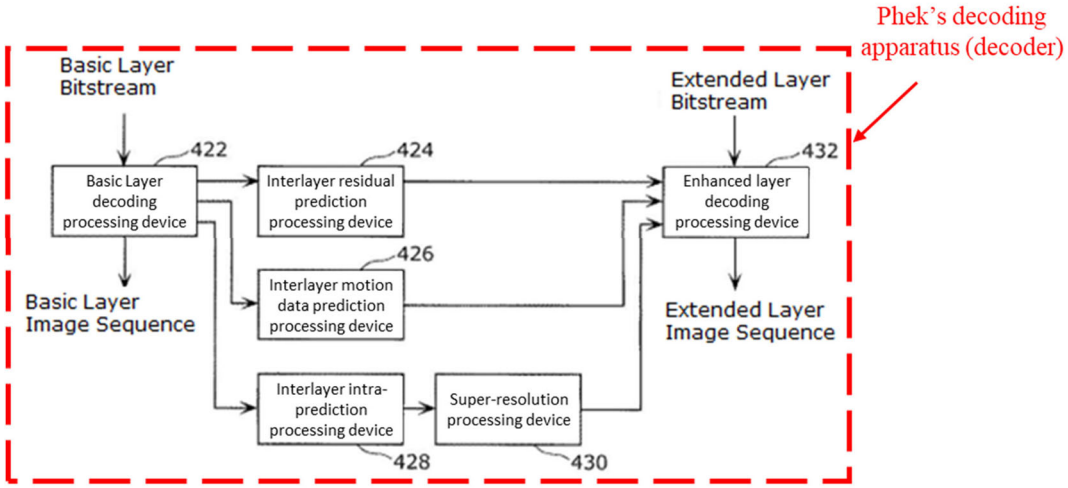


Illustration of Phek-Segall-Martins-He Combination

B. Claim 2

1. [2.P]

Phek's FIG. 4b shows a "decoding device (decoder)" in an "[i]nterlayer predictive" coding system. Ex.1005, [0014], FIG. 4b; *see also id.*, [0011]-[0015], FIGs. 3a-3b, 4a. In a decoder, an "image sequence is reconstructed from the encoded video bitstream." Ex.1005, [0012].



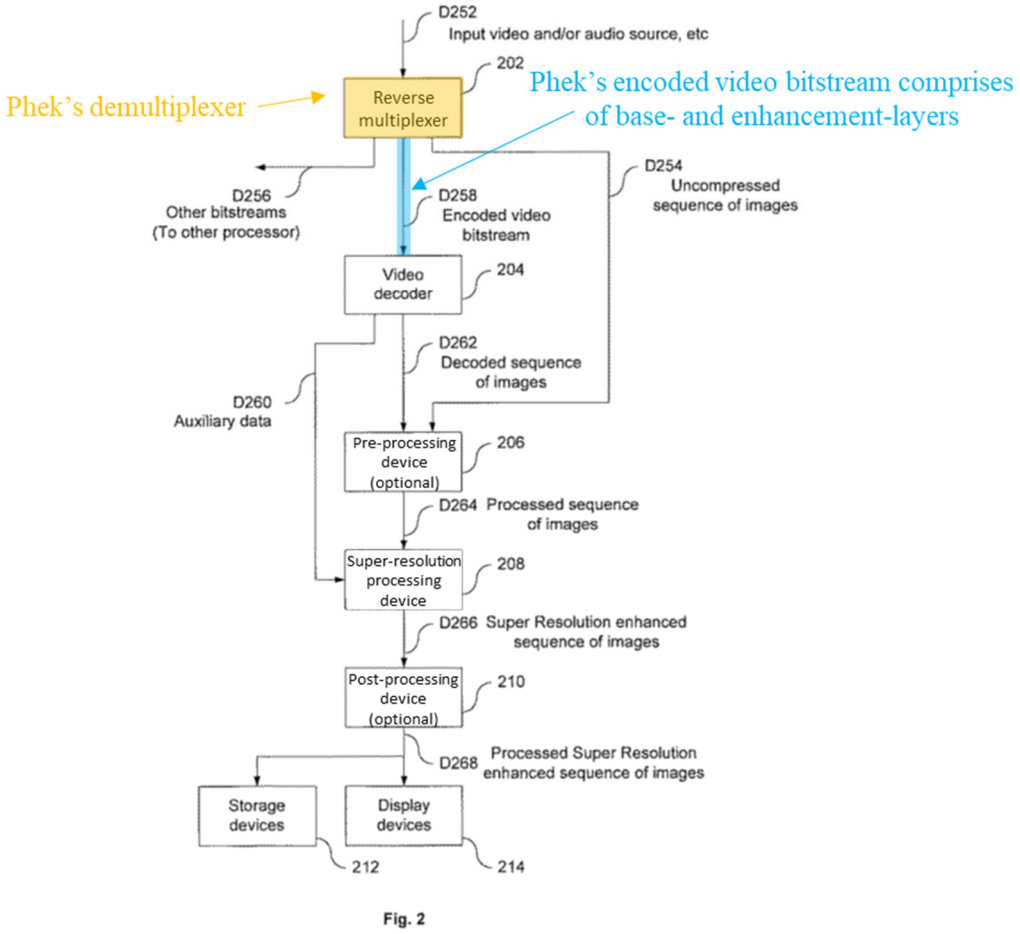
Ex.1005, FIG. 4b (modified).

Phek's FIG. 4b decoder includes (i) basic layer decoding processing device 422 that processes a basic-layer bitstream to decode a basic layer image sequence and (ii) enhanced-layer decoding processing device 432 that processes an enhancement-layer bitstream to decode an enhancement-layer image sequence. Ex.1005, [0014]-[0015], FIG. 4b. The decoder mirrors the encoder's operations in "us[ing] interlayer prediction to reduce redundancy between the enhanced layer and the basic layer, thereby improving encoding efficiency compared to encoding performed using only an enhanced layer." *Id.*, [0014]; *see also supra*, §IV.A.1; Ex.1003, ¶¶246-247.

2. [2.1]

Phek's "reverse multiplexer 202" (also referred to as "inverse multiplexer" or "demuxer") receives "input D252 [] composed of video and/or audio sources, etc."

and processes input D252 to "separate[]" "the video source…from the other sources." Ex.1005, [0010], FIG. 2; Ex.1003, ¶¶248-249.



Ex.1005, FIG. 2 (modified).

Further, Phek discloses that the encoder (FIG. 4a) encodes separate bitstreams for the basic- and enhancement-layers, while the decoder (FIG. 4b) processes the basic-layer and enhancement-layer bitstreams in respective decoding paths. Ex.1005, [0014]-[0015], FIGS. 4a-4b. A POSITA thus would have found it obvious

that when implementing layered hierarchical coding, Phek's encoded video bitstream D258 (FIG. 2) would include both the basic- and enhancement-layer bitstreams as FIGs. 4a-4b show. Phek does not expressly disclose multiplexing basic- and enhancement-layer bitstreams. Ex.1003, ¶249.

Segall teaches that scalable video streams are treated "from a systems ***multiplex***…perspective" in which "data might often be handled…as distinct hierarchically related streams of content that are coordinated using decoding timestamps or other such mechanisms." Ex.1006, 3. Segall acknowledges a POSITA would have known that scalable and hierarchical video bitstreams are commonly multiplexed into a composite signal for transmission and demultiplexed upon receipt. Ex.1003, ¶250 (citing Ex.1010, FIG. 12; Ex.1018, [0079], FIG. 10); *supra*, §IV.A.2.

In combination, Phek's FIG. 4a encoder multiplexes respective basic- and (first) enhancement-layer bitstreams into a single composite signal for transmission, and further multiplexes a second enhancement-layer bitstream into the composite signal when a second enhancement layer is provided as described above. Ex.1003, ¶251; *supra*, §IV.A.5.c. A demultiplexer (e.g., Phek's demultiplexer 202) implements a process for prescribed demultiplexing of the composite signal and outputs respective bitstreams for processing by corresponding layers of the decoder.

Id. The basic-layer bitstream constitutes a "***first sequence of encoded bits***" while the (first) enhancement-layer bitstream constitutes a "***second sequence of encoded bits***." *Id.*

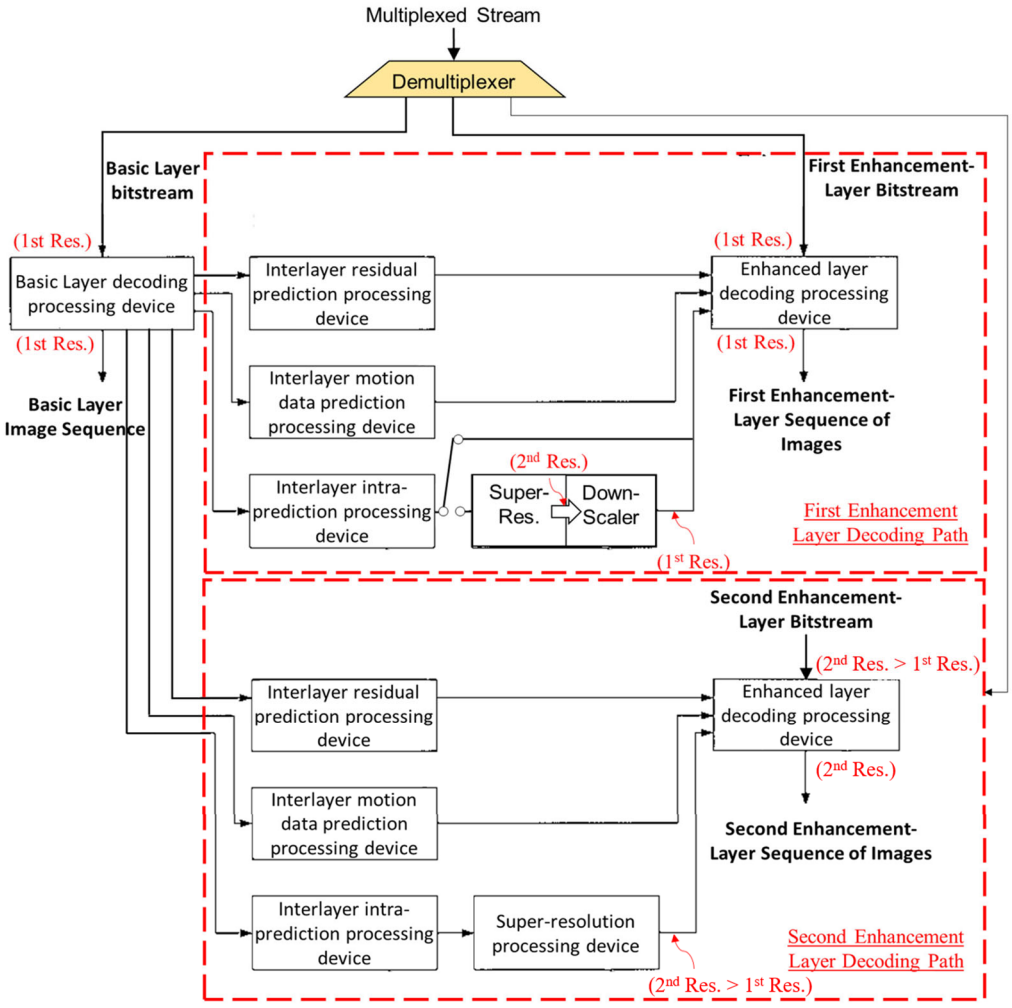


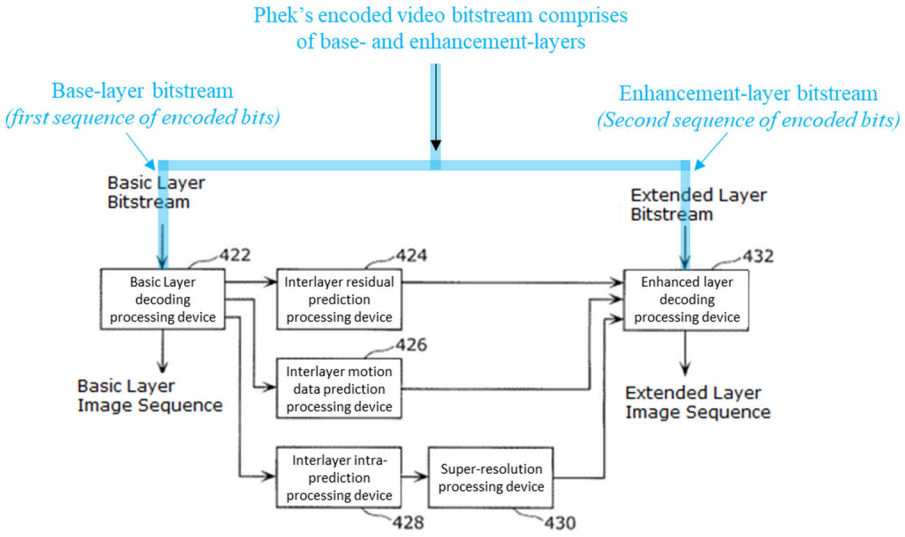
Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶251.

3. [2.2]

Phek's FIG. 4b discloses a decoder for hierarchically encoded video streams configured to decode image sequences from each of a basic-layer bitstream and enhancement-layer bitstream. Ex.1005, [0014]-[0015], FIG. 4b. "In the decoder, the basic layer bitstream is decoded using a basic layer decoding processing device 422." *Id.*, [0015]; Ex.1003, ¶¶252-253.

In combination, basic layer decoding processing device 422 ("***first decoder***") acquires the basic-layer bitstream ("***first sequence of encoded bits***") from the demultiplexer to implement thereon a process for a prescribed first decoding to create a basic layer image sequence ("***sequence of decoded pictures***"). Ex.1003, ¶253; *supra*, §IV.A.5.c. The basic layer video stream is encoded and decoded with a first spatial resolution ("***standard resolution***"). *Supra*, §IV.A.5.a; Ex.1003, ¶253.



Ex.1005, FIG. 4b.

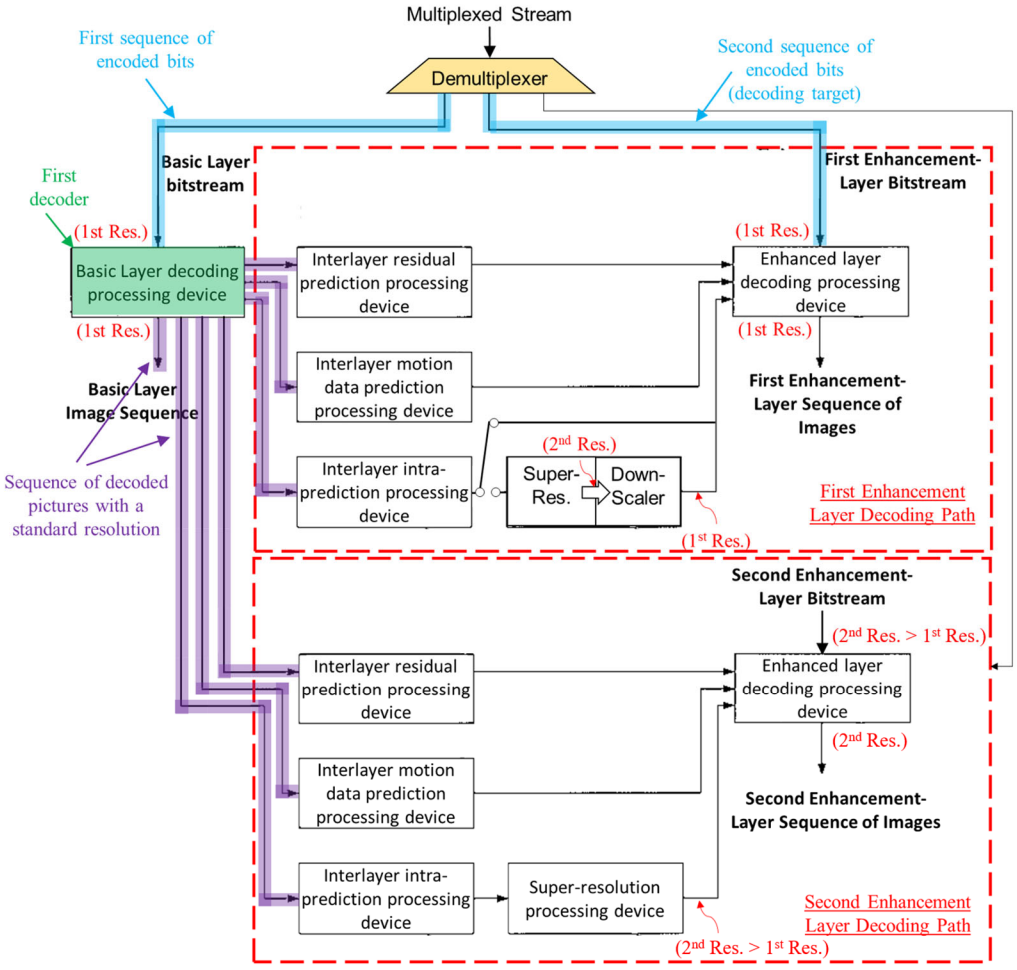


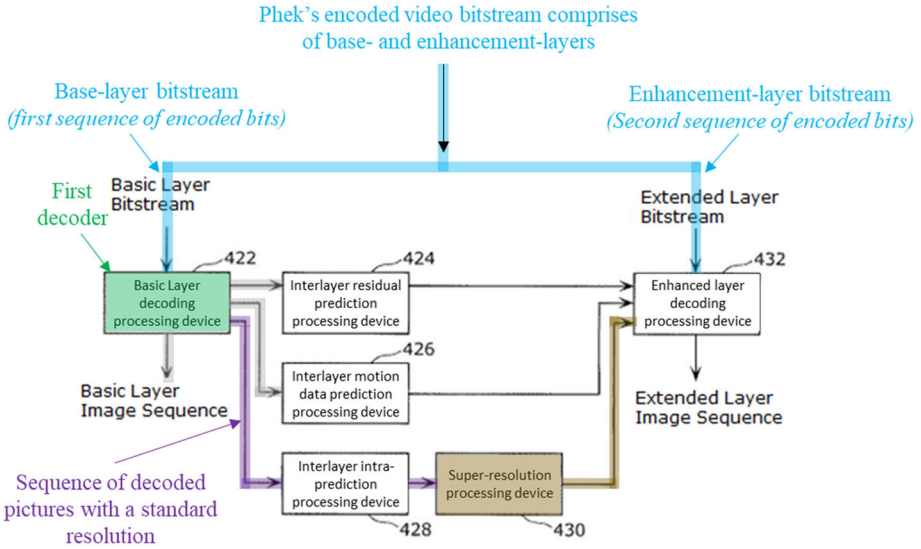
Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶253.

4. [2.3]

Phek's FIG. 4b includes "super-resolution processing device 430…used to improve the reference picture from the inter-layer intra-prediction processing device 428" and "performs residual interpolation." Ex.1005, [0015], FIG. 4b. Ex.1003, ¶¶254-258. "[S]uper-resolution processing…improve[s] encoding efficiency or

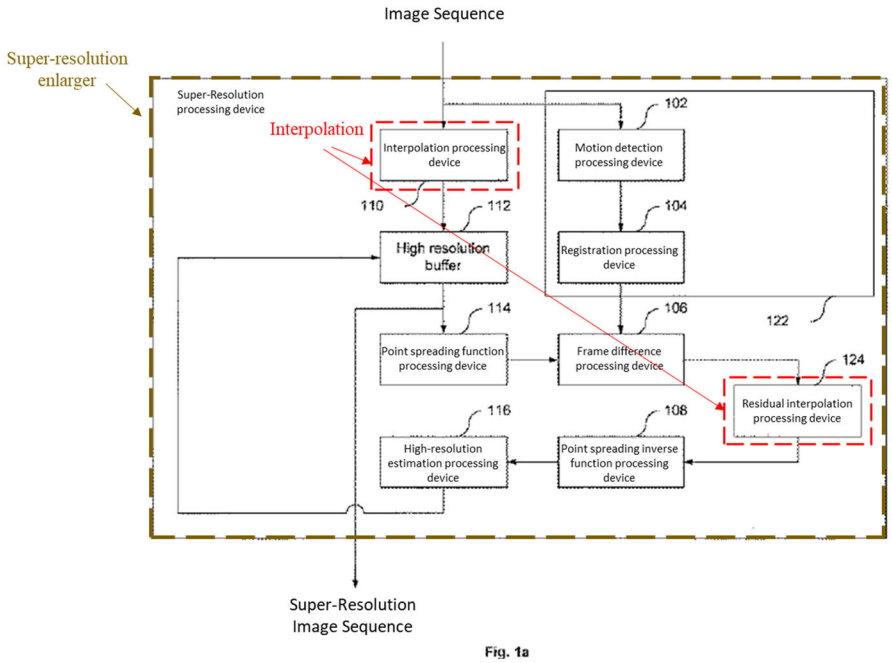
picture quality by improving the reference frames used for motion detection, motion compensation prediction, and display." Ex.1005, [0013], [0014] ("uses…a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer"). Super-resolution processing device 430 thus constitutes a "***first super-resolution enlarger***" configured to acquire the basic layer image sequence ("***sequence of decoded pictures created with a standard resolution***") at basic layer decoding processing device 422 ("***first decoder***") to work on the basic layer image sequence to implement interpolation of pixels with a first enlargement to create a sequence of "***super-resolution enlarged decoded pictures with a first resolution higher than a standard resolution***." Ex.1003, ¶254.



Ex.1005, FIG. 4b (modified).

In more detail, Phek's super-resolution processing creates a super-resolution enlarged image from a lower-resolution image sequence. The super-resolution enlarged image has both improved quality (e.g., more resolvable detail and higher SNR) and higher spatial resolution (e.g., more pixels) than any individual lower-resolution image used as input to super-resolution processing. *See* Ex.1005, [0002]; *supra*, §IV.A.1; Ex.1003, ¶255.

Phek's FIG. 1a illustrates "the processing devices (functional blocks) in the super-resolution processing device." Ex.1005, [0016]. In this device, "interpolation processing device 110 interpolates frames subject to super-resolution" using "[w]ell-known interpolation methods" e.g., "bilinear interpolation, bicubic interpolation, and Lanczos interpolation." *Id.* Moreover, Phek discloses "residual interpolation to prevent grid artifacts" "by adding smoothing constraints to high-resolution estimated images," generating "sharper images." *Id*., [0006], [0008]; Ex.1003, ¶256.



Ex.1005, FIG. 1a (modified); *see also id.*, FIG. 1b.

Martins also discloses super-resolution processing and enlargement techniques that would have been obvious to apply additionally or as an alternative to Phek's super-resolution algorithm in the Phek-Segall-Martins-He combination. *Supra*, §§IV.A.3, IV.A.5.d. For example, Martins' FIG. 2 illustrates a super-resolution enlarger that enlarges an image using "spatial interpolation." Ex.1007, 5, FIG. 2. In Martins' algorithm, "quality and spatial resolution of video can be improved by combining multiple pictures to form a single super-resolution picture." *Id.*, Abstract. "[S]uper-resolution pictures obtained by the algorithm are of much higher visual quality and have lower [mean-squared error] than super-resolution pictures obtained by simple spatial interpolation." *Id.* "The algorithm is based on

motion-compensated spatial upsampling from multiple images and decimation to the desired format." *Id.*

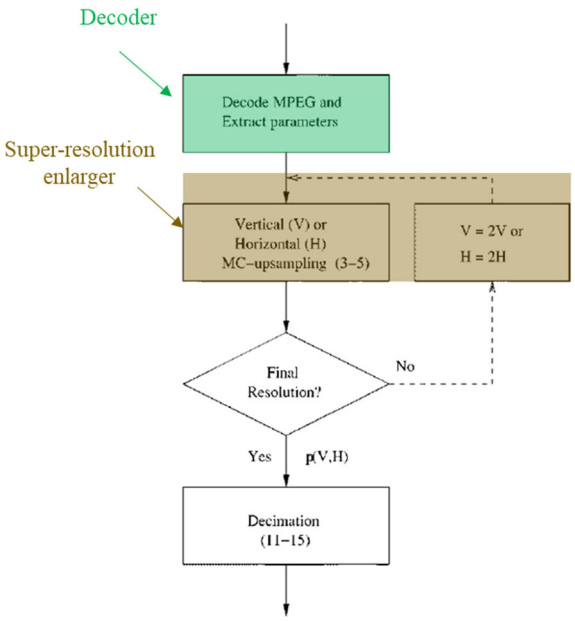


Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007, FIG. 2; Ex.1003, ¶257.

The first enhancement layer decoding path's super-resolution processor corresponds to the "***first super-resolution enlarger***":

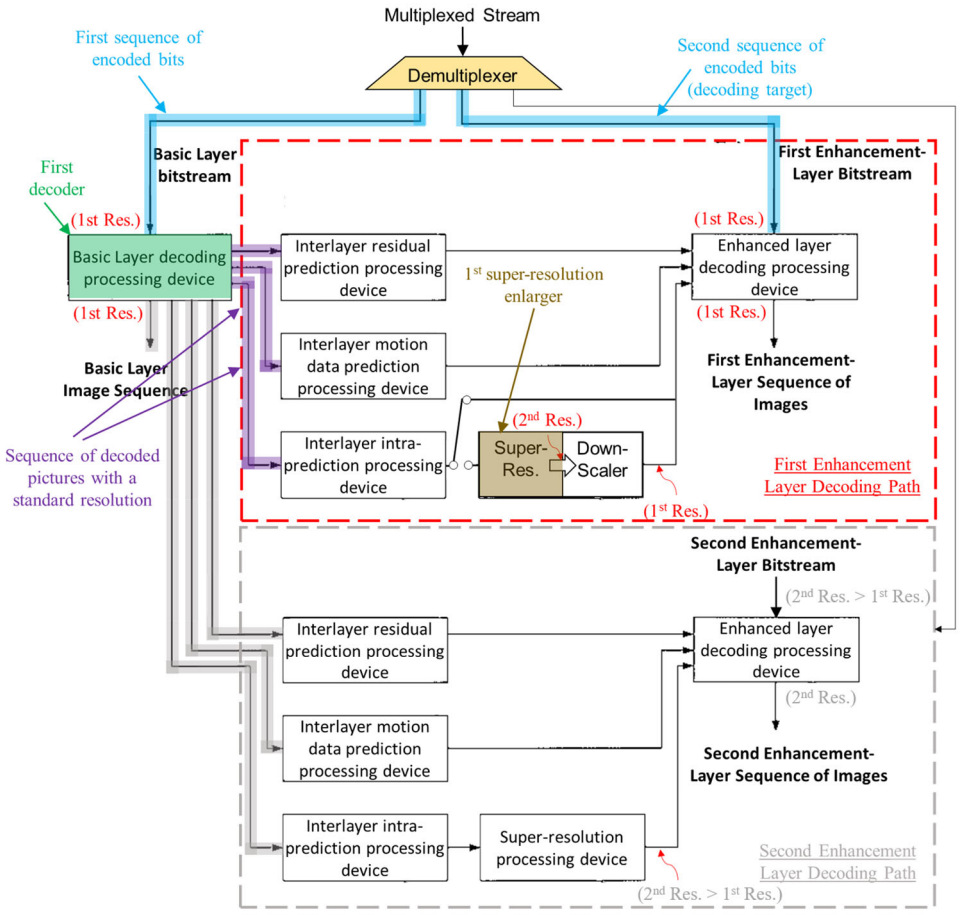


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶258.

5. [2.4]

It would have been obvious to apply a decimation technique like Martins' to the sequence of super-resolution enlarged decoded pictures generated by the first super-resolution enlarger, reducing the size (spatial resolution) of the super-resolution enlarged reference pictures to the first (standard) resolution of the basic-

and enhanced-layer pictures. Ex.1003, ¶¶259-263; Ex.1007, §IV; *supra*, §§IV.A.5.d.

Martins' FIG. 2 depicts picture enhancement process operations, which include "decimation" operation downscaling the super-resolution enlarged image to lower spatial resolution: "the super-resolution picture is decimated to the desired format." Ex.1007, Abstract, 2; *see also id.*, 1 ("obtain a sequence of higher fidelity…in the same format as the directly decoded one"), 3, 9, §IV.

Different "degree[s] of decimation" can be applied depending on operational context, e.g., to perform image "restoration, chrominance upsampling, deinterlacing, or resolution enhancement[]…" Ex.1007, 2. Image restoration enhances decoding quality without changing the restored image's size. *Id.*; Ex.1003, ¶¶260-261 (citing Ex.1011, 2 ("does not change the size")).

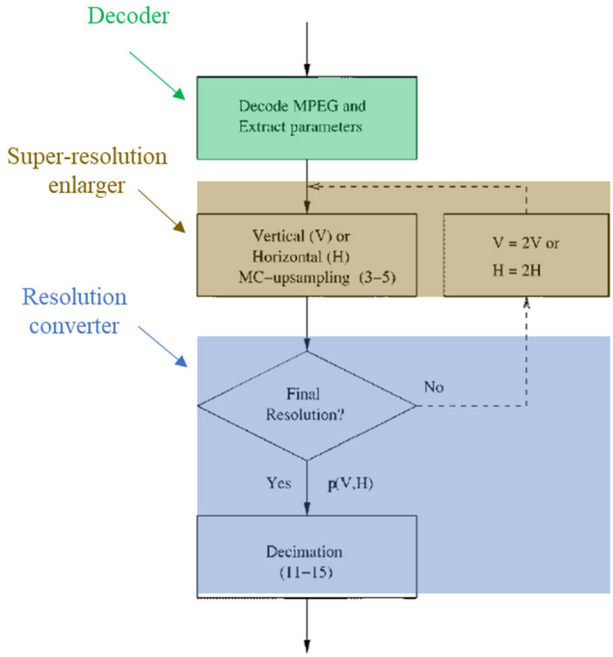


Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007, FIG. 2

FIGs. 11-12 illustrate results of Martins' techniques. Ex.1007, 8-9. FIG. 11 illustrates "the result of using simple upsampling of a directly decoded sequence." *Id.*, 8-9. FIG. 12 shows "results of our enhancement to progressive SDTV…images[]…." *Id.*, 8-9. Here, the video signal has been upsampled to enhance quality while retaining the original SDTV resolution.

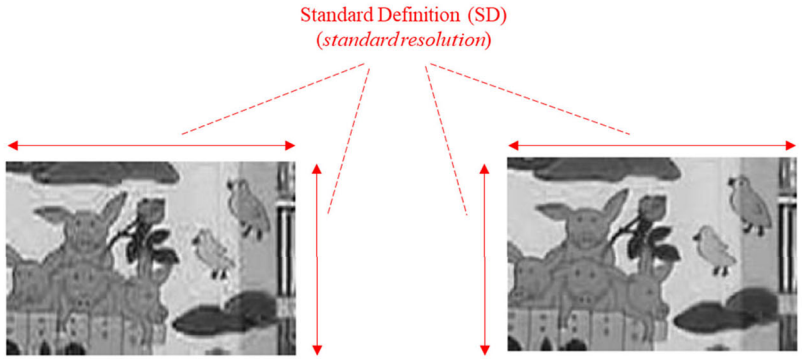


Fig. 11. Direct progressive SDTV. Mobcal at 5 Mbs/s. Part of the I-picture (frame 24, top field). PSNR = 28.5 dB.

Fig. 12. Enhanced to progressive SDTV. Mobcal at 5 Mb/s. Part of the I-picture (frame 24, top field). PSNR = 30.1 dB.

Original picture

Enhanced picture created using Martins' algorithm

Ex.1007, FIGs. 11 and 12.

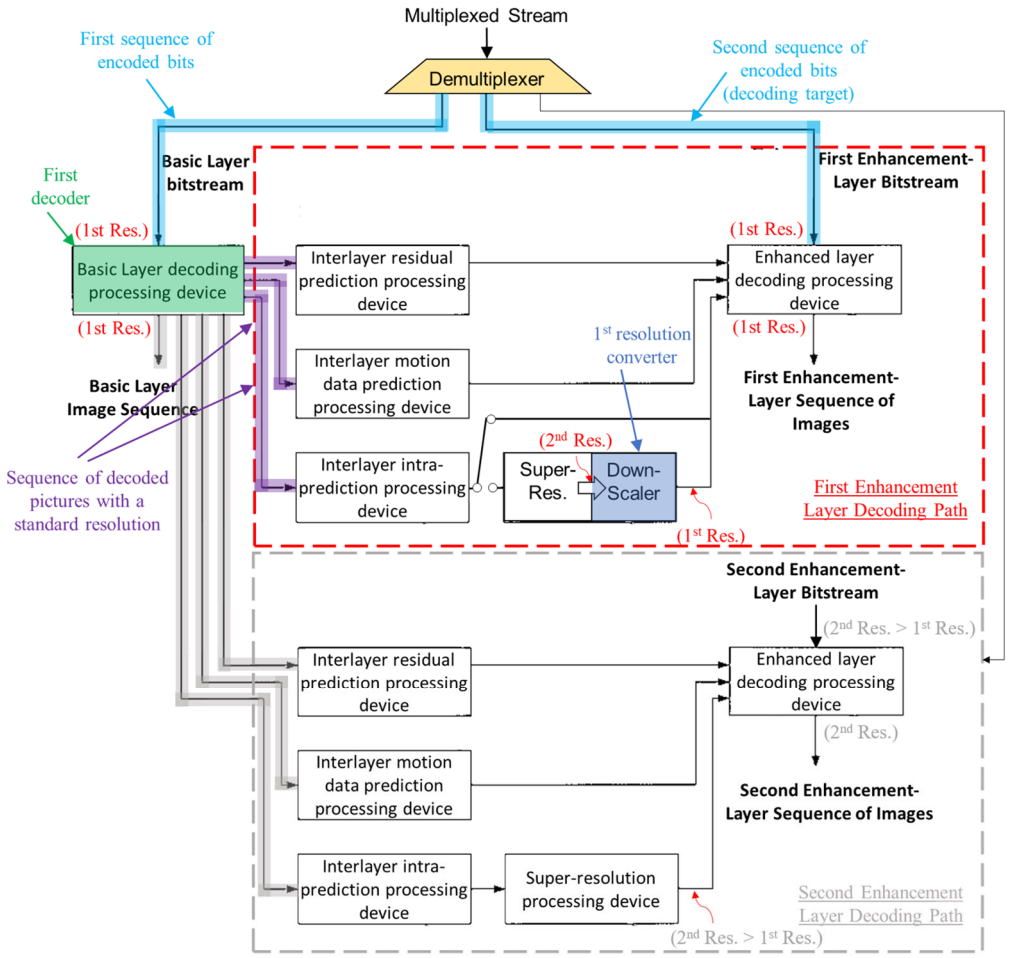


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶262.

The Phek-Segall-Martins-He decoder includes a decimation processor ("***first resolution converter***") based on Martins that acquires the sequence of super-resolution enlarged decoded basic-layer pictures created at the super-resolution processor ("***first super-resolution enlarger***") to work on these pictures by decimating them down to the first (lower or standard) spatial resolution ("***a process for a prescribed resolution conversion to create a sequence of super- resolution***

decoded pictures with a standard resolution"). *Supra*, §§IV.A.5.d., IV.A.3; Ex.1003, ¶263.

6. [2.5.i]

The combination includes Phek's enhanced layer decoding processing device 432, interlayer residual prediction processing device 424, interlayer motion data prediction processing device 426, and interlayer intra-prediction processing device 428 collectively providing a "***second decoder***."[12] Ex.1005, [0014]-[0015], FIG. 4b. Each processing device 424, 426, 428, and 432 facilitates decoding the (first) enhancement layer bitstream apart from decoding the basic layer bitstream and thus collectively provide a "***second decoder***." *Id.*, [0015]; Ex.1003, ¶¶264-270.

In combination, Phek teaches first enhancement layer decoding processing device 432 acquires a first enhancement layer bitstream ("***second sequence of encoded bits obtained with a standard resolution at the demultiplexer***"). Ex.1005, [0014]-[0015], FIG. 4b; Ex.1003, ¶265. It would have been obvious for the Phek-Martins-Segall-He first enhancement layer to encode/decode the enhancement layer sequence of images at the same (standard) spatial resolution as the base layer

[12] The SR-D processors are also in the encoding loop and can optionally be considered part of the "***second decoder***." Ex.1003, ¶265.

sequence of images by implementing known forms of layered, hierarchical coding in which both base and enhancement layers have the same spatial resolution (e.g., multi-view coding, quality scalability, and/or temporal scalability). *Supra*, §IV.A.5.a; Ex.1003, ¶265. Enhanced layer decoding processing device 432 acquires the first enhancement layer bitstream as "***a set of decoding targets***" that allow the second decoder to decode the enhancement layer stream (e.g., using residual values representing differences between predicted and target pictures in the stream). *Id.*, [0012], [0015]; *cf.*, Ex.1001, 15:30-54, 52:40-53:46 (describing similar decoding targets).

Phek-Segall-Martins-He further renders obvious that components corresponding to the "***second decoder***" (i) acquire the sequence of decoded pictures created with the standard resolution from basic layer decoding processing device 422 ("***first decoder***") as a set of first reference pictures and (ii) acquire the sequence of super-resolution decoded pictures created with the standard resolution from the output of decimation/downscaling processor ("***first resolution converter***") as a set of second reference pictures. Ex.1003, ¶266; Ex.1005, [0014]-[0015], FIG. 4b; Ex.1007, §IV, FIG. 2; *supra*, §§IV.A.5.b. At Phek's encoder, "interlayer intra-prediction processing device 408 uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding

the image of the enhanced layer" and "super-resolution processing device 410 is used to improve these additional reference images." Ex.1005, [0014]. At the decoder, "[a]s in the encoder, a super-resolution processing device 430 is used to improve the reference picture from the inter-layer intra-prediction processing device 428." *Id.*, [0015]. Based on fundamental decoding techniques, it would have been obvious that inter-layer intra-prediction processing device 428 obtains decoded base layer images as reference pictures for the interlayer prediction of enhancement layer images. Ex.1003, ¶266; Ex.1007 ("uses…reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer").

Further, He teaches that SR-D processors, which form a loop filter analogous to He's deblocking loop filter, can be selectively applied to filter/enhance reference pictures in an enhancement layer of a layered video coding (e.g., SVC) system. Ex.1008, [0034]. *See also, id.*, [0006]-[0008], [0023]-[0035], [0086]-[0087], FIGs. 1-3, 7; *supra*, §IV.a.5.e. When applying SR-D filtering to an image sequence, enhanced layer decoding processing device 432 acquires SR-D filtered versions of the decoded base layer sequence of images ("***sequence of super-resolution decoded pictures created with the standard resolution at the first resolution converter***") as a set of second reference pictures. Ex.1005, [0014]-[0015], FIG. 4b; *supra*, §IV.a.5.e. When SR-D filtering is not applied (bypassed) to an image sequence,

enhanced layer decoding processing device 432 uses the decoded base layer sequence of images (rather than SR-D filtered versions) acquired by inter-layer intra-prediction processing device 428 as a set of first reference pictures. *Id.* The combination's "***second decoder***" thus acquires the SR-D filtered (second) set of reference pictures when SR-D filtering is enabled and acquires the decoded base layer sequence of images as a set of first reference pictures both when SR-D filtering is enabled (as needed to perform SR-D filtering) and when it is not. Ex.1003, ¶267.

He also teaches that loop filtering can be selectively applied to different images in a stream such that both the first and second sets of reference pictures would be acquired by the first enhancement layer decoder to decode different sequences of images in the first enhancement layer stream. *Id.*; Ex.1008, [0027]-[0028], [0034]-[0037], [0053]-[0054]. Analogously, it would have been obvious to acquire both the first and second sets of reference pictures to decode different slices of enhancement layer images with reference to either SR-D filtered or non-SR-D filtered reference pictures as explained below. *Infra*, [11]; Ex.1003, ¶268.

The following figure illustrates using the non-SR-D filtered version of decoded base layer images as a set of first reference pictures when the SR-D filter is disabled:

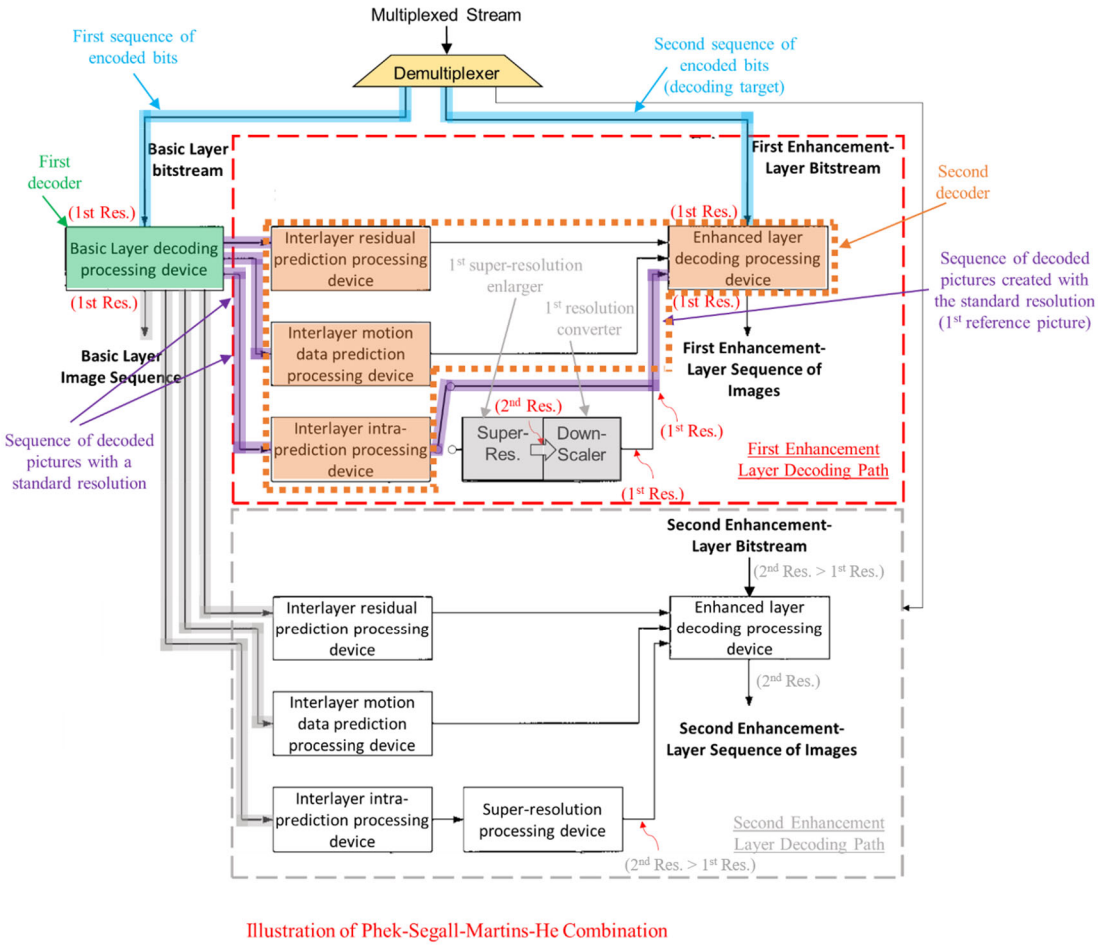


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶269 (first reference picture in purple).

By contrast, the following figure illustrates using the SR-D filtered version of decoded base layer images as a set of second reference pictures when the SR-D filter is enabled:

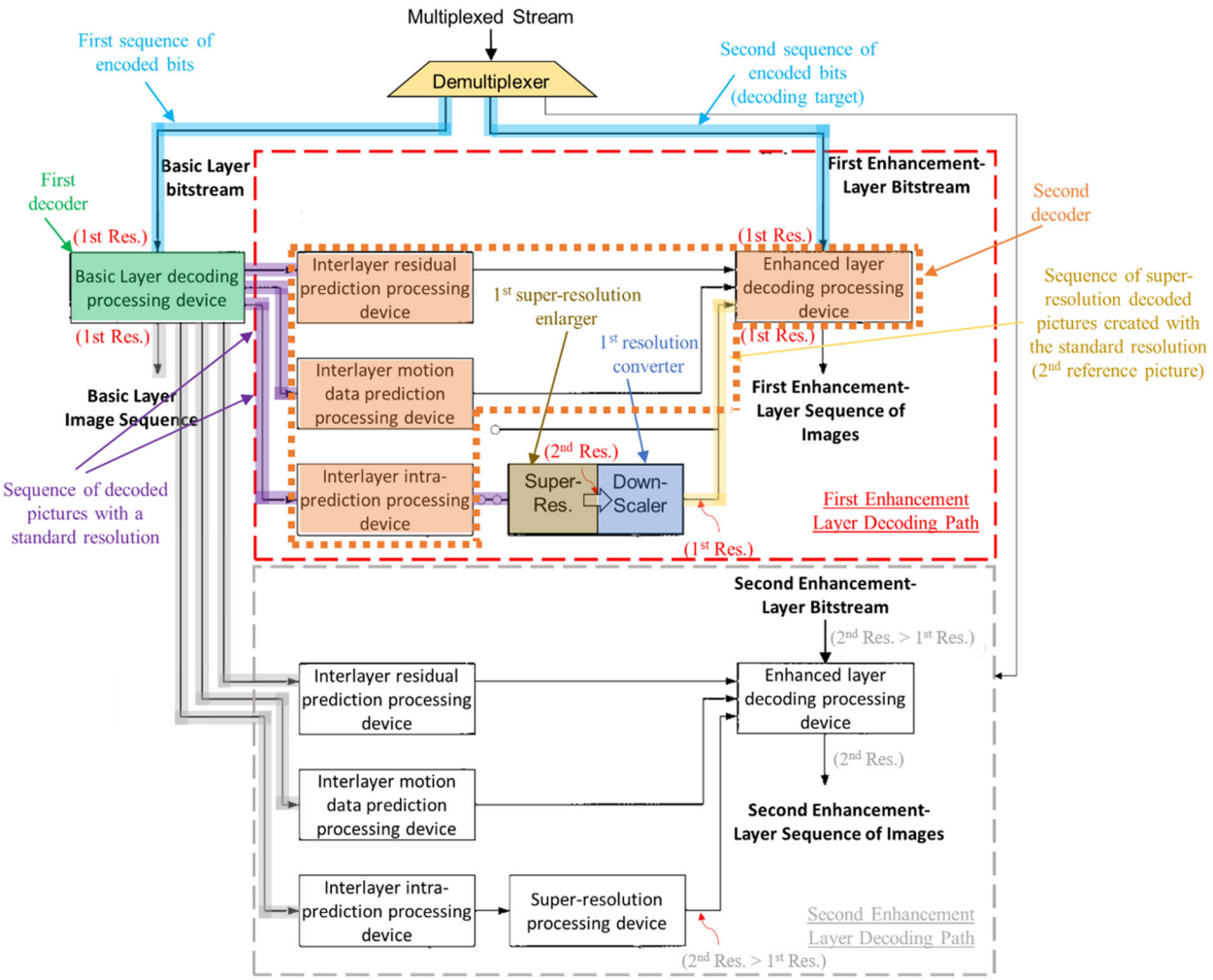


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶270 (second reference picture in yellow).

7. [2.5.ii]

As discussed above (*supra*, [2.5.i]), the combination's enhancement layer decoding processing device 432 decodes pictures in the enhancement layer bitstream using either decoded base layer pictures that have not been enhanced/improved by SR-D filtering ("***first set of reference pictures***") or decoded base layer pictures that

have been enhanced/improved by SR-D filtering ("***second set of reference pictures***"). *See also supra* §§[2.5.i]; Ex.1003, ¶¶271-279. Enhancement layer decoding processing device 432 (and/or other components in the first enhancement layer decoding path corresponding to the "***second decoder***") selects whether to use SR-D filtered or non-SR-D filtered reference pictures based on control information transmitted in the multiplexed signal from the encoder, based on He's teachings. *Supra*, §IV.A.5.e.

For example, He discloses "a deblocking control flag of scalable video system for indicating presentation of deblocking parameters for multiple layers" in a bitstream. Ex.1008, [0002]. "The deblocking parameters determine whether deblock filtering is enabled or disabled" where the parameter is "set to '0' to enable deblock filtering and to '1' to disable deblock filtering." Ex.1008, [0007]. He's FIG. 7 illustrates "deblocking filter control information provided in an output bitstream BTS…or provided in an input bitstream BTS (e.g., IBTS) to the video decoder 603…bitstream BTS includes at least one video information sequence 701…video information sequence 701 further includes one or more video frames including an exemplary video frame 705…video frame 705, for example, includes a frame header 707 at the front of the video frame 705 which further includes a

deblocking filter control present flag 709, labeled 'F'." Ex.1008, [0052]; Ex.1003, ¶272.

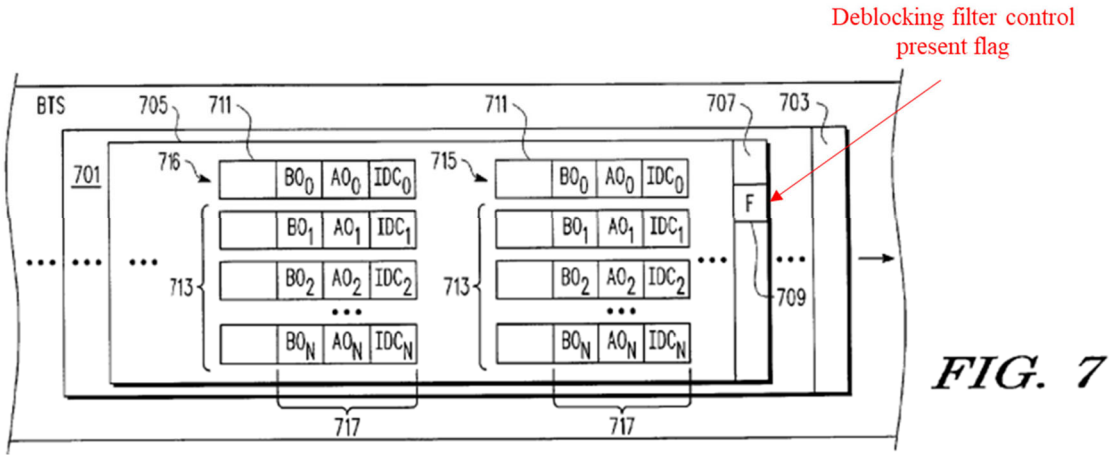


FIG. 7

Ex.1008, FIG. 7

He's FIG. 2 discloses deblock control circuit 221 which, based on the deblocking control flag, issues control signals to deblocking filter 131 to selectively enable/disable deblock filter use on reconstructed (decoded) frames. Ex.1008, FIG. 2, [0006]-[0008]. When enabled, the filtered (enhanced) reconstructed frame is used as a reference picture for prediction of a subsequent frame; when disabled, the original (unfiltered or non-enhanced) reconstructed frame is used as the reference picture for prediction of a subsequent frame. *Id.*, [0034]; Ex.1003, ¶273.

He's "deblocking filter 211 is configured in a similar manner as the deblocking filter 131, and includes a similar decoder filter control DFC signal provided by a deblock control circuit 221….if the deblocking filter 211 is disabled, then the

unfiltered reconstructed frame information URF' is passed as the 'filtered' reconstructed information RF′ in which deblock filtering is bypassed." Ex.1008, [0034]; Ex.1003, ¶274.

It would have been obvious to implement an analogous "control circuit" in the Phek-Martins-Segall-He combination to selectively enable/disable SR-D filtering of base-layer frames used to predict enhancement-layer frames. Ex.1003, ¶275. It would have been obvious for the control circuit to be configured to selectively enable/disable the super-resolution enlarger and downscaler circuits, which form a loop filter analogous to He's deblock filter. *Supra*, §IV.A.5.e; Ex.1003, ¶275.

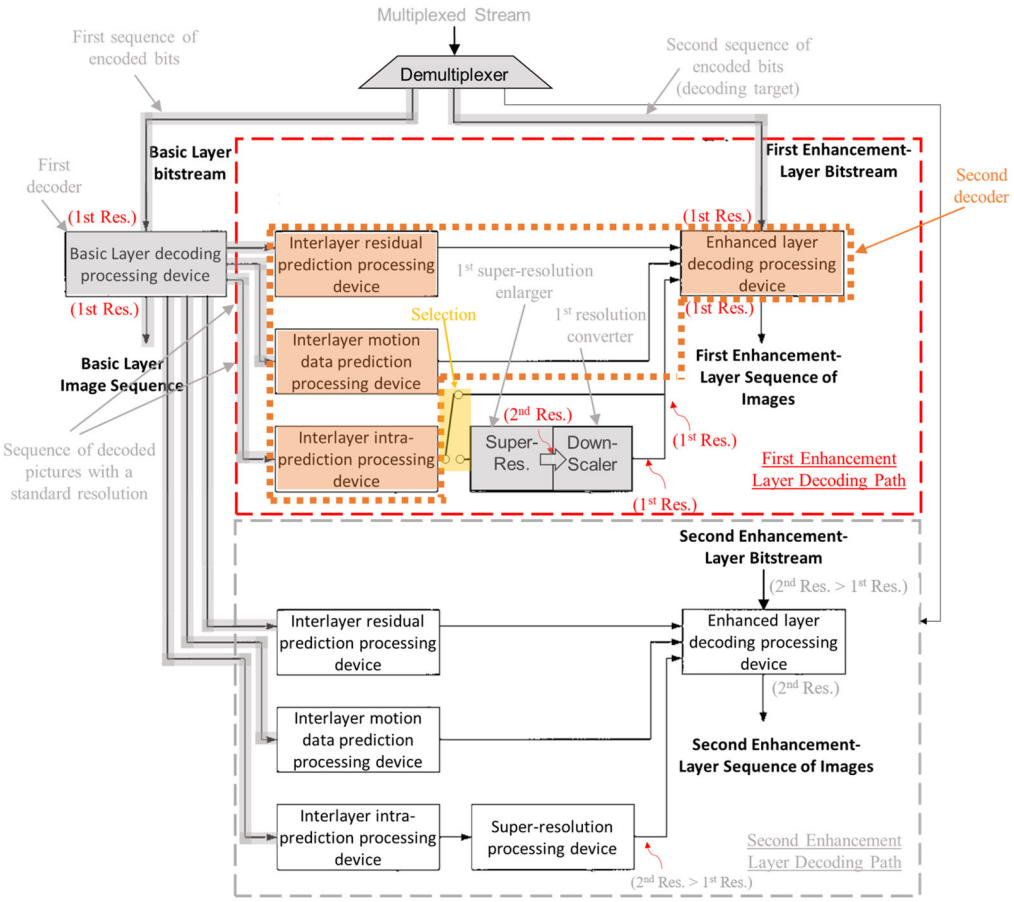


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶275.

This figure illustrates selectivity of the SR-D filter. As illustrated, the "***second decoder***" in the Phek-Segall-Martins-He combination selects either base-layer decoded pictures sequence (when SR-D filter is disabled) as a set of first reference pictures or the super-resolution enhanced and downscaled base-layer decoded pictures sequence (when SR-D filter is enabled) as the set of second reference pictures. Control signals like He teaches (e.g., He's "deblocking enable/disable

parameter" and "deblocking control flag") would provide the first enhancement-layer decoding path with "***reference picture selection information***." Ex.1003, ¶276.

Further, based on Phek, it would have been obvious for enhancement layer decoding processing device 432 to implement processes for prescribed prediction and second decoding with an extension of the standard resolution to create a sequence of super-resolution pictures decoded with the standard resolution based on the set of decoding targets and set of selected reference pictures. Ex.1003, ¶277. Phek's super-resolution processing improves the "reference picture." Ex.1005, [0015]. A POSITA would have found obvious that a "reference picture" in decoders is used to predict at least a portion of another picture in the same or related video stream. *Supra*, §III.A.1; Ex.1008, FIG. 2 and [0033] (describing use of reference pictures for motion compensated predictions). Phek also renders obvious a prescribed second decoding with an extension of the standard resolution by disclosing that "enhanced layer decoding processing device 432 decodes the enhanced layer bitstream." Ex.1005, [0015]; *see also id.*, [0012]. A POSITA would have known that complete decoding of a picture in a compressed video stream involves additional operations beyond prediction from a reference picture, including adding residual information to correct differences between the predicted picture and

the picture targeted for decoding. *Supra*, §IV.A.2; Ex.1003, ¶277; Ex.1008, FIG. 2, [0032] ("combiner 207 adds selected prediction information P' to R' to generate unfiltered reconstructed information").

It would have been obvious for the combination's second decoder to decode and create a sequence of (first) enhancement layer pictures by further adding residual information from the (first) enhancement layer bitstream ("***set of decoding targets***") to pictures predicted from the selected set of reference pictures, consistent with video processing techniques like those Phek identifies. Ex.1003, ¶278. A POSITA would have been motivated to add residual information to the predicted pictures to improve quality of the decoded pictures in the first enhancement layer stream. Ex.1003, ¶278. Moreover, it would have been obvious for the first enhancement layer to decode the first enhancement layer stream at the same, standard resolution as the base layer stream. *Supra*, §IV.A.5.a. First enhanced decoding processing device 432 thus decodes with an extension of the standard resolution and outputs pictures with the standard spatial resolution as claimed, albeit with greater detail than the base layer stream because of the super-resolution processing. *Id.*; *supra*, §IV.A.5.b. Further, the combination's second decoder can select a set of either SR-D filtered or non-SR-D filtered reference pictures by fully enabling/disabling the loop filter as He teaches for picture sequences in the video stream, or by individually determining

that each picture in a set should be predicted using SR-D or non-SR-D filtered references, creating the claimed sequence of super-resolution pictures decoded with the standard resolution. Ex.1008, [0077], [0027]-[0028], [0037]-[0038]; Ex.1003, ¶¶278-279.

8. [2.6]

Phek's "Interlayer Predictive Encoding and Decoding Devices" can be implemented using "scalable" video coding techniques, and Segall teaches a common type of SVC was "spatial" scalability. Ex.1006, 1 ("mechanism for reusing an encoded lower resolution version of an image sequence for the coding of a corresponding higher resolution sequence"), 2-3 ("lowest resolution video data in a spatially scalable system is sometimes referred to as the *base layer*…higher resolution video data is often referred to as the *enhancement layer*. Processes that determine or predict the value of enhancement layer data from previously reconstructed data of a lower resolution layer at the same time instance are referred to as *inter-layer* prediction processes[] …"), FIG. 1. Segall explains known scalable architectures allowed for "multilayer scenarios" so different enhancement layers could encode/decode video streams each with different spatial resolutions. *Id.*, 3; Ex.1003, ¶¶280-282.

As the combination overview explains, it would have been obvious to augment Phek's FIG. 4b decoder with a second enhancement layer implementing spatial scalability so that it would use decoded base layer pictures to predict higher resolution pictures encoded in a second enhancement layer stream. *Supra*, §IV.A.5.b; Ex.1003, ¶279. The second enhancement layer thus provides a "***second resolution converter***" configured to acquire the sequence of decoded pictures with the standard resolution from the basic layer decoding processing device 422 ("***first decoder***") to work on the sequence of decoded pictures to implement an interpolation of pixels with a second enlargement to create a sequence of enlarged decoded pictures as the output of the second enhancement layer with a high resolution higher than the standard resolution. Ex.1006, 4 ("up-sampling of the decoded lower resolution picture"), 4 ("lower resolution picture array to be up-sampled"), 6 ("interpolation filter" and "interpolating kernel"), 7 ("bilinear interpolation"); Ex.1003, ¶281.

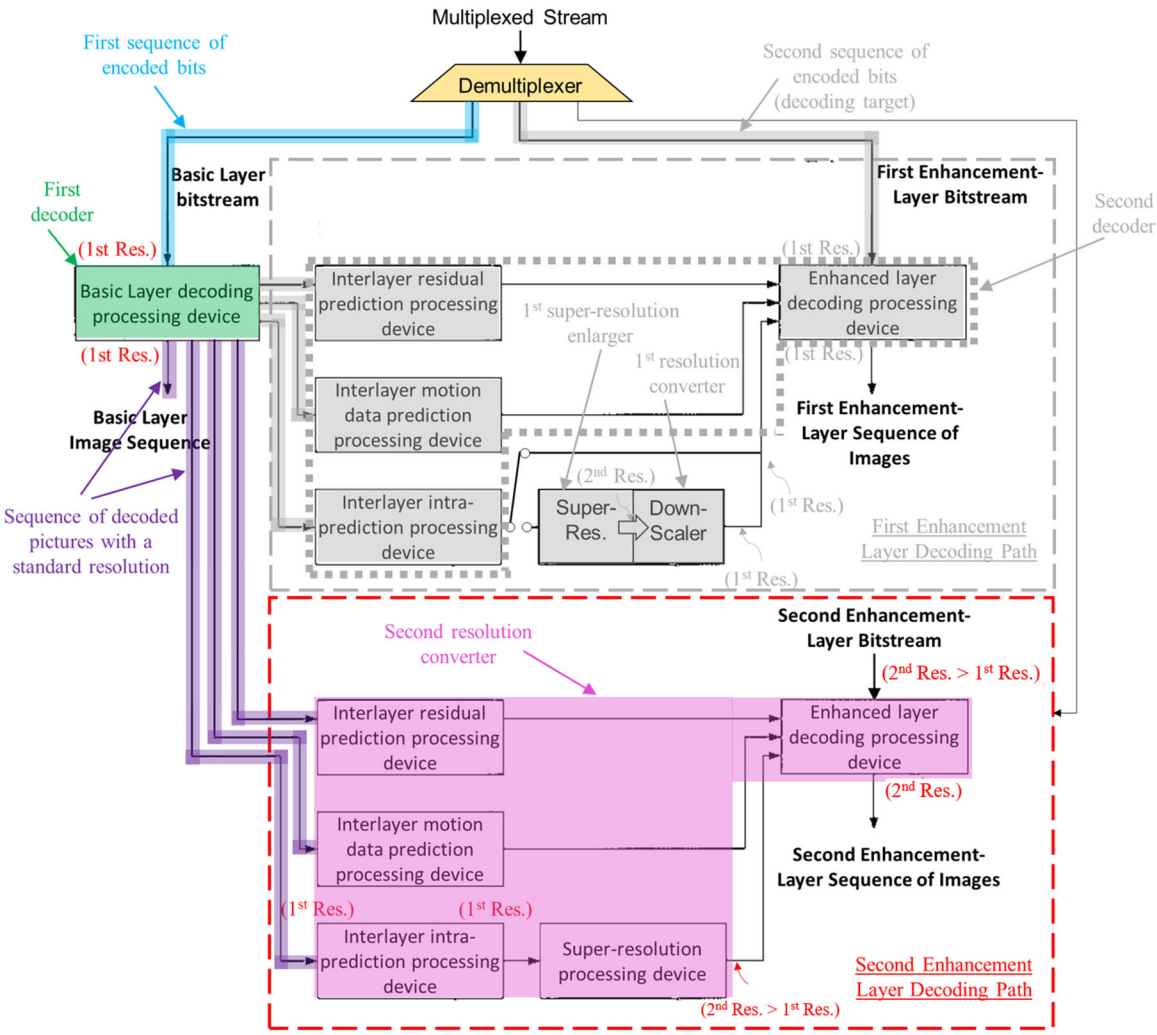


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶281. First mapping of second resolution converter.

Moreover, Phek's super-resolution processing uses a set of lower-resolution images to generate a super-resolution image of higher quality and spatial resolution than individual lower-resolution images. Ex.1005, [0002]-[0003]. Super-resolution processing device 430 that implements super-resolution processing to enhance reference pictures in the second enhancement layer provides another

structure rendering obvious the "***second resolution converter***." Ex.1003, ¶282. Second super-resolution processing device 430 works on decoded base-layer pictures at the standard resolution to create a sequence of enlarged decoded pictures with a high resolution as a second resolution higher than the standard resolution. Ex.1005, [0014]-[0016], FIGS. 1a-1b, 4a-4b.

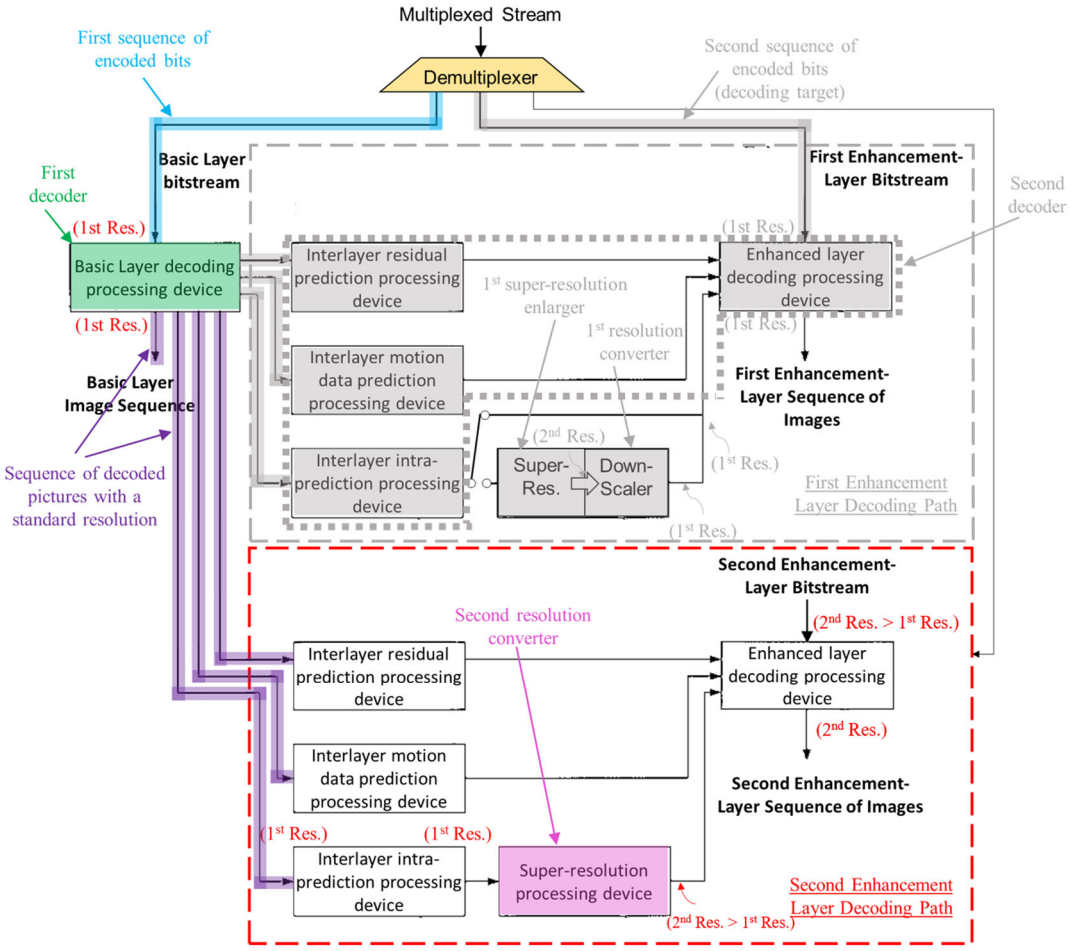


Illustration of Phek-Segall-Martins-He Combination

Ex.1003, ¶282. Second mapping of second resolution converter.

9. [2.7]

It would have been obvious to implement the combination's decoder such that the set of decoding targets in the first enhancement layer, the set of first reference pictures (e.g., decoded base layer pictures without SR-D filtering), and the set of second reference pictures (e.g., decoded base layer pictures with SR-D filtering) have the same spatial resolution. *Supra*, §IV.A.5.d; Ex.1003, ¶283.

C. Claims 3-4 Limitations Correspondence to Claim 2 Limitations

[3.P]	[4.P]	*See* [2.P]
[3.1]	[4.1]	*See* [2.1]
[3.2]	[4.2]	*See* [2.2]
[3.3]	[4.3]	*See* [2.3]
[3.4]	[4.4]	*See* [2.4]
[3.5.i]	[4.5.i]	*See* [2.5.i]
[3.5.ii]	[4.5.ii]	*See* [2.5.ii]
[3.6]	[4.6]	*See* [2.6]
[3.7]		*See* [2.7]
	[4.7]	*See* [2.6]
	[4.8]	*See* [2.6]
	[4.9]	*See* [2.7]

D. Claim 11

Claim 11's limitations simply reflect the common practice of individually encoding/decoding different blocks or other regions of a picture in a stream rather than pictures as a whole. *Supra* §IV.A.2, [2.5.i] and [2.5.ii]; Ex.1003, ¶¶368-370; *see also* Ex.1030, 8 ("basic coding unit for spatial redundancy reduction in MPEG is based on an 8 x 8 block, motion-compensation is usually based on a 16 pixel by 16 line *macroblock*"); *supra*, §IV.A.4. Phek teaches that enhancement layer pictures can be coded by macroblocks. Ex.1005, [0014] ("motion data include the partition size and type used of each macroblock, as well as the prediction direction and motion vector").

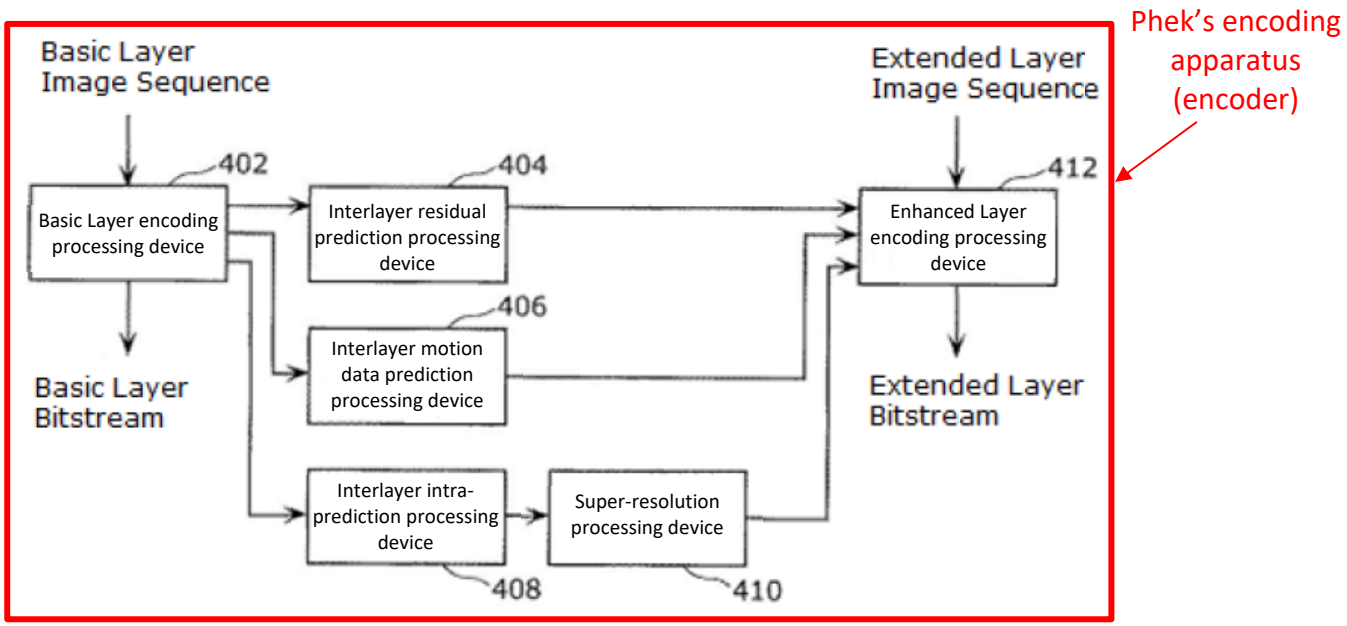
He likewise discloses that "[e]ach video frame is subdivided" into 16x16 pixel macroblock (MB) levels. Ex.1008, [0004], [0034], [0086]. "[V]ideo decoder 200 generally processes information one macroblock at a time." *Id.*, [0034]. He further teaches that separate loop filter (e.g., deblocking filter) controls can be applied to each "slice" in a coded stream, where "each slice incorporates one or more macroblocks." *Id.*, [0053]-[0054]. Loop filtering can be selectively enabled or disabled for each slice. *Id.*

It would have been obvious to apply He's teachings regarding slice-level selective loop filtering in the combination so that the second decoder (e.g., processors 424/426/428/432) in the first enhancement layer would divide the decoding target in the first enhancement layer stream into a plurality of slices (regions) and select whether to use an SR-D filtered version or a non-SR-D filtered version of a reference picture for prediction of each divided region. Ex.1003, ¶370. A POSITA would have implemented slices to reduce the coding unit's size and allow for more variance in the prediction processes applied to decode different regions of the target picture, which allows for greater compression and improved prediction quality for individual slices. *Id.*; Ex.1030, 8. A POSITA would have had a reasonable expectation of success given that slices were commonly implemented in MPEG systems and He already proposed selective, slice-by-slice loop filtering. Ex.1003, ¶368.

E. Claim 1

1. [1.P]

Phek's FIG. 4a shows an "encoding device" in an "[i]nterlayer predictive" coding system. Ex.1005, FIG. 4a, [0014]; *see also id.*, [0011]-[0015], FIGs. 3a-3b, 4b. In the encoder, "the image sequence is intra-coded or inter-coded." Ex.1005, [0011].



Ex.1005, FIG. 4a.

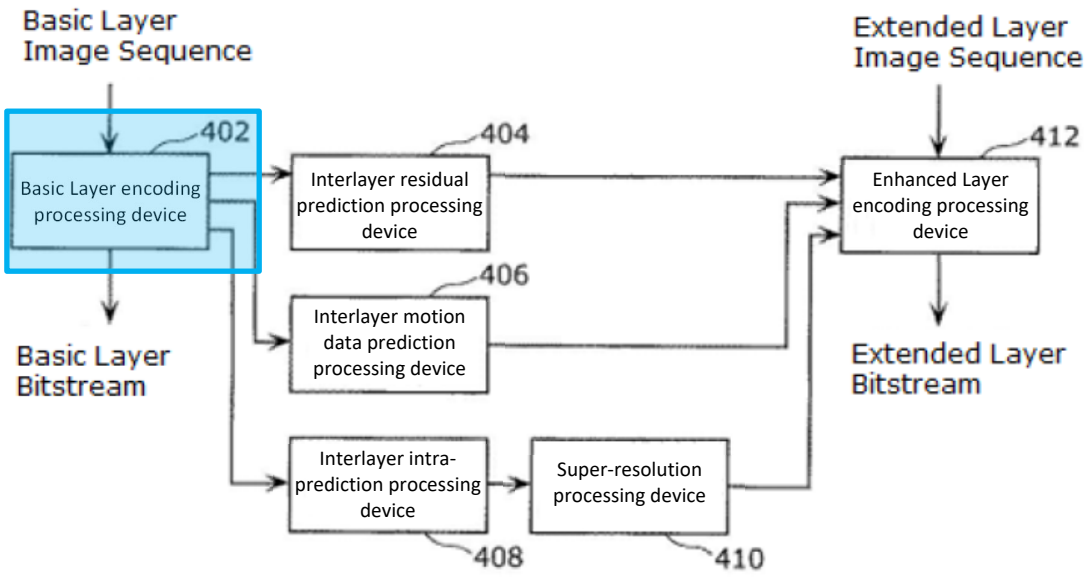
Phek's encoder includes (i) basic layer encoding processing device 402 processing a basic-layer bitstream to encode a basic layer image sequence, (ii) interlayer residual prediction processing device 404 reducing redundancy between enhanced layer residuals and basic layer residuals, and (iii) interlayer motion data prediction processing device 406 reducing redundancy between enhanced layer residuals and basic layer residuals. *Id.*, [0014]. The encoding device also includes (i) interlayer intra-prediction processing device 408 that "uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer," and (ii) super-resolution processing device 410 that "performs residual interpolation" and "improve[s]" the

"additional reference images." *Id.* The encoder mirrors the decoder's operations. *Supra* §IV.A.1; Ex.1003, ¶¶304-305.

2. [1.1]

Phek's FIG. 4a encoder for "hierarchical encoding" "uses interlayer prediction" where "[a] basic layer image sequence is encoded using a basic layer encoding processing device 402." Ex.1005, [0014].

Basic layer encoding processing device 402 ("***first encoder***") processes the basic-layer image sequence ("***sequence of moving pictures with a standard resolution***") to implement a process of encoding to create a bitstream for the basic-layer image sequence ("***first sequence of encoded bits***") and decoding to create a decoded set of the basic-layer image sequence ("***set of decoded pictures with the standard resolution***"). Ex.1003, ¶¶304-306; *supra*, §IV.A.1. The basic layer image sequence is encoded and decoded with a first spatial resolution ("***standard resolution***"). *Supra*, §IV.A.1; Ex.1005, [0014]; Ex.1003, ¶¶306-308.



Ex.1005, FIG. 4a.

3. [1.2]

Martins' FIG. 2 discloses "combining multiple pictures to form a single super-resolution picture" using an algorithm "based on motion-compensated spatial upsampling from multiple images." Ex.1007, Abstract. Through upsampling, pixels of the current picture are aligned with pixels of a reference picture using "motion estimation." *Id.*, 5. Resulting information "from all the reference pictures" are combined "to form a single superresolution picture." *Id.* Martins also discloses its super-resolution enlarger may enlarge an image using "spatial interpolation."

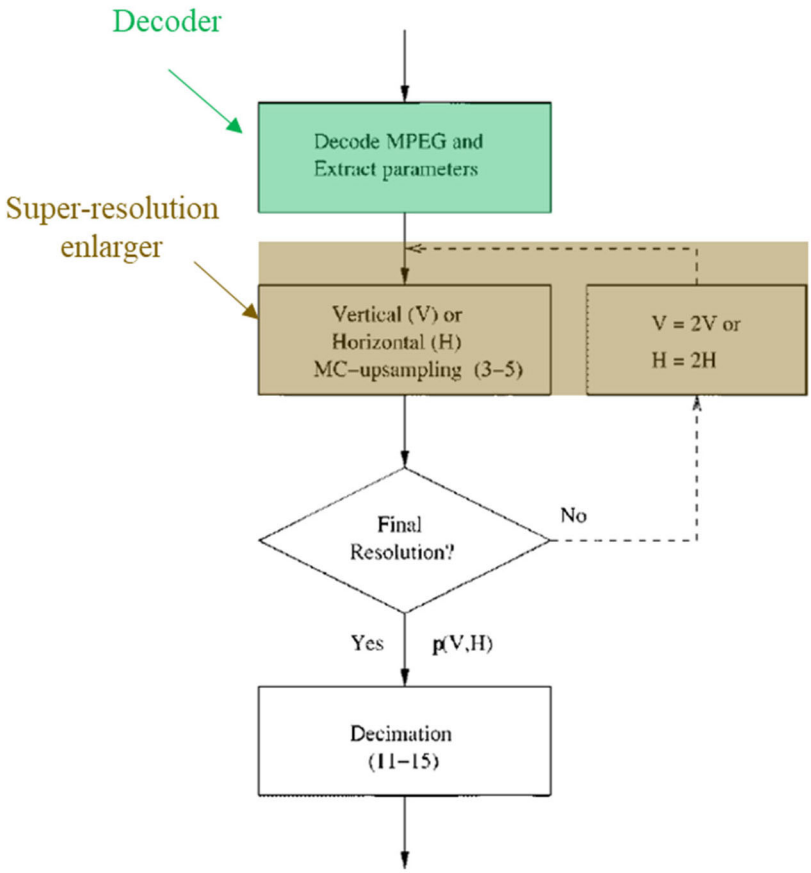
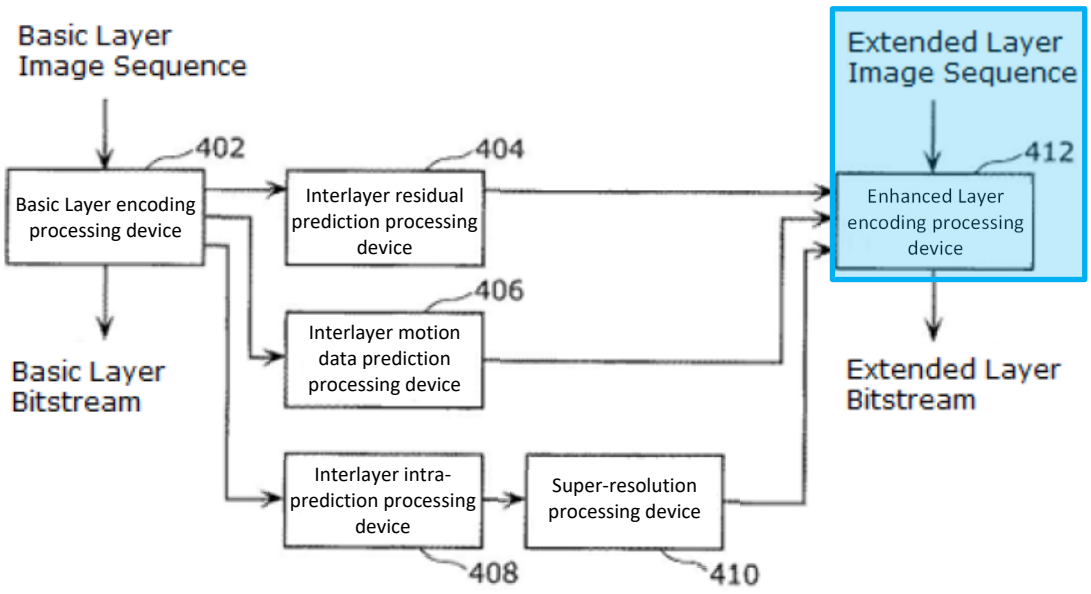


Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007, FIG. 2.

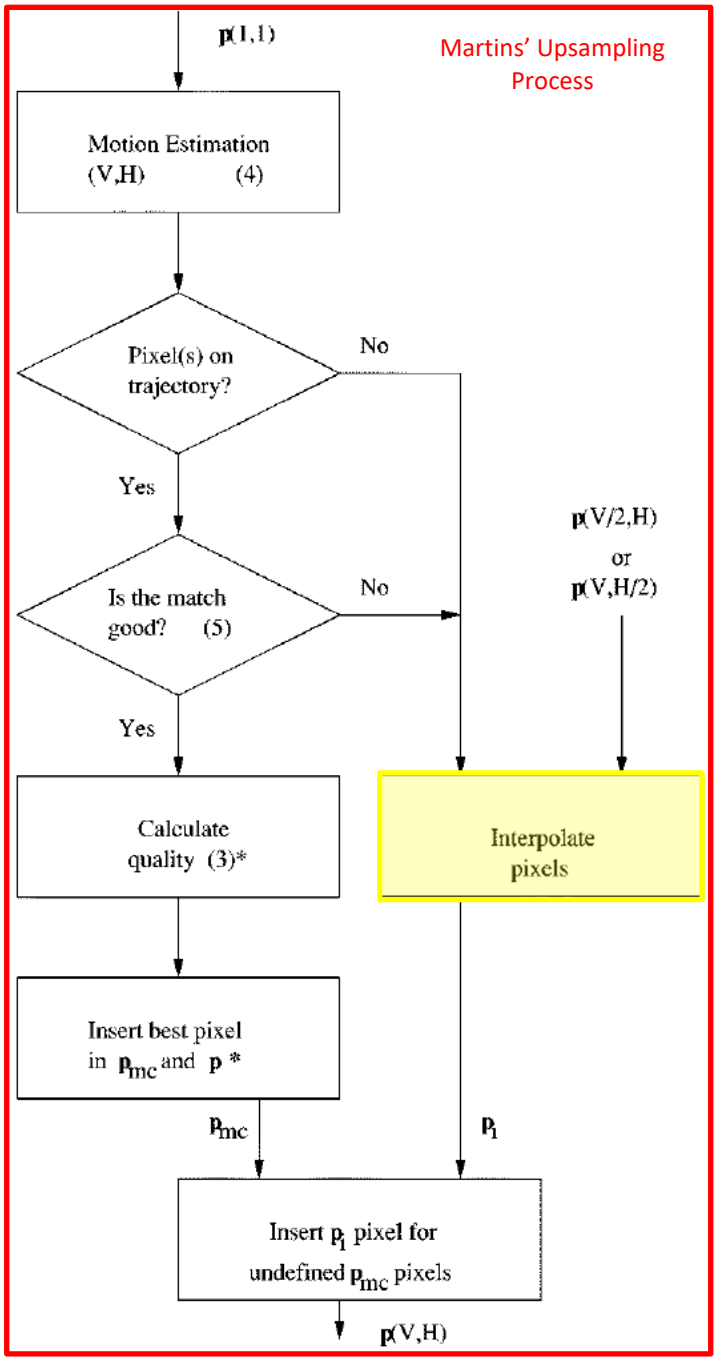
In combination, Martins' upsampling technique is applied to the extended layer image sequence ("***subsequence of the sequence of moving pictures with the standard resolution***") in the enhancement layer "to form a single superresolution picture" ("***a set of super-resolution enlarged pictures with a first resolution higher than the standard resolution***"). *Id.* A POSITA would have understood that Martins'

upsampling technique can be applied to the extended layer image sequence of Phek prior to the encoding processing device 412's processing. Ex.1003, ¶¶309-312.



Ex.1005, FIG. 4a

As part of the enhancement layer upsampling process, Martins' spatial interpolation algorithm creates the set of enlarged super-resolution pictures, "based on motion-compensated spatial upsampling from multiple images" ("***an interpolation of pixels with a first enlargement***"), as FIG. 3 shows. Ex.1007, Abstract, 5; Ex.1003, ¶312; *see also* [2.3].


p(1,1)
Martins' Upsampling Process
Motion Estimation (V,H) (4)
Pixel(s) on trajectory?
No
Yes
p(V/2,H) or p(V,H/2)
Is the match good? (5)
No
Yes
Calculate quality (3)*
Interpolate pixels
Insert best pixel in p_{mc} and p *
p_{mc}
p_i
Insert p_i pixel for undefined p_{mc} pixels
p(V,H)

Ex.1007, FIG. 3

4. [1.3]

Martins' FIG. 2 "decimation" technique downscales the super-resolution enlarged image resulting from Martins' upsampling to a lower spatial resolution to the first (standard) resolution of the basic layer image sequence ("***a set of super-resolution enlarged and converted pictures with a standard resolution***"). Ex.1003, ¶¶313-315.

FIG. 2 depicts Martins' picture enhancement process including the "decimation" technique that decimates the super-resolution picture "to the desired format." Ex.1007, Abstract, 2, 3, 9, §IV. Martins' decimation technique can be used for image restoration, enhancing decoding quality without changing the restored image's size, in the enhancement layer. *Supra* §IV.B.5[2.4]; *id.*, 2; Ex.1003, ¶314. To do so, Martins' decimation operation downscales the super-resolution enlarged image to a lower spatial resolution, as described above with respect to [2.4]. Ex.1007, Abstract, 2, 3, 9, §IV; *see also supra*, §IV.A.3 (describing decimation). Accordingly, in the combination, the set of super-resolution enlarged pictures generated through Martins' upsampling are then decimated to the (first) standard resolution to produce a set of super-resolution enlarged and converted pictures with the standard resolution.

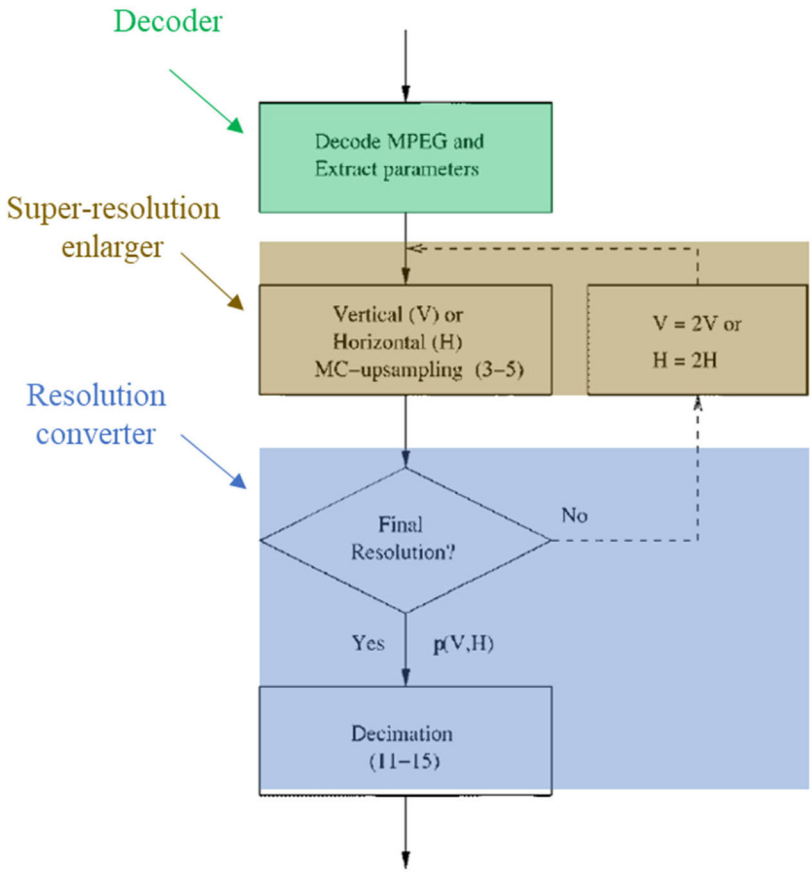


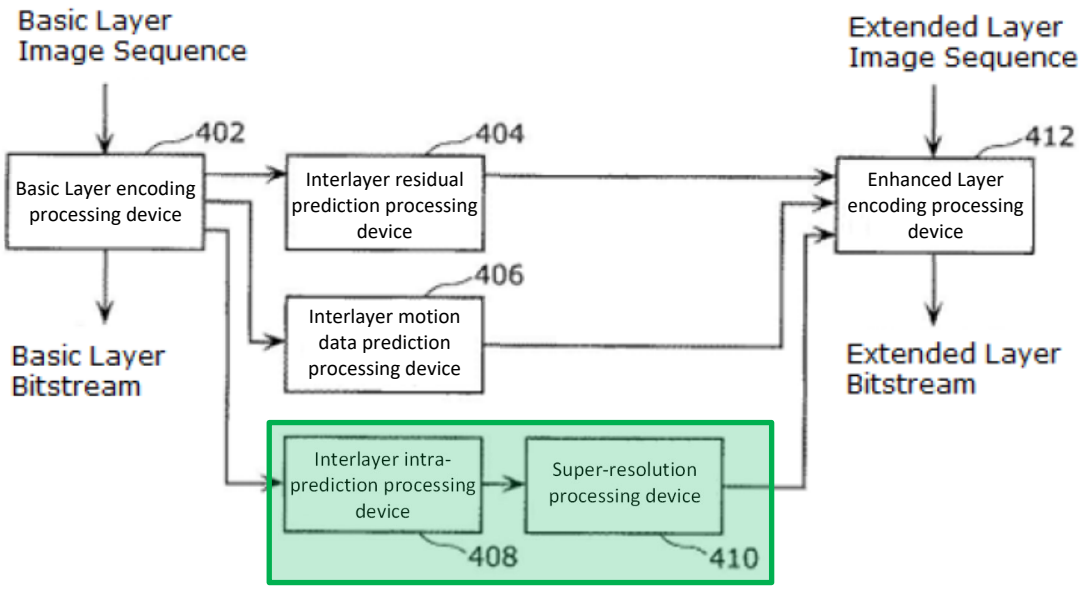
Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007, FIG. 2; Ex.1003, ¶314.

5. [1.4]

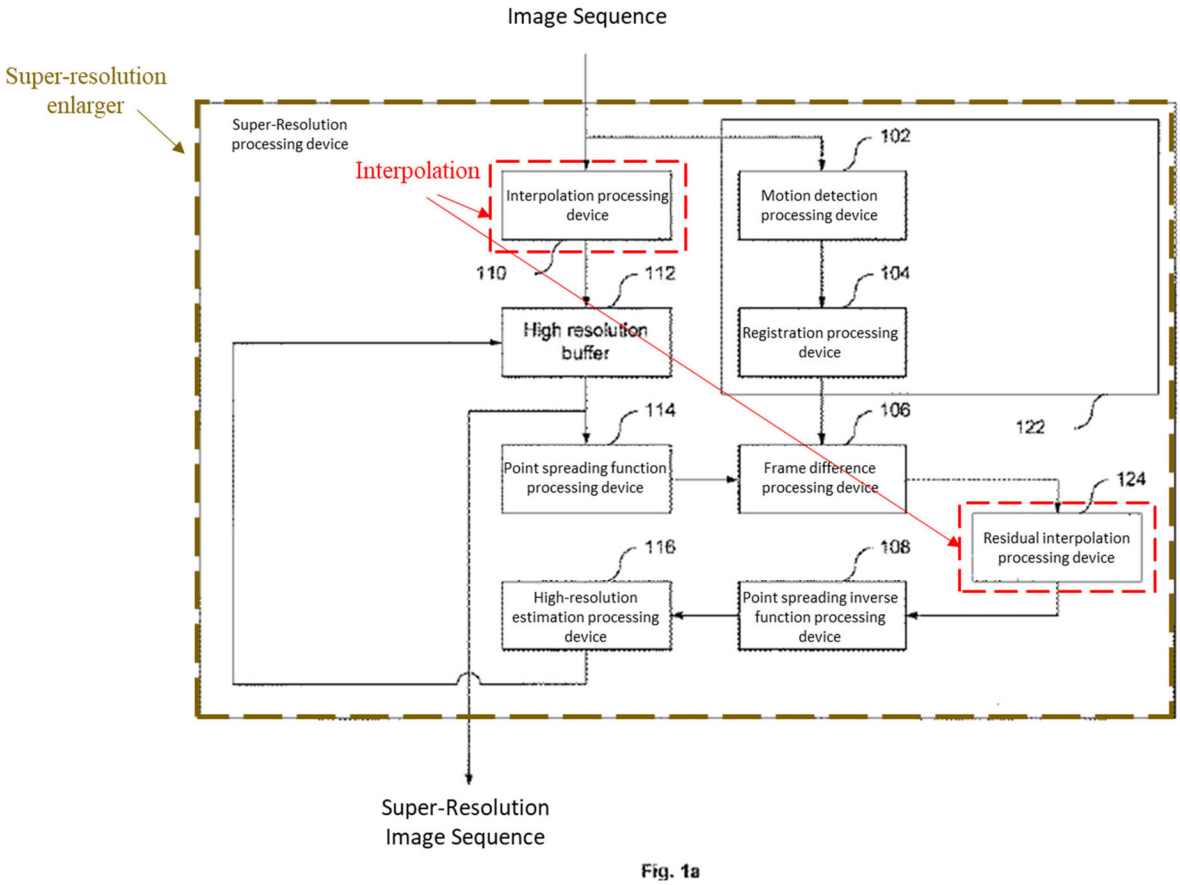
In FIG. 4a, Phek describes "super-resolution processing device 410…used to improve the[] additional reference images" provided by interlayer intra-prediction processing device 408, which "uses an intra image of the basic layer or a reconstructed image of the basic layer for encoding the image of the enhanced layer"

based on the basic layer image sequence received by basic layer encoding processing device 402. Ex.1005, [0015], FIG. 4a; Ex.1003, ¶¶315-319. Super-resolution processing device 410 "performs residual interpolation." Ex.1005, [0015]; Ex.1003, ¶¶315-319. Phek's "super-resolution processing…improve[s] encoding efficiency or picture quality by improving the reference frames used for motion detection, motion compensation prediction, and display." Ex.1005, [0013], [0014]. Super-resolution processing device 410 is thus a "***second super-resolution enlarger***" configured to acquire reference images ("***set of decoded pictures with the standard resolution***") from basic layer encoding processing device 402 ("***first encoder***"), which encoded the basic layer image sequence to produce the set of decoded reference images, through interlayer intra-prediction processing device to implement an interpolation of pixels with a second enlargement to create "***a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution***." Ex.1003, ¶¶316-318.



Ex.1005, FIG. 4a

Phek's super-resolution processing creates a super-resolution enlarged image with improved quality and higher spatial resolution than individual lower-resolution images used as input. *Supra*, §IV.A.1; Ex.1005, [0002]; Ex.1003, ¶¶316-318. Further, Phek's super-resolution processing includes "interpolation processing device 110" that "interpolates frames subject to super-resolution" using well-known methods. *Supra*, §IV.A.1; Ex.1005, [0006], [0008]; Ex.1003, ¶¶316-318.



Ex.1005, FIG. 1a (modified); *see also id.*, FIG. 1b.

Thus, Phek's super-resolution processing device 410 outputs "a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution." Ex.1003, ¶¶316-319.

Martins' super-resolution processing and enlargement techniques would also have been obvious to apply in addition to or as an alternative to Phek's super-resolution algorithm. *Supra*, §§IV.A.5.d, IV.A.3. Martins' super-resolution enlarger uses "spatial interpolation" to enlarge an image, improving "quality and

spatial resolution of video" by combining multiple images. *Supra*, §IV.E.3; Ex.1007, Abstract, 5, FIG. 2. Martins obtains the super-resolution images using an algorithm "based on motion-compensated upsampling from multiple images and decimation to the desired format." *Supra*, §IV.A.3; Ex.1007, Abstract; Ex.1003, ¶319.

6. [1.5]

It would have been obvious to apply a decimation technique like Martins' to the sequence of super-resolution enlarged decoded pictures generated by super-resolution processing device 410 to reduce spatial resolution of the super-resolution enlarged pictures to the standard resolution of basic-layer pictures. Ex.1003, ¶¶320-322; Ex.1007, Abstract, 1, FIG. 2; *supra* §IV.E.4. Through Martins' picture enhancement process, "the super-resolution picture is decimated to the desired format." Ex.1007, Abstract, 2; *see also id.*, 1, 3, 9; *supra*, §IV.A.3.

Accordingly, the combination's encoding device includes a decimation processor ("***second resolution converter***") based on Martins configured to work on the sequence of super-resolution enlarged pictures from Phek's super-resolution processing device 410 ("***second super-resolution enlarger***") by decimating them to the standard resolution ("***a process for a second resolution conversion to create a set of super-resolution enlarged and converted decoded pictures with a standard resolution***"). Ex.1003, ¶¶320-322.

7. [1.6.i]

The combination includes Phek's interlayer residual prediction processing device 404, interlayer motion data prediction processing device 406, interlayer intra-prediction processing device 408, and enhanced layer encoding processing device 412 that collectively provide a "***second encoder***." Ex.1005, [0014]-[0015], FIG. 4a. Each processing device 404, 408, and 412 facilitates encoding the basic layer image sequence received and decoded by basic layer encoding processing device 402. Ex.1003, ¶324.

8. [1.6.ii]

Phek teaches enhanced layer encoding processing device 412 acquires the extended layer image sequence ("***set of super-resolution enlarged and converted pictures***") from Martins' upsampling and decimation processes performed on the basic layer image sequence. Ex.1005, [0014], FIG. 4a; Ex.1003, ¶¶325-327. Enhanced layer encoding processing device 412 acquires the extended layer image sequence as "***a set of encoding target pictures***" allowing the second encoder to encode the enhancement layer stream (e.g., using residual values that represent differences between predicted pictures and target pictures in the enhancement stream). Ex.1005, [0014], FIG. 4a; Ex.1003, ¶¶325-327. The combination's second encoder is also configured to acquire (i) "***the set of decoded pictures from the first***

encoder as a set of first reference pictures" from the basic layer encoding processing device ("***first encoder***"), and (ii) "***the set of super-resolution enlarged and converted decoded pictures…as a set of second reference pictures***" from the super-resolution processing device 410. Ex.1005, [0014], FIG. 4a; Ex.1003, ¶¶325-327. For example, at the encoder, Phek's "interlayer intra-prediction processing device 408 uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer" and "super-resolution processing device 410 is used to improve these additional reference images." Ex.1005, [0014], FIG. 4a; Ex.1003, ¶¶325-327. A POSITA would have found it obvious that interlayer intra-prediction processing device 408 obtains decoded base layer images as reference pictures for the interlayer intra-prediction of enhancement layer images. Ex.1007, [0014]; Ex.1003, ¶¶325-327.

9. [1.6.iii]

The combination's enhancement layer encoding device 412 encodes pictures in the enhancement layer bitstream using either encoded/decoded base layer pictures not enhanced by SR-D filtering ("***set of first reference pictures***") or decoded/encoded base layer pictures enhanced by SR-D filtering ("***set of second reference pictures***"). Ex.1003, ¶¶328-331. Enhancement layer encoding processing

device 412 (and/or other components in the enhancement layer corresponding to the "***second encoder***") selects whether to use SR-D filtered or non-filtered reference pictures based on control information, such as a MINSAD (minimum sum of absolute differences) value used to measure complexity of the video content, motion vectors, or a power saving factor (PSF), based on He. Ex.1008, [0029]-[0031], FIG. 1; *see also supra*, §IV.A.4.

He discloses "a deblocking control flag" in a bitstream. *Supra*, §IV.B.7; Ex.1008, [0002]. Deblocking parameters "determine whether deblock filtering is enabled or disabled." *Supra*, §IV.B.7; Ex.1008, [0007]. "Deblocking filter control information" is provided in an input bitstream to a decoder or as an output bitstream BTS from the encoder. *Supra*, §IV.B.4; Ex.1008, [0052]; Ex.1003, ¶330. FIG. 2 includes deblock control circuit 221 issuing control signals to enable/disable use of a deblock filter. *Supra*, §IV.B.4; Ex.1008, [0006]-[0008]. When enabled, the filtered (enhanced) frame is the reference picture. When disabled, the unfiltered (non-enhanced) frame is the reference picture. *Supra*, §IV.B.7; Ex.1008, [0027]; Ex.1003, ¶330.

Based on He, a POSITA would have found it obvious to implement a "control circuit" in the Phek-Martins-Segall combination configured to selectively enable/disable SR-D filtering of base layer frames used to predict enhancement layer

frames. Ex.1003, ¶331. Specifically, it would have been obvious to use the control circuit to selectively enable/disable the super-resolution enlarger and downscaler units, which form a loop filter like He's deblock filter. *Supra*, §IV.A.5.e; Ex.1003, ¶¶328-331.

10. [1.6.iv]

He discloses "video encoder 100…for generating and transmitting video information in a bitstream (BTS)." Ex.1008, [0023]. Video information can be "any of various types," such as "any video or image or image sequence information." *Id.*, [0022]. He describes a "deblocking enable/disable parameter" and "deblocking control flag" ("***reference picture selection information***") reflecting the selection between SR-D filtered frames and non-SR-D filtered frames determined by the control circuit of the second encoder, as Phek and He teach. *Supra*, §IV.E.9; Ex.1003, ¶¶332-333.

He further teaches that prediction information (e.g., a "deblocking enable/disable parameter" and "deblocking control flag") is provided to a combiner outputting "residual information R to the input of a transform circuit 119." Ex.1008, [0025]. After performing a block transform on that information, transfer circuit 119 outputs "transform result B," which is provided to "a quantization (Q) circuit 121 [that] outputs quantized transform coefficients X." *Id.* These X coefficients are

provided to "output processing circuit 125, which provides a compressed bitstream (BTS) for transmission or storage." *Id.* Adding X coefficients into the bitstream BTS may be performed through encoding. *Id.*; Ex.1003, ¶¶332-333. Accordingly, it would have been obvious to configure the second encoder to also encode in the enhancement layer bitstream ("***a third process for encoding***") a deblocking enable/disable parameter and deblocking control flag information ("***reference picture selection information***").

11. [1.7]

It would have been obvious to implement the Phek-Segall-Martins-He combination's encoding device such that the set of encoding targets, set of first reference pictures (e.g., encoded/decoded base layer pictures without SR-D filtering), and set of second reference pictures (e.g., encoded/decoded base layer pictures with SR-D filtering) have the same resolution. *Supra*, §§IV.B.9; Ex.1003, ¶334.

F. Claim 5

1. [5.P]

Supra [1.P], [1.2]. Ex.1003, ¶33.

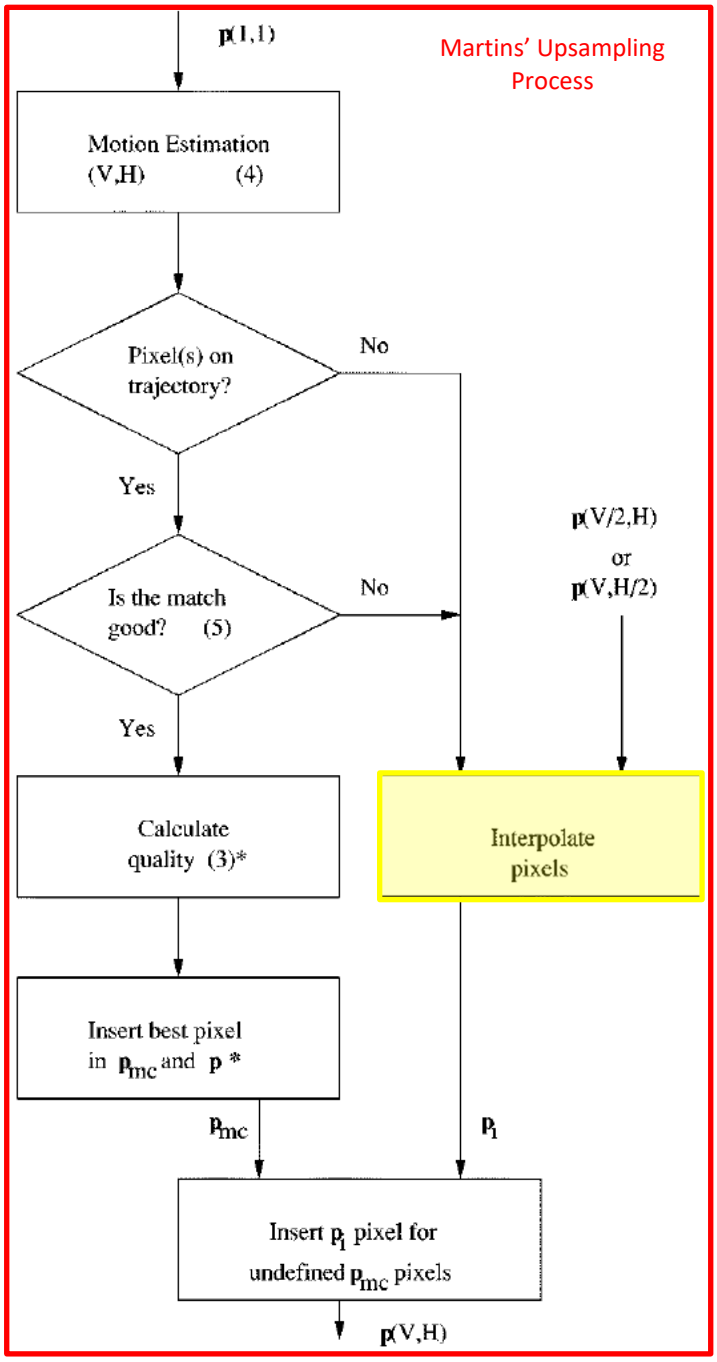
2. [5.1]

Martins "align[s] pixels of the current picture frame at time *t* with pixels of one of the reference pictures using motion estimation," and "combine[s] the information from all the references pictures to form a single superresolution picture

at *t*" ("***positioner configured to work...on a combination of a base picture...and one or more observation pictures...to make a position to pixel positions of a desirable high resolution***"). Ex.1007, 4. The resulting superresolution picture of Martins is an "MC [motion-compensated] upsampled high-resolution image" ("***nonhomogenous high resolution pictures***"). *Id.* Martins forms the superresolution picture "by mapping pixels from each of the references pictures" using a "block-based motion-compensation scheme" accounting for pixels at varying quality. *Id.*, 4-5. A POSITA would have understood that Martins' super-resolution enlarger included a "positioner" for performing this functionality. Ex.1003, ¶336.

3. [5.2]

Martins uses "spatial interpolation" to "[c]omplete the [s]uperresolution [p]icture" by "alternat[ing] between horizontal and vertical 2:1 upsampling." Ex.1007, 6, FIG. 3. A POSITA would have understood Martins' super-resolution enlarger included an "interpolator" for performing this functionality. Ex.1003, ¶337.


p(1,1)
Martins' Upsampling Process
Motion Estimation (V,H) (4)
Pixel(s) on trajectory?
No
Yes
Is the match good? (5)
No
Yes
p(V/2,H) or p(V,H/2)
Calculate quality (3)*
Interpolate pixels
Insert best pixel in p_{mc} and p *
p_{mc}
p_i
Insert p_i pixel for undefined p_{mc} pixels
p(V,H)

Ex.1007, FIG. 3.

4. [5.3]

Following the process described in [5.2], Martins outputs a resulting superresolution picture at a particular high resolution. Ex.1007, 6; Ex.1003, ¶338. Each pixel of the resulting superresolution picture has an "attribute" showing "whether it was determined by motion compensation" or "interpolation." Ex.1007, 6. A POSITA would have understood Martins' super-resolution enlarger included an "estimated pictured creator" for performing this functionality. Ex.1003, ¶338.

5. [5.4]

Martins' upsampling algorithm determines whether the resulting superresolution picture requires further repetitions to reach the prescribed "Final Resolution," including for both a high resolution picture upsampled using motion compensation ("***a nonhomogenous high resolution picture from the positioner***") and a high resolution picture upsampled using interpolation ("***a nonhomogenous estimated picture from the estimated picture creator***"). Ex.1007, 4, Fig. 2; Ex.1003, ¶339. If the "Final Resolution" has not been met, Martins' upsampling process repeats until the desired resolution is achieved. Ex.1007, 4, Fig. 2. Otherwise, repetition is unnecessary. Ex.1007, 4, Fig. 2. A POSITA would have found it obvious that Martins' super-resolution enlarger included a "repetition determiner" for performing this functionality. Ex.1003, ¶339.

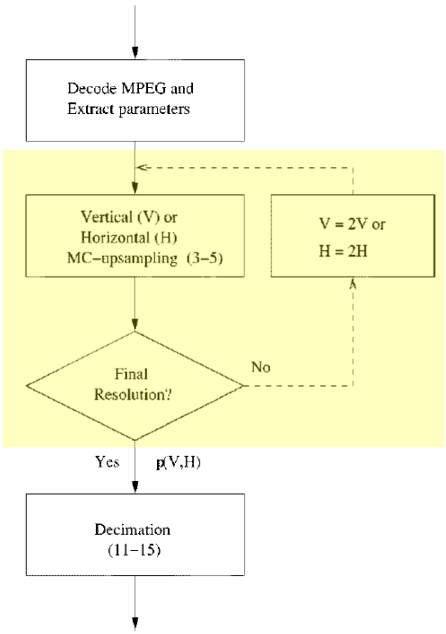


Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007, FIG. 3

G. Claim 6

1. [6.P]

Supra [1.P].

2. [6.1]

Phek's encoder encodes separate bitstreams for the basic- and enhancement-layers. Ex.1005, [0014], FIG. 4a. Phek's decoder processes the basic- and enhancement-layer bitstreams in separate paths. *Id.*, [0015], FIG. 4b. Phek's FIG. 2 describes "reverse multiplexer 202" receiving "input D252 [] composed of video and/or audio sources, etc." and separating "the video source...from the other sources."

Id., [0010], FIG. 2. Thus, a POSITA would have understood and it would have been obvious that, when implementing layered hierarchical coding, Phek's encoded video bitstream D258 from FIG. 2 would include both the basic- and enhancement-layer bitstreams shown in FIGs. 4a-4b. Ex.1003, ¶341. However, Phek does not expressly disclose that the basic- and enhancement-layer bitstreams are multiplexed. *Id.*

A POSITA would have known that scalable and hierarchical video bitstreams are commonly multiplexed on the encoder side into a composite signal for transmission. Ex.1003, ¶342; Ex.1009, FIG. 12; Ex.1017, [0079], FIG. 10; *supra*, §IV.A.5.c. For example, Segall teaches, "from a systems multiplex...perspective," "data might often be handled differently—as distinct hierarchically related streams of content that are coordinated using decoding timestamps or other such mechanisms." Ex.1006, 3.

In combination, Phek's encoding device (FIG. 4a) multiplexes respective bitstreams from the basic- and enhancement-layers into a single composite signal for transmission, along with video information (e.g., "***encoding parameters of respective types used in the encoding***"). Ex.1003, ¶343.

H. Claims 8-9 Limitations Correspondence to Claim 1 Limitations

[8.P]	[9.P]	*See* [1.P]
[8.1]	[9.1]	*See* [1.1]

[8.2]	[9.2]	*See* [1.2]
[8.3]	[9.3]	*See* [1.3]
[8.4]	[9.4]	*See* [1.4]
[8.5]	[9.5]	*See* [1.5]
[8.6.i]	[9.6.i]	*See* [1.6.i]
[8.6.ii]	[9.6.ii]	*See* [1.6.ii]
[8.6.iii]	[9.6.iii]	*See* [1.6.iii]
[8.6.iv]	[9.6.iv]	*See* [1.6.iv]
[8.7]	[9.7]	*See* [1.7]

I. Claim 10

Claim 10's limitations simply reflect the common practice of individually encoding/decoding different blocks or other regions of a picture in a stream, rather than pictures as a whole. *Supra*, §IV.A.5.e; Ex.1003, ¶¶366-367.

It would have been obvious to apply He's teachings regarding slice-level selective loop filtering in the combination so the second encoder (e.g., processors 404, 406, 408, 412) in the enhancement layer would divide the encoding target picture into a plurality of slices (regions) and select whether to use an SR-D filtered version or a non-SR-D filtered version of a reference picture to predict each divided

region. He teaches "[e]ach video frame is subdivided and encoded at the macroblock (MB) level, where each MB is a 16x16 block of pixels." Ex.1008, [0004], [0023], [0086]. "[E]ncoder 100 is typically configured to process information one macroblock at a time." *Id.*, [0023]. He also teaches that loop filter controls can be applied to slices that incorporate "one or more macroblocks," and the filter can be selectively enabled/disabled for each slice. *Id.*, [0053]-[0054]; *supra*, §IV.D. A POSITA would have sought to implement slices to reduce the coding unit's size of the coding unit and to allow for greater compression and improved prediction quality. *Supra*, §IV.A.5.e.

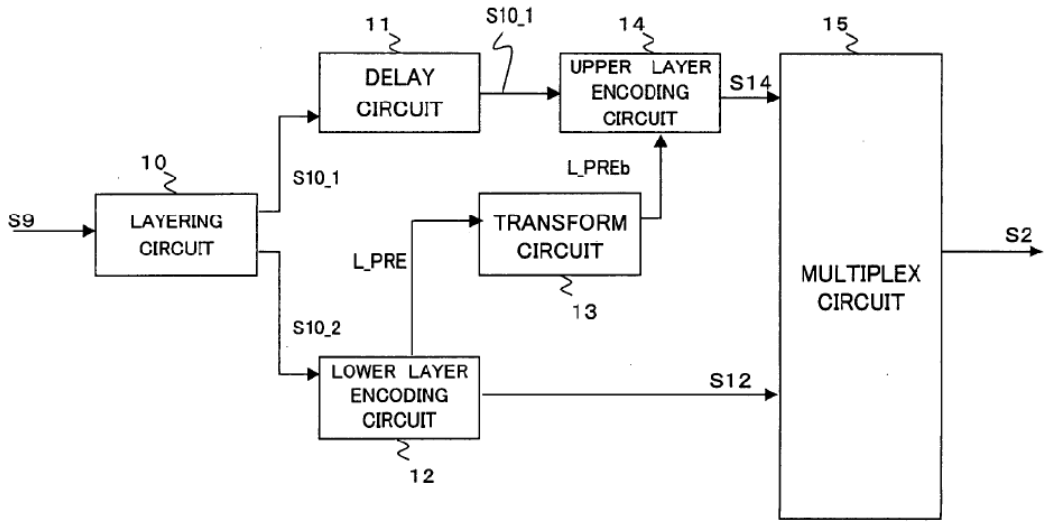
V. GROUND 2: Phek, Segall, Martins, He, and Wang Render Obvious Claim 7

A. Prior Art

1. Wang

Wang's encoding/decoding system processes "progressive image data" and "interlace image data" using layered encoding. Ex.1044, [0006], [0009]; Ex.1003, ¶¶371-374. Wang's Figure 2 encoding apparatus includes "layering circuit 10, delay circuit 11, lower layer encoding circuit 12, transform circuit 13, upper layer encoding circuit 14, and multiplex circuit 15." *Id.*, [0018]. Layering circuit 10 "generate[s] the progressive image data" and "the interlace image data" based on image data to be encoded. *Id.*, [0019]. Lower layer encoding circuit 12 "encode[s]

the interlace image data" received from the layering circuit 10 "to generate the lower layer encoded data S12 and output[s]" that data "to the multiplex circuit 15." *Id.*, [0021].

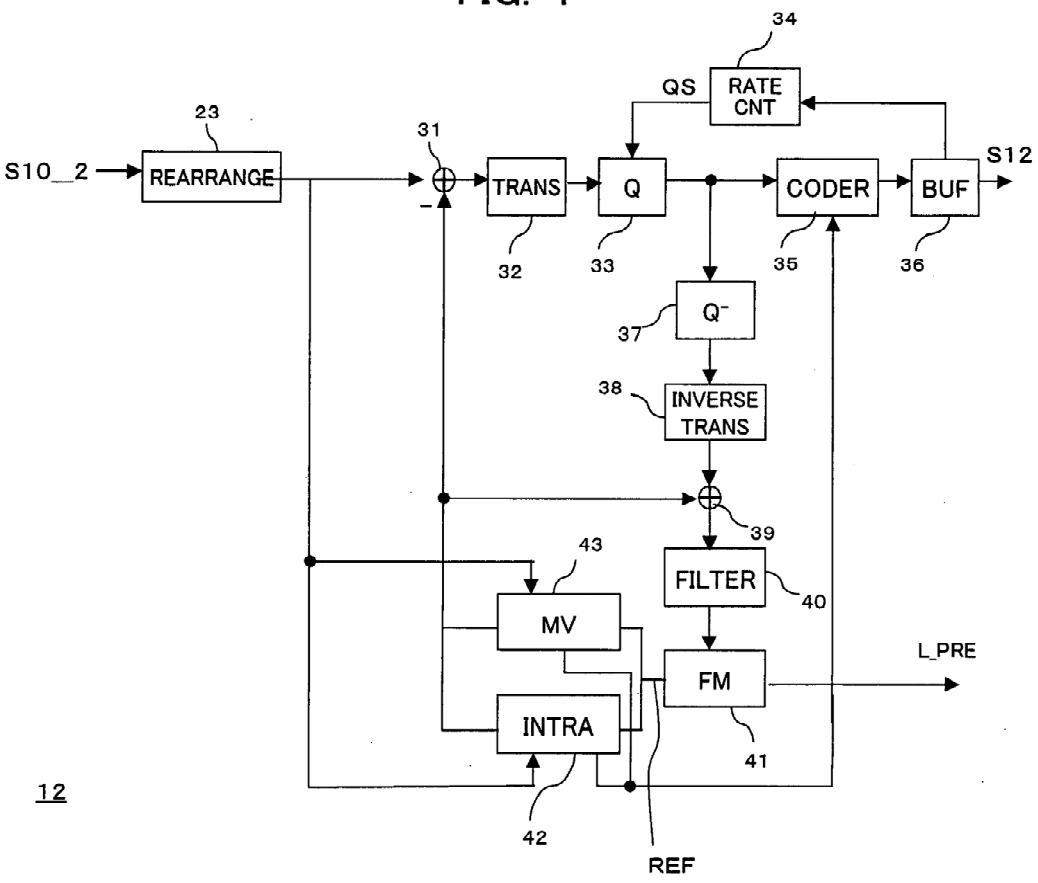


Ex.1044, FIG. 2.

Wang's Figure 4 shows "screen rearrangement circuit 23" which "rearrange[s] the progressive image data" received from layering circuit 10 "to a sequence of encoding in accordance with a GOP (Group Of Pictures) structure," which is output to the processing circuits ("processing circuit 31, intra-prediction circuit 42, and motion prediction and/or compensation circuit 43"). *Id.*, [0023]. A "quantization circuit 33" "quantize[s] the image data" "based on a quantization scale QS input from the rate control circuit 34 to generate image data indicating the transform coefficient after the quantization." *Id.*, [0024]. Wang explains the rate

control circuit 24 "generates the quantization scale" based on image data from the "buffer memory 36." *Id.* Figure 4 also shows "frame memory 41" which outputs predicted image data to "transform circuit 13" in Figure 2. *Id.* [0026].

FIG. 4



Ex.1044, FIG. 4

Wang also describes "transform circuit 13," which "interpolate[s] the lower predictive image data L_PRE" "from the lower layer encoding circuit 12" and "generate[s] the lower predictive image data L_PREb having the same resolution...as

that of the progressive image data S10_1" from "upper layer encoding circuit 14." *Id.*, [0028]. Resulting image data is output to "upper layer encoding circuit 14." *Id.*

2. Phek-Segall-Martins-He-Wang

Phek's layered coding system is ripe for improvement. *Supra* §IV.A.5. It would have been obvious to combine Phek's FIG. 4a encoding device with Wang, which prepares image data for encoding by generating progressive and interlace image data. Ex.1044, [0018]-[0024]. Phek lacks disclosures regarding picture storage between processing steps. A POSITA would have been motivated to incorporate teachings of Wang's system into the Phek-Segall-Martins-He combination to introduce a "buffer," including means for managing that buffer and pictures therein. Specifically, Wang discloses "layering circuit 10" that accumulates "image data S9 to be encoded." *Id.*, [0018]-[0019]; Ex.1003, ¶¶375-376.

A POSITA would have implemented Wang's rearrange circuit 23, prior to Phek's basic layer encoding processing device 402 to accumulate "pictures with the standard resolution" before encoding or resolution processing, consistent with a POSITA's knowledge (*see* Ex.1003, ¶378). In combination, a POSITA would have implemented Wang's "lower layer encoding circuit 12," including "frame memory 41," Ex.1044, [0021]-[0022], between Phek's basic layer encoding processing device 402 and Phek's "***second encoder***" (provided by Phek's devices 404, 406, 408,

and 412) to "***accumulate sets of decoded pictures from the first encoder***" during processing steps. Further, to manage these buffers and Phek's first/second encoders, a POSITA would have incorporated Wang's "quantization circuit 33" that processes image data based on "input from the rate control circuit 34," wherein the rate control circuit 34 determines processing scale "based on the image data read out from the buffer memory 36" (Ex.1044, [0024]), which involves controlling buffer sizes and adjusting encoding parameters. Ex.1003, ¶¶375-376. Accordingly, the Phek-Segall-Martins-He-Wang combination is nothing more than combining known elements according to known methods (e.g., incorporating a buffer as taught by Wang) with a reasonable expectation of success, given that buffers were well-known. Ex.1003, ¶¶375-376.

B. Claim 7

1. [7.P]

Supra [1.P]. Ex.1003, ¶375.

2. [7.1]

Phek's FIG.4a encoding device receives "[a] basic layer image sequence" for encoding by encoding processing device 402 ("***first encoder***"). Ex.1005, [0014]; *supra* §[1.1]. Martins discloses "motion-compensated spatial upsampling from multiple images" (Ex.1007, Abstract), which, in the combination, provides "***a first super-resolution enlarger***." *Supra* §[1.2]. In combination, Wang's "rearrange

circuit 23" is implemented to accumulate pictures "***with the standard resolution***" prior to encoding or picture processing. Ex.1003, ¶378; Ex.1044, [0018]-[0019]. For example, Wang's rearrange circuit 23 is "configured to rearrange the picture data of the progressive image data…." Ex.1044, [0022]. A POSITA would have understood that Wang's rearrange circuit 23 stores this unprocessed image data ("***subsequences of the sequence of moving pictures with the standard resolution***") in a buffer to provide said data to Phek's encoding processing device 402 and Martins' spatial upsampling process. Ex.1003, ¶378.

3. [7.2]

Phek's FIG.4 encoding device creates a decoded set of the basic-layer image sequence that is encoded with a first spatial resolution ("***sets of decoded pictures from the first encoder***"). *Supra*, §IV.E.2; Ex.1005, [0014]; Ex.1003, ¶¶306-308. That decoded set is acquired by Phek's second encoder. *Supra*, §IV.E.8. A POSITA would have recognized the need for a buffer between the first and second encoders to accumulate the decoded set of the basic-layer image sequence before it is received by Phek's second encoder. Ex.1003, ¶379. Thus, a POSITA would have incorporated Wang's "lower layer encoding circuit 12," including the "frame memory 41" therein, to accumulate the decoded set of the basic-layer image sequence. Ex.1003, ¶379; Ex.1044, [0021]-[0022]. Although Wang does not

describe "frame memory 41" in detail, a POSITA would understand, in view of the myriad disclosures of buffers prior to the Critical Date, that Wang's frame memory 41 may be used to temporarily store sets of decoded pictures received from Phek's encoding processing device 402 until it is determined that those sets of decoded pictures need be provided to Phek's second encoder for further processing. Ex.1003, ¶379.

4. [7.3]

To manage picture data flow between Wang's first and second accumulation buffer and Phek's first and second encoders, a POSITA would have incorporated Wang's "delay circuit 11" to "***detect[]***" and "***control***" "***buffer accumulation amounts***." Ex.1003, ¶380. The "delay circuit 11" is "configured to delay the picture data…by…exactly the processing time in the lower layer encoding circuit 12 and transform circuit 13 and output the result to the upper layer encoding circuit 14." Ex.1044, [0020]. Thus, the delay circuit synchronizes frames between upper layer encoding circuit 14 and lower layer encoding circuit 12, and accordingly Wang's delay circuit teaches controlling buffer accumulation amounts. Ex.1003, ¶¶380-381.

5. [7.4]

Wang's "rate control circuit 34" "generate[s] the quantization scale QS based on the image data read out from the buffer memory 36 and outputs this to the

quantization circuit 33." Ex.1044, [0024]. Rate control circuit 34 informs the combination's first and second encoder regarding the proper code rate "based on the image data read out from the buffer memory" ("***control[s] code rates at the first encoder and the second encoder***"). Ex.1003, ¶382.

VI. THE BOARD SHOULD REACH THE MERITS

A. *Advanced Bionics* Favors Institution

This Petition presents new art and arguments; the file history is silent regarding any previous consideration of the presented combination in evaluating the '995 patent. Segall, Martins, He, and Wang were never considered by the examiner, and thus the Petition's obviousness grounds were not considered and are not cumulative of any art and arguments presented during examination.

Although the Office previously considered Ground 1's teachings in instituting IPR2024-00372, an institution decision is not a final patentability determination. Further, the Petition simply requests review of the '995 patent under much of the same evidence and argument the Board already found to meet the institution threshold. Exercising discretion under §325(d) is unwarranted.

B. *General Plastic* Favors Institution

Petitioner has not previously filed a petition directed to this patent. Samsung Electronics Co., Ltd. ("Samsung") previously filed an IPR petition challenging this patent, but Petitioner is neither "the same party" as Samsung, "nor possess[es] a

significant relationship" with Samsung, and thus, "under *Valve*, *General Plastic* factor one necessarily outweighs the other *General Plastic* factors." *Vivendum Prod. Sols., Inc. v. Rotolight Ltd.*, IPR2023-01218, Paper 12 at 6 (PTAB Apr. 19, 2024). Petitioner "had no involvement with, much less input into, [Samsung's] IPR." *Ford Motor Co. v. Neo Wireless LLC*, IPR2023-00763, Paper 28 at 5 (PTAB Mar. 22, 2024). Petitioner and Samsung do not have a "significant relationship" and are defendants in "different district court proceedings" involving "different allegedly infringing products." *Id.* at 9-10.

C. *Fintiv* Favors Institution

The district court litigation is at an early stage. A claim construction hearing will not occur until September 16, 2025, and expert discovery will not close until November 18, 2025. Ex.1045, 4. Trial is not scheduled to begin until March 2, 2026. Ex.1045, 1. The Petition challenges all claims of the '995 patent, while the district court litigation will address only claims 1-6 and 8-11; thus, this Petition will address the patentability of claims (claims 5-7) that the district court will not address, reducing any overlap. Finally, the merits are strong; the Board already instituted as to certain challenged claims on similar evidence. *See Samsung Elecs. Co., Ltd. v. Advanced Coding Techs. LLC*, IPR2024-00372, Paper 9 (PTAB Jul. 18, 2024).

Petitioner reserves the right to file an opposition to Patent Owner's discretionary arguments submitted pursuant to the Interim Processes for PTAB Workload Management.

VII. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8(a)(1)

A. Real Party-In-Interest

Google LLC is the petitioner and real party-in-interest.[13]

B. Related Matters

Advanced Coding Technologies LLC v. Google LLC, No. 2:24-cv-00353-JRG (E.D. Tex.).

Advanced Coding Technologies LLC v. Apple Inc., No. 2:24-cv-00572-JRG (E.D. Tex.).

Apple Inc. v. Advanced Coding Technologies LLC, IPR2025-00984

Petitioner is not aware of any other disclaimers, reexamination certificates, or IPR petitions addressing the '995 patent.

[13] Google LLC is a subsidiary of XXVI Holdings Inc., which is a subsidiary of Alphabet Inc. XXVI Holdings Inc. and Alphabet Inc. are not real parties-in-interest to this proceeding.

C. Grounds for Standing

Petitioner certifies that the '995 patent is available for IPR and Petitioner is not barred or estopped from requesting same.

D. Counsel and Service Information

Pursuant to 37 C.F.R. §§ 42.8(b)(3), 42.8(b)(4), and 42.10(a), Petitioner designates the following lead counsel:

- Raghav Bajaj (Reg. No. 66,630), raghav.bajaj@lw.com, Latham & Watkins LLP; 300 Colorado Ave. Suite 2400; Austin TX 78701; 737.910.7370.

Petitioner also designates the following backup counsel:

- Doug Lumish (*pro hac vice* motion to be filed), doug.lumish@lw.com, Latham & Watkins LLP; 140 Scott Drive; Menlo Park, CA 94025; 650.463.2633.
- Patricia Young (*pro hac vice* motion to be filed), patricia.young@lw.com, Latham & Watkins LLP; 1271 Avenue of the Americas; New York, NY 10020; 212.906.1661.
- Linfong Tzeng (Reg. No. 65,140), linfong.tzeng@lw.com, Latham & Watkins LLP; 140 Scott Drive; Menlo Park, CA 94025; 650.463.3092

- Joseph H. Lee (*pro hac vice* motion to be filed), joseph.lee@lw.com, Latham & Watkins LLP; 650 Town Center Drive, 20th Floor; Costa Mesa, CA 92626; 714.755.8046.

Pursuant to 37 C.F.R. § 42.10(b), a Power of Attorney from Petitioner is attached. Petitioner consents to electronic service.

E. IPR Fee

The Director is authorized to charge the fee specified by 37 C.F.R. § 42.15(a) to Deposit Account No. 506269.

VIII. CONCLUSION

For the reasons set forth above, Petitioner Google respectfully requests *inter partes* review of the Challenged Claims of the '995 patent.

Respectfully submitted,

Dated: May 12, 2025

By: / Raghav Bajaj /

Raghav Bajaj (Reg. No. 66,630)
raghav.bajaj@lw.com
Latham & Watkins LLP
300 Colorado Ave. Suite 2400
Austin, TX 78701
Telephone: 737.910.7370
Fax: 737.910.7301

Counsel for Petitioner
Google LLC

CERTIFICATE OF COMPLIANCE WITH 37 C.F.R. § 42.24

I hereby certify that this Petition complies with the word count limitation of 37 C.F.R. § 42.24(a)(1)(i) because the Petition contains a total of 13,999 words, as calculated by Microsoft Word's word-count feature. This total excludes the cover page, signature block, and the parts of the Petition exempted by 37 C.F.R. § 42.24(a)(1).

By: / Raghav Bajaj /

Raghav Bajaj (Reg. No. 66,630)
raghav.bajaj@lw.com
Latham & Watkins LLP
300 Colorado Ave. Suite 2400
Austin, TX 78701
Telephone: 737.910.7370
Fax: 737.910.7301

Counsel for Petitioner
Google LLC

CERTIFICATE OF SERVICE

The undersigned certifies that a complete copy of this Petition for *Inter Partes* Review of U.S. Patent No. 10,218,995 and all Exhibits and other documents filed together with this Petition were served on the official correspondence address for the patent shown in Patent Center:

Nath, Goldberg & Meyer
112 S. West Street
Alexandria, VA 22314

via FEDERAL EXPRESS next business day delivery, on May 12, 2025.

By: / Raghav Bajaj /

Raghav Bajaj (Reg. No. 66,630)
raghav.bajaj@lw.com
Latham & Watkins LLP
300 Colorado Ave. Suite 2400
Austin, TX 78701
Telephone: 737.910.7370
Fax: 737.910.7301

Counsel for Petitioner
Google LLC